Proteomics and Transcriptomics of the Betz cell in the Human Brain.



Connor Scott

Somerville College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy - Clinical Neurosciences

Hilary 2021

This thesis is dedicated entirely to my mother,
Jacqueline Scott,
whose work-ethic, determination, and benevolent nature inspires me every day.

I will be forever thankful for the sacrifices you have made for me.

Acknowledgements

First and foremost, I am extremely grateful to my supervisors, Prof. Olaf Ansorge and Prof. Roman Fischer for their invaluable advice, continuous support, and patience during my D.Phil. I am particularly thankful to Prof. Ansorge for taking the time to educate me over the years on the perplexing nature of the human brain and giving me the opportunities to develop my skills as a scientist. I would like to thank Prof. Zoltán Molnár for taking the time to be an examiner for both my Transfer of Status and Confirmation of Status. Prof. Zoltán Molnár's insights into subpopulations of L5 pyramidal neurons, L5 biomarkers, and his thoughts on the Betz cell have significantly impacted my thinking on the nature of the Betz cell. I would also like to thank Prof. Benedikt Kessler for facilitating and encouraging LCM proteomic research, and for helpful advice during my Transfer and Confirmation of Status. Additional thanks are given to Prof. Benedikt Kessler for also taking the time and agreeing to be my examiner for my D.Phil *viva voce*. Similarly, I would like to extend my gratitude to Prof. Patrick Hof for also taking the time to be my examiner for my D.Phil *viva voce*. I would also like to state my appreciation to all the individuals in John Radcliffe's Neuropathology department, with special thanks to Carolyn Sloan, Hannah Brooks, and Matthew Nolan for helping me over the years. A special thank you goes to Simon Davis who has given crucial advice and expertise on the practicalities of LC-MS/MS and proteomics; and whose input has been fundamental to the progression of this D.Phil thesis. In addition to this, Simon Davis has also been critical to the development of the LCM LC-MS/MS workflow outlined in Chapter 4 of this thesis along with the development of the high-throughput LC-MS/MS approach used in Chapter 6 and described in Appendix E. Footnotes have been placed in the relevant areas of this thesis to highlight his contributions.

I would like to extend my deep appreciation to Somerville College, especially the Middle Common Room community. It would not be possible to name everyone whose company and conversations have inspired me during my time at Oxford and I am fortunate enough to be surrounded by many people who have greatly supplemented my studies, formed life-long memories with; and provided a refreshing break from stresses of doctoral research. In particular, I would like to thank Madison Eisler for her continued support and encouragement during the writing of this thesis,

and to my mother, Jacqueline Scott for distilling the importance of education and self-growth, and for always believing in me, no matter the circumstances.

Finally, I would like to issue my deepest condolences to the families and immeasurable gratitude to all the individuals who have donated their tissue for research. Without them, many projects like my own would not be possible and without this invaluable resource, we would not be able to learn about the intricacies of the human brain or understand the pathophysiology of neurodegenerative disease, in the hopes of finding a therapeutic for these disorders.

Abstract

Very little is known about the enigmatic gigantopyramidal neuron in the Primary Motor Cortex. These Betz cells are the largest neurons in the human brain, display unique morphology, play a fundamental role in voluntary movement, and are selectively vulnerable to a range of neurodegenerative diseases, including Amyotrophic Lateral Sclerosis (ALS). Yet, there has been no targeted studies to investigate the molecular profile of these neurons and how that may be affected in disease. Here, for the first time, laser microdissected human Betz cells were subjected to extensive proteomic analysis via a workflow that has been optimised to produce leading proteomic performance from individually pooled neurons. The proteome of the Betz cell has been comparatively analysed to similar neurons to reveal key insights into their physiology. This thesis molecularly confirms previous assumptions about their biology and discovers cellular mechanisms that are enriched within these neurons. To complement proteomic findings, laser microdissected Betz cells were also subjected to RNA-sequencing to investigate their gene expression patterns, to validate recently generated proteomic observations, and for the first time, to reveal the transcriptome of the Betz cell. To ease and improve future Betz cell research, the proteomic/transcriptomic data were screened to search for Betz cells biomarkers – for which there are currently none that immunochemically target Betz cells without also binding to surrounding neurons. Several proteins were identified that could selectively target Betz cells. Finally, proteomic/transcriptomic analysis was also applied to microdissected Betz cells from ALS patients to determine how/why these neurons degenerate in ALS. The findings herein, align with the excitotoxicity hypothesis of ALS, i.e. that upper motor neurons become over-activated by excess glutamate activity. These data also provide clues to other pathways that may cause the degeneration of Betz cells in ALS e.g. DNA damage, oxidative stress, dysregulated energy production, and impaired axonal guidance.

Contents

List of Figures **xv**

List of Abbreviations **xxi**

1 Introduction **1**
1.1 Motivation . . . 1
1.1.1 Brief Introduction to Betz Cells . . . 1
1.1.2 Why Focus on Betz Cells? . . . 2
1.2 Objectives . . . 4
1.3 Outline of Thesis . . . 5

2 Background **7**
2.1 Neurocytology . . . 8
2.1.1 Neuronal Diversity . . . 9
2.1.2 Classifying Neurons . . . 9
2.2 Technological Advances in Neurocytology . . . 16
2.2.1 Biomarkers . . . 16
2.2.2 Transcriptomics of the Brain . . . 17
2.2.3 Proteomics of the Brain . . . 20
2.2.4 Laser Capture Microdissection . . . 21
2.2.5 LCM Transcriptomics . . . 23
2.2.6 LCM Proteomics . . . 27
2.3 Betz cells: Magnocellular Neurons of Motor Control . . . 29
2.3.1 The Primary Motor Cortex . . . 29
2.3.2 Corticospinal Tract . . . 29
2.3.3 Microcircuitry of the Adult Primary Motor Cortex . . . 32
2.3.4 The Cellular Diversity of the Primary Motor Cortex . . . 35
2.3.5 Layer 5 of the Primary Motor Cortex . . . 36
2.3.6 Betz Cells . . . 40
2.4 Amyotrophic Lateral Sclerosis . . . 54
2.4.1 Betz Cell Pathology in ALS . . . 54
2.5 Aims of this Thesis . . . 59

3 Methodology **61**
3.1 Brain Dissection 62
3.1.1 Fresh Brain Dissection & Tissue Processing 62
3.1.2 Tissue Processing for FFPE Material 63
3.2 Sectioning 64
3.2.1 FFPE Tissue for Immunohistochemical/Histological Staining 64
3.2.2 Frozen Tissue for Immunohistochemical/Histological Staining 64
3.2.3 Sectioning for Laser Capture Microdissection 64
3.3 Histological Staining 65
3.3.1 Cresyl Violet Staining 65
3.3.2 Haematoxylin & Eosin Staining 66
3.4 Immunohistochemistry 66
3.4.1 Antibody Selection 66
3.4.2 Antibody Validation and Optimisation 67
3.4.3 Immunohistochemistry on FFPE Sections 68
3.4.4 Immunohistochemistry on Frozen Sections 68
3.5 Laser Capture Microdissection 69
3.5.1 Purkinje Cells 69
3.5.2 Molecular Layer 70
3.5.3 Granule Cells Area 70
3.5.4 Betz Cells 70
3.5.5 Cerebellar Dentate Nucleus Neuron (DNN) 71
3.6 Mass Spectrometry - Sample Preparation 71
3.6.1 Standard Sample Preparation / In-Solution Digestion 71
3.6.2 In-Cap Digestion 72
3.6.3 SP3 Method 72
3.6.4 RWS approach (RIPA Buffer/Wet cap/Custom SP3) 73
3.7 Mass Spectrometry 74
3.7.1 LC-MS/MS (Lumos Orbitrap) - Purkinje Cell Titration . . . 74
3.7.2 LC-MS/MS (Lumos Orbitrap) - For Subsequent Experiments 74
3.7.3 LC-MS/MS (TimsTOF) - Evo-sep (Standard High-Throughput Approach) 75
3.7.4 LC-MS/MS - Dionex Nanoflow Coupled with TimsTOF (Optimised high-throughput approach) 75
3.7.5 MaxQuant Settings 76
3.7.6 Fragpipe / MSFragger 76
3.7.7 Proteomic Quality Control 77
3.7.8 Data Analysis 77
3.8 Digital Pathology 78

3.8.1 Whole Slide Imaging . . . 78
3.8.2 Image Analysis . . . 78
3.9 RT-qPCR . . . 79
3.9.1 Primer Selection . . . 79
3.9.2 Primer Design . . . 79
3.9.3 Primer Optimisation . . . 80
3.10 LCM-seq . . . 80
3.10.1 Preparing and Validating SPRI Beads . . . 80
3.10.2 Library Construction . . . 81
3.10.3 RNA-Sequencing . . . 83
3.10.4 Data Pre-processing and Data Analysis . . . 84

4 Developing a sensitive, scalable method for spatial, cell-type-resolved microproteomics of the human brain. 85
4.1 Aims of this Chapter . . . 86
4.2 Introduction . . . 86
4.2.1 Breakdown of this Chapter . . . 87
4.3 Results . . . 88
4.3.1 LCM LC-MS/MS can Identify Proteins from Minute Amounts of Post-Mortem Human Brain Tissue . . . 88
4.3.2 LCM LC-MS/MS of Post-Mortem Purkinje Cells successfully Identified Several Proteins relating to Purkinje Cell Activity 92
4.3.3 Partial Recovery of Protein leads to a Marked Decrease in Protein Identifications . . . 96
4.3.4 Sample Preparation Dramatically affects the Number of Identifiable/Quantifiable Proteins . . . 99
4.3.5 Effect of Histology Staining on LCM Proteomics . . . 102
4.3.6 Effect of immunohistochemistry on LCM Proteomics . . . 106
4.4 Discussion . . . 110
4.4.1 Main Conclusions . . . 118

5 Explorative LCM Proteomics of the Human Betz cell 119
5.1 Aims of this Chapter . . . 119
5.2 Introduction . . . 120
5.2.1 Breakdown of this Chapter . . . 121
5.3 Results . . . 122
5.3.1 Building a Suitable Betz Cell Research Cohort . . . 122
5.3.2 Betz Cells and Purkinje Cells - Initial Data . . . 125
5.3.3 Relative Proteomics - Betz Cells vs. Purkinje Cells . . . 125
5.3.4 Digital Pathology . . . 132

5.3.5 LCM Proteomic-Directed Immunohistochemistry 137
5.3.6 Comparative LCM Proteomics of the Betz Cell, L5 Motor Neuron, Purkinje Cell, and Cerebellar DNN 151
5.3.7 The Search for Preliminary Betz Cell Biomarkers 165
5.4 Discussion 180
5.4.1 Main Conclusions 191

6 LCM Proteomics and Transcriptomics of the Human Betz cell in the healthy and ALS brain. 193
6.1 Aims of this Chapter 193
6.2 Introduction 194
6.2.1 Breakdown of this Chapter 194
6.3 Results 197
6.3.1 Developing a High Throughput Micro-Proteomic Workflow . 197
6.3.2 Case Selection for LCM Proteomics and LCM-seq 197
6.3.3 LCM Proteomics of the Human Betz cell 198
6.3.4 LCM-seq 203
6.3.5 LCM-seq Analysis of the Human Betz cell 208
6.3.6 LCM-seq Analysis of ALS Neurons 219
6.4 Discussion 225
6.4.1 The Betz Cell in the Healthy Brain. 225
6.4.2 The Betz Cell in the ALS Brain. 227
6.4.3 Main Conclusions 234

7 Conclusions 235
7.1 Achieved Objectives 235
7.2 The Human Betz cell 236
7.3 Limitations of this thesis 237
7.4 Future Work 239
7.4.1 Data from this Thesis 239
7.4.2 Future Betz Cell Research 240
7.5 Final Thoughts 242

Appendices

A Accurately identifying a Betz cell 247
A.1 Accurately Identifying a Betz Cell - 247

B Additional information about the Betz cell **253**
B.1 Comparative Neuroanatomy of Betz Cells . . . 253
B.2 Electrophysiology of Betz Cells . . . 258
B.3 Molecular Composition of Betz Cells . . . 260
B.4 Motor Cortex / Betz Cells in Development . . . 262
B.5 Betz Cells in Ageing and Disease - . . . 265
B.5.1 TDP-43 Pathology in Betz Cells - . . . 268

C Chapter 4 - Supplementary Data **271**
C.1 Reproducibility of LCM Proteomics . . . 271
C.2 Comparing Buffers . . . 274
C.3 LCM Proteomics of FFPE Purkinje Cells . . . 274
C.4 LCM Proteomic Comparison: Frozen & FFPE Granule Cells . . . 276
C.5 Effect of Staining on LCM Proteomics . . . 285

D Chapter 5 - Supplementary Data **291**
D.1 Functional Classification . . . 292
D.2 Absolute Proteomics - Betz Cells . . . 294
D.3 Absolute Quantification - Betz Cells, Purkinje Cells, DNNs, and L5 Motor Neurons . . . 299
D.4 Relative Proteomics - Betz Cells, L5 Neurons, DNNs, and Purkinje Cells. . . . 302
D.5 Betz Cell Comparisons - 1D Annotation Enrichment . . . 310
D.6 Additional IHC Data . . . 316
D.7 Excessive Contamination of Neuropil Proteins in L5 Neuronal Samples 319

E Chapter 6 - Supplementary Data **321**
E.1 Developing a High-Throughput LCM Proteomic Workflow . . . 321
E.2 LCM-seq Optimisations . . . 327
E.3 LCM-seq Quality Control . . . 333
E.4 LCM-seq Analysis . . . 338
E.4.1 Healthy Betz Cells vs. Healthy Purkinje Cells . . . 338
E.4.2 Healthy Betz Cells vs. Healthy L5 Neurons . . . 344
E.4.3 Healthy Betz Cells vs. ALS Betz Cells . . . 351
E.5 Microdissected Betz Cells / Allen Institute Clusters - Comparison . 360

F Betz-Purkinje Research Cohort **365**
F.1 Full Post-Mortem Cohort used in this Thesis . . . 365
F.2 Additional Information on Cohort . . . 368

Bibliography **371**

List of Figures

2.1 A Pyramidal Neuron Stained with the Golgi Stain. 10
2.2 Diversity of Neuronal Morphology. 11
2.3 Example of the Allen Institute's Transcriptomic Clusters Taxonomy. 14
2.4 Laser Capture Microdissection (LCM). 24
2.5 The Corticospinal Tract. 31
2.6 The Assumed Microcircuit of the Human Primary Motor Cortex. . 34
2.7 The Cortical Layers in the Human Primary Motor Cortex. 38
2.8 Assumed Molecular Classification of L5 Pyramidal Neurons. 39
2.9 Betz Cell Morphology differs from other Pyramidal Neurons. 43
2.10 Overview of Betz Cell Morphology. 46
2.11 The Varying Morphology of Betz Cells across the Primary Motor Cortex. 49
2.12 Structural Degeneration of Betz Cells in ALS. 57
2.13 Reactive Gliosis is a Betz Cell-Specific ALS Pathology. 58

4.1 Systematic for the LCM LC-MS/MS Workflow Implemented in this Thesis. 88
4.2 Purkinje Cell LCM - Before/after/in-cap. 90
4.3 Purkinje Cell Titration - Soma Size. 91
4.4 Purkinje Cell Titration - No. of Identifiable Peptides. 92
4.5 Purkinje Cell-Associated Proteins within Purkinje Cell Titration. . 93
4.6 Validation of Purkinje Cell Protein IDs using the Human Protein Atlas. 95
4.7 Our LCM LC-MS/MS Workflow Enriches for Cell-Type-Specific Proteins. 97
4.8 Purkinje Cell Peptides/Proteins - *In Situ* Digestion Method. 98
4.9 Permutations for Optimising LCM LC-MS/MS Sample Preparation. 100
4.10 Optimising LCM LC-MS/MS Sample Preparation. 101
4.11 Effect of Different Granule Cell Staining on Number of Identified Proteins. 104
4.12 Effect of Different Granule Cell Staining on Protein Abundance. . . 105
4.13 Effect of IHC on No. of Identified Proteins. 107

4.14 Volcano Plot - IHC vs. Nissl Stained Purkinje Cells. 109
4.15 Comparison of LCM Proteomic Approaches Reported in the Literature. 113

5.1 Proposed LCM Approach for this Chapter. 122
5.2 Preparing Fresh Post-Mortem Brains to Access Entire Betz Cell Population. 124
5.3 Representation of Microdissecting Betz cells - Before/After/In-cap. 126
5.4 Betz Cells vs. Purkinje Cells - No. of Identifiable Proteins. 127
5.5 Volcano plot - Betz Cells vs. Purkinje Cells. 127
5.6 Enriched Gene Ontology Terms in Betz cells. 130
5.7 1D Annotated Volcano plot - Betz Cells vs. Purkinje Cells. 131
5.8 Validating LC-MS/MS data - Immunostaining for ITPR1. 134
5.9 Validating LC-MS/MS data - Immunostaining for ARHGAP26. . . 135
5.10 Quantitative IHC can Validate LC-MS/MS Findings. 136
5.11 AE-1 Staining of the Primary Motor Cortex. 138
5.12 ATPB - Validating Proteomic Data. 140
5.13 ATPB IHC - Cerebellum. 141
5.14 ATPB - Quantitative IHC. 141
5.15 CYCS - Validating Proteomic Data and Quantitative IHC. 143
5.16 VDAC1 - Validating Proteomic Data and Quantitative IHC. 145
5.17 DST (Low Magnification - Motor Cortex) - Validating Proteomic Data. 147
5.18 DST (High Magnification - Motor Cortex) - Validating Proteomic Data. 147
5.19 DST (Cerebellum) - Validating Proteomic Data and Quantitative IHC. 148
5.20 ESYT1 - Validating Proteomic Data and Quantitative IHC. 150
5.21 Betz, Purkinje, L5, and DNNs - No. of Identifiable Proteins. 152
5.22 Venn Diagram of Proteins Identified in the 4-Cell Type Comparison. 152
5.23 Hierarchical Clustering Revealed Ten Unique Clusters. 155
5.24 Enriched GO Terms in the Betz Cell Cluster (Cluster 4). 159
5.25 Enriched GO Terms in the Betz Cell/DNNs Cluster (Cluster 3). . . 160
5.26 Enriched GO Terms in the DNNs Cluster (Cluster 6). 161
5.27 1D Annotated Volcano plot - Betz Cells vs. DNNs. 162
5.28 1D Annotated Volcano Plot - Betz Cells vs. Purkinje Cells. 163
5.29 1D Annotated Volcano plot - Betz Cells vs. L5 Neurons. 164
5.30 Mosaic Pattern of CRYM in Betz Cells. 166
5.31 Mosaic Pattern of CRYM in Betz Cells - More Examples. 167
5.32 REEP1 Selectively Stains for Betz Cells. 169
5.33 REEP1 Selectively Stains for Betz Cells Cont. 170
5.34 KIF1A Highlights Betz Cell Morphology. 172
5.35 CAPRIN2 is Enriched within Betz Cells. 174

5.36 Transcripton Factor OTX1 Selectively Targets Betz Cells. 176
5.37 Transcripton Factor HDAC1 Highlights Betz Cell Morphology. . . . 178
5.38 Transcripton Factor HDAC1 Highlights Betz Cell Morphology Cont. 179

6.1 LCM *"-omics"* of Healthy and Diseased Neurons. 196
6.2 Summary of Cells Microdissected for Large-Scale LCM Proteomics. 199
6.3 Large-Scale LCM Proteomics - No. of Identifiable Proteins per Sample. 200
6.4 Pearson Correlation Coefficient Heatmap - Large LCM Proteomic Study. 201
6.5 Manhattan Plot: Summary of DEPs in Betz Cells. 202
6.6 Summary of Cells Microdissected for LCM-seq Analysis. 204
6.7 Principal Component Analysis - LCM-seq. 205
6.8 MuSiC Results of LCM-seq Cohort - Outliers Removed. 206
6.9 LCM-seq - Sample-to-Sample Distance Heatmap. 207
6.10 Volcano Plot - Betz Cells vs. Purkinje Cells (LCM-seq). 209
6.11 mRNA Expression of Previously Identified Biomarkers. 211
6.12 Volcano Plot - Betz Cells vs. L5 Neurons (LCM-seq) 216
6.13 Differentially Expressed Genes from All Cell-Type Comparisons. . . 217
6.14 Treemap of the 754 Betz Cell-Enriched Transcripts. 218
6.15 Volcano Plot - Healthy Betz Cells vs. ALS Betz Cells. 220
6.16 Volcano Plot - Healthy L5 Neurons vs. ALS L5 Neurons. 221
6.17 Volcano Plot - Healthy Purkinje Cells vs. ALS Purkinje Cells. . . . 223
6.18 Differentially Expressed Genes found in Multiple Cell-Types. 224

A.1 Basic Stains for Identifying Betz Cells. 248
A.2 Immuno-labelling for Identifying Betz Cells. 249
A.3 Neurofilament Heavy Highlights Betz Cell Morphology. 250

B.1 Large Layer V Neurons in the Developing Primary Motor Cortex. . 264

C.1 No difference in LC-MS/MS performance was seen in longer LCM experiments. 273
C.2 Pearson Correlation Coefficient of RIPA and TFE. 274
C.3 FFPE Purkinje Cell Titration - No. of Identifiable Proteins. 275
C.4 Purkinje Cell-Related Proteins within FFPE Purkinje Cell Titration. 276
C.5 FFPE vs. Frozen Material - No. of Identified Proteins. 277
C.6 Pearson Correlation Coefficient of FFPE vs. Frozen material. 278
C.7 Overlap in Proteins between FFPE vs. Frozen Samples. 279
C.8 Histone Proteins were Reduced in FFPE Preparations for LCM Proteomics. 280
C.9 Volcano Plot of FFPE vs. Frozen Material. 281

C.10 Overall Abundance of Proteins in FFPE vs. Frozen Granule Cells. . 282
C.11 Pearson Correlation Coefficient Heatmap - Different Staining Methods. 285
C.12 Principal Component Analysis (PCA) - Different Staining Methods. 286
C.13 Hierarchical Clustering - Different Staining Methods. 287
C.14 Pearson Correlation Coefficient Heatmap - IHC vs. Nissl. 288
C.15 Venn Diagram - Protein Overlap (IHC vs. Nissl). 289

D.1 Venn Diagram - Betz Cells vs. Purkinje Cells. 292
D.2 Proteomap of Proteins found in Betz Cells. 293
D.3 1D Annotation Enrichment of Betz Cell iBAQ Values. 295
D.4 2D Annotation Enrichment of Betz/Purkinje cell iBAQ Values. . . . 296
D.5 Oxidative Phosphorylation- and Endoplasmic Reticulum-Related Proteins are among the most abundant Proteins in Betz Cells. . . . 297
D.6 ATP-, NADH-, TCA-, and Metabolism-Related Proteins are among the most abundant Proteins in Betz Cells. 298
D.7 2D Annotation: Betz Cells vs. L5 Neurons, DNNs, and Purkinje Cells. 300
D.8 Multi-scatter Graph - LCM Proteomics of 4 Cell-Types. 301
D.9 PCA - Betz Cells, L5 Neurons, DNNs, and Purkinje Cells. 302
D.10 Hierarchical Clustering - Betz, L5, Purkinje and DNNs. 303
D.11 Protein-Protein Interaction Network of 116 Betz DEPs from Cluster 4. 304
D.12 1D Annotation Enrichment - Betz Cells vs. Purkinje Cells (BP). . . 310
D.13 1D Annotation Enrichment - Betz Cells vs. Purkinje Cells (MF & CC). 311
D.14 1D Annotation Enrichment - Betz Cells vs. DNNs (BP). 312
D.15 1D Annotation Enrichment - Betz Cells vs. DNNs (MF & CC). . . 313
D.16 1D Annotation Enrichment - Betz Cells vs. L5 Neurons (BP). . . . 314
D.17 1D Annotation Enrichment - Betz Cells vs. L5 Neurons (MF & CC). 315
D.18 ARHGAP26 Staining of the Primary Motor Cortex. 316
D.19 ESYT1 Staining of the Cerebellum. 317
D.20 ER81 Staining of the Primary Motor Cortex. 318
D.21 Manhattan Plot: Summary of Proteins Up-Regulated in L5 Neurons. 319

E.1 Lumos / TimsTOF comparison. 323
E.2 TimsTOF Optimisation. 324
E.3 TimsTOF Optimisation 2. 326
E.4 RIN changes depending on the size of the Material Microdissected. . 328
E.5 Purkinje Cell Titration - cDNA Yield from LCM-seq. 330
E.6 Optimising Lysis buffer - cDNA Yield from LCM-seq. 331
E.7 cDNA Yield Before and After Optimisation. 332
E.8 PCA - LCM-seq Results (with Outliers). 333
E.9 Correlation Matrix of LCM Samples. 334

E.10 MuSiC Results of LCM-seq Cohort - Outliers Only. 335
E.11 Proportion of All RNA Species Identified. 335
E.12 Venn Diagram - All Protein-Coding Genes in Data-Set. 336
E.13 Top-Weighted PCA Transcripts. 337
E.14 REViGO Output - Betz Cells vs Purkinje Cells - MF. 338
E.15 REViGO Output - Betz Cells vs Purkinje Cells - BP. 339
E.16 REViGO Output - Betz Cells vs Purkinje Cells - CC. 340
E.17 Betz Cells vs. Purkinje Cells LCM-seq Comparison: Top 60 Betz DEGs. 341
E.18 The Top GO Enriched Pathways from GSEA - Betz Cells vs. Purkinje Cells. 342
E.19 The Top GO Enriched Pathways from GSEA - Betz Cells vs. Purkinje Cells. 343
E.20 Betz Cells vs. L5 Neurons Comparison: Top 60 Betz DEGs. 344
E.21 Top Enriched GOBP Terms - Betz Cells vs. L5 Neurons. 345
E.22 Top Enriched KEGG Terms - Betz Cells vs. L5 Neurons. 346
E.23 Top Enriched Reactome Terms - Betz Cells vs. L5 Neurons. 347
E.24 Top Enriched WIKI Terms - Betz Cells vs. L5 Neurons. 348
E.25 Top Enriched GOMF Terms - Healthy Betz Cells vs. ALS Betz Cells. 351
E.26 Top Enriched GOBP Terms - Healthy Betz Cells vs. ALS Betz Cells. 352
E.27 Top Enriched CC Terms - Healthy Betz Cells vs. ALS Betz Cells. . 353
E.28 Top Enriched KEGG Terms - Healthy Betz Cells vs. ALS Betz Cells. 354
E.29 Top Enriched WIKI Terms - Healthy Betz Cells vs. ALS Betz Cells. 354
E.30 Top Enriched Hallmark Terms - Healthy Betz Cells vs. ALS Betz Cells. 355
E.31 Top Enriched Reactome Terms - Healthy Betz Cells vs. ALS Betz Cells. 355
E.32 2D Annotation of Allen Institute and Microdissected Betz Cells. . . 362
E.33 ASGR2 & CSN1S1 IHC on Betz Cells. 363

List of Abbreviations

ALS . . . . . Amyotrophic Lateral Sclerosis.

AMPA . . . . α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid.

BA . . . . . Brodmann area.

Benj Hoch . . Benjamini–Hochberg False Discovery rate.

CNS . . . . . Central Nervous System.

CST . . . . . Corticospinal Tract.

DAB . . . . . 3,3'-Diaminobenzidine.

DEGs . . . . . Differentially expressed genes.

DNN . . . . . Cerebellar Dentate Nucleus neuron.

DEPs . . . . . Differentially expressed proteins.

ER . . . . . Endoplasmic Reticulum.

ET neurons . Extratelencephalic neurons.

F / ISH . . . . Fluorescent in-situ hybridization.

FACS . . . . . Fluorescence-activated Cell Sorting.

FDR . . . . . False Discovery rate.

FFPE . . . . . Formalin-Fixed Paraffin-Embedded.

FTD . . . . . Frontotemporal dementia.

GABA . . . . . Gamma-aminobutyric acid.

GO . . . . . Gene Ontology.

GOBP . . . . . Gene Ontology Biological Process.

GOCC . . . . . Gene Ontology Cellular Component.

GA . . . . . Golgi apparatus.

GOMF . . . . Gene Ontology Molecular Function.

GSEA . . . . . Gene Set Enrichment Analysis.

H&E . . . . . Haematoxylin and eosin stain.

HPA . . . . . Human Protein Atlas.

iBAQ . . . . . Intensity based absolute quantification.

IDs . . . . . Identifications.

IHC . . . . . Immunohistochemistry.

IT neurons . . Intratelencephalic neurons.

KEGG . . . . . Kyoto Encyclopedia of Genes and Genomes.

LogFC . . . . . Log2 Fold Change.

LC-MS/MS . Liquid Chromatography with tandem mass spectrometry.

LCM/LMD . Laser capture microdissection / Laser microdissection.

LPC . . . . . Laser pressure catapulting.

M1 / PMC . . Primary Motor Cortex.

MALDI . . . . Matrix-assisted laser desorption/ionization.

MTG . . . . . Middle temporal gyrus.

nanoPOTS . . Nanodroplet Processing in One pot for Trace Samples.

NMDA . . . . N-methyl-D-aspartate.

ORA . . . . . Over Representation Analysis.

PANTHER . . Protein ANalysis THrough Evolutionary Relationships.

PCA . . . . . Principal component analysis.

PCC . . . . . Pearson's correlation coefficient.

PCP . . . . . Pyramidal cell plate.

PMI . . . . . Post-mortem interval.

PPI . . . . . Protein-protein interaction.

PT neuron . . Pyramidal Tract neuron.

PTMs . . . . . Post-translational modification.

RIN . . . . . RNA integrity number.

RT-qPCR . . . Quantitative reverse transcription polymerase chain reaction.

RWS . . . . . RIPA-Wet Cap-SP3 method.

scRNA-seq . . Single cell RNA-sequencing.

SMART-seq2 Switching Mechanism At the end of the 5'-end of the RNA Transcript-sequencing.

SP3 . . . . . Single-pot solid-phase-enhanced sample preparation.

SPRI . . . . . Solid-phase reversible immobilization.

TF . . . . . Transcription factor.

TIMS . . . . . Trapped Ion Mobility Spectrometry.

TMT . . . . . Tandem mass tag.

TOF . . . . . Time-of-flight.

"Instead, consider the possibility that any man could, if he were so inclined, be the sculptor of his own brain, and that even the least gifted may, like the poorest land that has been well cultivated and fertilized, produce an abundant harvest."

– Santiago Ramón y Cajal, 1897

1 Introduction

Contents

1.1 Motivation 1
1.1.1 Brief Introduction to Betz Cells 1
1.1.2 Why Focus on Betz Cells? 2
1.2 Objectives 4
1.3 Outline of Thesis 5

1.1 Motivation

1.1.1 Brief Introduction to Betz Cells

Since their discovery in the late 19thcentury, research into the neurobiology of the Betz cell has been neglected compared to other neuronal cell-types; which is surprising, considering their prominent position as the defining cell of the human primary motor cortex and implication in neurodegenerative diseases such as Amyotrophic Lateral Sclerosis (ALS). They are the largest neuron in the Human Central Nervous System (CNS) and have a critical role in the generation and regulation of voluntary motor control. While 20thcentury investigations yielded descriptive data on their unique morphology, little has been done to explore their molecular composition with the tools now available in the 21stcentury. Although the comparative analysis of the

healthy and diseased human brain is of paramount importance in understanding and potentially treating pathophysiological mechanisms; from a basic science perspective, it is remarkable that such a morphologically striking neuron has not been extensively studied to provide deeper insights into healthy human physiology. This is confounded by the fact that common animals models (e.g. rodents) do not possess classically defined Betz cells and therefore, studying non-human models in the context of the motor cortex may lead to results that are of limited value for a better understanding of human physiology. To date, Betz cells have not been investigated with the same rigour as other neuronal cell-types using common laboratory techniques such as RNA-sequencing (RNA-seq), quantitative Polymerase Chain Reaction (qPCR), Fluorescence-activated Cell Sorting (FACS), or *in situ* hybridization (ISH); resulting in little information on their molecular composition.

1.1.2 Why Focus on Betz Cells?

With the recent development and popularity of next-generation *-omics* technologies i.e. transcriptomics, proteomics, and metabolomics, it is now possible to generate large amounts of precision data, offering novel insights into the expression/abundance of macromolecules that are used to define function, discover biomarkers, and investigate biochemical alterations in disease. There has been only one - very recent (2020) - study (preprint) that explores the transcriptomic diversity of cell-types in the motor cortex and specifically attempts to uncover the molecular signature of the Betz cell. However, this study could not identify a Betz cell-specific profile, and at the time of writing, the molecular profile of the human Betz cell is largely unknown with information on their physiology limited to non-molecular derived analyses. Moreover, no *-omics* technology has been applied to the Betz cell in a targeted and systematic manner in an attempt to reveal their potentially unique molecular signature; making observations and conclusions drawn from large-scale unbiased approaches based on bulk tissue analytics difficult to interpret/validate.

The little information that is currently known about the molecular composition of Betz cells can be attributed to the lack of robust biomarkers to facilitate

their accurate identification. Without this, it is difficult to advance Betz cell research; additionally, Betz cells are not as numerous as other neuronal cell-types and routine sampling from brain biobanks usually result in minimal access to Betz cell populations. While some stains/immunohistochemical-markers (H&E, Nissl, Non-phosphorylated neurofilament antibodies (SMI32/SMI311)) exist to visualise Betz cells based on size, location, and somatodendritic processes, it is not currently possible to highlight Betz cells from surrounding neurons solely using immunohistochemical techniques, making immuno-based isolation/enrichment techniques problematic. Transcription factors in neurons help maintain their post-mitotically differentiated features. As such, it is plausible that the large size and unique morphology of Betz cells are maintained by a collection of transcription factors and perhaps, Betz cells contain a specific molecular profile distinct from their neighbouring motor neurons. Staining for this profile would allow Betz cells to be easily discerned from other neurons and accelerate Betz cell research. Unfortunately, the existence of such TFs or Betz cell-specific markers is unknown, and research into discovering them is further hampered by the difficulty of isolating Betz cells from the surrounding neuronal populations in large numbers and limited access to suitable human material.

More generally, there has been an ambitious concerted effort in recent years to characterise and classify all cells in the human body and their interactions with each other. With the invention and development of single-cell/single-nuclei RNA-sequencing (scRNA-seq/snRNA-seq) technologies, this mammoth task has become a realistic end-goal. However, many of these technologies still rely on *a priori* knowledge cell-type-specific biomarkers/unique molecular fingerprints, either to isolate cells upstream for analysis or to later guide cell deconvolution algorithms to classify cell-type-specific clusters *in-silico*. If Betz cells are to be included in these global endeavours, targeted explorative studies must be performed to complement the findings and hypotheses generated by these large-scale, unbiased studies.

ALS is a fatal neurodegenerative disease that is defined by the selective degeneration of motor neurons, including Betz cells. In post-mortem ALS studies, it has been shown that Betz cells exhibit both morphological and cytological alterations that differ from the surrounding L5 motor neurons, suggesting a Betz-specific ALS pathology that only manifests in these gigantopyramidal neurons. This ALS Betz pathology has not been extensively studied, most likely due to Betz cells commonly being grouped with other upper motor neurons in studies; often noted simply as *'larger'* counterparts of the surrounding L5 motor neuron population rather than their own neuronal subtype. This, coupled with the lack of ALS animal models in species with gigantopyramidal motor neurons, has resulted in very few experiments investigating the particular pathways affected in Betz cells in the ALS brain.

1.2 Objectives

This thesis was designed to develop a targeted strategy for the investigation of the molecular profile of the human Betz cell using a combination of transcriptomics and proteomics. To the best of my knowledge, this has never been done before. The data generated should (1) molecularly confirm or refute observations concerning the nature of Betz cells, which have previously only been speculated on with traditional histological or stereological examinations, (2) compare the biochemical composition of Betz cells to other neurons, in an attempt to reveal their distinct features, and potentially find biomarkers to improve Betz cell identification and therefore accelerate further Betz cell research, (3) discover differences between Betz cells and their neighbouring motor neurons, to clarify whether Betz cells are simply large motor neurons or represent their own unique sub-type. Finally, (4) I hope to use the data generated to begin to explore why Betz cells are selectively vulnerable in ALS and define biochemical pathways that characterise diseased Betz cells in human brains affected by ALS.

1.3 Outline of Thesis

The following chapter will outline the importance of studying the neuronal diversity of the human brain, along with the significant and nuanced challenges that are associated with it. Next, the various technological advances that have accelerated the field of neurocytology will be discussed, with particular emphasis placed on modern approaches that generate data on both the protein and gene expression of cell-types. As it will play a significant role in this thesis, the literature surrounding laser microdissection (LCM) on the human CNS will then be discussed, again with a emphasise on analytical pathways to study RNA and protein expression at cell-type-specific resolution. This will be followed by a literature review on the human Betz cell, focusing on their morphology, molecular composition, and latest findings. Finally, there will be a brief introduction to ALS and the neuropathology of Betz cells in this disease.

Chapter 3 will contain a comprehensive description of the methods used in this thesis. Chapter 4 aims to set the foundation for subsequent chapters by describing the development and optimisation of an LCM proteomic approach to identify a large number of proteins from pooled quantities of individual neurons. This chapter will also showcase the performance of the workflow and test the outcome of several histological practicalities that may be implemented in future studies. The ability of this workflow to be used to compare the proteome of different neurons will be shown in Chapter 5. This chapter will also explore the proteomic nature of the Betz cell and compare its proteome to several different neuronal subtypes, using immunohistochemistry (IHC) to validate LC-MS/MS findings.

Chapter 6 will build upon the work from Chapter 5, by examining the proteome of the Betz cell on a large scale to confirm previous findings and discover new insights using a bespoke cohort of cases designed specially to maximise high-quality RNA. Betz cells from individuals with ALS will also be investigated in parallel, to begin to unravel the selectively vulnerable nature of the Betz cell in ALS. To gain insight into their gene expression, and to complement the proteomic data generated,

Betz cells from healthy and individuals with ALS will be transcriptomically targeted with a complementary approach (LCM-seq), which is a more mature technology than LCM proteomics.

2

Background

Contents

2.1 Neurocytology 8
2.1.1 Neuronal Diversity 9
2.1.2 Classifying Neurons 9
2.2 Technological Advances in Neurocytology 16
2.2.1 Biomarkers 16
2.2.2 Transcriptomics of the Brain 17
2.2.3 Proteomics of the Brain 20
2.2.4 Laser Capture Microdissection 21
2.2.5 LCM Transcriptomics 23
2.2.6 LCM Proteomics 27
2.3 Betz cells: Magnocellular Neurons of Motor Control . 29
2.3.1 The Primary Motor Cortex 29
2.3.2 Corticospinal Tract 29
2.3.3 Microcircuitry of the Adult Primary Motor Cortex . . . 32
2.3.4 The Cellular Diversity of the Primary Motor Cortex . . 35
2.3.5 Layer 5 of the Primary Motor Cortex 36
2.3.6 Betz Cells 40
2.4 Amyotrophic Lateral Sclerosis 54
2.4.1 Betz Cell Pathology in ALS 54
2.5 Aims of this Thesis 59

2.1 Neurocytology

The cell is the most fundamental unit of life; first named by Robert Hooke in 1665 [1], it was not until the mid-19thcentury where cell theory came into fruition, and the critical role cells play in human physiology and disease became to be appreciated. Although every cell in an organism contains essentially the same genome, there exist subtle or large differences in the expression of the genome that cause the diverse functions seen in human physiology. Disease, in many cases, is due to aberrant changes in this expression; as such, it is of paramount importance to understand both normal and abnormal cellular phenotypes to identify key features that can be used for therapeutic interventions or aid in developing biomarkers to provide a means of selectively investigating vulnerable or resistant cells.

The human brain is arguably the most complex organ in the animal kingdom, while no systematic investigations have been done to quantify the total number of cells present, current estimates based on primate cellular scaling rules [2] predict that an average human brain of 1.5kg should contain 93 billion neurons (25 billion in the neocortex & 70 billion in the cerebellum) and 112 billion non-neuronal cells [3, 4]. However, it is not known if humans fall strictly into this scaling system or if human brains are unique, as such, different groups have estimated varying numbers of cells in the brain [5–8]. As a result, the total number of neurons and glial cells is not explicitly known. Nevertheless, these studies show that the brain represents an incredibly complex cytoarchitecture. To add to this complexity, the brain is comprised of an interwoven mesh of dendrites, axons, and glial processes resulting in over 100,000km of interconnections with trillions of electrical and chemical synapses [9] and a storage capacity of 1.25 x 10^{12} bytes [10]. Understanding the relationship of these cells and their connections remains a significant challenge in modern neuroscience and must be explored and investigated to better comprehend the structure, function, and evolutionary position of the human brain, including its vulnerability to diseases not seen in non-human primates.

2.1.1 Neuronal Diversity

Neuroanatomy and neurocytology have advanced considerably in recent years, which can be contributed to the development of sophisticated neuroimaging, cytoarchitectural/myeloarchitectural, and molecular biology techniques [11]. Collectively, these techniques have shown that the human brain can be divided into ≈180 neuroanatomical areas [12, 13]. It is now known that all areas of the brain exhibit a laminar structure [14]; which differs in the thickness and number of layers across individual areas. Defining the boundaries of a particular neuroanatomical area is a challenging task, and very few regions of the human brain can be classified into discrete regions [11]. There do exist some examples, where the boundary of an area can be determined purely based on the existence of a particular cell-type e.g. the von Economo neurons of the anterior cingulate cortex and fronto-insular cortex or the Betz cells of the primary motor cortex [11]. The ability to define neuroanatomical boundaries, their underlying cellular diversity and resulting physiological function is critical for understanding the progression and pathophysiology behind neurodegenerative disease. While the large size of von Economo neurons and Betz cells makes their existence obvious in their respective regions, it is difficult to identify other individual neuronal cell-types in the human brain; especially when neurons can possess an incredible range of connectional, functional, and molecular properties. Historically, morphology alone was used to group neuronal subtypes, however, this is also incredibly diverse - see Figure 2.1 for an example of a typical pyramidal neuron and 2.2 for examples of neuronal diversity.

2.1.2 Classifying Neurons

Many neurodegenerative diseases selectively target specific cell-types in the CNS [15]. As such, being able to define these cell-types and their underlying physiology is critical in understanding the dysfunction caused by these diseases. Additionally, as the fundamental unit of brain physiology, accurately defining nervous cell-types, unravelling their function, and studying how they interact with one another will

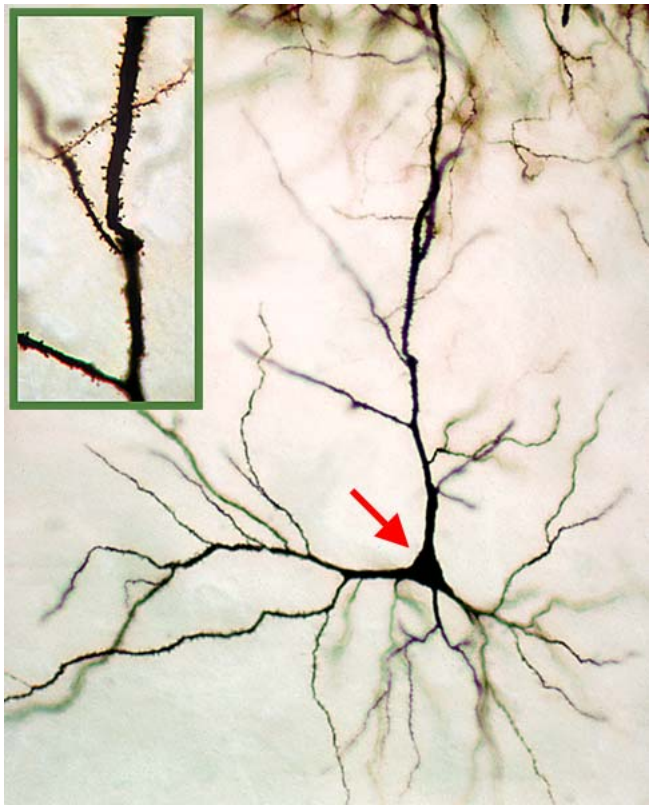


Figure 2.1: A Pyramidal Neuron Stained with the Golgi Stain.
Red arrow denotes soma and inset highlights dendritic spines morphology. *(With permission from Prof. Tom Fletcher.)*

increase our understanding of the human brain. A common criterion for grouping neurons is to classify them based on their principal neurotransmitter. Biophysical studies have helped classify cortical neurons into two general classes: glutamatergic excitatory neurons and GABAergic inhibitory neurons. However, several subclasses exist within each group, further characterised by axon destination and/or molecular characteristics. For example, there are 5 main classes of glutamatergic excitatory neurons: cortico-cortical projection neurons[1], subcerebral projection neurons[2], locally projecting L4 neurons, and locally projecting L6b subplate neurons [16–19]. This hierarchical system of neuronal classification continues further, with L4 neurons being divided into spiny stellate cells or star pyramidal cells depending on the presence of an apical dendrite. Intratelencephalic (IT) neurons are found across neuronal layers but are abundant in L2, L3, and L5a and can be defined by their

[1] Also known as callosal projection neurons or intratelencephalic neurons.
[2] Also known as extratelencephalic neurons.

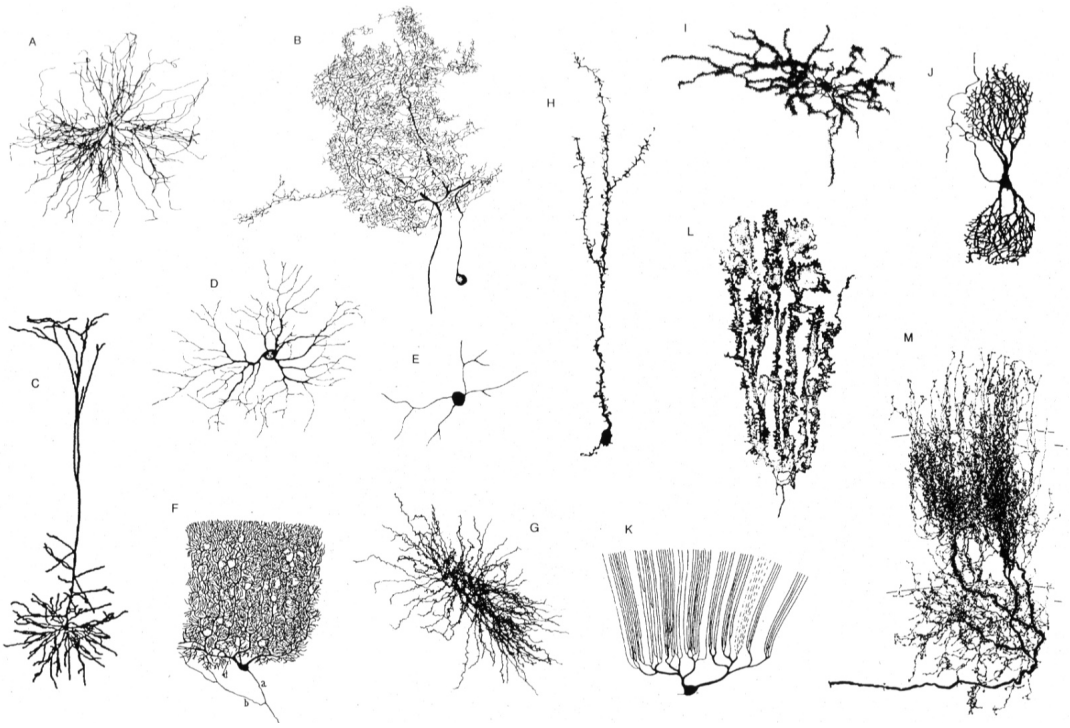


Figure 2.2: Diversity of Neuronal Morphology.
A: α-motoneuron of the spinal cord (Cat). B: Spiking interneuron in the mesothoracic ganglion of a locust. C: Layer 5 pyramidal neuron (Rat). D: Retinal ganglion cell (Cat). E: Amacrine cell of the Retina (Salamander). F: Cerebellar Purkinje cell (Human). G:Relay neuron in the ventrobasal thalamus (Rat). H: Olfactory bulb granule cell (Mouse). I: Spiny Projection neuron in the Striatum (Rat). J: Nerve cells in the Nucleus of Burdach (Human fetus). K: Purkinje cell (Fish). L: Golgi epithelial cell in the cerebellum (Mouse). Axonal arborization of isthmotectal neurons in the turtle. *Adapted from: Information Processing in Dendritic Trees by Bartlett W. Mel in Neural Computation Vol. 6, No. 6, November 1994, reprinted courtesy of The MIT Press.*

thin-tufted dendrites and can project ipsilaterally and contralaterally to multiple cortical areas. Extratelencephalic (ET) neurons are located in L5b and as opposed to IT neurons contain thick-tufted dendrites and project to the striatum, thalamus, midbrain, hindbrain, and spinal cord. Finally, the cortico-thalamic neurons in L6 provide feedback loops to the thalamus. Within the inhibitory class of neurons, there exist four main subclasses which are named according to the molecular markers they express, i.e. parvalbumin-expressing (PVALB+) cells, somatostatin-expressing (SST+) cells, vasoactive intestinal peptide-expressing (VIP+) cells, and 5-hydroxytryptamine receptor 3A-expressing but VIP negative (HTR3A+ VIP-)

cells. PVALB+ cells can be divided into basket or chandelier cells, SST+ cells can be classed as Martinotti or non-Martinotti cells, VIP+ cells are classified as bipolar or multipolar cells, and finally, HTR3A+ VIP- cells are grouped as either bouquet cells or neurogliaform cells [20–23].

Despite this robust system of classes, it is unclear how many cell-types are present within each of these excitatory and inhibitory subclasses. Conceptually, it is also difficult to define criteria that distinguish neuronal types from each other – what variables should be used to define one neuronal type from another? While morphological distinctions are useful for the initial analysis of neurons, they offer limited values in many experiments. Scientists have debated different ways to organise these cell-types; organising based on projection targets produces large variations due to the different combinations that any one neuron can project to, making the classification unrealistically diverse [24–27]. Whereas classifying based solely on electrophysiology leads to very little variation among neurons, for example, in a study on the electrophysiology and morphology of >14,000 rat somatosensory cortical neurons; the authors assigned all excitatory neurons to a single class, while inhibitory neurons could be divided into 10 different classes [28]. In contrast, the authors managed to find 55 neuronal classes based solely on morphology. Single-cell RNA sequencing (scRNA-seq) and single-nuclei RNA-seq (snRNA-seq) has provided another approach to classify neuronal cell-types, this time, based on mRNA expression. For example, the Allen Institute for Brain Science performed snRNA-seq on the adult mouse primary visual cortex and identified 23 GABAergic classes, 19 glutamatergic classes, and 7 non-neuronal classes, all of which represented the major cell-types reported in the literature, and even identified new types of cells or further subdivided existing cell-types [29]. The Allen Institute for Brain Science has taken up the mammoth task of transciptomically defining neural cell-types in the mouse and primate brain. Using laser microdissection and single-nuclei sorting coupled with RNA-sequencing, they have managed to classify molecular clusters of cell-types according to relative mRNA expression and neuronal layer. For example, in their attempt to find conserved cell-types in the mouse and human middle temporal gyrus

(MTG), they identified 6 non-neuronal cell-types, 24 excitatory neuronal types, and 45 inhibitory neuronal types in the human MTG [30]. The Allen Institute has also applied its approach to the lateral geniculate nucleus, primary visual cortex, anterior cingulate cortex, and most recently the primary motor cortex. Through this, they have generated large public data-sets that allow for data exploration and analysis. Moreover, they have constructed a hierarchical system that combines their transcriptomic clusters with the previously proposed system of classifying excitatory and inhibitory neurons in an attempt to create a better neuronal taxonomy- see Figure 2.3. They have also introduced nomenclature to assign unique identifiers to each neuron represented by the major cell class, layer enrichment, subclass marker genes, and cluster-specific marker genes [30]. For example, the transcriptomic cluster *Inh L2-5 PVALB SCUBE3* represents an inhibitory neuron that is found in layers 2 to 5, is a parvalbumin interneuron and differentially expresses SCUBE3, validation with *in-situ* hybridization and independent transcriptomic profiling has identified this cluster to belong to chandelier neurons. The use of molecular/location-derived nomenclature is believed to create a comprehensive consensus of all neural subtypes. While snRNA-seq is an incredibly powerful technique and has been the major contributor to recent neuronal classification projects [31], it is still the case that the resulting transcriptomic cell-types/clusters still do not align perfectly with morphology and electrophysiology data.

It is generally accepted that the most influential criteria for defining a neuronal type are: morphology (dendritic architecture, axonal shape, and soma size), physiology (resting potential, firing rate, and biophysical properties) and biochemical composition (genome, methylome, transcriptome, proteome, metabolome, and glycome). It is important to note that connectivity is also relevant but at present is significantly harder to assess. Collectively, these variables could be used to create a comprehensive taxonomy that would aid in cell identification and increase confidence in reproducibility between studies. The Allen Institute is currently developing multi-modal characterisation schemes to merge transcriptomics with morpho-electrical data using PATCH-seq to solve the large discrepancies between

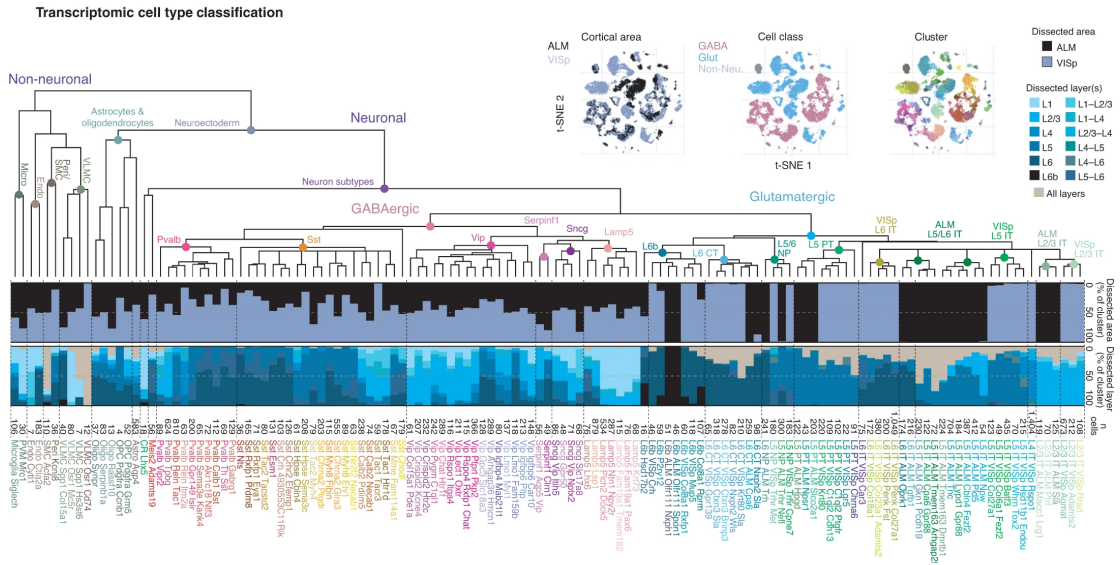


Figure 2.3: Example of the Allen Institute's Transcriptomic Clusters Taxonomy. The Allen Institute has developed a taxonomy for cell-types in the brain. Cells are grouped as neuronal or non-neuronal. Neuronal subtypes are further separated by their primary neurotransmitter, major cell class, layer enrichment, subclass marker gene, and cluster-specific marker gene. *(Permission from Springer Nature).*

transcriptomic class and electrophysiology classes. However, this approach still does not incorporate potential cell-type-specific differences in protein, metabolite, and glycan expression; most likely because the field of research around these biomolecules is not as advanced as sequencing techniques.

While morphology, molecular composition, electrophysiology, and connectivity are critical in defining a neuronal cell-type, in practice it is difficult to determine all of these variables, in such cases, the use of molecular biomarkers takes interim precedence. The reasoning behind this is that some molecular features stay static throughout the lifetime of a neuron, whereas other variables are more prone to change in response to stimuli. However, it is important to note that not all molecular markers are static, for example, genes encoding for the unique morphology of a cell may only be transiently expressed during development – therefore, selecting static molecular markers are fundamental to successful classification. As previously mentioned, the brain has a dizzying amount of cellular heterogeneity, so it is not surprising that there remain many neurons whose biochemical composition,

electrophysiology, and connectivity are either not clear or unknown; either due to small populations, difficulty physically accessing tissue for research, or due to the prevailing dogma grouping cell-types together despite no molecular basis. This remains a significant problem, while the invention and use of canonical molecular biomarkers have increased the confidence that a particular neuronal type has been sampled and investigated in subsequent experiments there are, unfortunately, many cell-types that lack such markers making some studies difficult to validate or reproduce. The invention of a robust "neuron ontology" - similar to what the Allen Institute is proposing - would greatly aid in the accurate identification of neuronal cell-types and continues to be an ongoing/major goal of modern neuroscience.

As seen, defining cell-types requires a tremendous amount of work; often requiring specialised equipment, access to appropriate biological material, complex procedures to independently study macromolecules, sophisticated *in silico* analysis and well-designed valuation methods. To ease the workload by pooling resources, an international effort has begun to map all cells in the human body, the ambitious project: the Human Cell Atlas aims to produce a comprehensive reference catalogue of all human cells based on their transient features and properties along with their abundance and location. As previously mentioned, many species of biomolecules need to be studied to gain the full molecular profile of a cell i.e. lipids, glycans, genes, transcripts, proteins. However, arguably, the study of nucleic acids and proteins contributes the most to this cellular signature. Next-generation sequencing has become the tool of choice to study cell-types and individual cells in the brain. Transcriptomic data is commonly used to infer protein expression, due mostly to the ease and lowered cost of quantifying mRNA compared to their protein counterparts. However, the relationship between mRNA and protein expression is still debated, with some suggesting only a modest correlation (PCC = 0.32) [32]. Furthermore, unlike transcriptomics, proteomics can characterise the expression, structure, function, interactions, and modification of proteins, all of which may vary at any stage during the lifetime of a neuron or in response to external stimuli [33]. Despite the modest correlation, proteomics can also be

used to infer mRNA expression and is the gold standard tool to measure post-translational modifications (PTMs), leading some to believe it is a superior technique in characterising a biological system [34]. In light of this, where possible, both proteomic and transcriptomic workflows should be used when studying a cell-type, allowing for data to be analysed independently and complementary rather than extrapolating information from either entity. By combining proteomic and transcriptomic workflows, it should be possible to gain the majority component of the total molecular profile of a cell.

2.2 Technological Advances in Neurocytology

2.2.1 Biomarkers

Immunohistochemistry and the development of mono- and polyclonal antibodies made it possible to detect the presence of individual proteins within cells and revealed that morphologically indistinguishable cells can vary dramatically on the molecular level [35]. This allows 'identical' cell-types to be intricately classified based on protein expression [36]. There has been a push to define and categorise cells based on their molecular phenotype rather than their apparent morphology, and studies are lacking that attempt to tie molecular composition with cellular morphology. Underpinning this molecular signature has been key in defining cell-types; moreover, the discovery of canonical biomarkers can aid in the confidence that these cells have been appropriately sampled in cell-type-specific experiments. Furthermore, the molecular categorisation of cells has been for the most part been *ad hoc* using biomarkers discovered by accident or chosen for convenience rather than using a systematic and comprehensive approach. A significant milestone was the discovery of cell-type-specific surface markers allowing nervous tissue cell-types to be visualised or isolated using antibody-based FACS or laser microdissection. Historically, large-scale low-throughput ISH experiments were performed to define single-cell gene expression data in a spatial manner [37–40] and although these

archaic databases remain a valuable resource for validation/data analysis; Next-generation sequencing is now the method of choice in generating large amounts of biochemical data from brain tissue.

Traditionally, individual cell-types are isolated/enriched from their cellular environment using knowledge of cell-specific biomarkers. Now, it is becoming popular to study all cell-types in a selected tissue, omitting isolating steps, and instead, using sc/snRNA-seq to investigate a large number of individual cells at once. This approach still requires *a priori* knowledge to validate unbiased *in silico* deconvolution algorithms or guide manual segmentation of individual cells into groups/clusters. Results from these snRNA-seq experiments can vary dramatically depending on the type of clustering used, changes in data analysis workflow or amount of technical noise; making computer-guided segmentation challenging [41, 42]. Furthermore, unbiased approaches may struggle to separate closely related cell-types, either due to cell-types sharing many molecular similarities or due to the reduced expression data generated when investigating at a single-cell resolution.

In the context of understudied neurons, like the Betz cell, it may be advisable to apply targeted *-omic* approaches to first generate preliminary data that can then be used to complement/validate the findings of large unbiased snRNA-seq experiments. Targeted isolation of cell-types with no known biomarkers is difficult with FACS, whereas, in the absence of biomarkers, laser microdissection provides a robust tool to isolate cells from their environment based on cellular morphology. Unfortunately, the use of tissue sections limits microdissection to cell-slices rather than whole cells or nuclei, resulting in lower starting materials. However, the increasing development of highly sensitive equipment and efficient workflows have allowed microdissection to be sufficient in generating *-omic* quantities of data.

2.2.2 Transcriptomics of the Brain

Subtle differences and elegant control of spatiotemporal gene expression are responsible for the diverse nature of neural cell-types, the development of distinct

neuroanatomical regions, connections and functions of neural circuits, and ultimately control the phenotype and behaviour of an organism. By studying the transcriptome of a tissue region, or cell-type, large amounts of information can be generated. Transcriptomics provides a means of identifying a significant part of the molecular profile of cells, giving insights into their structure, function, and potentially discovering robust biomarkers. By utilising high-throughput quantitative transcriptomic approaches such as microarray analysis or RNA-seq, it is also possible to study differentially expressed genes across cell-types in health and disease or in response to therapeutics. Recent advances in sequencing technology have led to decreased cost and increased use of mRNA analysis and as a result, has produced publicly available, comprehensive brain transcriptomic databases that contain high-quality spatiotemporal cell-type-specific data [43]. Kang et al., (2011) remain one of the largest human brain transcriptome studies to date; the authors analysed 16 brain regions from 57 post-mortem brains (1,340 tissue samples in total) ranging from 5.7 weeks after conceptions to 82 years of age - essentially covering the entire human lifespan [44]. These data showed that the vast majority (>90%) of brain-expressed genes are spatiotemporally regulated and that the brain transcriptome is arranged into detailed coexpression networks. While these data were generated by exon microarrays the same samples were also analysed by RNA-seq, the results of which are available from the BrainSpan Consortium website.

Transcriptomics has also been used to investigate the connectional, structural, and molecular changes that have occurred in our recent evolution, Sousa et al., (2017) compared the transcriptome between the chimpanzee, macaque, and human across 16 brain regions and found that while the majority of genes contained conserved patterns approximately 1/4 of protein-coding genes were differentially expressed between at least two species in one or more regions [14]. There now exists a plethora of public resources that provide the means of exploratory data analysis of the human brain - see Table 1 from [43]. The current goal of transcriptomic-based neuroscience investigations is to reach the same level of quantitative depth in cell-type-specific or single-cell experiments. Since scRNA-seq was introduced by Tang

et al., (2009) the technology has developed rapidly and has been extensively applied to many biological systems [45]. Investigating cellular composition, heterogeneity, and disease-relevant cell-types have been the primary goal of scRNA-seq.

As with most new biotechnologies, large-scale scRNA-seq classification was first performed on mouse models. In one study, Zeisel et al., (2015) discovered that 9 major brain cell-types (Interneurons, S1 pyramidal neurons, CA1 pyramidal neurons, mural, endothelial cells, microglia, ependymal cells, astrocytes, and oligodendrocytes) from the somatosensory cortex and hippocampal CA1 region could be further grouped into 49 subpopulations based on the unbiased clustering of their unique transcriptome profile [46]. This not only highlighted the power of transcriptomic-based taxonomy of neural cells but this study also managed to discover biomarkers that aided in the successful identification of these sub-cell-types in future studies wishing to directly target said cells. Since then, a significant amount of work has been put in adopting this approach to other mouse regions [47–50]; and these approaches have started to be implemented to classify the cell-types of the human brain [30, 51–60]. Moreover, several studies have used next-generation sequencing to compare cell-type-specific transcriptomes to diseased counterparts [61–65]. scRNA-seq has a proven history of being an incredibly powerful resource for classifying neural cell-types, understanding cell-type-specific mechanisms, and the role they play in health and disease. However, the isolation of cells for downstream scRNA-seq still requires the use of cell-type-specific markers, which for many cells are still not known. Without *a priori* knowledge, rare or understudied cell-types may not group neatly into distinct clusters, making biomarker discovery or molecular profile determination difficult. Moreover, although extremely useful, investigating the transcriptome of individual cells in the brain is limited to answering questions on gene expression, and while protein expression is often inferred from sequencing experiments, studying protein expression directly via proteomic investigations is likely to yield more accurate data on protein abundance, enriched pathways, and post-translational modifications (PTMs).

2.2.3 Proteomics of the Brain

The enormous network of cells in the cerebrum reflects its complexity and are responsible for the magnitude of physiological actions. Detrimental changes at the cellular/molecular level of the human brain are responsible for a variety of degenerative, developmental, and psychological disorders making investigative research into understanding this complexity of paramount importance; and remains one of the central challenges in modern neuroscience. In recent years, the human brain has been studied using a wide variety of techniques, such as *in situ* hybridization, magnetic resonance imaging, positron emission tomography, electron microscopy, optogenetics and transcriptomics. However, large-scale proteomic experiments have not been as readily applied to the brain, which is surprising considering that most of the functions of the brain are ultimately mediated by proteins [66]. This neglect can be attributed to the inability to exponential increase the amount of starting material as seen with nucleic acids. This has made it particularly difficult to develop equipment and methods to study proteins in a sensitive, quantitative, and unbiased manner as seen in the RNA sequencing field. However, these restrictions are slowly being resolved; recent advances in mass spectrometry proteomics have provided workflows that are now capable of generating quantitative data to a similar depth of RNA-sequencing [67–70].

The proteome is significantly more complex than the genome or transcriptome. While it is estimated that the human genome contains $\approx$ 20,000 protein-coding genes [71] corresponding to $\approx$ 100,000 transcripts; the effect of different PTMs results in the generation of >1,000,000 proteins entities [72]. Furthermore, 5% of the proteome is responsible for these PTMs and can take the form of ligases, transferases, kinases, proteases, and phosphatases which are responsible for a wide array of cellular functions. Overall, there are >200 types of PTMs that can occur at any stage of the life cycle of a protein, and understanding their normal or aberrant mechanisms is critical for investigating neurodegenerative diseases [73]. Despite this, much of the focus in understanding protein function and

interaction is inferred from transcriptomic datasets; which is problematic due to the complex relationship between RNA expression and protein expression [74, 75]. The downside of proteomics is that a considerable amount of CNS proteins is involved in membrane-associated processes e.g. synaptic docking and release, cell-cell adhesion, or neurotransmitter receptor signalling; unfortunately, the low abundance and amphipathicity of these membrane proteins make lysis and digestion inefficient, potentially biasing data. Nevertheless, advances in proteomics are likely to lead to deeper insights into cell biology.

There exist many approaches to investigate proteins *ex vivo*, however, enzymatic digestion followed by high-performance liquid chromatography electrospray ionization mass spectrometry has become the general workflow when investigating biological tissues and has become continuously refined over the years [76]. Input material can be enriched to contain particular cell-types, organelles, protein-protein interactions, or posttranslational modifications. The Human Protein Atlas (HPA) has proposed that 74% of all human proteins are expressed in the brain - approximately 14,000 proteins - of which, many are expressed in a temporospatial manner, demonstrating alternative splicing in different anatomical regions/cellular populations suggesting a very complex, highly compartmentalized brain proteome [77–79]. Spatial proteomics of the human brain has the potential to define the complex architecture in health and disease [66] which warrants developing methodologies that can probe and investigate the spatial dimension of the "neuroproteome".

2.2.4 Laser Capture Microdissection

Despite these technological enhancements making proteomic assays more sensitive and powerful, most *ex vivo* approaches necessitate large amounts of bulk tissue ranging from 10s to 100s micrograms for untargeted proteomics [80–83] or upwards of 0.5 – 2 grams of tissue for experiments requiring enrichment of phosphoproteins, insoluble proteins, or immuno-precipitated proteins [84, 85]. As the name suggests, ‘bulk tissue’ commonly contains a minimum of thousands of cells and contains the

combination of several different cellular phenotypes, which is a consequence of the complex heterogeneity found within most tissue samples [86]. While still useful in proteomic research, the data generated from bulk tissue are often limited due to resulting data representing the accumulation of all phenotypes present in the sample, and almost certainly concealing important information relating to subtle changes found in individual cell populations [87]. This 'averaging effect' is a common problem in most *-omics* technologies requiring bulk tissue [88–90] and has led to a huge growth of bioinformatic tools (both proteomic and transcriptomic) to help separate cellular populations into cell-specific clusters *in silico* based on known biomarkers and/or unbiased clustering after molecular information has been captured [91–94].

To resolve this 'averaging effect' when using bulk tissue, several techniques have been developed to isolate particular cell-types upstream to *-omics* analysis, however, the extremely heterogeneous nature of the neocortex adds significant technical difficulties in separating pure cellular populations from their respective microenvironment. A common technique used to overcome this is Fluorescence-Activated Cell Sorting (FACS), which uses light scattering and fluorescent probes conjugated to antibodies targeting proteins of interest to sort heterogeneous biological samples into two or more collecting containers, one or more cells at a time. Bulk brain tissue is typically mechanistically or enzymatically homogenised before FACS to shear cellular processes and isolate cell bodies. FACS provides an excellent method of separating cells of interest from a biological mixture to be analysed in downstream applications and has been used in several approaches [95–97]. However, despite homogenisation, the interwoven nature of the brain and presence of large cells can introduce significant obstacles for successful cellular isolation, furthermore, this application relies on *a priori* biochemical knowledge of the cell/s of interest, including known unique biomarkers and robust antibodies to target them. Moreover, appropriate starting material containing a minimum of thousands of cells is needed for FACS to be effective, which is not possible for cell-types that originate from rarer populations or are difficult to enrich for [98]. There are also physical limitations, such as the size of the cell of interest. It is recommended to have a FACS diameter

4-5 times larger than the cells being sorted, and while this is a technique that will undoubtedly improve over the upcoming years, at present, the human brain contains several cell-types whose size exceed that of the most advanced FACS diameter, most notably, the Betz cell with an impressive volume of $\approx$ 86,000 µm^3 [99].

Laser Capture Microdissection (LCM) or Laser Microdissection (LMD) is a promising approach to isolate cells from their microenvironment for downstream biochemical analysis [100]. There are slight variations in LCM technology, however, the following multi-step approach is commonplace: First, an epifluorescence or bright-field microscope is used to visualise the cells/areas of interest. Next, *in silico* annotations are drawn outlying areas/cells to be microdissected. Then, a laser is fired along the optical path of the above annotation, this ablation causing the cells/areas of interest to be isolated from the surrounding microenvironment. Finally, the laser delivers a defocused pulse which catapults the region or cell into a collection device in a contact-free manner to be used for downstream applications. Compared to FACS, LCM is a low-throughput and laborious technique. However, LCM is a "simpler" technique and confidence of correct sampling can be achieved by visually inspecting cells/areas of interest before and after microdissection. Furthermore, as with FACS, molecular markers can be used to microdissect cells based on distinct biochemical features; further, LCM can also be used to pool rare cell-types that would be difficult to isolate by gating channels alone. More relevantly, LCM remains one of the most effective methods to isolate cells for downstream analysis when no molecular markers are known; as annotations can be directed solely based on cell morphology. To date, as with FACS, LCM has proved popular in studies where areas or individual cells are isolated for downstream RNA analysis - see the next section. Nervous tissue is difficult to dissociate due to long neuronal processes and high glial content; however, cells can be removed from their microenvironment with relative ease using LCM.

2.2.5 LCM Transcriptomics

The scarcity of appropriate tissue, small cellular populations, and lack of molecular markers increases the need to develop techniques that allow for minimal input

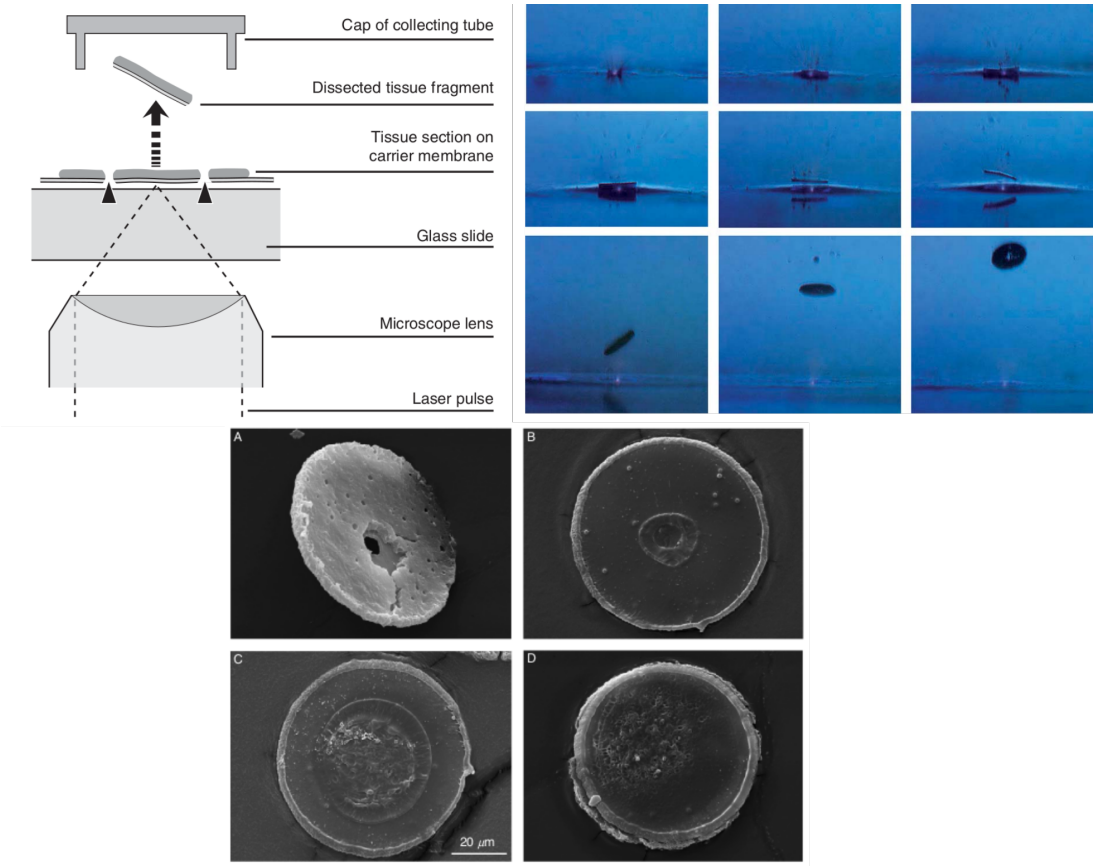


Figure 2.4: Laser Capture Microdissection (LCM).
TL: Principle an LCM microscope. Tissue sections are placed on a UV-absorbing polymer that is mounted onto a routine microscope slide. The region of interest is microdissected by melting the polymer. An isolated region of interest is then catapulted into the cap of a tube. TR: A 30 nanosecond time-series of a histological material being catapulted into a collection device. Bottom: Scanning electron micrographs of histologic specimens that have been catapulted. *(Adapted from [101] with permission from Elsevier).*

material without sacrificing the quality and depth of generated data. There have been some attempts to satisfy the increasing need to assess cell-type-specific/region-specific transcriptomic data; and there now exist several large data-sets that are the result of LCM transcriptomic investigations [14, 44, 102–104]. However, most studies combining LCM with RNA-seq often use LCM to infer a neuronal layer-specific transcriptome or, instead, used LCM to reduce the complexity of nervous tissue by isolating individual microdissected layers for FACS and subsequently sequencing sorted cells rather than directly sequencing the individual cells obtained by LCM. In a recent study, one group coined the term "hippo-seq" whereby they

used LCM to dissect either the dorsal trisynaptic loop or the ventral trisynaptic loop of the hippocampus [105]. Critically, once microdissected, the authors implemented a manual sorting process to dissociate hippocampal tissue and sort cells into populations of either dentate gyrus granule cells, mossy cells, CA1, CA2, or CA3 pyramidal neurons and generated transcriptomic datasets for each cell-type studied. Using LCM to simplify nervous tissue before sorting and then sequencing is a valid tool in assessing regionally transcriptomic differences. However, being able to generate equivalently large data-sets from individual or pooled microdissected neurons rather than microdissecting whole areas that then need to be sorted would significantly reduce labour, allow for more intricate cell-specific LCM experiments, and would be invaluable in examining cells with small populations or those with no known biomarkers, which would render FACS ineffective.

Microarray analysis on microdissected neurons has already been described [106–111] and technologies have now evolved to allow RNA-seq on microdissected material [112–114]. The ability to use RNA-seq to topographically map areas of the brain via microdissection has been achieved; by microdissecting the medial ganglionic eminence at a resolution of 2500 μm^2 [115]. However, these experiments still required a large amount of input (200-4000 isolated cells) to retrieve sufficient amounts of analysable mRNA. Recently, a procedure has been established that couples LCM with RNA-seq on nervous tissue and only requires a small amount of input material (1-100 cells) [116, 117]. Coined "LCM-seq" this approach utilises the robust ability of SMART-seq2 [118] to amplify small amounts (picograms) of mRNA in conjunction with an LCM procedure that minimises RNA loss and degradation. Critically, this protocol avoids the RNA extraction steps seen in standard RNA preparations and instead proceeds straight to cDNA synthesis, arguing that the pipette steps for RNA extraction dramatically reduce the already minuscule yield and that amplifying mRNA before RNA isolation provides superior results. When comparing LCM-seq prepared cDNA to cDNA prepared with the traditional workflow (i.e. with an independent RNA isolation step) the authors found they were able to achieve a much higher yield with their LCM-seq approach (147.4 ng $\pm$ 15.7

vs. 39.7 ng ± 11.9), however, it is important to note that the samples that had an RNA isolation step managed to achieve a slightly higher mapping ratio to the mouse genome (93.7% ± 0.8 vs. 89.8% ± 0.7). Interestingly, the LCM-seq approach was able to achieve the same cDNA yield with 30 motor neurons then what was possible with 120 motor neurons with the traditional approach, highlighting the effective ability to amplify mRNA from minuscule amounts of material when not lost by pipetting. This paper also showed that LCM-seq strikingly improved the reproducibility of gene identification and that acceptable data variability can be achieved by microdissecting as little as 1 neuron[3]. Furthermore, ≈ 14,000 genes were identified with the standard approach and ≈ 16,000 with the LCM-seq approach; impressively the LCM-seq experiments with only 1 neuron still managed to achieve a depth of ≈ 10,000 genes. Most relevantly, Nichterwitz et al., (2016) showed this technique could be applied to post-mortem human nervous tissue that has been partially degraded due to elongated times between tissue preservation and death. For post-mortem tissue, the authors extracted slightly more neurons (179.7 ± 30.8 spinal motor neurons, 169± 17.6 midbrain neurons from the ventral tegmental area, and 157.0±11.5 midbrain neurons from the substantia nigra pars compacta) and found that 90.53% of sequenced reads could be mapped to a human reference genome (hg38) across 13 samples, with an average of 14,893 genes being identified. Finally, Nichterwitz et al., (2016) showed that their LCM-seq procedure, is cheaper, easier to implement, improves sensitivity and reproducibility than traditional approaches; and could also be scaled down to 10, 5, and even 1 human motor neuron.

LCM-seq has been used by the same authors to study the selectively vulnerable and resistant neurons in Parkinson's Disease [119], Spinal Muscular Atrophy [117, 120, 121], Amyotrophic Lateral Sclerosis [121, 122], and has also been used to study glioblastoma [123], schizophrenia [124], the development of the peripheral

[3]It is important to note that when using LCM it is often the case that a cell in its entirety is not microdissected, rather only a segment of the cell is microdissected according to the thickness of the tissue section. However, with LCM studies it is common to use "cell number" even though this understates how much material is being analysed.

nervous system [125], and has been used to study the prefrontal layer V pyramidal neurons in individuals with abused suicides [126].

2.2.6 LCM Proteomics

While proteomics has been performed using FACS as the chosen method for cell isolation [127–129] there has been a slower response in attempts to couple LCM with proteomics even though proteins are more stable than nucleic acids and less prone to LCM degradation. A major hurdle with proteomic analysis using LCM is the difficulties with working with minute amounts of biological material. The vast majority of LCM is manually dictated, with automated LCM equipment not yet commonplace. This makes LCM laborious and unable to realistically isolate the same order of magnitude of cells seen in FACS. Moreover, unlike nucleic acids, proteins cannot have their abundances exponentially duplicated, therefore data has to be generated on trace amounts - typically in the picogram range - of proteins found in the raw sample, requiring extremely sensitive equipment and highly efficient workflows that minimise degradation and sample loss. As there becomes an increasing need to study rare or scare cell types, so does the demand for techniques to work with increasingly smaller amounts of input material. The biggest issue faced with LCM proteomics is trying to identify low abundant proteins when input material is minimal. As seen with LCM-seq, one considerable technological factor to be considered is the fact that the act of pipetting can drastically lower the already low amounts of identifiable proteins in small samples due to the material adhering to pipette tips and being lost in-between experiment steps [130, 131]. Recently, Zhu et al., (2018) have established a workflow for minimising material lost in mass spectrometry preparation. Their approach coined nanoPOTS (nanodroplet processing in one pot for trace samples) implements a semi-automated platform that uses nanodroplets to perform the necessary protein digestion in very small quantities ($<$200 nL) within a single nanowell; removing the many pipetting steps that occur in traditional mass spectrometry preparations [132]. This technique is highly sensitive to extremely small amounts of material, achieving an impressive depth of $\approx$ 1,500 to

≈ 3,000 proteins from as few as 10 to 140 HeLa cells, respectively. The nanoPOTS technique has proved popular and has been used by several studies [132–136]. Another study implemented a different approach; in this, they showed that protein lost by pipetting can be dramatically reduced by exploiting the hydrophobic and hydrophilic binding capabilities of protein/peptides to paramagnetic beads [137, 138]. By trapping proteins/peptides onto paramagnetic beads bound to a magnet it is possible to minimise loss by pipetting during SDS removal, protein digestion, and sample clean-up steps in mass spectrometry preparation. Termed "Single-Pot Solid-Phase-enhanced Sample Preparation" (SP3), this method has been shown to increase proteomic performance in samples with a small amount of starting material.

LCM proteomics is slowly becoming a popular field of research, attributed to advanced sample preparation strategies, the growing access to highly sensitive mass spectrometers, and the increasing need to enhance, complement, or validate established RNA-sequencing techniques. However, post-mortem nervous remains a difficult material to work with. Apart from the vast heterogeneity of the tissue, uncertainty around pre-death conditions and post-mortem delay can result in low-quality samples and few suitable experimental replicates. Nevertheless, there have been attempts to apply LCM proteomics to nervous tissue, although the methodologies behind these studies vary considerably in their collection, lysing, and digestion approaches – making comparisons difficult. While they show promising results, many have only been tested on rodent brains [139–144] with experiments on humans nervous tissue yielding low proteomic depth [145–147]. Despite using LCM, these studies still isolate relatively large areas of tissue, rather than individual cells, ranging from 5,7000,000 μm^3 to 15,000,000 μm^3 and although being significantly smaller than the several 10s to 100s micrograms used in bulk tissue proteomics the labour involved with microdissecting this amount of material still require a considerable amount of laboratory work and sample preparation; often involving microdissecting whole areas rather than individual cells which cause the averaging effect mentioned previously. Furthermore, while these approaches result in satisfactory data output they do not compete with the high numbers of transcripts quantified

in transcriptomic studies, with only ≈ 1,000-2,000 proteins being identified in LCM proteomic experiments with human samples [133, 148]. One human-based study managed to achieve an impressive proteomic depth of 2,300, however, this experiment required a huge amount of paediatric nervous tissue (460,000,000 μm^3) highlighting the need for optimised workflows and optimum equipment [147]. Moreover, these previous studies use long chromatography gradients to reduce the complexity of the cell lysate with times ranging from 1.5 – 2 hours making experiments lengthy and unsuitable for large scale projects and/or clinical studies.

2.3 Betz cells: Magnocellular Neurons of Motor Control

2.3.1 The Primary Motor Cortex

Locomotion is a deceptively simple physiological function despite being one of the most thoroughly studied behaviours in the animal kingdom [149] the complexity behind its mechanism of action is still unclear. The primary motor cortex - known as Brodmann Area 4 (BA4) or sometimes referred to as M1, is a large structure located in the posterior portion of the frontal lobe and is part of the precentral gyrus. It posteriorly borders the precentral sulcus, anteriorly borders the central sulcus, begins medially at the medial longitudinal fissure, and ends at the lateral sulcus. It is also found in the frontal lobe of eutherian mammals [150]. It is widely accepted that M1 is the principal driver behind generating motor signal transmission to initiate movement, while Brodmann Area 6 assists in this function. BA6, which is located anteriorly to BA4, is commonly divided into the premotor area (PMA) (anterior to M1) and supplementary motor area (SMA) (dorsal to the premotor area).

2.3.2 Corticospinal Tract

The corticospinal tract is a white matter tract that starts in the cerebral cortex and terminates at lower motor neurons or interneurons in the spinal cord. Briefly, efferent axons from M1 merge with the axons from BA2, BA3, BA6, and BA8 to form the corona radiata [151]. Fibres originating from the corona radiata project

through the posterior limb of the internal capsule and arrive at the cerebral peduncle where they enter the basilar portion of the pons. At the pontomedullary junction, axons continue and form part of the medullary pyramid; at the base of the pyramids, 75-90% of axons decussate and merge with the lateral funiculus of the spinal cord, forming the lateral corticospinal tract and synapsing with their intended targets [152]. The remaining axons at the base of the pyramids descend into the ipsilateral anterior funiculus of the spinal cord to form the anterior corticospinal tract where axons decussate shortly before reaching their targeted neurons [151, 152] - see Figure 2.5. To facilitate movement, neurons within M1 passes an action potential along the corticospinal tract to innervated interneurons, or directly to spinal motor neurons; from there, signals are passed into the peripheral nervous system where they elicit muscle contraction via neuromuscular junctions.

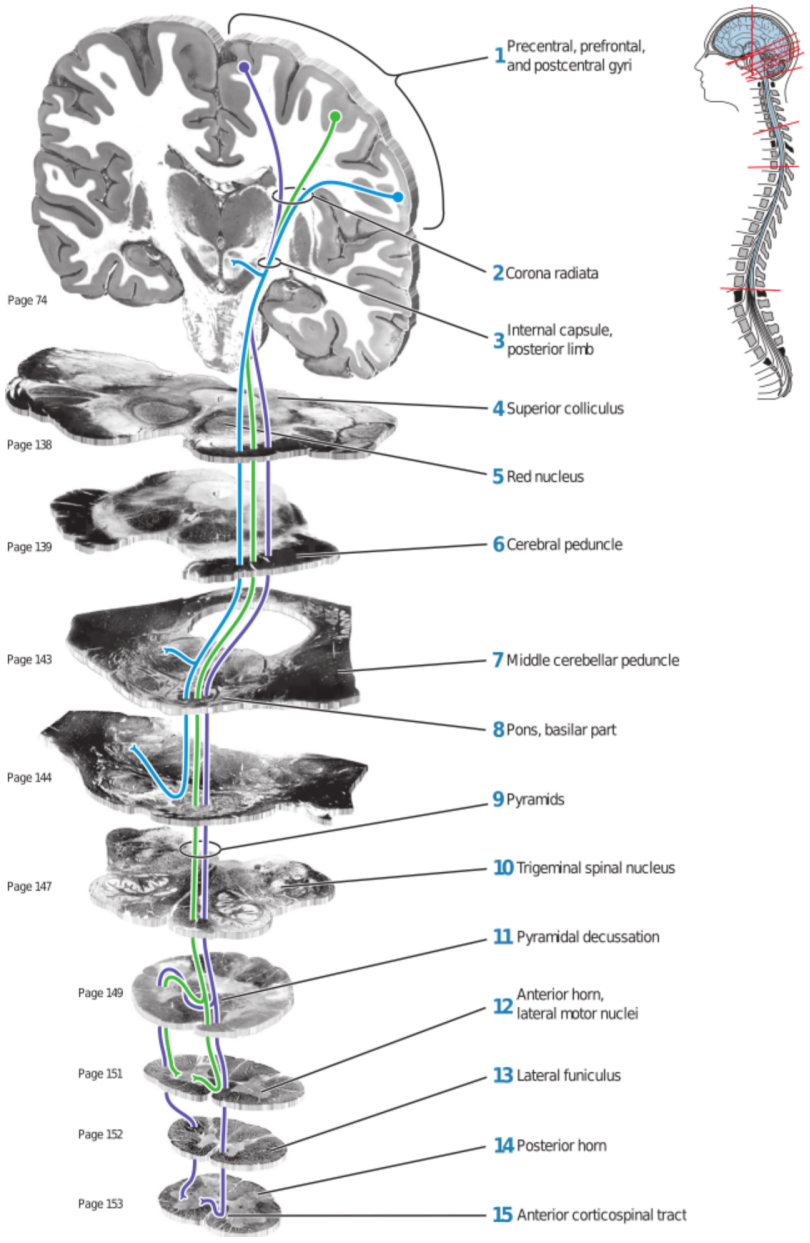


Figure 2.5: The Corticospinal Tract.
The different coloured projection maps represent the different routes motor neuron axons project to, depending on their location in the primary motor cortex. Blue: Motor neurons controlling the face, mouth, jaw, tongue, and throat. Purple: Projection map of neurons that lower extremities. Green: Projection map of neurons that control the upper extremities. *(From [153], with permission from John Wiley and Sons*).

2.3.3 Microcircuitry of the Adult Primary Motor Cortex

Like other areas, the primary motor cortex is comprised of a multitude of intra- and inter- cortical afferent and efferent projections. The principal role of this cortex is to generate the signals needed to initiate movement, the circuitry for which will be briefly explained here. In general, pyramidal neurons can be broadly divided into categories based on the destination of their axonal projections. For example, within M1, intratelencephalic (IT) neurons are found in all neuronal layers and project ipsilaterally and contralaterally to the striatum, including the putamen, caudate nucleus, and globus pallidus; forming a looped circuit that regulates basal ganglion circuitry that ultimately controls the inhibition of the thalamus to the primary motor cortex, allowing for feedback control over motor generation [154]. Another group of pyramidal neurons are extratelencephalic (ET) neurons that project outside of the telencephalon. Importantly, within the motor cortex are a subclass of ET neurons that project to the spinal cord, these Pyramidal Tract (PT) neurons form the corticospinal tract discussed previously in this chapter.

Gilbert and Wiesel (1983) provided the first account of cortical circuitry by studying excitatory neurons of the visual cortex [155]. They proposed that thalamic input is received by L4 where excitatory neurons project to superficial neurons in layers 2 and 3 which in turn project to L5. Finally, L5 neurons project to L6 neurons which then project both to L4 [156] and the thalamus [16, 157]. This concept was expanded by Douglas and Martin (2004) who implemented the circuity of inhibitory neurons into Gilbert's and Wiesel's model. Furthermore, Douglas and Martin determined that this circuitry was canonical across cortical areas [16, 158], see panel B in Figure 2.6. According to these models, thalamocortical input is typically received in L4, which is problematic within BA4, as the existence of an L4 in the primary motor cortex remains controversial [154, 159, 160]. Several theories have been proposed to explain the microcircuit of the agranular BA4. Some believe there exists a small buried L4 making the microcircuitry of BA4 canonical to other cortical areas, while others believe M1 has a unique microcircuit, where extrinsic

input is not received by the thalamus but instead from neighbouring somatosensory areas [161]. Nevertheless, once excited by thalamocortical input the primary motor cortex produces a signalling cascade that begins in the superficial layers of the cortex and results in processed information exiting M1 via cortico-striatal and corticospinal projections, resulting in the generation of movement, reflective of the circuitry model originally proposed by Douglas and Martin. The signalling cascade to generate movement initiating signals is as follows, excitatory input[4] is received by IT neurons in L3 and are processed by both L3 and L2 IT neurons via mutual projections forming a superficial cortical loop [154]. L2 IT neurons then project to L5a IT neuron and L5b PT neurons, L3 also project to these L5 neurons. Equally, some L5a IT neurons also project upwards back to L2/L3 neurons to regulate signal processing. As with the superficial loop, L5 IT neurons form a deep cortical loop with both divisions of L5. This deep cortical loop includes unidirectional projections to PT neurons; see panels C and D in Figure 2.6. PT neurons primarily project to the corticospinal tract to facilitate movement but will also send axonal collaterals to the ipsilateral L5 or striatum, see panel A in Figure 2.6. IT neurons in L2/L3 can also project contralaterally as well as ipsilaterally, as such IT neurons in L5a can receive intra- and inter-hemispheric inputs from L2/L3 IT neurons, while PT neurons only receive their input from the ipsilateral L2/3.

Within L5 there is an asymmetry in its circuitry, IT neurons have a strong relationship with intracortical layer IT neurons and PT neurons, while studies in the mouse brain show that PT neurons have minimal recurrent connectivity among themselves with few projections to IT neurons [162]. This results in PT neurons being downstream to IT neurons making PT neurons in L5 responsible for the main output of the primary motor cortex. In addition to sending signals via the corticospinal tract, neurons in L5 also project to L6 which continue to transmit processed information to the thalamus either directly or via connections to L4, however, this L6-L4-thalamic route is still debated in M1 due to the L4 controversy.

[4]either originating from the thalamus or associated sensory neocortical regions

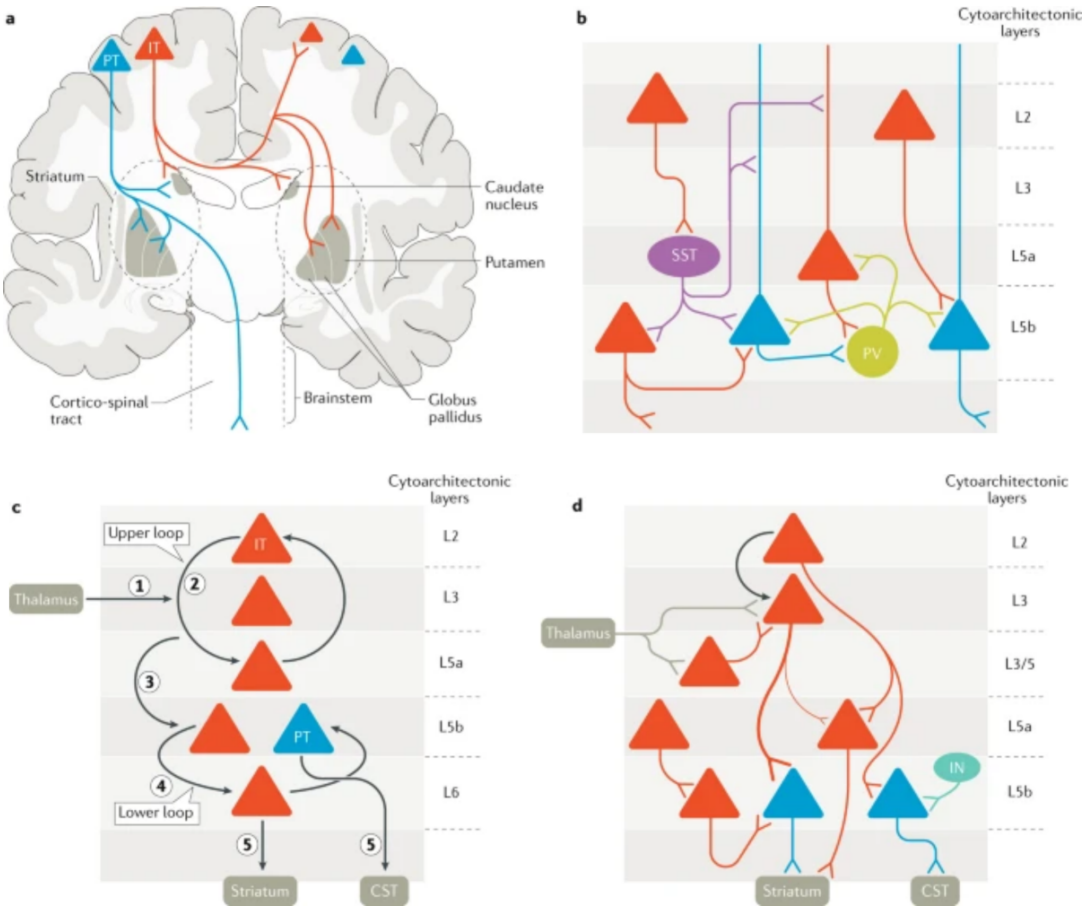


Figure 2.6: The Assumed Microcircuit of the Human Primary Motor Cortex. a: IT and PT neurons have specific long-range connectivity. PT neurons send axons to the corticospinal tract, however, PT neurons can also project to various targets within the telencephalon. IT neurons connect to both the ipsilateral and contralateral striatum and cortex. b: Simplified diagram of inhibitory interneurons within the primary motor cortex. c: Layer specific information processing within the primary motor cortex. d: Cell-specific information processing within the primary motor cortex. Orange triangle: IT neurons, Blue triangle: PT neurons (including Betz cells), Circle nodes: inhibitory neurons (INs). SST: Somatostatin, PV: Parvalbumin. The thickness of axons represents the weight of information processing. *(Taken from [154] with permission from Springer Nature.)*

Along with the pyramidal neurons there exist a magnitude of inhibitory interneurons that allow for feedback and regulate control over the cascading signals in M1 [154].

2.3.4 The Cellular Diversity of the Primary Motor Cortex

The Allen Institute for Brain Science has recently applied elegant and complex next-generation single nuclei sequencing to explore the evolution of cellular diversity in the primary motor cortex in the mouse, marmoset, macaque, and human brain [163]. This data is only available as a preprint and was not available at the beginning of our research. This provides us with the unique opportunity to compare cell-sorting derived neuronal taxonomy data with our targeted data. This body of work implemented high-throughput transcriptomic (SMART-Seq v4, SSv4, Cv3, and Illumina RNA-sequencing technology) and epigenomic (snmC-seq2, SNARE-seq2) techniques on nuclei isolated from individual neuronal layers of M1. In total, they were able to profile >450,000 single nuclei. In the human M1, they were able to identify and cluster 116 transcriptomically discrete neuronal cell-types and 11 non-neuron cell-types. The authors found that although the molecular identity of cells in M1 was conserved across species there remained several species specialisation. In humans, the glutamatergic projection/GABAergic interneuron ratio in M1 was 2:1, whereas this increased to 3:1 in the marmoset and 5:1 in the mouse while the relative proportion of GABAergic cell-types was similar across species; the authors argue that the relative ratio difference seen may be a result of a shift in the size of the GABAergic progenitor pool in the ganglionic eminences or, an elongated period of neurogenesis and migration.

The authors also note, as with their previous study performed on the human middle temporal gyrus [30], that subtypes of glutamatergic neurons were specific to neuronal layers. In their M1 study, they found that GABAergic neurons at times also displayed laminar patterning across the cortex. For example, they observed a subset of caudal ganglionic eminence (CGE)-derived GABAergic neurons expressing LAMP5/SNCG that were restricted to L1. Whereas, they also noted medial ganglionic eminence (MGE)-derived GABAergic cell-types expressing somatostatin

or parvalbumin showed laminar distributions, and those expressing vasoactive intestinal polypeptide were less specific to neuronal layers. It was also reported that 3 astrocyte subtypes were found in varying abundance and had specific layer distributions. Consistent with other reports [164] the authors found that 33% of cells identified were GABAergic, compared to the 23% and 16% in the M1 of marmosets and mice, respectively. Although the overall proportion was uniform in each species. They also found significantly more L2 and L3 IT neurons in the human M1. Interestingly, they observed that L5 ET neurons were found at a significantly smaller proportion in primates than in mice. The authors also found that L6 cortico-thalamic neurons in primates were found in numbers less than half of that compared to the mouse, however, this could be a result of the relative dilution of cell populations with the expansion of the neocortex in primates. Fascinatingly, the authors found a unique oligodendrocyte population in humans (Oligo L2-6 OPALIN MAP6D1) that appear to contain a variety of glutamatergic neuron-specific transcripts which was validated with FISH. Overall, gradual changes in individual cellular genotypes aligned with evolutionary distance. For example, human M1 was more similar to the macaque than to the marmoset, and were most dissimilar to the mouse, representing the divergence from our distant common ancestor $\approx$ 25 million, $\approx$ 40 million, and $\approx$ 70 million years ago, respectively.

2.3.5 Layer 5 of the Primary Motor Cortex

Layer 5 is the executive layer of the primary motor cortex and is responsible for sending motor initiating signals to the spinal cord. Layer 5 is commonly divided into two sublayers: L5a and L5b. L5a, like the superficial layers, contains IT neurons that project across the cortex and striatum, while L5b contains both IT neurons and PT neurons that project to the brain stem and spinal cord, whose axon collaterals innervate the striatum and thalamus [165]. These PT neurons provide a functional link between M1 and motor centres in the brainstem and spinal cord [166]. It has been suggested that the preparatory activity that occurs prior to generating movement is triggered by the thalamocortical loop in the motor cortex

[167, 168]. In the rodent vibrissal motor cortex (vM1), it has been shown that the motor thalamus projects to L1, L2/3, L5a, L5b, and L6 [165] and that L2/3, L5a, and L5b received the strongest input from the motor thalamus while L6 receives significantly less. In this study, the authors found that thalamic input to L5a was perisomatic, while L2/3 neurons received thalamic input both via basal dendrites below their soma and from their apical dendrites in L1. The authors also found that the secondary motor cortex (M2) and orbital cortex strongly project to PT neurons in L5b. Moreover, this study showed that within L5b there existed two distinct PT neuronal populations that received differential input. Both received anterior thalamic inputs, however, upper L5b PT neurons received strong input from L2/3 while excitatory input from M2 was strongest in lower L5b PT neurons. The authors concluded that upper L5b PT neurons were most influenced by upstream circuits in their local circuit, whereas, lower L5b PT was most influenced by upstream circuits from other cortical areas and were involved in cognitive aspects of motor behaviour.

The possibility of multiple types of PT neurons in L5 was confirmed in the rodent brain by snRNA-sequencing [167, 169] and clustered with the following transcriptomic groups: Slco2a1, Npsr1, and Hpgd. In this dataset, the authors found that PT neurons either innervated the thalamus or innervated the motor centres of the medulla. All PT innervated cell-types examined innervated the superior colliculus. Thalamus projecting PT neurons were found in the upper L5b, while lower L5b PT neurons projected to the medulla; this was uniform in both the primary and secondary motor cortex. 100% of medulla projecting PT neurons mapped to the Slco2a1 cluster, the remaining PT neurons projected to the thalamus, 20% mapped to the Npsr1 cluster and 76% mapped to the Hpgd cluster. Finally, the authors found that PT neurons in upper L5b showed early and persistent preparatory activity for generating movement. Lower L5b PT neurons instead contained late preparatory activity which is involved in controlling movement initiation and termination. This segregation of PT neurons may exist in other mammals [170, 171] but has not been extensively confirmed in humans.

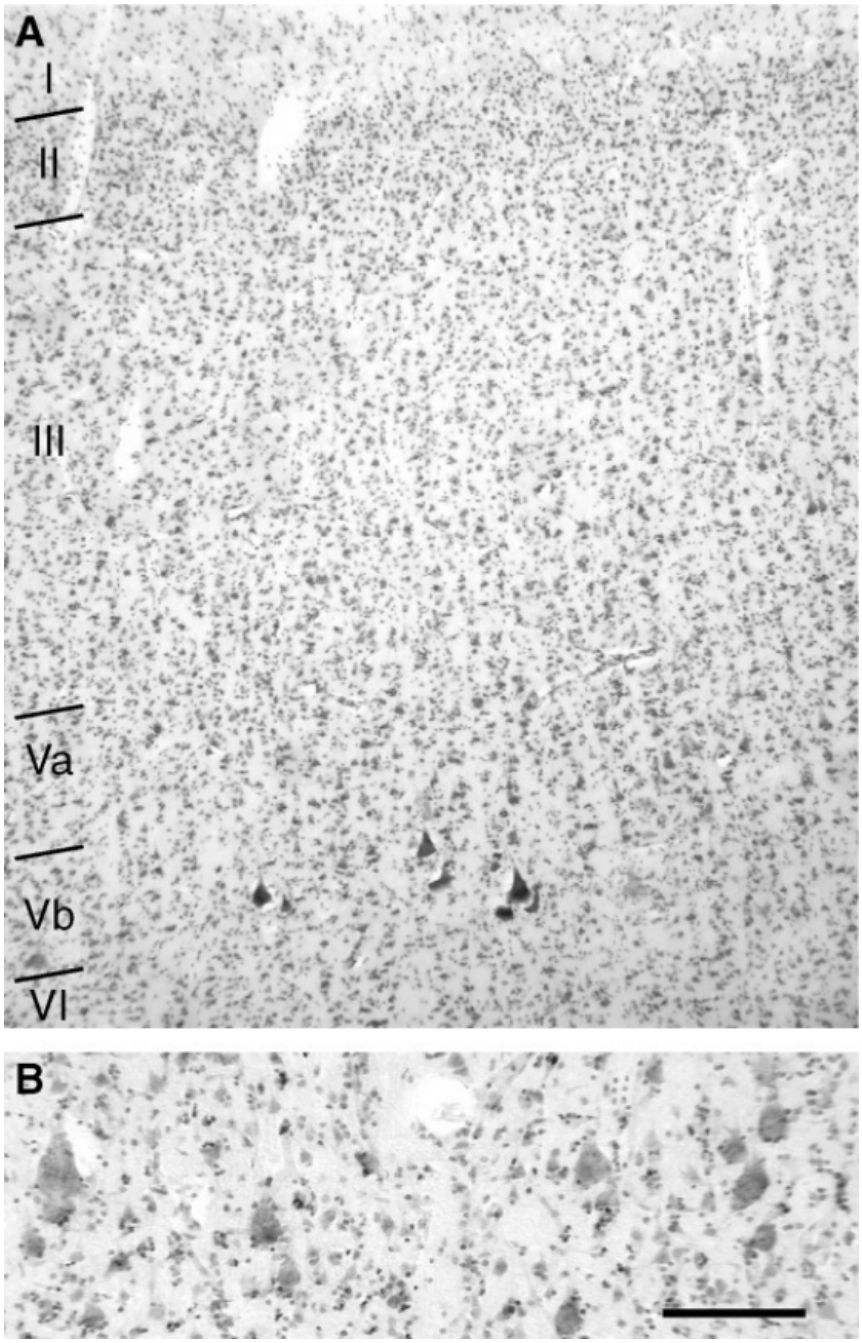


Figure 2.7: The Cortical Layers in the Human Primary Motor Cortex.
A: Nissl preparations of the primary motor cortex reveals its laminar cortical layers, including its assumed lack of layer IV. Betz cells with Layer Vb are the most obvious cytological feature in the primary motor cortex. B: Betz cells are found scattered throughout layer Vb, and can be identified by their size, location, and distinct morphology. (*From [99] with permission from John Wiley & Sons*).

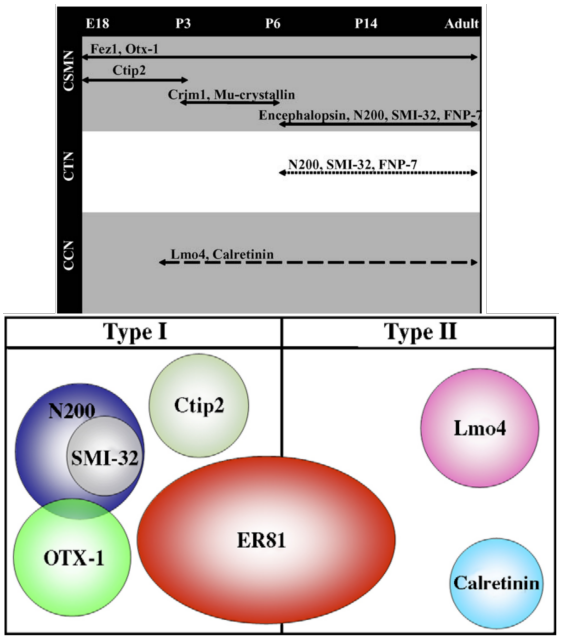


Figure 2.8: Assumed Molecular Classification of L5 Pyramidal Neurons. A collection of molecular markers that can immunochemically identify sub-populations of L5 pyramidal neurons across the neocortex, and the overlap of expression of these markers between sub-populations. (*From [172] with permission from Elsevier*).

In addition to the above molecular definitions, L5 pyramidal neurons can also be grouped into multiple groups across the neocortex based on morphology [172]. Type I L5 neurons project to the basal pons, spinal cord, or superior colliculus and can be defined by their thick tufted apical dendrites and fire in bursts [173] whereas, type II L5 projection to the contralateral hemisphere, or the ipsilateral striatum. Type II neurons have an apical dendrite which is slender with few branches and lack terminal tufts, their apical dendrite is normally found in L2/3, and they never fire in bursts [173]. The key to studying these subpopulations of Layer 5 pyramidal projection neurons (or any neuronal population) is to identify biochemical entities (biomarkers) that may be used to classify rather than relying solely on morphology which may be similar among transcriptomically distinct populations. Reliable biomarkers exist for GABAergic interneurons [172] and they can be aggregated into three groups depending on their expression of parvalbumin, somatostatin, or 5HT3a [174] while markers for glutamatergic, especially those that are neuronal layer-specific are rarer; Molnár & Cheung 2006 provides a detailed review of known L5 markers in the

neocortex - see Figure 2.8. For example, it has been shown that within Type I-Layer 5 neurons there exist at least two neurochemical subpopulations that either express N200 or ER81, but not both. The discovery of new biomarkers and their co-localisation with known biomarkers will be imperative to further delineate the heterogeneity of L5 neurons. In several species, L5b of the primary motor cortex contains the gigantic cells of Betz. These Betz cells are morphologically unique, the principle neuron of M1, crucial for generating motor movements, and degenerate in motor neuron diseases. Despite this, little is known about them, likely due to their seeming non-existence in common animal models (mice/rats). Furthermore, there currently exists no robust biomarker that selectively targets Betz cells, limiting basic research, and leading some to believe that Betz cells are simply large PT neurons and not a unique subpopulation of neurons. The markers proposed by Molnár & Cheung 2006 may provide insights into the biochemical nature of the Betz cells, however, surprisingly many of these markers have not been tested in the human motor cortex with the exception of neurofilament, which while useful in non-selectively highlighting Betz cells immunochemically fails to capture unique biochemical differences between Betz cells and surrounding neurons.

2.3.6 Betz Cells

Vladimir Betz was a Ukrainian scientist who studied many aspects of physiology, including the field of neurophysiology. In his career, Betz's studied $\approx$ 9,000 brains and spinal cords from healthy and diseased individuals along with infant and animal specimens [175]. His greatest accomplishment was the discovery of giant pyramidal neurons in the motor cortex:

> "The sulcus of Rolando divides the cerebral surface into two parts; an anterior in which the large pyramidal nerve cells predominate... They are predominantly in the fourth cortical layer [5] and are from 0.05-0.06mm wide and from 0.04-0.12mm long... undoubtedly these cells have all the attributes of so-called 'motor cells' and definitely continue as cerebral nerve fibres". [175]

[5] The concept of neuronal layers had not been fully established during Betz's time; we now know that Betz cells are found in Layer V. Betz's erroneous conclusion is likely due to the agranular nature of the primary motor cortex.

The literature surrounding Betz cells mostly matured in the mid to late 20thcentury. At times Betz cell research can be sparse and spans several scientific disciplines with limited attempts to collate findings. Unfortunately, the most recent Betz cell review was performed in 1942 [176]. A comprehensive review of the Betz cell literature in the light of modern techniques/observations has yet to be reported, and a full review of the literature is not possible within this thesis. The most relevant aspects of Betz cell physiology will be discussed below e.g. morphology and known biochemical data, however, additional information on electrophysiology, comparative neurology, and the role of Betz cells in non-ALS diseases can be found in Appendix B.

Betz Cell Morphology -

Betz cells or "cells of Betz" are large presumed glutamatergic pyramidal neurons located almost exclusively in the primary motor cortex, and are one of the rare examples where a single neuronal cytological feature can be used to define a specific neuroanatomical region. The term 'Betz cell' is traditionally used to describe the gigantopyramidal motor neurons of the human and has been extended to the large gigantopyramidal neurons in the motor cortex of other primates. While many other mammals also contain gigantopyramidal motor neurons with the same or similar morphology, they are not always referred to as such, potentially adding to the dogma that Betz cells are not their own independent cell-type. Nevertheless, some non-primate animal models have adopted the Betz cell nomenclature [177–179]. The lack of biochemical data and extensive electrophysiology data has made it difficult to confirm if non-primate gigantopyramidal neurons and primate Betz cells represent the same neuronal niche. Despite this, for convenience, when referring to species containing gigantopyramidal neurons with strict Betz cell morphology, they will be named as such in this thesis. Moreover, there exist several investigations in the literature that study and compare the morphology of Betz cells with gigantopyramidal motor neurons of other species, as such a concise summary of these similarities and differences can be seen in Appendix B.1.

The Betz cell is the largest cell in the human central nervous system and can dwarf their surrounding neurons by a factor of 20 [99]; although estimates of Betz cell size vary, Rivara et al. (2003) noted that Betz cells are slightly smaller than what Vladimir Betz proposed at 53 x 106 µm; however, it is important to note that Betz cell sizes differ considerably even within the same coronal plane, and may explain this difference in results. The reasoning behind their large size has not been established; it is believed that, like other neurons, their soma size is reflective of the metabolic and biosynthetic requirement of the whole cell, including its axon [180]. Some believe Betz cell size is proportional to their axonal length or size of the targeted muscle and supported by the fact that some of the largest Betz cells are found in the leg/foot area of the homunculus [99]. Although plausible, many species possess disproportionately large Betz cells despite their size and in the example of the giraffe brain - whose Betz cell would be expected to be the largest in the animal kingdom due to axonal length - their Betz cell sizes are not disproportionally larger than expected [181, 182]. Others have shown that there is a correlation between human height and Betz somal size, but not with bodyweight [183]. The axons of Betz cells project into the corticospinal tract, where they directly synapse with alpha motoneurons in the spinal cord [184]. Historically, it was thought that the axons of the CST solely originated from Betz cells, however, it is now known that Betz cells axons only represent 2-3% of the 30% CST fibres coming from the primary motor cortex [185, 186]. While not solidified, it is assumed that Betz cells are likely to be the origin cell-type responsible for the largest fibres in the CST [172, 185–187]. It is also theorised that Betz cells are not purely PT neurons and may also have axonal collaterals that project to other regions, however, there is little literature on this and has not been molecularly investigated.

Surprisingly, despite their prominent appearance in the neocortex there is relatively little known about Betz cells compared to other well-studied large neurons, such as the Purkinje cell, Meynert cell, or von Economo neurons. Although they have a striking appearance in the primary motor cortex, the most defining feature of the Betz cell is their unique somatodendritic morphology – containing dendrites

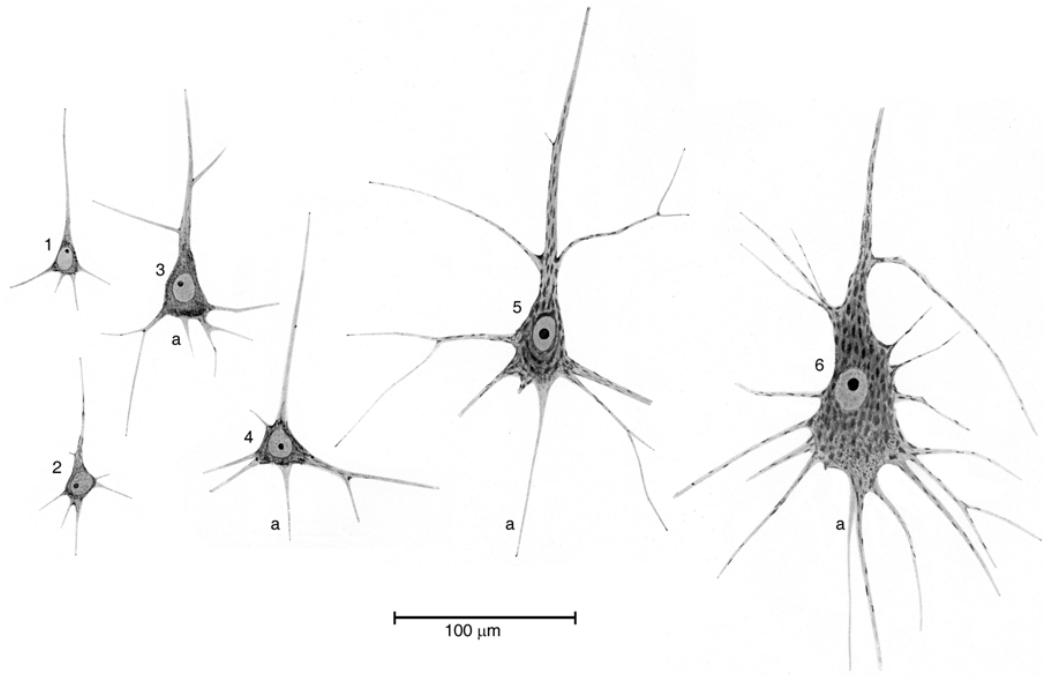


Figure 2.9: Betz Cell Morphology differs from other Pyramidal Neurons. 1. The largest pyramidal neuron from L3 of the Primary Visual Cortex. 2. Stellate cell from L6 of the Visual Association Area. 3. Large pyramidal neuron from L3 of the Visual Association Area. 4. Meynert cell from L5 of the Primary Visual Cortex. 5. The largest pyramidal neuron in L3 of the primary motor cortex. 6. The Betz cell; note their increased total size compared to other neurons also called "giant". In addition to the increase in basilar dendrites, note that somatodendrites extend around the entirety of the Betz cell perikaryon - a defining feature of the Betz cell. a = axon. *(Taken from [188] who adapted it from from [189], with permission from Springer.)*

that arise around the circumference of the perikaryon rather than localised solely to the basal part of the soma [190–192] - see Figure 2.9. However, it can still be challenging to correctly identify Betz cells from surrounding neurons as L5 also contains large neurons with non-Betz cell morphology [176, 177, 185, 191, 193–195]. These non-Betz cell PT neurons are of a similar or overlapping size to the smallest Betz cells, and without robust Betz cell markers, these neurons are commonly misidentified as Betz cells, confusing Betz cell-specific research, or limiting Betz cell investigations to only studying the largest Betz cells in the neocortex. Betz cells are commonly grouped with other motor neurons as the umbrella term of *corticospinal upper motor neurons*, which while true, often means they are overlooked in studies, or that conclusions drawn from such studies are reflective of a merged phenotype

between Betz cells and other L5 neurons [196, 197]. This is particularly problematic in ALS research, where Betz cell-specific pathology has been shown to differ from other motor neurons [184]. A possible reason behind this confusion is that most studies exploring the cytoarchitecture of the brain use a Nissl stain (see Figure 2.11) to stain for neurons. Since this highlights neuronal soma and Nissl-dense dendrites rather than the full dendritic tree of a neuron; the most defining feature of Betz cells are not visualised. As such, other morphological features need to be observed to confirm the identification of Betz cells. These Betz cell-specific features can be subtle and are often ignored; instead, Betz cells are commonly identified solely based on their soma size and location in L5 of the motor cortex, which when used as the sole criterion can be deceptive, for the reasons outlined above. For an overview of the morphological features that best discern Betz cells from other neurons, see Appendix A.

Another key identifying feature of Betz cells in the presence of large accumulations of intracellular lipofuscin [198] (see Figure 2.10) and the common observation that Betz cells - at least in the human brain - tend to cluster to form cumulative nests of up to 5-6 cells [199] (see Figure 2.10). Other important features include: having a smaller nuclear size in proportion to their cytoplasmic volume [186] and containing a higher number of dendrites (>7) than neighbouring neurons which typically possess 4-7 [176, 191, 200, 201].

In humans, Betz cell dendrites are not symmetric and are orientated parallel to the main axis of the gyrus [201] forming a dense array of dendrites around the base of the Betz cell [190, 191, 201, 202]. Interestingly, while still unique, total Betz cell morphology varies across the motor cortex; showing slight inter-regionally variation with some Betz cells appearing to be bipolar (see Figure 2.11) with their basal dendrites entering the white matter [201]. In a variety of species, including humans, Betz cells have also been seen to occasionally form long taproot dendrites that descend deep into the white matter [190, 192, 203], which has commonly been proposed as a feature of fast-firing PT neurons [204].

Histologically, Betz cells can be identified by their abundance of Nissl bodies and large accumulation of lipofuscin [184]. Lipofuscin accumulation is commonly found in neurons in the ageing or diseased brain but is commonly seen in Betz cells of patients under 65 [192] and in the ageing primate brain [195]. When investigating the ultrastructural features of macaque Betz cells using electron microscopy several features were discovered; Betz cells appear to have a complex indentation around the circumference of their nuclear membrane, often accompanied by clusters of endoplasmic reticula and mitochondria; along with an abundance of neurofilament [177, 198, 205]. The authors also found an abundance of Nissl bodies, which were highly concentrated in the nuclear membrane folds, beneath the plasma membrane, and at the base of dendrites. It has been suggested that Betz cells prolifically produce ubiquitous neurofilaments and that this should be used as a criterion for correctly identifying Betz cells. Another ultrastructural investigation of the macaque Betz cell found novel inclusion bodies that the authors concluded contained 10 nm filament as their fundamental unit, and whose function was unknown [206]. Within the domestic cat, it has been found that glial cells border the soma of Betz cells, more so than the base of their dendrites [177]. These glial cells are found in high concentration around Betz cells; the authors found that in any 1 µm section they routinely observed 2-3 neuroglial cells around Betz cell soma, which contributed to equal numbers of astrocytes and oligodendrocytes. It was also noted that the non-synaptic surface of Betz cells is covered by their processes although astrocytic processes contribute to a higher proportion [177]. These astrocytic processes also tend to form sheets that extend around Betz perikaryal axon terminals to isolate them from potentially interfering chemicals in the neuropil.

Another prominent feature of Betz cells is their predominate apical dendrite, unlike other pyramidal neurons, the Betz apical dendrite often bifurcates near the soma into daughter branches of similar diameter [177, 204] that extends towards the superficial layers of M1 [199, 201]. Dendritic spines are seen across Betz cell processes and bulbous protrusions have been described around their perikaryon [199, 201], although the shape, size, and the number of Betz cell dendritic spines vary

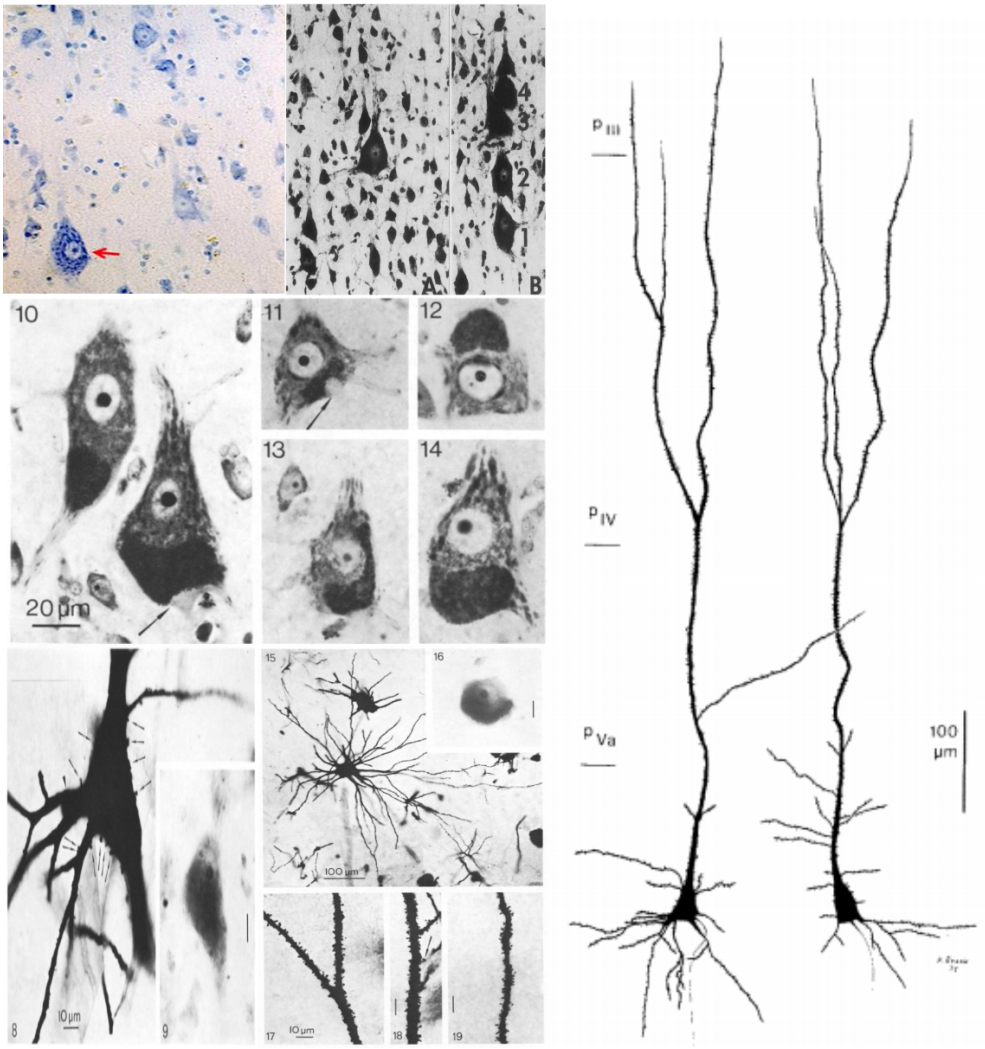


Figure 2.10: Overview of Betz Cell Morphology.
TL: Betz cells can be identified by Nissl preparations due to their abundance of Nissl bodies. TM: Betz cells can also be identified by their large size in respect to surrounding neurons, they are also found in characteristic 'nests' or clusters. 8: Betz cell soma is decorated with spines and proximal dendrites. Basal dendrites are broad initially and then branch abruptly. 9: lipofuscin accumulations are often found next to or below the nucleus. 10-14: highlights several features of Betz cell morphology; including an abundance of lipofuscin, the axon hillock devoid of lipofuscin, dense Nissl bodies throughout the cytoplasm that become elongated and align with the ascending apical dendrite. Moreover, the prominent nucleolus and nucleus seen in these micrographs are also defining features of Betz cells. 15: View beneath a Betz cell, showing the large area covered by their long dendrites. 17-18: the apical dendrites of Betz cells are densely populated with spines of various shapes, including unusually large spines (arrow), the bifurcating nature of the Betz cell apical dendrite can also be seen. 19: Densely heterogeneous spines can also be seen on the basal dendrite of Betz cells. Large camera lucida drawings: showing bifurcating apical dendrite, circumferential somatodendrites, and long basilar dendrites. *(Taken from [207] with permission from Springer).*

considerably between and within individuals [201]. In the cat, Kaiserman-Abramof and Peters (1972) commented that Betz cell dendritic spines are similar to other pyramidal neurons although they note that some Betz cells have dendritic spines that are unusually large – with an appearance similar to a large axon terminal with multiple postsynaptic densities. Conversely, a study of giant pyramidal neurons of the motor cortex in several mammals [190] found that Betz cells contained fewer spines than other pyramidal neurons in the motor cortex but warned of methodological errors in their approach that may underestimate total dendritic spine numbers. In the giraffe brain, Betz cells are highly spiny [182] whereas a feline study by Deschenes et al., (1979) showed the paucity of spines in the apical dendrite of Betz cells [204], which differs from the numerous spines seen in apical dendrites in other pyramidal neurons, which is a feature also seen in the human motor cortex [201]. Historically, it has been observed that the number of dendrite spines in the apical dendrite is reduced $\approx$ 50-150 µm from the soma [208] and this reduction is a result of the early bifurcation seen in the Betz cell apical dendrite. Additionally, Deschenes et al., (1979) also noted strange glial fragments along the apical shaft of the Betz cell, which may refer to the Betz dendritic spine variation noted by Meyers (1987) in the human brain. In this, Meyers states that the dendritic spines of Betz cells are considerably variable and suggested that Betz cells display a higher degree of individuality than other pyramidal neurons. Although, the complexity of the autolysis period of post-mortem brains may have had a pronounced negative effect on spine numbers in Golgi preparations [209] and may underrepresent the true number of spines in the human neocortex.

Betz cells also contain a large number of somato-axonal terminals, however, these are not homogenously dispersed around the perikaryon, instead, it has been shown - in felines - that these terminals may cluster and cover between 10-80% of the Betz perikaryon in a particular area [177]. The authors crudely calculated that on average $\approx$ 24% of the Betz cell soma is covered by axon terminals and determined that the average Betz cell perikaryon bears approximately 870 axon terminals. Interestingly, in one study investigating the ultrastructural nature of Betz cells, the authors found

large bundles of filaments (20-25nm) thicker than that of neurofilaments in healthy human Betz cells, however, they were unable to resolve their identity [198]. These authors also noted the existence of lamellar structures, that have previously been seen in Betz cells and anterior horn spinal motoneurons with ALS. Other neurons such as Purkinje cells also contain these lamellar bodies, along with non-neuronal cells such as oocytes, pituitary adenomas, germinomas, and other neoplasms [210, 211] although their role in health and disease is not well established.

Rivara et al., (2003) performed one of the most comprehensive stereological studies of the Betz cells from 6 human cases and found a total mean of 125,290 Betz cells with significant inter-case variability which is larger than the estimates of approximately 25,000 – 40,000 per hemisphere [212, 213]. It is important to note that the distribution of Betz cells across M1 is not uniform, but follow a mediolateral gradient around the precentral gyrus. The largest Betz cells are found in the foot, toe, and leg regions of the homunculus and become slightly smaller and densely populated within the upper limbs of the homunculus. There is a sharp decrease in size and number as M1 curves ventrolaterally towards the lateral fissure - see Figure 2.11. The majority of Betz cells (75%) are found in the dorsal third of M1 (lower limb representation), 18% in the medial third (upper limbs), and 7% in the inferior third (face region) [99, 213]. Lassek also proposed that the majority of Betz cells (93%) are found in the cortical surface within the first third of the motor strip [213]. It has been stated that motor neuron size is proportional to the size of the muscle it controls, or the size of the muscle movement however, while the Betz cells in the leg area are considered the largest, this is only true when the average cell size is compared across the motor strip. Within the leg area, Betz cell size varies greatly, despite presumably containing the longest axons [213]. The morphological differences across M1 suggest the existence of several Betz cell subpopulations, each with potentially specific functional/metabolic features which could suggest potential molecular differences.

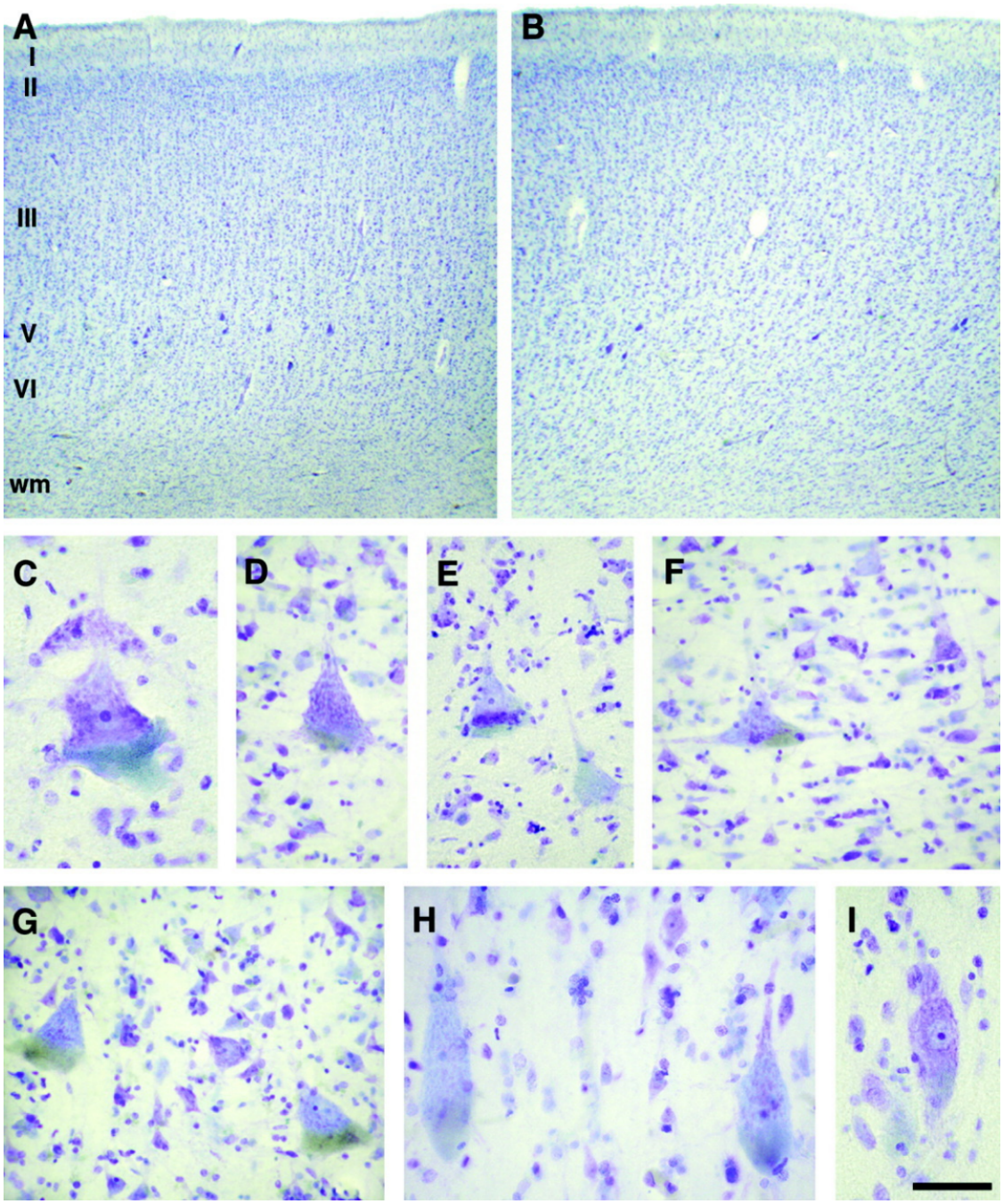


Figure 2.11: The Varying Morphology of Betz Cells across the Primary Motor Cortex. At the depth of the central sulcus (A), near the border with BA3a, Betz cells appear rounded with compressed perikarya (C-F), however, at hand-knob (G) they appear triangular, and at the rostral boundary of BA4 (B) near the border of BA6 (H and I) they become elongated. *(Taken from [99] with permission from John Wiley & Sons).*

The Golgi stain highlights the full morphology of the Betz cell, making identification significantly easier. However, Golgi staining has fallen out of fashion due to the cumbersome nature of the chemicals used in the staining process and difficulties with consistency and quantification. Instead, it has become commonplace to use Nissl staining to highlight just the Betz cell perikaryon and use size and location as sole determinates of Betz cell identification. As previously mentioned, this had led to some studies grouping all pyramidal neurons in Layer 5 as Betz cells or upper motor neurons, despite morphological differences and potentially different molecular composition and function. This is further complicated by extrapolating from rodent studies of brain structure, motor function, and neurodegenerative diseases since, as classically defined, rodents do not contain Betz cells [214, 215] although this is not a prevailing dogma as some studies use the term corticospinal motor neurons/upper motor neuron and Betz cell interchangeably, while others do refer to a subpopulation of motor neurons in the rodent brain as Betz cells [216]. Vague descriptions of how Betz cells are studied in human motor cortex investigations, lack of biochemical data, and no unifying view of Betz cells as a separate motor neuron subclass have resulted in many believing the Betz cells are simply larger counterparts of surrounding PT neurons rather than their own subtype. This concept can confuse findings when extrapolating observations to the human brain from rodent models; ultimately making the literature surrounding Betz cells blurred and vague.

Molecular Composition of Betz Cells -

One of the largest factors impacting why relatively little is known about the Betz cell is the lack of next-generation *-omics* directed investigations that has been directly applied to reveal their biochemical composition. In 2019, we published our research in developing a workflow for directly targeting and isolating CNS neurons for high-depth proteomics, and although primarily a technical paper, it contains the first and – at the time of writing - only data into the proteomic nature of the Betz cell [217]. The data has been made publicly available at ProteomeXchange -

PXD012101[6]. Interestingly, during the time of writing this thesis, there has also been an impressive preprint in 2020 by the Allen Brain Institute, that extensively applied cutting-edge transcriptomics to the mouse, marmoset, and human primary motor cortex [163], which among other questions, aimed to identify the transcriptomic profile of the Betz cells. For a small description of the little molecular information that is known about the Betz cell - see Appendix B.3.

In their cellular diversity of M1 preprint, one of the Allen Institute's aims was to use unbiased transcriptomic approaches to identify a Betz cell-specific transcriptomic profile/cluster within their data. The authors state that they were not able to identify a Betz specific cluster, instead they found clusters of L5 glutamatergic neurons that they believe also contained cells with Betz cell morphology-based on PATCH-seq comparative analysis of macaque and human Betz cells. The authors used this to confirm the presence of two human Betz cell-containing clusters, rather than Betz cell-specific clusters: Exc L5 FEZF2 ASGR2 and Exc L5 FEZF2 CSN1S1. When comparing ET neurons and IT neurons, they found ≈ 50 differentially expressed genes that were conserved across all 4 species. Interestingly, within this gene set, they found several enriched genes encoding for ion channel subunits in the ET neurons, potentially representing evolutionally conserved ET physiological properties. The authors found primate-specific genes that encoded for potassium and calcium channels and argued that this could represent Betz cell or primate-specific ET physiology[7]. From their L5 ET vs IT comparison, the authors also found a collection of genes that gradually increased in abundance and ET specificity as they became more related to humans. To explore this further, they mapped genes as a function of evolutionary distance from humans, which resulted in identifying 131 differentially expressed genes (DEGs). Within this gene set, there was a high enrichment of genes associated with axon guidance – including genes from the Ephrin Slit, and Robo

[6]The methodology, results and discussion that lead to this paper will be explained in detail in a later chapter.

[7]Importantly, the authors grouped L5 neurons as either IT neurons or ET neurons, therefore, the results from ET neuron analysis is the accumulation of Betz cells, non-Betz PT neurons, and remaining ET neurons.

families, which the authors argue could represent genes that regulate the cortico-motoneuronal connections allowing increasing dexterous fine motor control across species and may help explain the biological consequences of Betz cells in primates directly contacting alpha motoneurons and not spinal interneurons as in rodents.

While a significant milestone in underpinning the complexity of the primary motor cortex, there remain some unanswered questions. One of the Betz cell-containing clusters (Exc L5 FEZF2 ASGR2) also contains L3 neurons which were not addressed by the authors[8]. This may be because L3 neurons, L5 non-Betz cells, and Betz cells may contain strong transcriptomic similarities that are not possible to divide with an unbiased approach or *a priori* molecular knowledge of each neuronal population. The authors elegantly compared the results from PATCH sequencing to the single nuclei sequencing in an attempt to validate their results by matching neuronal morphology with individual transcriptomics. However, there is limited information on how they identified the Betz cells used for PATCH-seq. This is particularly important in this study, where the only PATCH-seq performed was on presumed Betz cells from surgically resected premotor cortex and not from the primary motor cortex. There is limited literature on the existence of Betz cells in this region, so the validation of correct Betz cell identity is critical; in light of no unique biomarkers, there was also no mention of what morphological features were used to identify the presumed Betz cells. As PATCH-seq had to be performed on organotypic brain slices, the authors incubated slices from the human premotor cortex with recombinant adeno-associated virus vectors to fluorescently label Betz cells for subsequent analysis [218]. However, the lack of Betz cell biomarkers, meant that these viral constructs either generically stained all neurons or enriched for L5 ET neurons. The difficult task of identifying Betz cells from surrounding neurons, the uncertainty of Betz cells in the premotor cortex, and the lack of information for targeting the unique morphology of the Betz cell leads to doubts regarding the PATCH-seq validation.

[8] At least in the pre-print version of this study.

Theories Behind the Betz cells Unique Morphology -

It has been hypothesised that there remains a direct correlation between perikaryon size with body size, axon length, and size of targeted muscle [219–223]. Some have suggested that perikaryon size is related to brain size, size of the cerebrum, and in the case of motor neurons, digital dexterity [224]. It is tantalising to agree with these observations in the context of the Betz cell. They are the largest pyramidal neuron and contain the longest axons in the human brain; their size is at their largest in regions that are furthest away from their targeted neurons (hip/leg area), which are also some of the largest muscles in the human body. Humans also possess remarkable digital dexterity. However, there are some irregularities with these theories. Sherwood et al., 2003 showed there was no relationship between Betz cell soma size and digital dexterity [200], furthermore, caniforms and feliforms that rank low in digital dexterity have been reported to possess the largest Betz cells, more so than primates that are considered to contain impressive digital dexterity [225]. Turning attention to other examples of the animal kingdom, the giraffe with its uniquely long neck (up to 2.6m) does not contain disproportionally large Betz cells (600 μm^2 to 1200 μm^2) [181, 190, 226]. The unique neck of the giraffe, however, may explain its impressive dendritic volume of 73,742 μm^3 which is considerably larger than the other artiodactyls studied. Conversely, it has been stated that similar to other ungulates the corticospinal tract of giraffes may terminate at the level of the cervical cord, which would make the corticospinal tract $\approx$ 1.5m which while long is not excessive and may justify their relative Betz cell size. Furthermore, unlike humans, the Betz cell perikaryon in giraffes does not change in size across the motor cortex, nor are they found in the absence of Betz cell clusters. Jacobs et al., (2017) concluded that the very large Betz cells in feliforms are most likely attributed to the high priority need for explosive speed, strength, and fast reactions due to their predatory nature, especially since hunting is mostly limited to prey of similar size or bigger. The increased need for explosive movement may place a higher pressure on Betz cells to relax flexor muscles immediately prior to movement as well as adjusting the excitability threshold of surrounding pyramidal neurons.

This predatory need is less critical for primates, however, there may be increased pressure in possessing specialised developed cortico-motoneuronal-musculoskeletal systems. The increased degree of dendritic branching/bundling along with their neuronal clustering appears to allow Betz cells to better synchronize their output and control on targeted α-motoneurons [191, 201], which in the hand region is particularly enriched. Others have suggested that Betz cells are at least indirectly involved in striated muscle contraction velocity and power [227, 228] and able to initiate initial downstream locomotion quickly [229].

2.4 Amyotrophic Lateral Sclerosis

Amyotrophic Lateral Sclerosis (ALS) is a fatal neurodegenerative disease that affects motor neurons with an incidence of $\approx 1/100{,}000$. ALS is the most common phenotype within the umbrella term of Motor Neuron Diseases (MND), which also include Primary Lateral Sclerosis, Progressive Muscular Atrophy, Progressive Bulbar Palsy, and Pseudobulbar Palsy. ALS has a median survival of 30 months from symptom onset [230], progressing rapidly as compared with other neurodegenerative disorders [231, 232]. The progressive paresis that clinically characterises ALS often leads to death from respiratory failure. The vast majority of cases (90-95%) of cases are sporadic (SALS) while the remaining are familial (FALS). Both sporadic and familial ALS is characterised by the progressive loss and/or degeneration of upper (cortical) and lower (spinal) motor neurons. Symptoms of ALS include cramping, twitching, and muscle weakness that advances to dyspnea and dysphagia.

2.4.1 Betz Cell Pathology in ALS

Betz cells play an undoubted role in motor neuron diseases (MND) [233], for example, in Primary Lateral Sclerosis (PLS) Betz cells are progressively degenerated as shown in neuropathological examinations where the absence of Betz cells in patients with PLS is commonplace [234]; however, it is not known if they are destroyed or have undergone advanced atrophy to an extent that they are not recognised within the L5 neuronal populations of the primary motor cortex. Betz cell neuropathology

is also a clear feature of ALS but less understood. While patients with ALS exhibit upper motor and/or lower motor clinical features, the mechanism of action is unclear. *In vivo*, neuronal damage/loss is seen with patients with ALS [235]. However, two stereological studies found no difference in mean perikaryon volume, mean neuronal nuclear volume, total perikaryon volume, total neuronal nuclear volume, or mean neuronal number between ALS and control upper motor neurons [236, 237]. This surprising result often leads to the conclusion that ALS motor cortex dysfunction is caused by regional metabolic changes rather than cytological alterations. It is important to note, however, that these studies were not designed to specifically examine Betz cell numbers. Despite this, loss/atrophy of Betz cells and Betz cell pathophysiology has been reported in several other studies, correlating ALS severity with Betz cell loss [202, 238–240]. The discrepancy between studies potentially being attributed to the previously mentioned challenges of studying Betz cells. Conversely, while Betz cell atrophy or lost Betz cell soma have been observed in ALS, the swelling of Betz soma due to excessive lipofuscin has also been reported as a feature of ALS pathology [202].

Dendritic degeneration is also a clear feature of ALS Betz cell pathology [202, 241–243] - see Figure 2.12. For example, the prominent apical dendrite of Betz cells profoundly degenerates in ALS. Genc et al., (2017) investigated the neurodegeneration of Betz cells in familial ALS, sporadic ALS, ALS with frontotemporal dementia, and Alzheimer's Disease (AD) and found widespread dendritic vacuolation and somal damage in all but AD. In this study, CA1 pyramidal neurons of the hippocampus - which are destroyed in AD - did not undergo apical dendritic pathology and suggests that apical dendrite degeneration is likely to be a Betz cell-specific feature of ALS. Furthermore, they like others, observed shrunken Betz cells in patients with ALS but state their small sample size contributed to the lack of statistical significance. Dendritic spine loss is a common underlying cause of neuronal vulnerability and is seen in many neurodegenerative diseases [244]. Betz cell spine loss has been revealed by Golgi staining studies in ALS patients [202]. In older studies, it has also been suggested via Golgi staining that Betz cell dendritic

degeneration proceeds soma size shrinkage in ALS [202, 242, 245]. The corticospinal tract has also been shown to undergo degeneration in ALS [239]. Interestingly, it has been noted that Betz cell ALS dendritic pathology can be seen in the absence of visual erosion of the corticospinal tract, although this could be explained by the small proportion of Betz cell axons within it. Another feature of Betz cell ALS pathology is the acute accumulation of reactive astrocytes around the Betz perikaryon [202] - see Figure 2.13. This gliosis can at times completely shroud Betz cells making it particularly difficult to confirm their identity and can even blanket and migrate up their apical dendritic. Hammer et al., (1979) noted that the gliosis may be a response to the declining health of basilar dendrites, while others, who have observed this glial phenomenon in spinal muscular atrophy (SMA) have suggested that reactive astrocytes in the spinal cord may cause the neurodegeneration seen in this disorder [246]. Interestingly, in normal-appearing Betz cells in ALS, there is an association with stellate cells axons and the degeneration of Betz cell synapses [247]. It has been argued that alterations in stellate neuron pre-synapses on Betz cells cause dysfunction, by limiting the inhibitory control of Betz cells[9] resulting in an electrophysiological modification which decreases motoneuronal control and function.

It has been concluded that ALS is a disease that accelerates ageing, and that Betz cell ALS pathology mimics that of Betz cells in the ageing brain [192, 202]. It is important to note, that although it appears that rodents do not possess Betz cells as defined with current technologies, their upper motor neurons undergo similar pathology in animal models of ALS [233, 248–252]. Despite the clear role Betz cells play in ALS pathology there has been a lack of detailed cortical studies to better understand this disease. This can be attributed to the arguments surrounding the direction of disease spread, in the historical dogma of the "dying-back" hypothesis [253] i.e. ALS progresses from neuromuscular junctions to the spinal cord and finally to the motor cortex. This rationale would make Betz cells the last site of neuronal pathology and therefore do not contribute to the vulnerability or progression of the disease. This hypothesis has led to early cerebral dysfunction being underestimated

[9]For more information on how Betz cells may exhibit an inhibitory effect on neurons, see B.2.

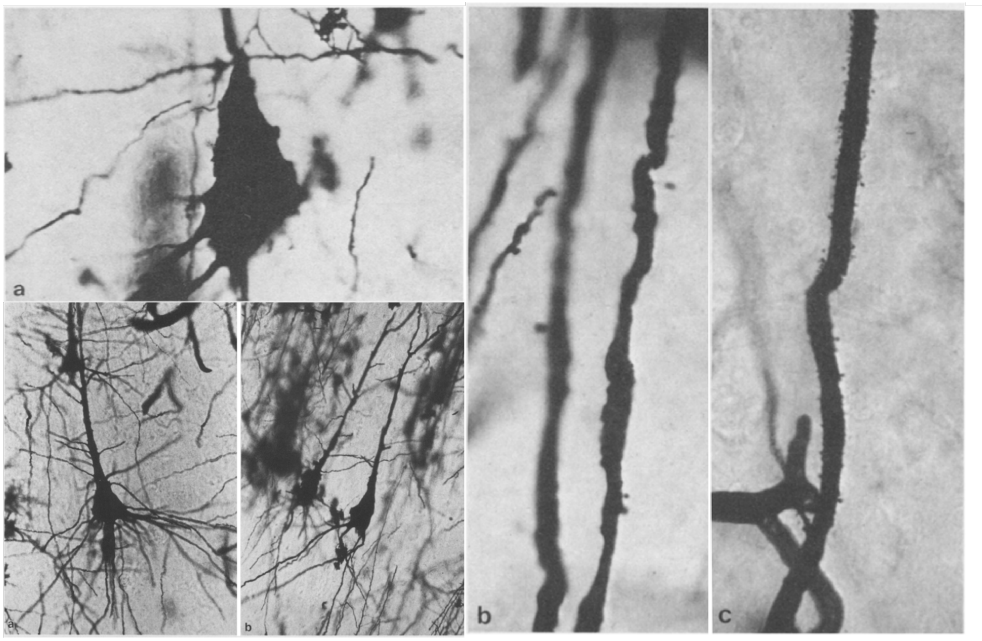


Figure 2.12: Structural Degeneration of Betz Cells in ALS.
TL: Irregular distorted surface of a Betz cell in an ALS patient. BL1: A healthy Betz cell - dendritic arbour is normal and has a dense basilar array. BL2: A Betz cell in a patient with MND; many dendrites have disappeared, with only the apical dendrite and several basal dendrites remaining. NR: Apical dendrite of a Betz cell in ALS; the dendrite is irregular, with a distorted rough surface, with abnormal swelling. FR: Apical dendrite of a healthy Betz cell with uniform spines and consistent thickness. TL = Top left, BL = Bottom Left, NR = Near Right, FR = Far Right. *(Taken from [242] with permission from Springer.)*

in patients, although it has now recently been shown that early motor cortex dysfunction and Betz cells degeneration may begin in the early stages of the disease, strengthening the concept of the "dying-forward" hypothesis [254, 255], yet, studies investigating the selective vulnerability of upper motor neurons are still lacking – likely due to the difficulty of obtaining the appropriate human tissue and the belief that most ALS post-mortem brains represent late-stage disease.

Perhaps the most interestingly pathological feature in Betz cells is that they do not exhibit the same pathology that is found in 97% of ALS patients - the hallmark insoluble cytoplasmic inclusions of TDP-43 - a summary of TDP-43 pathology in Betz cells can be seen in Appendix B.5.1.

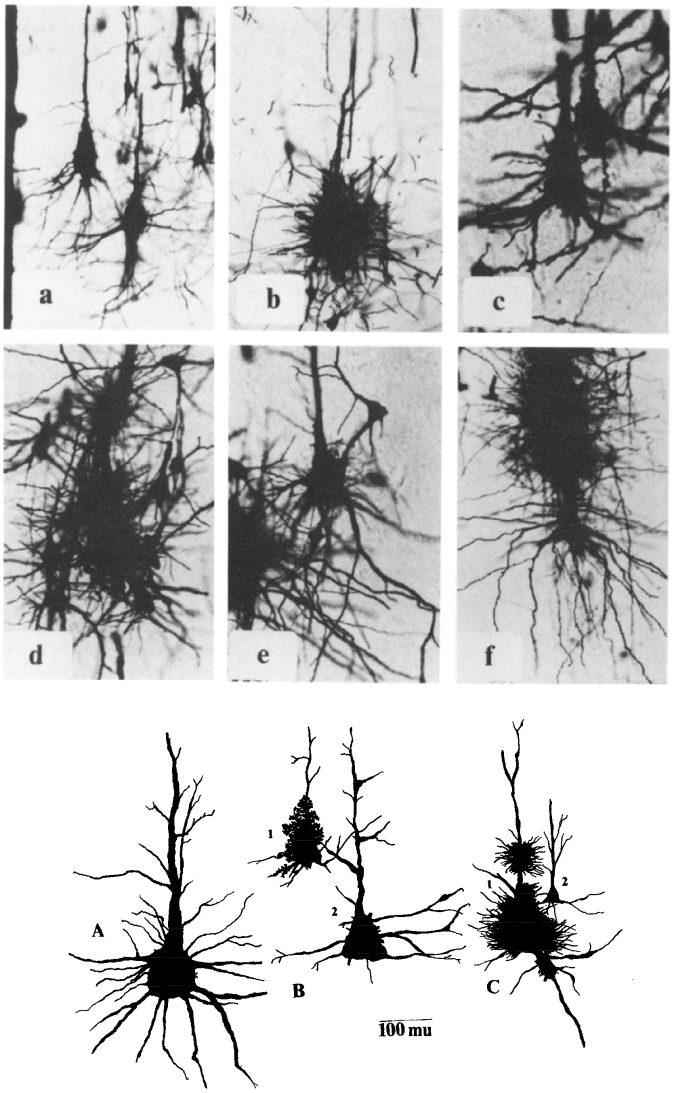


Figure 2.13: Reactive Gliosis is a Betz Cell-Specific ALS Pathology.
a: Slight degeneration in two Betz cells; the left Betz cell exhibits basilar dendrite "stumping" while the right Betz shows the beginnings of gliosis, as glial fibres can be seen radiating from Betz cell soma. b: Reactive gliosis completely covers Betz cell soma in ALS, few basilar dendrites remain. c: Distortion and degeneration of dendrites can be seen in this Betz cell. d-f: show the reactive gliosis that has surrounded Betz cell somas. Fibrous gliosis extends up the apical shaft, importantly, surrounding pyramidal neurons are spared from glial involvement. Camera lucida drawings showing Betz cells changes in the healthy, ageing, and ALS brain; A: Normal Betz cell with the characteristic circumferential dendrites. B: Ageing Betz cells, B1: astrocytes surround a Betz cell that is swollen showing advanced degenerative changes. B2: Betz cell which has lost much of its circumferential dendrites, and marked dendritic swelling. C: Motor neurons in ALS. C1: Reactive gliosis surrounding a Betz cell with advanced degeneration. C2: A normal-appearing L5 pyramidal neuron with no reactive gliosis. *(Taken from [202] with permission from Elsevier).*

2.5 Aims of this Thesis

Many questions arise when attempting to define the role Betz cells have in health and disease. These enigmatic neurons are among the most visually striking in the human neocortex, yet basic information regarding their physiology is still unknown. Likely due to a lack of biomarkers and subtle morphology leading them to be labelled as *simply* large motor neurons, with a molecular composition no different to surrounding L5 motor neurons.

This thesis aims to:

- Develop a high-performance approach that will allow LCM proteomics to be performed on Betz cells.
- Reveal and investigate the proteome and transcriptome of the human Betz cell.
- Understand how the molecular composition of Betz cells is affected in ALS.
- Investigate if Betz cells have a unique molecule signature or are just larger versions of their L5 neighbours.
- Discover Betz cell biomarkers to ease and enhance future Betz cell research.

3 Methodology

Contents

3.1 Brain Dissection 62
3.1.1 Fresh Brain Dissection & Tissue Processing 62
3.1.2 Tissue Processing for FFPE Material 63
3.2 Sectioning 64
3.2.1 FFPE Tissue for Immunohistochemical/Histological Staining 64
3.2.2 Frozen Tissue for Immunohistochemical/Histological Staining 64
3.2.3 Sectioning for Laser Capture Microdissection 64
3.3 Histological Staining 65
3.3.1 Cresyl Violet Staining 65
3.3.2 Haematoxylin & Eosin Staining 66
3.4 Immunohistochemistry 66
3.4.1 Antibody Selection 66
3.4.2 Antibody Validation and Optimisation 67
3.4.3 Immunohistochemistry on FFPE Sections 68
3.4.4 Immunohistochemistry on Frozen Sections 68
3.5 Laser Capture Microdissection 69
3.5.1 Purkinje Cells 69
3.5.2 Molecular Layer 70
3.5.3 Granule Cells Area 70
3.5.4 Betz Cells 70
3.5.5 Cerebellar Dentate Nucleus Neuron (DNN) 71
3.6 Mass Spectrometry - Sample Preparation 71
3.6.1 Standard Sample Preparation / In-Solution Digestion . 71
3.6.2 In-Cap Digestion 72
3.6.3 SP3 Method 72
3.6.4 RWS approach (RIPA Buffer/Wet cap/Custom SP3) . . 73

3.7 Mass Spectrometry 74
3.7.1 LC-MS/MS (Lumos Orbitrap) - Purkinje Cell Titration 74
3.7.2 LC-MS/MS (Lumos Orbitrap) - For Subsequent Experiments 74
3.7.3 LC-MS/MS (TimsTOF) - Evo-sep (Standard High-Throughput Approach) 75
3.7.4 LC-MS/MS - Dionex Nanoflow Coupled with TimsTOF (Optimised high-throughput approach) 75
3.7.5 MaxQuant Settings 76
3.7.6 Fragpipe / MSFragger 76
3.7.7 Proteomic Quality Control 77
3.7.8 Data Analysis 77
3.8 Digital Pathology 78
3.8.1 Whole Slide Imaging 78
3.8.2 Image Analysis 78
3.9 RT-qPCR 79
3.9.1 Primer Selection 79
3.9.2 Primer Design 79
3.9.3 Primer Optimisation 80
3.10 LCM-seq 80
3.10.1 Preparing and Validating SPRI Beads 80
3.10.2 Library Construction 81
3.10.3 RNA-Sequencing 83
3.10.4 Data Pre-processing and Data Analysis 84

3.1 Brain Dissection

3.1.1 Fresh Brain Dissection & Tissue Processing

Whole human post-mortem brain tissue was collected by the Oxford Brain Bank; a research ethics committee (REC) approved, HTA regulated research tissue bank (REC reference 15/SC/0639). Whole brains from human donors were dissected fresh. If present, the dura was removed from the cerebrum and the cerebral hemispheres severed by a large incision along the longitudinal fissure. The primary motor cortex was identified using neuroanatomical landmarks such as the paracentral lobule, central sulcus, and characteristic ‘hand knob’ region. The whole primary motor cortex was removed as one piece including the extremities; to aid in correct identification in subsequent sectioning, the somatosensory cortex was also sampled with the cortex. This ’M1 strip’ was further dissected into discrete portions

representing the motor map of the homunculus and snap-frozen in liquid nitrogen vapour. After freezing, tissue blocks were kept at -80°C. Photographs were taken at the time of sampling and the orientation of the M1/S1 blocks were recorded using the standard anatomical terms of location. The corresponding primary motor cortex on the adjacent hemisphere was also removed with the same methodology and instead placed in 10% neutral buffered formalin. To confirm that the primary motor cortex had been successfully dissected, the aforementioned photographs were consulted and the likely spot of the primary motor cortex was sectioned. H&E staining was then performed systematically along the cortex, the dissection was deemed successful if the distinguishing Betz cells were present.

The cerebellum was detached from the cerebrum at the level of the fourth ventricle and the brain stem was removed. The cerebellum was divided along the posterior Cerebellar notch with one Cerebellar hemisphere placed in its entirety in 10% neutral buffered formalin. The other hemisphere was dissected into 1cm thick slabs and snap-frozen in liquid nitrogen vapour.

3.1.2 Tissue Processing for FFPE Material

Tissue that had been placed in neutral buffered formalin were incubated for 24-48 hours, like its frozen counterpart the fixed motor cortex was divided into pieces representing the homunculus. Fixed tissue was sliced transversely to reduce thickness and placed onto a Leica TP1020 Automatic Benchtop Tissue Processor on a short 12-hour protocol to preserve tissue quality and reduced RNA degradation (data not shown). Tissue was submerged and progressed through increasing concentrations of alcohol, xylene, and paraffin wax baths. Finally, tissue was embedded using a Leica EG1150 Modular Tissue Embedding Center into paraffin wax blocks and stored at 4°C.

3.2 Sectioning

3.2.1 FFPE Tissue for Immunohistochemical/Histological Staining

FFPE tissue blocks were pre-chilled at 4°C. Unless specified otherwise, tissue was cut at 6µm using a Leica RM2235 microtome at room temperature, transferred to a water bath containing distilled water heated to 56°C to facilitate wax smoothing and mounted onto Superfrost Adhesion Slides. Slides were air-dried overnight.

3.2.2 Frozen Tissue for Immunohistochemical/Histological Staining

Tissue blocks were placed inside a Thermo Sandon Cryotome FE Cryostat for 20 minutes to adjust to the ambient temperature of the cryostat, which was set to -18°C. Tissue blocks were mounted on top of pre-cooled cryostat mounts using OCT. The cryobar was set to -20°C and the specimen holder set to -14°C. Unless specified, sections were cut at 10µm and placed onto Superfrost Adhesion Slides and air-dried at room temperature for several minutes until completely dry.

3.2.3 Sectioning for Laser Capture Microdissection

All work surfaces including equipment, blades, and accessories were cleaned with diluted sodium hypochlorite and 90% alcohol. Sectioning was performed as previously described, however, sections were instead mounted on 1.0 PEN membrane slides (Zeiss, 10221810). Before sectioning, all 1.0 PEN membrane slides were UV irradiated at a wavelength of 254 nm for 30 minutes to eliminate contaminates and increase the hydrophilicity of the membrane. FFPE sections were baked overnight at 37°C before staining to increase tissue adhesion to PEN membrane slide.

3.3 Histological Staining

3.3.1 Cresyl Violet Staining

FFPE & Frozen samples-

Sections were either deparaffinized in xylene for 15 minutes (x3 5-minute washes and rehydrated through a series of baths containing 100%, 90%, 70% alcohol, for 30 seconds each or placed immediately into 100% alcohol, and the rehydrated as above. Sections were then incubated in 1% (w/v) cresyl violet in purified water for 1 minute, and dehydrated with 70%, 90%, and 100% alcohol and cleared using xylene.

Staining for Laser Capture Microdissection (Protein) -

PEN membrane slides were treated the same as outlined in 3.2.3 and followed the same staining process as 3.3.1, however, xylene steps were omitted to avoid disruption of the PEN membrane. Frozen sections were air-dried for immediate staining or stored overnight at -80°C.

Staining for Laser Capture Microdissection (RNA) -

PEN membrane slides were treated the same as outlined in 3.2.3 and followed the same staining process as 3.3.1, however, xylene steps were omitted to avoid disruption of the PEN membrane. Sections were incubated in ethanol for 1 minute to minimise ribonuclease activity. To further reduce RNA degradation, Cresyl violet was dissolved in ethanol rather than water, and the entire staining procedure was performed on dry ice. Once stained, sections were placed in falcon tubes, sealed, and stored at -80°C. For LCM, falcon tubes containing the stained PEN membrane slides were slowly warmed to room temperature and once acclimatized, the lids were opened and placed on the microdissection microscope to avoid moisture from the air condensing on the LCM slide due to temperature differential; resulting in RNA degradation. All solutions were RNA-free.

3.3.2 Haematoxylin & Eosin Staining

FFPE sections —

Staining was automated using a Thermo Shandon Linistain; in short, slides were transferred through chemical baths containing: xylene, 100% alcohol, 70% alcohol, tap water, Harris' haematoxylin, tap water, 0.4% acid alcohol, tap water, Scot's tap water, 0.25% acetic acid eosin, 70% alcohol, 100% alcohol and xylene.

Frozen sections —

For frozen sections, samples were treated similarly but were placed directly into 70% alcohol at the beginning of the protocol and then processed as normal.

Staining for Laser Capture Microdissection —

Tissue sections were placed on 1.0 PEN membrane slides and treated the same as 3.2.3, however, xylene steps were omitted to avoid disruption of the PEN membrane. Frozen sections were air-dried for immediate LCM use or stored for short-term use at -80°C.

3.4 Immunohistochemistry

3.4.1 Antibody Selection

For any particular protein of interest, the literature was first screened for well-documented examples of antibodies being used for IHC, with a preference for studies studying similar tissue/cells. Failing that, the Human Proteome Atlas (HPA) was used to screen for the protein of interest and select antibodies that have already been extensively characterised or show promisingly immunoreactivity based on histology images. Antibodies were selected based on host species, species reactivity, tested applications, and clonality.

3.4.2 Antibody Validation and Optimisation

All antibodies were subjected to quality control and optimisation steps before interpretation and analysis. For each antibody, the datasheet was used to identify the manufacturer's recommended dilutions and appropriate positive control tissue. A serial dilution was tested around these recommendations to determine the optimal dilution for our tissue. Microscopic examinations were performed to determine the presence of nonspecific staining, endogenous peroxidase activity, and expected immunoreactivity. Where-ever possible, the presence of internal controls (both positive and negative) within the tissue sections were used to increase the confidence of specific immunoreactivity.

Antibody	**Cat #**	**Dilution**
Anti-AE-1	AB215892	1:100
Anti-ARHGAP26	17747-1-AP	1:500
Anti-ASGR2	HPA014899	1:250
Anti-ATPB	AB14730	1:1000
Anti-CALB1	CB38	1:4000
Anti-CAPRIN2	20766-1-AP	1:2000
Anti-CRYM	PA565072	1:500
Anti-CSN1S1	HPA035659	1:500
Anti-CYCS	MAB897	1:1000
Anti-DST	PA 5-56396	1:200
Anti-ER81	PA541484	1:400
Anti-ESYT1	HPA016858	1:100
Anti-HDAC1	HPA029693	1:100
Anti-ITPR1	HPA016487	1:500
Anti-KIF1A	612094	1:500
Anti-NeuN	MAB377	1:1000
Anti-OXT1	AB25985	1:400
Anti-REEP1	HPA058061	1:100
Anti-SMI311	837801	1:2000
Anti-SMI32	SMI-32P	1:1000
Anti-TDP43	10782-2-AP	1:4000
Anti-VDAC1	GTX114187	1:100

Table 3.1: A list of antibodies used in this thesis and their dilutions.

3.4.3 Immunohistochemistry on FFPE Sections

Samples were baked at 56°C for 20 minutes to remove residual water and improve tissue adhesion. Sections were then deparaffinized in xylene for 15 minutes (x3 5-minute washes) and rehydrated through a series of baths containing 100%, 90%, 70% alcohol and tap water. Rehydrated sections were incubated for 30 minutes in a 10% hydrogen peroxide solution diluted in phosphate-buffered saline (v/v) to eliminate endogenous peroxidase activity. Antigen retrieval was performed by autoclaving (121°C for 10 minutes) in sodium citrate buffer (pH 7.6). Sections were permeabilised with triton in tris buffered saline (TBS/T) and incubated in 10% normal goat serum for 30 minutes to inhibit non-specific binding. Primary antibodies were added to the assay with 5% normal goat serum in TBS/T and diluted according to the results of 3.4.2. Primary antibody incubation was overnight at 4°C. Staining visualization was performed using the Agilent's Dako EnVision Kit (Agilent, Cat. number k5007), briefly, a cocktail of mouse and rabbit secondary antibodies conjugated to horseradish peroxidase were added to the sections and incubated for 1 hour at room temperature and washed twice for 5 minutes with TBS/T. Chromogen development was achieved by incubating the sections in an x1 DAB working solution prepared from a ×50 concentrated 3,3'-diaminobenzidine stock and diluted with Dako substrate buffer for 5 minutes. Finally, sections were counter-stained with Harris' haematoxylin for 1 minute, washed in distilled water, and dehydrated in a series of increasing alcohol washes finishing in xylene. Coverslips were mounted to the sections using a compound mixture of Distyrene, a Plasticizer and Xylene (DPX).

3.4.4 Immunohistochemistry on Frozen Sections

After air-drying, sections were then fixed for 20 minutes in ice-cold 100% acetone. Endogenous peroxidase activity was removed with a 10% hydrogen peroxide solution diluted in phosphate-buffered saline (v/v) for 30 minutes. Antigen retrieval was omitted. The remaining steps followed the above section. For frozen IHC sections destined for LCM, xylene, DPX, and coverslip mounting were omitted.

3.5 Laser Capture Microdissection

Surrounding workspaces including the LCM instrument were cleaned with diluted sodium hypochlorite and 90% alcohol. The TubeCollector 2x200 CM II (Zeiss; 415101-2000-410) was UV irradiated at a wavelength of 254 nm for 30 minutes to remove contaminates and used for all LCM experiments. LCM settings were optimised using built-in software (PALMrobo V4.8), and the following were selected; for cutting: Laser Speed: 65, Laser Cut Energy: 43 and Laser Focus 55; For catapulting: LPC Energy: 20 and the Focus: Δ (LPC Energy, LPC Focus) = -15. An x40 objective lens was used for the initial Purkinje cell titration (4.3.1) and an x10 objective lens was used in all other LCM experiments. When the objective lens was adjusted, the above settings when scaled accordingly using built-in software. Several measurements, including the number of annotations, annotation area (µm2), XY coordinates, and total annotation area (µm2) were recorded.

Protein applications

Cells/tissue areas destined for LC-MS/MS were microdissected using the criteria in the next sub-section. The collection media used varied slightly depending on optimisations/established protocols, see 3.6 for more information.

RNA applications

Cells/tissue areas destined for LCM-seq were microdissected using the criteria in the next sub-section. Cells/areas were catapulted directly into 5µl 1% Triton X-100 containing RNase inhibitor (RNaseOUT™, 2 U/µl). For downstream preparations, see 3.10.2.

3.5.1 Purkinje Cells

The Purkinje cell layer was identified by its location between the molecular and granular layer of the cerebellum. Purkinje cells within this layer were recognised by their large perikaryon, regular spacing, the existence of Nissl bodies and if present, large dendritic root. Annotations were manually drawn around the cell

body of the Purkinje cell at either an x20 or x40 objective lens. For the titration experiment, LCM was performed on an x40 objective lens and cells were catapulted into 0.2ml adhesive cap tubes (Zeiss, 415190-9201-000). Future experiments used Lobind Protein or Lobind DNA/RNA 0.2ml tubes, depending on the application.

3.5.2 Molecular Layer

The molecular layer of the cerebellum was identified as the outmost Cerebellar layer adjacent to pia matter and foliae. Rectangular annotations were placed within the molecular layer corresponding to the size of 100 Purkinje cells (60,000µm^2). Areas were catapulted into caps (Zeiss, 415190-9201-000) with or without buffer/adhesive for LC-MS/MS sample optimisation.

3.5.3 Granule Cells Area

Granule cells were identified as small compact neurons located in the innermost structure of the cerebellum, known adeptly as the granular layer. Rectangular annotations were placed within the granular layer to save manually annotating the individual dense neurons. LCM was performed on an x10 objective lens. Areas were catapulted into Lobind Protein or Lobind DNA/RNA 0.2ml tubes, depending on the application.

3.5.4 Betz Cells

Betz cells were identified and selected by their defining existence within Layer Vb of the primary motor cortex, unique dendritic morphology, the abundance of lipofuscin, and copious rough endoplasmic reticulum (Nissl bodies) - See Appendix A. After manual annotation, the Betz cells were sorted by size in-silico and the bottom 10% annotations were deleted. LCM was performed originally on an x40 objective lens but later on an x10 objective cells to speed up the LCM process. To compensate for their larger size, the LPC energy was increased to 27. When drawing annotations, the LPC joint was positioned on the broadest part of the cell soma to avoid cells

unintentionally being removed prematurely. Cells were catapulted into Lobind Protein or Lobind DNA/RNA 0.2ml tubes, depending on the application.

3.5.5 Cerebellar Dentate Nucleus Neuron (DNN)

The Cerebellar Dentate Nucleus neuron (DNN) was identified in cerebellum sections by the prominent serrated edge of the Dentate nucleus. Within this large Cerebellar structure, DNNs were identified by their large size, occasional lipofuscin, and regular pattern throughout the Dentate Nucleus. LCM was performed on an x10 objective lens to speed up the LCM process. To compensate for their larger size, the LPC energy was increased to 27. When drawing annotations, the LPC joint was positioned on the broadest part of the cell soma to avoid cells unintentionally being removed prematurely. Cells were catapulted into Lobind Protein tubes.

3.6 Mass Spectrometry - Sample Preparation

3.6.1 Standard Sample Preparation / In-Solution Digestion

After cell lysis, the lysate was reduced with dithiothreitol (Final Concentration: 5mM) for 30 minutes at room temperature. Proteins within the lysate were then alkylated with iodoacetamide (Final Concentration: 20mM) for 30 minutes at room temperature. Proteins were purified by methanol-chloroform precipitation and the resulting pellet was resuspended in 6M urea (0.1m tris pH 7.8) and further diluted to 1M urea in ultrapure water. Proteins were digested by adding 25ng of trypsin and incubated overnight at 37°C. 1% formic acid (v/v) was used to acidify peptides and purified with C18 solid-phase extraction cartridges (SOLA HRP, ThermoFisher Scientific). Peptides were completely dehydrated under vacuum and resuspended in a 6µl solution comprising of 2% acetonitrile, 0.1% trifluoroacetic acid). Peptides were stored at -20°C.

3.6.2 In-Cap Digestion

The following steps took place within the LCM cap: after cell lysis, the lysate was diluted to 1M Urea or 10% TFE with 50mM ammonium bicarbonate. Proteins were digested by adding 25ng of trypsin and incubated overnight at 37°C. 1% formic acid (v/v) was used to acidify peptides and purified with C18 solid-phase extraction cartridges (SOLA HRP, ThermoFisher Scientific). Pipettes were completely dehydrated under vacuum and resuspended in a 6 µl solution comprising of 2% acetonitrile, 0.1% trifluoroacetic acid). Peptides were stored at -20°C.

3.6.3 SP3 Method

After cell lysis, the lysate was reduced with dithiothreitol (Final Concentration: 5mM) for 30 minutes at room temperature. Proteins within the lysate were then alkylated with iodoacetamide (Final Concentration: 20mM) for 30 minutes at room temperature. Reduction and alkylation occurred within the LCM cap. Paramagnetic beads were prepared as previously described [137] and added to the cell lysate. Acetonitrile (100% v/v) was added to the bead-lysate mixture for a final concentration of 70% (v/v). The mixture was incubated for 18 minutes at room temperature at 1000 rpm. Beads were then restrained against a magnet for 2 minutes and the supernatant transferred to 0.2ml tubes containing fresh beads. Next, the 70% (v/v) acetonitrile incubation was repeated to recover any unbound protein. Beads were immobilized on a magnet and washed with 70% (v/v) alcohol twice and once with 100% acetonitrile. 25ng trypsin in 50mM ammonium bicarbonate was used to resuspend the beads and digest proteins, this incubation was done overnight at 37°C. Next, beads were resuspended by ultrasonic sonication. Acetonitrile was added to a final volume of 95% (v/v) and incubated for 18 minutes at room temperature at 1000 rpm to rebind peptides. Bead mixtures were then placed on a magnet for 2 minutes and the supernatant discarded. 6µl of 2% DMSO was added to the tubes containing potential unbound peptides, these beads were then mixed with the beads from their corresponding "originally bound" counterparts. The beads were immobilised for the last time on the magnet for 5 minutes and the

supernatant containing all the peptides were carefully transferred to an LC-MS vial containing 600nl 1% trifluoroacetic and stored at -20°C.

3.6.4 RWS approach (RIPA Buffer/Wet cap/Custom SP3)

After catapulting the cells into 15µl RIPA buffer and incubating for 30 minutes at room temperature the cell lysate was centrifuged to bring the lysate to the bottom of the tube. A small amount of RIPA buffer (5µl) was added to the cap, pipetted mixed, and centrifuged to capture any remaining material. Benzonase Nuclease (1:10, Sigma-Aldrich) was added and incubated for 30 minutes to remove nucleic acids. Proteins within the lysate were reduced with dithiothreitol (Final Concentration: 5mM) for 30 minutes at room temperature. Proteins within the lysate were then alkylated with iodoacetamide (Final Concentration: 20mM) for 30 minutes at room temperature. Paramagnetic beads were prepared as above (excluding the second acetonitrile incubation) and added to the lysate. Acetonitrile (100% v/v) was added to the bead-lysate mixture for a final concentration of 70% (v/v). The mixture was incubated for 18 minutes at room temperature at 1000 rpm. Beads were then restrained against a magnet for 2 minutes and washed with 70% (v/v) alcohol twice (30 seconds each) and once with 100% acetonitrile (15 seconds). 25ng trypsin in 50mM ammonium bicarbonate was used to resuspend the beads and digest proteins, this incubation was done overnight at 37°C. Next, beads were resuspended by ultrasonic sonication. Acetonitrile was added to a final volume of 95% (v/v) and incubated for 18 minutes at room temperature at 1000 rpm to rebind peptides. Bead mixtures were then placed on a magnet for 2 minutes and the supernatant discarded. 6µl of 2% DMSO was used to resuspend the beads followed by a final incubation on the magnet for 5 minutes. The resulting supernatant containing the peptides were carefully transferred to an LC-MS vial containing 600nl 1% trifluoroacetic and stored at -20°C.

3.7 Mass Spectrometry

3.7.1 LC-MS/MS (Lumos Orbitrap) - Purkinje Cell Titration

Digested peptides were analysed by LC-MS/MS using a Dionex Ultimate 3000 fixed to an Orbitrap Fusion Lumos mass spectrometer (ThermoFisher Scientific). A 75µm × 500mm C18 EASY-Spray column (ThermoFisher Scientific) with 2µm particles was used at a flow rate of 250 nL/min. The Orbitrap was used with the "Universal Method" that has been described by ThermoFisher Scientific; MS1 scans were acquired between 400 and 1500 m/z at a resolution of 120,000 with an AGC target of 4×10^5. Selected precursors between 2+ and 7+ charge state and above the intensity threshold of 5×10^3 were chosen for collision-induced dissociation (CID) fragmentation, with a collision energy of 35%, an AGC target of 4×10^3, a maximum injection time of 300 ms, a dynamic exclusion window of 60 seconds, and a maximum duty cycle of 3 seconds. MS/MS spectra were acquired in the ion trap using the rapid scan mode.

3.7.2 LC-MS/MS (Lumos Orbitrap) - For Subsequent Experiments

Digested peptides from the above RWS approach were analysed by LC-MS/MS using a Dionex Ultimate 3000 fixed to an Orbitrap Fusion Lumos mass spectrometer (ThermoFisher Scientific). A 75µm × 500mm C18 EASY-Spray column (ThermoFisher Scientific) with 2µm particles was used at a flow rate of 250 nL/min. Liquid chromatography was performed on a 60-minute linear gradient from a 2% Buffer B (5% DMSO, 0.1% formic acid in water) to a 35% buffer B (5% DMSO, 0.1% formic acid in acetonitrile). The Orbitrap was used with the "Universal Method" that has been described by ThermoFisher Scientific; MS scans were acquired between 400 and 1500 m/z at a resolution of 120,000 with an AGC target of 4×10^5. Selected precursors between charge state 2+ and 7+ and above the intensity threshold of 5×10^3 were selected for higher-energy collisional dissociation (HCD) fragmentation, with a normalized collision energy of 28%, an AGC target

of 4×10^3, a maximum injection time of 80 ms, a dynamic exclusion window of 30 seconds and a maximum duty cycle of 1 second. MS/MS spectra were acquired in the ion trap using the rapid scan mode.

3.7.3 LC-MS/MS (TimsTOF) - Evo-sep (Standard High-Throughput Approach)

Digested peptides from the above RWS approach were loaded onto C18-containing pipette tip pre-columns (EvoSep) following the manufacturer's protocol. Briefly, the tips were activated in 1-propanol, conditioned in buffer B and then equilibrated in buffer A. Peptides were then loaded onto the tip and washed with buffer A. Finally, buffer A was added to the tips to avoid the C18 material dried-out before and during analysis. Buffer A: 0.1% formic acid in water. Buffer B: 0.1% formic acid in acetonitrile. Peptides were analysed using an EvoSep One (EvoSep) coupled to a timsTOF Pro (Bruker) using a fixed gradient of 11, 21, or 44 minutes using a 100, 100, or 150 µm x 80, 80, or 150 mm C18 column packed with 3 µm particles (EvoSep). The particular parameters were selected based on the number of samples to be run per day i.e. 100, 60, or 30 samples per day.

3.7.4 LC-MS/MS - Dionex Nanoflow Coupled with TimsTOF (Optimised high-throughput approach)

Digested peptides from the above RWS approach were analysed by a nano-UPLC-MS/MS using a Dionex Ultimate 3000 coupled to a timsTOF Pro (Bruker). A 75µm × 150mm column (IonOpticks) with 1.6µm particles was used at a flow rate of 400 nL/min. Liquid chromatography was performed on a 17-minute linear gradient from a 2% Buffer B to 30% Buffer B (A = 0.1% formic acid in water. B = 0.1% formic acid in acetonitrile). Both the TIMS accumulation and ramp times were set to 100 ms and mass spectra were selected from 100 - 1700 m/z with a 0.85 – 1.30 Vs/cm2 ion mobility range. Precursors were picked for fragmentation from an area of the TIMS-MS scan that removed the majority of ions with a charge state of 1+. These precursors were isolated via ion mobility dependent collision energy which increased linearly from 27 - 45 eV over the entire ion mobility range. Three PASEF

MS/MS scans were obtained per full TIMS-MS scan, producing a duty cycle of 0.53 seconds. Ions were included in the PASEF MS/MS scan if they met a target intensity threshold of 2000. These ions were sampled multiple times until a summed target intensity of 10,000 was reached. A dynamic window of 0.015 m/z by 0.015 Vs/cm^2 was used. Summed ions were excluded from reanalysis for 24 seconds.

3.7.5 MaxQuant Settings

MaxQuant [256] was used for the processing of the Lumos or timsTOF raw files. The following workflow was adopted: The raw data files obtained from the Orbitrap were searched against the Uniprot Reference Human Proteome database (UP000005640, Downloaded: 17/01/2018). Peptides were searched against a concatenated target-decoy database with reverse protein sequences [257] and given a probabilistic score using Andromeda [258]. Enzyme digestion specificity was set to the default "Trypsin/P" which mimics (*in silico*) trypsin's ability to cleave the carboxyl side of lysine and arginine except if found before proline. The missed cleavage threshold of 2 was set. Peptide and protein identifications are reported using the protein parsimony principles as described by Nesvizhskii and Aebersold [259]. Peptide and protein identifications were filtered to a 1% false discovery rate at both the Peptide Spectrum Match and Protein level. A unique peptide threshold of 1 was set to identify a protein group. "Cysteine carbamidomethylation" was set as a fixed modification while "methionine oxidation" and "N-terminal acetylation" were set as variable modifications, finally where specified the "match between run" algorithm was enabled, all other parameters were set to default. Protein and peptide quantification was performed using MaxLFQ and iBAQ [260].

3.7.6 Fragpipe / MSFragger

MSFragger [261–263] was used to process the 156 timsTOF raw files generated for the large proteomic study in chapter 6. LC-MS/MS raw data files were analysed using Fragpipe v15 with MSFragger v3.2. The default settings for MaxQuant analysis were replicated. Briefly, a precursor mass tolerance of 50 ppm and a fragment

mass tolerance of 20 ppm was used for peak matching and, mass calibration and parameter optimisation was performed. The missed cleavage threshold of 2 was set. Oxidation of methionine and n-terminal acetylation were allowed as variable modifications and carbamidomethylation of cysteine was set as a fixed modification. Post-processing of MSFragger results was performed using Philosopher [264] with default settings within the FragPipe pipeline. Peptides and proteins were filtered to a 1 % false discovery rate. Protein intensities were normalised and quantified using IonQuant within FragPipe using the MaxLFQ algorithm with a minimum of 2 ions. The match between runs option was enabled. Library samples were microdissected areas (x8, 1,000,000µm^2), pooled, and fractionated into 8 fractions with a Bravo HpH Assaymap (Agilent) to generate more protein identifications for MBRs.

3.7.7 Proteomic Quality Control

To highlight potential problems with sample handling, liquid chromatography, or mass spectrometry steps; text files generated by Maxquant were analysed using an automatic Proteomics Quality Control (PTXQC) package [265].

3.7.8 Data Analysis

Perseus was used for pipeline analysis of Maxquant's text files [266, 267]. Briefly, raw data were filtered for known contaminates, reverse sequences, and proteins only identified by site. Raw LFQ intensities were then LOG2 transformed, and if appropriate, normalised. Samples including replicates were designated into categorical groups; proteins within groups with less than 70% valid values were removed. The remaining NaN values were imputed, if appropriate, using NAguideR [268]. Where applicable, para- or non-parametric variance tests were performed using the default settings. For 1D and 2D enrichment analysis, GO annotations were imported using Perseus, and iBAQ or LogFold difference was used as input data. Gene list analysis from the PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System [269] was used for Gene Ontology investigations using default settings. REVIGO [270] was used when necessary to reduce GO terms

using default settings. For brief GO over-representation summaries, g:Profiler was used [271]. Unless otherwise stated the background list was the total population of proteins suitable for analysis. Results from data analysis were plotted using R.

3.8 Digital Pathology

3.8.1 Whole Slide Imaging

Whole slide imaging was achieved by scanning sections on an Aperio Scanscope Turbo AT with a 20x/0.75 NA Plan Apo objective lens coupled with an x2 optical magnification lens to achieve a total magnification of 40x and a resolution of 0.25 microns/pixel.

3.8.2 Image Analysis

Aperio's raw files (.svs) were opened in Qupath (0.2.0-m2) [272] for visualisation. To overcome the inherent difficulties of performing image analysis on whole digital slides, each slide was split into individual tiles comprising of a grid of 250-600 tiles, with x, y and z coordinates retained; a downsampling factor of 4 was applied. Each tile was then sent to Fiji for machine learning analysis: The Trainable Weka Segmentation tool [273] was used for image segmentation. In short, a classifier was trained for each primary antibody used; for each classifier 15 representative images were hand-annotated to define cytological features, this was also done for each cell-type. Once trained, the classifier was used to segment the entire tile cohort individually. After analysis, the resulting annotations corresponding to cell boundaries were transferred to the original .svs file. Size exclusion quality control steps were performed to remove artefacts due to misfolding tissue, faint staining, edge effect, dirt or debris. Positive pixel count was then performed on the new file which now contains the machine learning annotations and the non-reduced digital information. Scripts for the above pipeline were designed to progress through the workflow in a semi-automated manner.

Transcript	Forward Strand	Reverse Strand	Rationale
CALB1	GAGCCAACAATCCAGCCTTCT	ATTTCGACGCTGACGGAAGT	Neuronal
GADPH	AATGGGCAGCCGTTAGGAAA	GCCCAATACGACCAAATCAGAG	House-Keeping Gene
MAP2	CCAGTTTCTGCGCCCAGATTT	AGCTCCCAATCAATGCTTCCT	Neuronal
OLIG2	TCGCATCCAGATTTTCGGGT	GCAGAAAAAGGTCATCGGGC	Oligodendrocytic
RBFOX3	TTGGCTGTAGGCACACAGAG	TTTCAACCTCCAGGACCGAA	Neuronal
RPL13	CCTTTCCGCTCGGCTGTTTT	GGCCTTACGTCTGCGGATCT	House-Keeping Gene
S100B	GACAAGGAAGAGGATGTCTGAGC	TGATGAGCTCCTTCAGTTCGG	Astrocytic
TMEM119	AGCACGGACTCTCTCTTCCAG	GTGCCCCCAGGACCAGTTC	Microglial
UBE2D2	AGAGAATCCACAAGGAATTGAATGA	TAGGGACTGTCATTTGGCCC	House-Keeping Gene

Table 3.2: List of oligonucleotides used in this thesis.

3.9 RT-qPCR

Oligonucleotides were designed to ensure mRNA from microdissected post-mortem neurons were of sufficient quality for qPCR, and therefore, RNA-sequencing. Multiple oligonucleotides were generated to ensure cell-type-specific transcript expression could be resolved and was used to ensure microdissected samples contained enriched amounts of the cell-of-interest before sequencing.

3.9.1 Primer Selection

Oligonucleotides were designed to target neurons (CALB1, MAP2, RBFOX3), astrocytes (S100B), oligodendrocytes (OLIG2) and microglia (TMEM119). Additionally, primers for housekeeping genes were also designed - GADPH, RPL13, and UBE2D2.

3.9.2 Primer Design

All primer pairs were designed using Primer-BLAST ensuring that sequences covered human exon-exon boundaries to avoid amplifying gDNA and to selective target sequences arising from mRNA. Where ever possible, the primer pair was selected that captured the maximum number of isoforms, had a length of ≈21 base pairs, produced a product of 70-100 base pairs, and had a GC content of 50–60%. Primer pairs were confirmed to be specific to the desired transcript by performing BLAST and UCSC *in silico* PCR. Selected primer pairs were then analysed with IDT's OligoAnalyzer to ensure the minimal formation of hairpin structures, self-dimers, hetero-dimers or any other secondary structures. All primer pairs were reconstituted to 100 µM and stored in 10 µM aliquots using ultra-pure water in a lateral flow hood.

3.9.3 Primer Optimisation

Primer pairs of different concentrations were tested to select the optimal pair for qPCR. Concentrations ranged from 100 nM to 800 nM and were processed in triplicate using 5 nanograms of cDNA. Melting curve analysis was performed on all primer pairs and the primer pair with a single dominant peak and lowest Cq value were selected for future use. To ensure the correct transcripts were being amplified, electrophoresis of the PCR products was performed by placing them on a 5% polyacrylamide gel for 1 hour at 100V. Amplified products within the gel were stained with 3X GelRed (biotium, 10,000X stock solution) for 1 hour and visualised using a Gel Doc EZ Imager (BIO-RAD). Correct product sizing was confirmed using Quick-Load Purple Low Molecular Weight DNA Ladder (NEB) and BIO-RAD's Image Lab Software. Next, the efficiency of the selected primer pairs was determined by LinRegPCR. A 6 point standard curve was prepared ranging from 0.406 - 13 nanograms to investigate if primer pair efficiency varied with increasing input material. The Pfaffl method [274] was used analytically to compare Cq values between the transcript of interest, taking into account the different primer efficiencies. GeNorm [275] was used to select the best housekeeping genes in any given experiment.

3.10 LCM-seq

3.10.1 Preparing and Validating SPRI Beads

All procedures were performed in RNA-free conditions using diluted sodium hypochlorite, 90% alcohol, and RNAaseZAP. To clean and size-select nucleic acids an in-house SPRI bead mix was prepared (PEG 8000, 19.5% or 24% (w/v); NaCl, 1M; Sodium Azide, 0.05% (v/v); EDTA, 1mM; Tris-HCL, 19mM; Igepal CA630, 0.01% (v/v); magnetic SPRI beads, 1% (v/v)) following previously published methods [117, 276]. SPRI beads were cleaned before being added to the SPRI bead mix, following the same protocol. By comparing pre-purification and post-purification DNA (GeneRuler 50 bp DNA Ladder; ThermoFisher) it was possible to determine if the

SPRI bead mix could successfully capture and bind DNA during wash steps. It was also possible to calculate the percentage recovery rate of the SPRI by analysing the band intensity of the pre-purification and post-purification DNA. Overall, the in-house SPRI bead mix recovered 81.1% ±4.83% of nucleic acids (data not shown). Various ratios of SPRI bead mix (ThermoFisher) were added to GeneRuler 50 bp DNA Ladder (ThermoFisher) to determine the ratio needed to selectively capture nucleic acids of 200bp or more (data not shown). To ease cDNA clean up steps, the final PEG concentration was adjusted to ensure a 1:1 ratio of nucleic acids:SPRI bead mix would capture nucleic acids of 200bp or more.

3.10.2 Library Construction

Reverse Transcription and cDNA Amplification

Library preparation for sequencing on an Illumina Nova Seq 6000 sequencer was performed on a modified version of the Smart-seq2 based LCM-seq protocol [116]. All workspaces, equipment, and surrounding equipment cleaned with diluted sodium hypochlorite, 90% alcohol, and RNAaseZAP. Unless stated, all procedures were performed on ice. After microdissection, 1 µl 10mM dNTP (ThermoFisher, 10mM per nucleotide), 1 µl (10µM) oligo-dT primer (Sigma Aldrich, 5'- AAGCAGTG-GTATCAACGCAGAGTACT30VN -3'), and 0.1 µl of ERCC spike-ins (2.5 × 10^5 dilution from stock; Invitrogen) was added to the lysate and incubated at 72°C on a heat block for 3 minutes to open secondary structures and then snap cooled on ice. Reverse transcription was performed with the following mix: 2 µl SSRTII 5 × buffer (ThermoFisher), 0.5 µl 100mM DTT, 0.5µl SSRTII enzyme, 2 µl 5M betaine, (Sigma Aldrich), 0.1µl 1M MgCl2 (Sigma Aldrich), 0.25µl 40 U µl-1 RNase inhibitor and 0.1 µl 100µM TSO-LNA-oligo (Sigma Aldrich, 5'-AAGCAGTGGTATCAACGCAGAGTACrGrG +G-3'). The following reverse transcription protocol was used: 90 minutes at 42°C, 10 cycles of 2 minutes at 50°C and 2 minutes at 42°C, followed by 15 seconds at 70°C. Samples were kept at 4°C until cDNA amplification. For cDNA amplification, 25 µl of KAPA HiFi HotStart Ready mix was added with 0.2 µl of 10 µM ISPCR primers (Sigma Aldrich,

5'- AAGCAGTGGTATCAACGCAGAGT -3'), and 2.3 µl PCR grade water. The following protocol was used for PCR: 3 minutes at 98°C, 18 cycles of 20 seconds at 98°C, 15 seconds at 67°C, and 6 minutes at 72°C Finally, samples were incubated at 72°C for 5 minutes and then chilled at 4°C until the next step.

SPRI Bead Purification

The SPRI bead mix was prepared following 3.10.1. The bead mix was briefly mixed and incubated at room temperature for 20 minutes before the following steps. First, an equal ratio of cDNA to bead mix was added (25µl:25µl) to ensure only nucleic acids of 200bp or more were purified. After light mixing, the cDNA:bead mix was incubated at room temperature for 8 minutes and then placed on a magnet stand for 2 minutes. The supernatant was removed while on the magnet stand and 150µl of freshly made 80% ethanol was added for 1 minute to wash the bound cDNA. Residual ethanol was removed by either air-drying samples, or placing them on a 37°C heat block for 3 minutes until dry. Samples were considered dry when the characteristic *"hairline cracks"* [117, 276] was observed. Once dry, beads were rehydrated and resuspended in Buffer EB (17µl, QIAGEN) by pipette mixing. If necessary, samples were briefly spun down. Once rehydrated, samples were incubated for 10 minutes at room temperature, placed on a magnet stand for 2 minutes, and the supernatant was transferred to a new PCR tube. After bead purification, the electropherogram of the samples was inspected to determine the level of degradation and calculate the average fragment size using an Agilent 2100 Bioanalyzer using the High Sensitivity DNA kit. The concentration of the cDNA was measured using the Quant-iT™ PicoGreen™ dsDNA Assay and a FLUOstar Omega Plate reader. Standard curve and linear regression analysis were performed using custom scripts written in R.

Transposase-based Tagmentation

Prior to tagmentation, a subset of cleaned cDNA samples was subjected to qPCR to ensure capability with RNA-sequencing. Several cell-type-specific transcripts were selected for amplification, as a rough measure to confirm the enrichment of neuronal material (data not shown) and ensure material could be amplified. To

convert cDNA into sequencing libraries Illumina's Nextera XT DNA kit was used following manufacturers instructions. Briefly, 10 µl of Tagment DNA Buffer was added to 1 ng of cDNA, next 5 µl of Amplicon Tagment Mix was added and placed on a preheated thermocycler set at 55°C and incubated for 5 minutes and then held at 10°C. Once at 10°C, 5 µl Neutralize Tagment Buffer was immediately added to stop the reaction. To enable multiplexing of samples, dual sequencing indices (5 µl each, Nextera XT Index Kit v2, Set A) were added to each sample. To ensure correct colour balance across samples, combinations of i7 and i5 adapters were selected using Illumina's Experiment Manager. The Nextera PCR Master Mix was then added (15 µl) and the following program was run: 72°C for 3 minutes, 95°C for 30 seconds, followed by 12 cycles of 10 seconds at 95°C, 30 seconds at 55°C, and 30 seconds at 72°C. Finally, samples were incubated at 72°C for 5 minutes. The above bead clean-up was repeated using a 24% PEG bead mixture to remove unwanted primers and/or adapters.

3.10.3 RNA-Sequencing

Illumina's bead-based normalization was omitted due to incompatibility issues in downstream quality control checks; instead, all 39 individual libraries were pooled together manually ensuring a final library of 10nM. Library concentration was determined using Quant-iT™ PicoGreen™ dsDNA Assay and fragment sizes determined using Agilent's High Sensitivity DNA kit on a 2100 Bioanalyser. DNA Molarity was calculated using the following:

$$\frac{Concentration(ng/ul)}{(660g/mol \times averagelibrarysize(bp)} \times 10^6 = nM$$

Once pooled, the sequencing library was sent to the Oxford Genomics Centre at the Wellcome Trust of Human Genetics where independent quality control checks were performed before sequencing. RNA-sequencing was performed on a NovaSeq 6000 system using one lane of an S4 flow cell, reads were paired ends and at 150 base pairs.

3.10.4 Data Pre-processing and Data Analysis

Samples were demultiplexed and processed by the Oxford Genomics Centre and passed internal quality control. The associated .fastq files were screened with FastQC [277] to look for potential outliers or issues with library construction. The median read count was 79,687,430; lowest read count: 34,763,561; largest read count: 111,573,341. In total, 62.41% of sequences were human-specific and 35.49% were found across multiple genomes. There were very little mouse (0.16%), rat (0.13%), yeast (0.01%), bacterial (0.09%) specific sequences; 0.01% of sequences were flagged as contaminants. FastQC highlighted that approximately 60% of counts represented duplicated reads. The saturation of read depth showed that the full diversity of RNA in the microdissected samples was captured; overall, ≈30 million reads were unique across samples. Moreover, FastQC also highlighted several over-represented sequences and considerable adaptor content - likely a result of using the recommended amounts of primers rather than scaling down for microdissected samples. BLAST revealed that a large proportion of over-represented sequences were that of ribosomal RNA (rRNA) despite the use of Oligo(dT) primer to capture and amplify mRNA. Trimmomatic [278] was used to remove Illumina adapter and primer sequences from .fastq files. Next, RNA-seq reads were mapped to the human reference genome (hg38) using STAR with default settings [279]. FastQC was used on the resulting .bam files to ensure sufficient mapping to the reference genome. Next, the .bam files were then cross-referenced to a database containing known human rRNA sequences using the split_bam.py script from RSeqQC [280] to remove rRNA sequences. Finally, FastQC and RSeqQC were used to confirm the removal of rRNA and to screen for GC % content, gene body coverage, and inner distance. To generate a count matrix the number of uniquely mapped reads of each microdissected sample was calculated using featureCounts [281] with default parameters at the gene level. To quantify and normalise the expression of genes along with remove outliers, DESeq2 [282] was implemented using default settings. The resulting heatmaps, volcano plots, and box plots were generated in R. For GSEA, the fgsea package [283] was implemented and genes were ranked by DESeq2's test statistic.

4

Developing a sensitive, scalable method for spatial, cell-type-resolved microproteomics of the human brain.

Contents

4.1 Aims of this Chapter 86
4.2 Introduction 86
4.2.1 Breakdown of this Chapter 87
4.3 Results 88
4.3.1 LCM LC-MS/MS can Identify Proteins from Minute Amounts of Post-Mortem Human Brain Tissue 88
4.3.2 LCM LC-MS/MS of Post-Mortem Purkinje Cells successfully Identified Several Proteins relating to Purkinje Cell Activity 92
4.3.3 Partial Recovery of Protein leads to a Marked Decrease in Protein Identifications 96
4.3.4 Sample Preparation Dramatically affects the Number of Identifiable/Quantifiable Proteins 99
4.3.5 Effect of Histology Staining on LCM Proteomics 102
4.3.6 Effect of immunohistochemistry on LCM Proteomics . . 106
4.4 Discussion 110
4.4.1 Main Conclusions 118

4.1 Aims of this Chapter

- Determine the performance of LCM proteomics on in-house post-mortem material.
- Optimise a laser capture microdissection mass spectrometry (LCM LC-MS/MS) approach to maximise data output from minute amounts of cellular material.
- Investigate how a potential workflow would be affected by upstream histological procedures.

4.2 Introduction

A huge amount of information regarding cellular function and composition can be inferred from the vast and extremely complex human proteome. However, such proteomic experiments have been hindered by the limited access of human CNS tissue, difficulty obtaining enriched populations of cells, and the increased challenge of working with proteins that cannot be exponentially duplicated; requiring the use of extremely sensitive equipment and highly efficient workflows.

To investigate the proteome of individual cell-types the cellular heterogeneity of the brain needs to be simplified. The focus of this thesis is the Betz cell, which, to date, cannot be systematically targeted with cell isolation techniques like FACS due to its size and lack of biomarkers. Moreover, the Betz cell at present can only be identified by morphology, therefore, the obvious technique to isolate Betz cell from its microenvironment for proteomic investigation is laser capture microdissection. However, before comprehensive experiments can be performed on rare cell-types like the Betz cell, protocols need to be established and optimised to maximise proteomic data output with minimal sample input. As discussed in Section 2.2.6, LCM proteomics has been previously achieved on post-mortem neurons, however, sample input either consists of high material input (thousands of cells) making rare cell-type investigations cumbersome or requires whole tissue areas to be microdissected resulting in an averaging effect of cellular phenotypes which fail to achieve single

cell-type resolution, potential masking critical observations. This chapter aims to describe the development of a sensitive approach that addresses the above issues and allows for minute amounts of neurons to be laser microdissected to achieve spatial, cell-type resolved proteomics to an unprecedented depth.

4.2.1 Breakdown of this Chapter

1. An initial experiment will be carried out to see if proteomic data can be generated from microdissected Purkinje cells.
2. Attempts will be made to optimise an LCM LC-MS/MS workflow to maximise the number of proteins that can be identified from small numbers of cells.
3. The optimised workflow, referred to as RWS herein, will be tested on frozen and FFPE microdissected neurons.
4. Common staining techniques i.e. IHC and histological stains will be coupled with the RWS workflow to investigate how they perform.

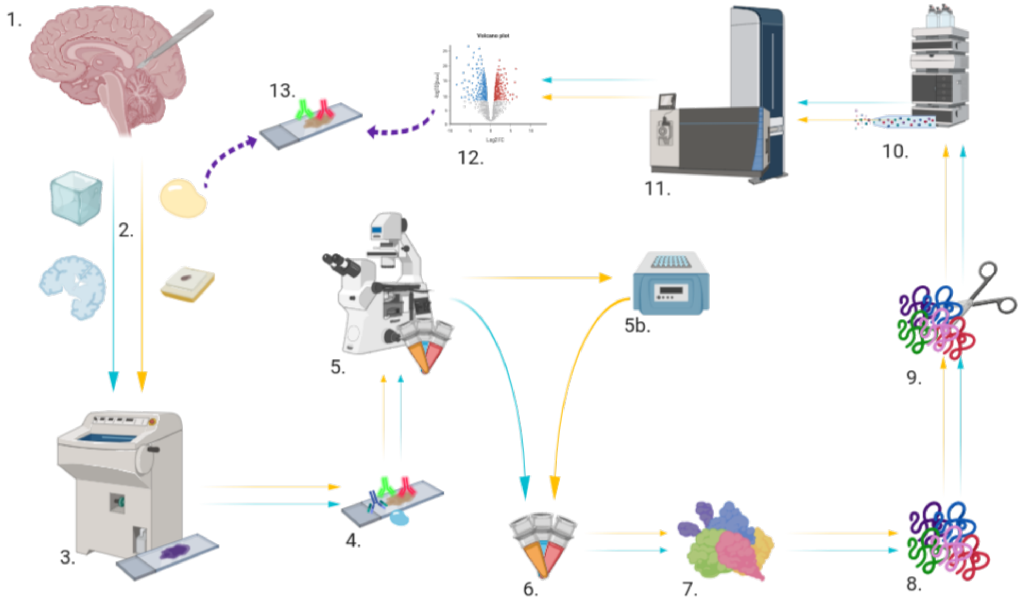


Figure 4.1: Systematic for the LCM LC-MS/MS Workflow Implemented in this Thesis. 1: Areas of interest are dissected from post-mortem human brain material. 2: Tissue is either frozen or turned into FFPE blocks with special attention made to minimise protein/RNA degradation. 3: Samples are sectioned and placed on LCM-compatible PEN membrane slides. 4: Cells of interest are visualised using standard histological dyes or IHC. 5: Cells of interest are microdissected from their cellular environment and captured into tubes for downstream processing. 5b. FFPE-derived sections are heated to reverse formalin cross-linking. 6: Microdissected cells are pooled and lysed. 7: Proteins are purified by removing nucleic acids. 8: Proteins are reduced and alkylated. 9: Proteins are digested to produce peptides. 10: Liquid-chromatography is used to separate heterogeneous populations of peptides. 11: Mass spectrometry is used to determine peptide sequence and abundances. 12: Peptides are mapped to the human proteome, identified, and data analysis is performed to determine protein expression of microdissected cells. 13: Results from the previous step is used to back-map promising proteomic findings to the original tissue source. This chapter aims to optimise steps 4-8 to minimise sample loss and produce high-performance proteomics.

4.3 Results

4.3.1 LCM LC-MS/MS can Identify Proteins from Minute Amounts of Post-Mortem Human Brain Tissue

To test whether it was possible to identify proteins from our post-mortem human tissue and to investigate the minimum amount of input material required for LCM proteomics, a proof of principle experiment was designed using frozen human

cerebellum (54-Year-Old, Male, PMI = 72 hours, Neuropathology Diagnosis = ALS). The Purkinje cell was selected for this proof of principle experiment as they are easily recognised with standard histology stains, have a defined anatomical location, exhibit a large size relative to surrounding neurons, and are found in vast numbers making identification and collection less labour extensive. Moreover, unlike Betz cells, many Purkinje cell-specific proteins are known and should be detectable in the generated LC-MS/MS data-set, aiding validation. Tissue sections of the cerebellum were stained with Haematoxylin and Eosin (H&E) to better visualise the Purkinje Cells before LCM. Purkinje cells were microdissected and pooled into 0.5ml tube caps containing an adhesive to capture catapulted cells. Increasing numbers of Purkinje cells (10, 50, 100, 200, 400, and 800) were pooled into separate tubes for downstream mass spectrometry. In addition to these pooled cell samples, a library sample was created using a single unstained cerebellum cryosection that was processed in parallel and used in conjunction with MaxQuant's Match Between Runs (MBRs) Algorithm [256] to potentially improve performance. All pooled cell samples were confirmed as successfully microdissected by reviewing the histology slide and inspecting the catapulted cells on the inside of the 0.5ml tube cap – see Figure 4.2.

The total area microdissected was recorded to help standardise the workflow and provide informed data for future experiments that may target cells that differ in size – Table 4.1. After successful capture, samples were spun down in 6M Urea, denatured, alkylated, and digested with trypsin before LC-MS/MS using standard protocols - see Section 3.6.1. The resulting data-set was mapped to a reference proteome (Homo sapiens, UP000005640) and features recognised using MaxQuant[1].

Within the library cerebellum sample, it was possible to identify 18,283 peptides which mapped to 3,218 protein groups. The results from this proof of principle experiment can be seen in Figure 4.4. Among the microdissected samples, a steep increase in identifiable peptides and proteins was seen between the 200, 400, and 800 Purkinje cell samples. Interestingly, a strong relationship between IDs and

[1]post-LCM LC-MS/MS sample preparation for this Purkinje Cell Titration Experiment was performed by Simon Davis - Target Discovery Institute, University of Oxford.

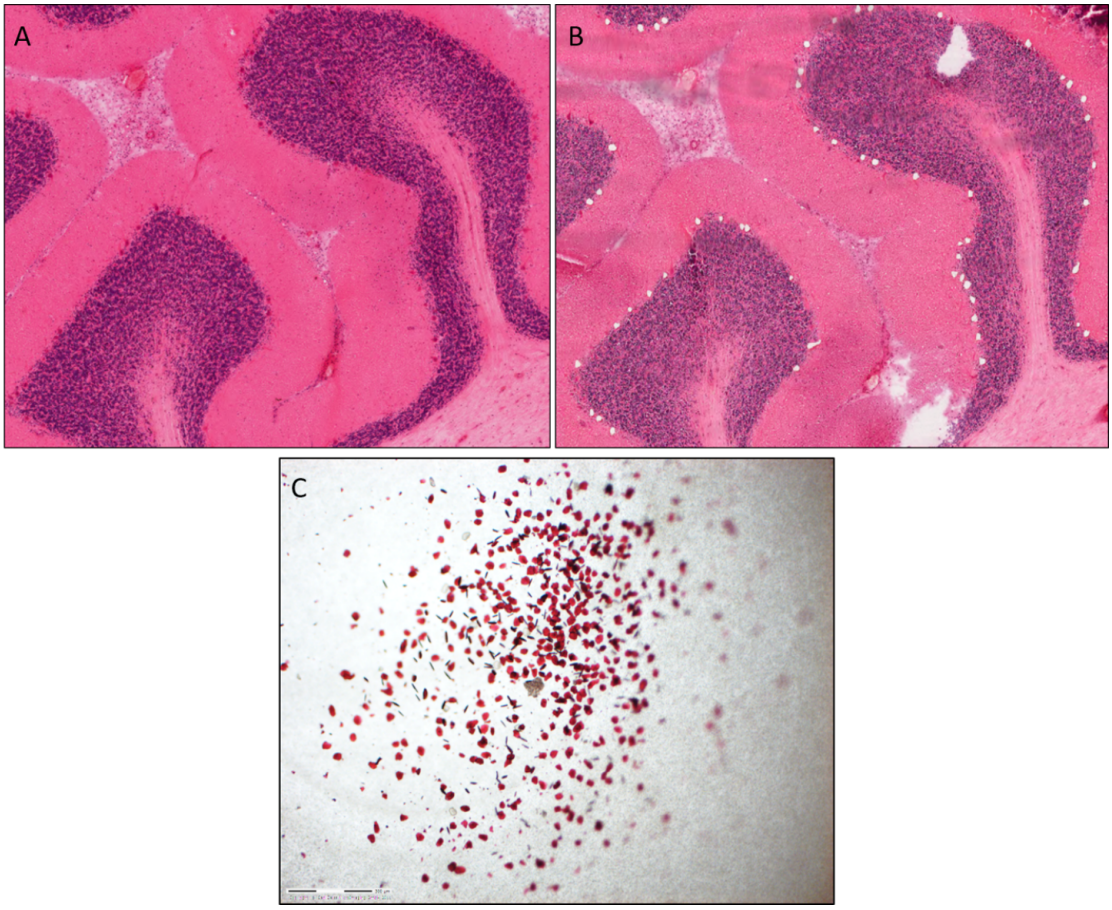


Figure 4.2: Purkinje Cell LCM - Before/after/in-cap.
A: Shows frozen cerebellum stained with H&E before LCM. B: Shows the aftermath of LCM, with Purkinje cells successfully removed from the Purkinje cell layer. C: Inspection of the LCM cap confirms that Purkinje cells have been successfully collected for downstream processing.

Distribution of Purkinje Cells - Titration

No. Purkinje Cells	Total Area Taken (μm^2)	Average Purkinje Cell Area (μm^2)	Average Purkinje Cell Area (SD)
10 (10)	8434	843.4	190.32
50 (51)	37773	740.65	171.07
100 (96)	74290	773.85	204.34
200 (190)	117722	619.59	187.23
400 (407)	308697	758.49	180.16
800 (781)	590213	755.72	213.05

Table 4.1: Summary statistics for initial Purkinje Cell Titration.
SD: Standard Deviation.

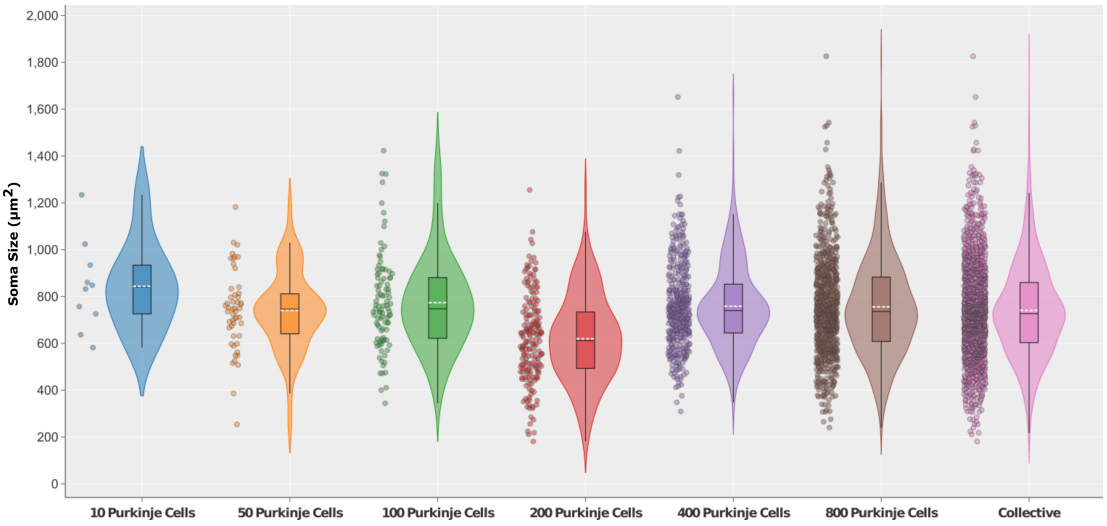


Figure 4.3: Purkinje Cell Titration - Soma Size.
Violin plots show the full distribution of the Purkinje cell titration experiment. Solid black like denotes median value, while dotted white line shows mean value.

input material was not seen despite the constant doubling of material from 50 Purkinje cells to 800 Purkinje cells. This could be attributed to a significant loss of material during the processing steps, which would be more detrimental to samples containing fewer cells; alternatively, the lower cell fractions may contain levels of peptides lower than the detection threshold of the mass spectrometer.

Matching to a library did increase the overall number of IDs for proteins identified by MBRs, contributing to 27.4 and 20.4% of the total proteins found within the 400 and 800 Purkinje cells, respectively. Within the 10 – 100 Purkinje cells, proteins found by matching to the library were higher than those found by MS/MS spectra alone - emphasising the beneficial effect of the MBRs algorithm. Combining the identifications from both MS/MS spectra and those from matching to the library cerebellum sample resulted in 125, 84, 108, 181, 4221, 6728 peptides being identified for the 10, 50, 100, 200, 400 and 800 pooled Purkinje cell samples, respectively; these peptides resulted in a total of 67, 53, 73, 107, 1517 and 2051 proteins being identified in their respective samples - see Figure 4.4.

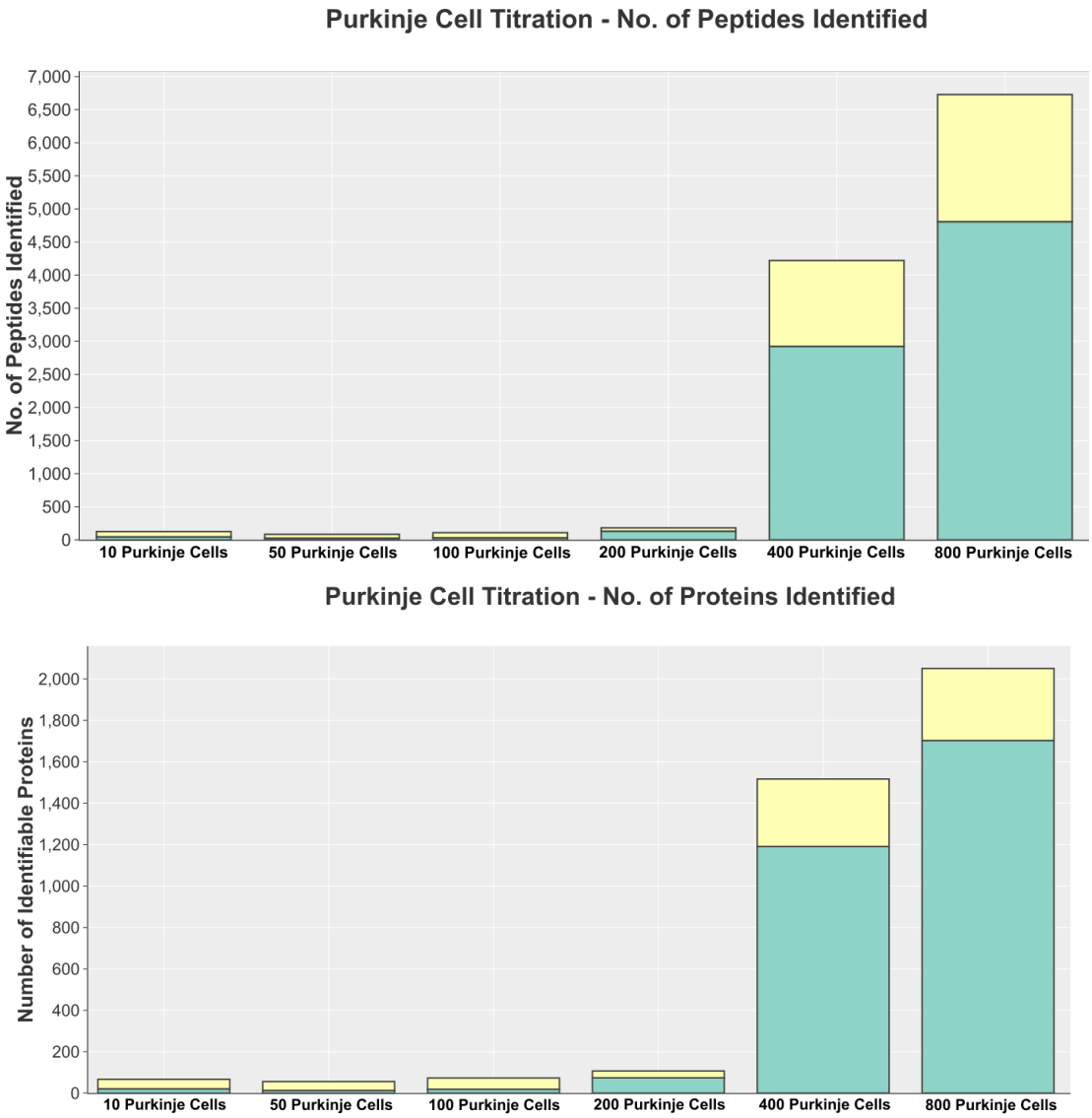


Figure 4.4: Purkinje Cell Titration - No. of Identifiable Peptides.
The results of the first Purkinje cell titration. Turquoise bars denote the number of peptides/proteins identified solely by MS/MS, with cream bars showing the number of peptides/proteins identified by MaxQuant's Matched-Between-Run (MBRs) algorithm.

4.3.2 LCM LC-MS/MS of Post-Mortem Purkinje Cells successfully Identified Several Proteins relating to Purkinje Cell Activity

To confirm that LCM LC-MS/MS of microdissected Purkinje cells could resolve individual proteins that play a role in Purkinje cell function, the full list of proteins identified in this experiment were screened for their associated Gene Ontology

(GO) annotations. Proteins that contained GO annotations involved in Purkinje cell activity were highlighted as "Purkinje cell-associated" proteins. Despite the shallow protein depth in the smaller cell fractions; all samples contained proteins that had Purkinje cell-associated GO annotations - a full list of which can be found in Table 4.2. These data show that increasing proteomic depth, unsurprisingly, gives rise to more Purkinje-associated proteins. However, this also suggests that there may exist a soft threshold whereby, the level of proteomic depth achieved is representative of a cell-type-specific molecular profile.

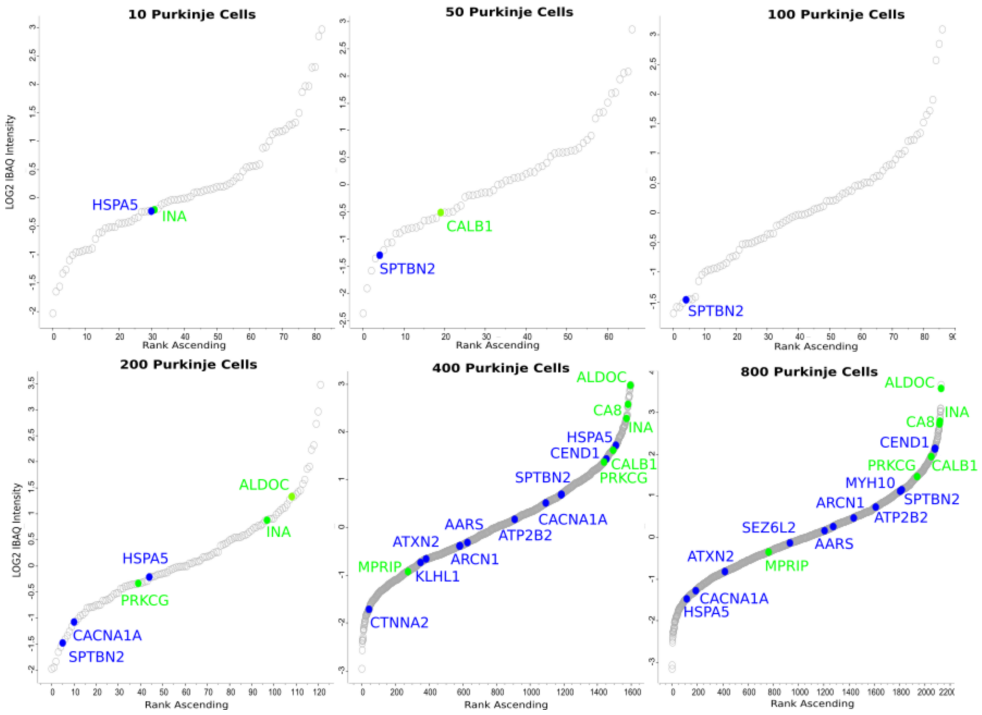


Figure 4.5: Purkinje Cell-Associated Proteins within Purkinje Cell Titration. Intensity based absolute quantification (iBAQ) highlights several proteins that are involved in Purkinje cell biology. Blue: Proteins with Purkinje cell-associated GO annotations. Green: Known canonical biomarkers for the Purkinje cell.

Purkinje cells	Proteins Identified	No. of Proteins
10	HSPA5, INA	2
50	SPTBN2, CALB1	2
100	SPTBN2	1
200	ALDOC, CACNA1A, HSPA5, INA, PRKCG, SPTBN2	6
400	AARS, ALDOC, ARCN1, ATP2B2, ATXN2, CA8, CALB1, CACNA1A, CEND1, CTNNA2, HSPA5, KLHL1, INA, MPRIP, PRKCG, SPTBN2	16
800	AARS, ALDOC, ARCN1, ATP2B2, ATXN2, CA8, CALB1, CACNA1A, CEND1, HSPA5, INA, MPRIP, MYH10, PRKCG, SEZ6L2, SPTBN2	16

Table 4.2: Summary of the Purkinje Cell-Associated Proteins found in this Experiment.Protein List of which Purkinje cell-associated proteins that were identified in each titration sample.

Although the above proteins are known to be involved in Purkinje cell physiology, it is important to confirm that these proteins are localised to Purkinje cells in the human brain. To validate these preliminary data, the above proteins were screened using the Human Protein Atlas database (HPA) [78]. Protein evidence score was used along with a visual interpretation of HPA's IHC data on human biopsied cerebellum to authenticate the proteomic data. Of the proteins 18 identified above, 16 were found to be located within human Purkinje cells - see Figure 4.6. The remaining 2 (KLHL1 & CACNA1A) had ambiguous staining, although there is some evidence that they are localised to Purkinje Cells in mice [284, 285]. Importantly, some of the proteins above were – in the context of the cerebellum – localised solely to Purkinje cells and not the surrounding microenvironment e.g. ARCN1 and MHY10. This is relevant as it cannot automatically be assumed that all proteins identified by laser microdissection originated from our intended target and not from the surrounding "contaminating" tissue which may consequently also contain proteins that have immuno-reactivity within Purkinje cells.

Encouragingly, along with proteins with Purkinje cell GO terms; several canonical biomarkers for Purkinje cells were also identified in this data-set. Purkinje Cell

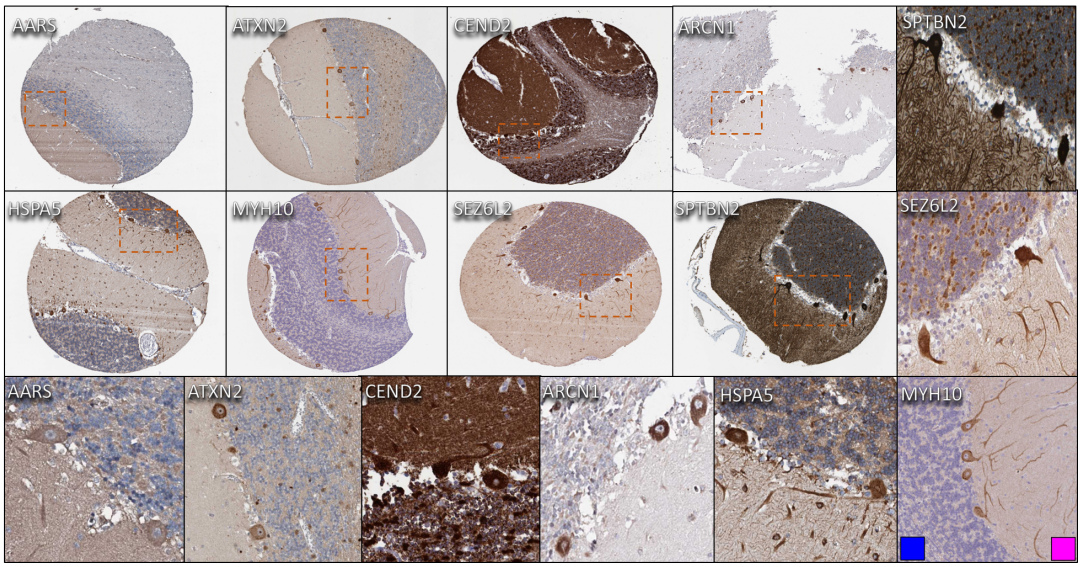


Figure 4.6: Validation of Purkinje Cell Protein IDs using the Human Protein Atlas. The publicly available IHC data on the HPA was used to screen for Purkinje cell-associated proteins identified in the Purkinje cell titration experiment to validate findings. IHC data for known Purkinje cell biomarkers were not shown as their localisation to Purkinje cells is well established. Purkinje cell immunoreactivity was seen in the vast majority of identified proteins. Overall immunoreactivity was varied, with ATXN2, ARCN1, MYH10, and SEZ6L2 being the most selective in visualising Purkinje cells, with the surrounding granular layer (blue box) and molecular layer (purple box) containing minimal positive staining.

Protein 2 (PCP2) and Purkinje Cell Protein 4 (PCP4) - aptly named for their high abundance within Purkinje Cells - were found in the 400 and 800 cell samples. Calbindin (CALB1), which is probably the most well-known Purkinje biomarker was found in 50% of the samples. Other biomarkers found in this initial data-set include ALDOC, PRKCG, MPRIP, INA, and CA8. These biomarkers tended to be some of the most highly expressed proteins; although cell-type-specific biomarkers may be present at unremarkable levels, the fact that the most abundant proteins, in this case, were known Purkinje cell biomarkers increases the confidence that unique biomarkers may be identified by LCM proteomics in the future i.e. in the context of the Betz cell.

Next, to confirm that laser microdissection does enrich for the cell of interest, the following comparison was performed, utilising the mammoth data generated by the HPA. IHC expression data from all proteins used by the HPA was downloaded and filtered to only show proteins which are highly expressed in Purkinje cells but had no

expression in the surrounding granular/molecular layers. This protein list was then filtered to only show the proteins which had the highest evidence of being found in Purkinje cells. The resulting protein list was used as a surrogate list of potential biomarkers of the Purkinje cell. Proteins within this list were matched to a list of proteins identified in the previous Purkinje cell experiment - including both Purkinje cells and cerebellum library samples, resulting in a small list of 21 proteins. Of the 21 proteins found in both Purkinje cell samples and library samples, 18 proteins were enriched in LCM Purkinje cells compared to the whole tissue cryosection, highlighting that LCM can enrich for proteins found in the microdissected material - see Figure 4.7. The remaining 3 proteins (TRIM3, ARRB1, and CMBL) were found in similar or higher amounts within the cryosection. However, these proteins are not solely located in the soma of Purkinje cells and also exist in their axons and dense arbour of their dendritic tree. As the dendritic arbor of Purkinje cells takes up a significant amount of the cerebellar molecular layer, it may be the case that for these proteins the enrichment of proteins from the perikaryon may be overshadowed by the large volume that the collective Purkinje cell dendritic arbor contributes to the cerebellar cortex. Moreover, Over-representation Gene Ontology Analysis (ORA) of the proteins identified in Purkinje cells but not in the cerebellum section, highlighted an increase in proteins found in the organelle subcompartment of cells (Adj. p value: 9.29e-6), mitochondria (Adj. p value: 0.02), Golgi apparatus (Adj. p value: 2.52e-2), endoplasmic reticulum (Adj. p value: 8.95e-4), and nuclear pore (Adj. p value: 1.16e-2); providing more evidence that microdissection of individuals cells can be used to enrich for intracellular proteins.

4.3.3 Partial Recovery of Protein leads to a Marked Decrease in Protein Identifications

It was noted that following centrifugation from the previous titration experiment, some material was still present on the 0.5ml cap; it was not clear if this material consisted of cellular remnants left from lysis (and thus unrecovered proteins) or residual dye from the H&E staining process. To investigate, the titration experiment

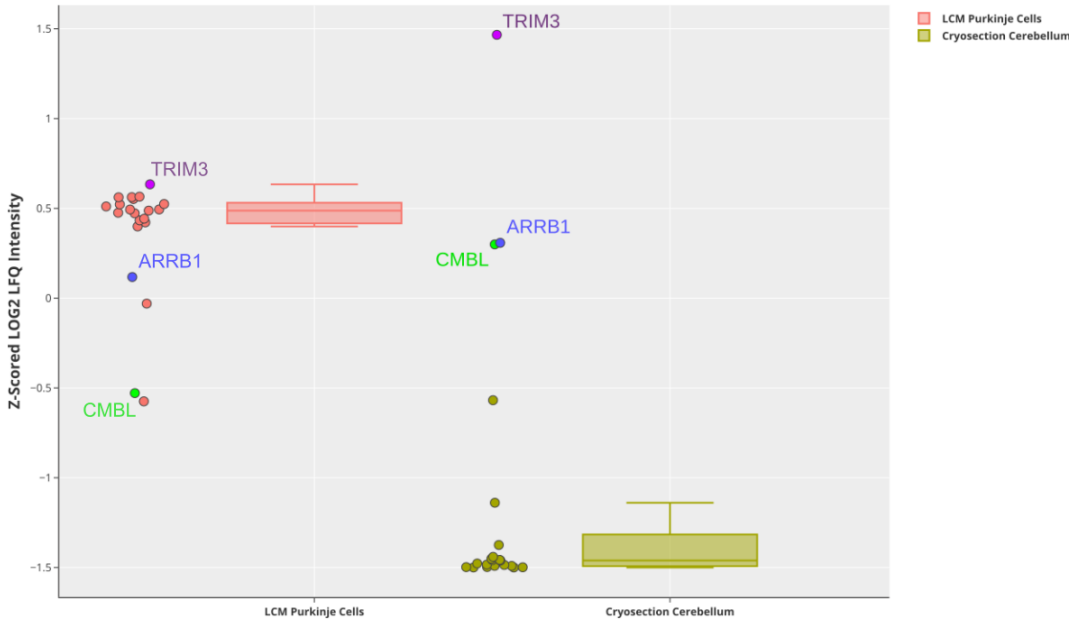


Figure 4.7: Our LCM LC-MS/MS Workflow Enriches for Cell-Type-Specific Proteins. The intensities of Purkinje cell-specific markers (identified by screening Human Protein Atlas' IHC data) found that LCM samples (pink) contained an enrichment of Purkinje cell-specific proteins compared to the LC-MS/MS output of analysing a single cerebellar tissue cryosection (green). Each point represents one of the Purkinje cell-specific markers found in both LCM and cryosection samples. Proteins that were comparatively enriched in cryosections have been colour coded.

above was repeated for 100, 200, 400, and 800 pooled Purkinje cells however all centrifugation, denaturing, and alkylating steps were omitted to minimise handling/pipetting steps (and therefore sample loss). Instead, Purkinje cells were digested with trypsin directly in the cap (*in situ* digestion). This approach impressively resulted in 3852, 8921, 8855, 8858 peptides being identified, mapping to 1497, 2224, 2360, and 2282 proteins, respectively (Figure 4.8). This translated to a remarkable 20.5- and 20.8-fold increase in protein identification in the smaller 100 and 200 cell fractions, and a modest 1.6 and 1.1-fold increase from the larger 400 and 800 cell fractions. Of all peptides identified, 75% had no missed cleavages, 22% had one, and 3% had two missed cleavages demonstrating good digestion efficiency. However, there was a clear bias in this data-set resulting from the lack of reduction and alkylation steps; with <1% of peptides containing a cysteine residue compared to the 8% found in our total data-set.

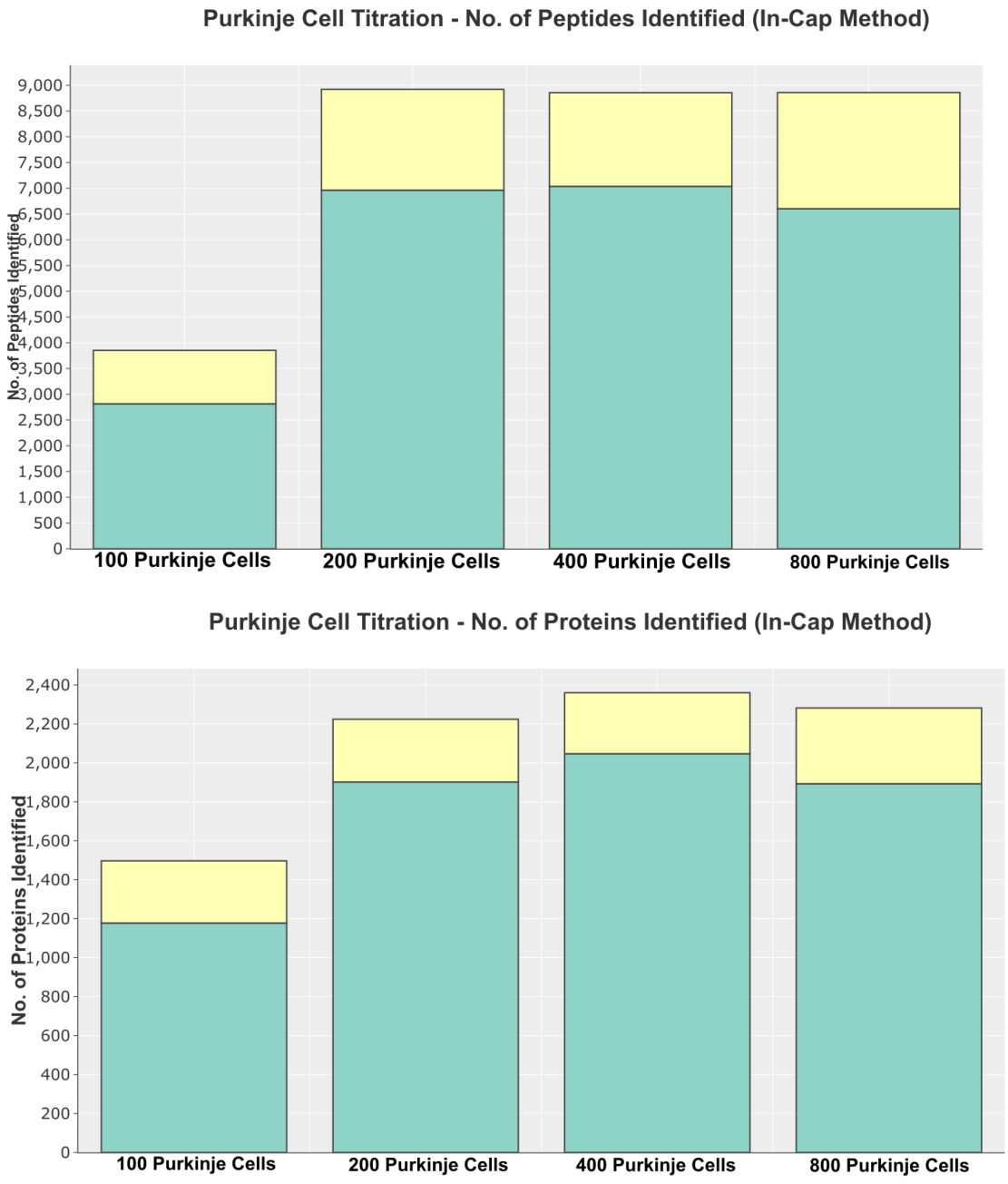


Figure 4.8: Purkinje Cell Peptides/Proteins - *In Situ* Digestion Method. The results of *in situ* digestion directly on microdissected neurons. Turquoise bars denote the number of peptides/proteins identified by MS/MS, cream bars show the number of peptides/proteins identified by the MBRs algorithm.

4.3.4 Sample Preparation Dramatically affects the Number of Identifiable/Quantifiable Proteins

Knowing that remnants left behind in the adhesive caps represented a large proportion of proteinous material meant it was critical to develop a workflow that captured as much of this material as possible. Moreover, for proteomics to be achieved on individual pooled cells, it was also critical to minimise sample loss. To address these concerns, a series of experiments were performed to optimise the LC-MS/MS sample preparation for laser microdissected material[2]. The LCM LC-MS/MS workflow can be split into three parts: cell collection, cell lysis and sample preparation. To find the optimal protocol, all 3 parts would need optimising, and for each part, there were several approaches to take. As testing multiple permutations would result in a high number of samples to be analysed, it was decided to perform the optimisations on the molecular layer of the human cerebellum which is easily identified, quick to microdissect, and should not differ drastically in tissue quality from neighbouring Purkinje cells. This also allowed all permutations to be microdissected on the same tissue slice reducing batch bias and importantly, eliminated the need to manually annotate and microdissect thousands of individual cells. Since it was previously demonstrated that a meaningful amount of proteins can be resolved from as little as 100 Purkinje cell slices or $\approx$ 60,000µm^2 it was decided to use this amount of starting material for all permutations. Two lysis approaches were used: collecting LCM targets into (1) adhesive caps as done with the previous Purkinje cell titration or (2) directly into a lysis buffer with regular 0.2ml tubes. Four lysis buffers were tested: 6M Urea, 6M Urea in 30% acetonitrile, 50% trifluoroethanol (TFE), and RIPA buffer. Finally, three methods of sample preparation were explored: (1) The "in-cap" *in situ* digestion method used on the previous Purkinje cell titration (Figure 4.8), (2) a scaled-down version of the standard LC-MS/MS sample preparation i.e. in-solution digestion (ISD), and (3)

[2]The data in this subsection was generated purely by Simon Davis - Target Discovery Institute, University of Oxford.

Lysis Buffer:	Collection Method:	Collection Method:	Protein Recovery and reduction:
6 M Urea	Buffer	20 µL 6 M urea, 0.1 M tris	Add DTT to 5 mM; Incubate 30 min, RT
	Cap	-	Add 10 µL 6 M urea, 0.1 M tris, 5 mM DTT; Incubate 30 min, RT
6 M Urea, 30% acetonitrile	Buffer	20 µL 6 M urea, 100 mM Tris, 30% acetonitrile (v/v)	Add DTT to 5 mM; Incubate 30 min, RT
	Cap	-	Add 10 µL 6 M urea, 0.1 M tris, 5 mM DTT, 30% acetonitrile (v/v); Incubate 30 min, RT
50% Trifluoroethanol	Buffer	20 µL Ultrapure water	Add 20 µL trifluoroethanol. Add DTT to 5 mM; Incubate 60 min, 60 °C
	Cap	-	Add 10 µL 50% trifluoroethanol (v/v), 0.1 M tris, 5 mM DTT; Incubate 60 min, 60 °C
RIPA Buffer	Buffer	20 µL RIPA	Add DTT to 5 mM; Incubate 30 min, RT
	Cap	-	Add 10 µL RIPA (Pierce #89900), 5 mM DTT; Incubate 30 min, RT

Figure 4.9: Permutations for Optimising LCM LC-MS/MS Sample Preparation.

and a modified SP3 protocol [137, 138]. All permutations were done in triplicate using a H&E stained cryosection of the cerebellum.

Surprisingly, the various sample preparation approaches had stark differences (Figure 4.10); identifications ranged from 38 – 6399 peptides which corresponded to 17 – 1521 protein groups. Molecular layer areas collected directly into buffer containing 50% TFE and then processed with the SP3 method yielded the most peptide/proteins (1,521 proteins). The next best performing method was catapulting directly into RIPA buffer with the SP3 method (1,316 proteins); followed by catapulting directly into the adhesive cap, centrifuging into RIPA buffer and then processing with the SP3 method (1,260 proteins). There was a 74% overlap in proteins (found in at least two replicates) between the two highest-performing methods (data not shown).

Overall, catapulting directly into a buffer outperformed catapulting into an adhesive cap and subsequently recovering the protein. Despite being the best buffer, when used with the "in-cap" method 50% TFE identifications were reduced compared to other "in-cap" lysis buffers, most likely due to proteins not being efficiently

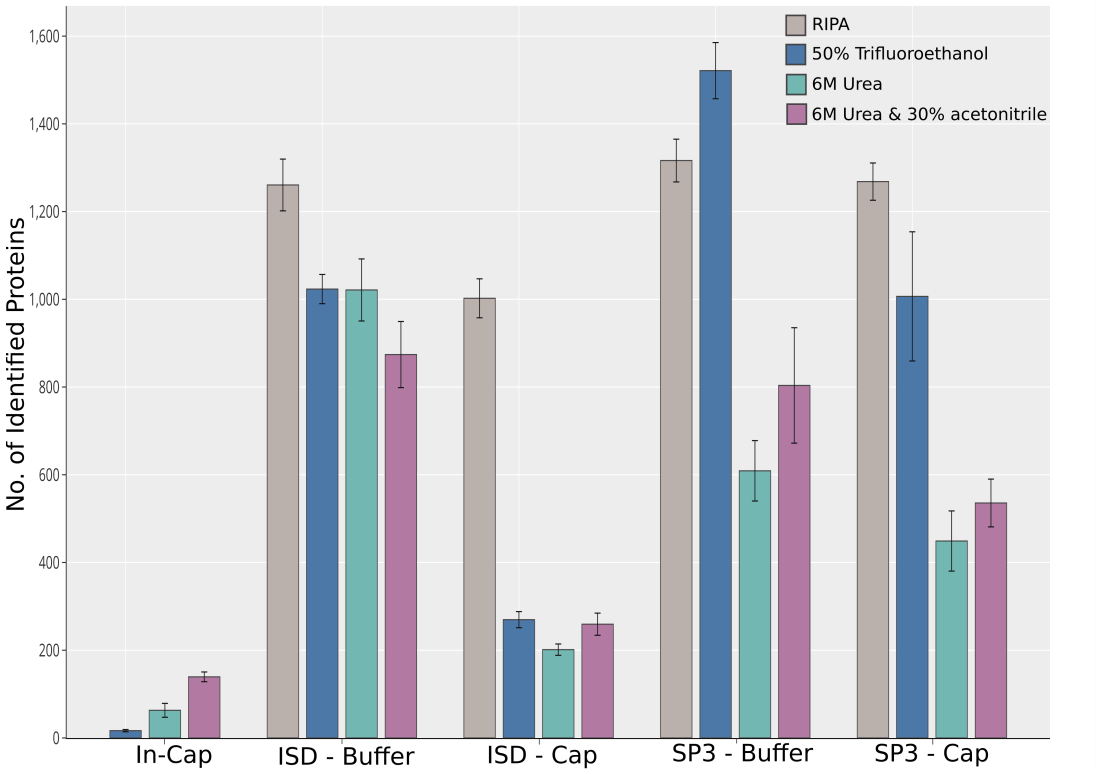


Figure 4.10: Optimising LCM LC-MS/MS Sample Preparation.
The results from optimising sample preparation steps. Overall, RIPA and TFE outperformed other lysis buffers tested. Bars: standard deviation. In-cap: *in situ* digestion performed directly on cells in the cap. ISD: In-solution digestion, indicating the standard sample preparation steps were performed. SP3: Samples were prepared using the magnetic bead approach from Hughes et al., 2014. Cap: Areas were catapulted directly into dry LCM caps, as done in the previous Purkinje cell experiments. Buffer: Areas were catapulted directly into LCM caps prefilled with lysis buffer.

recovered from the adhesive material in the cap[3]. SP3 protocols outperformed the "in-cap" approach in all cases. Combining the SP3 protocol with RIPA or 50% TFE outperformed using the same lysis buffers for the in-solution digestion. The stark differences between "ISD - Cap" and "ISD - Buffer" highlighted the improved benefit of capturing directly into lysis buffer, rather than capturing into an adhesive cap and subsequently centrifuging into the same lysis buffer. To test the quantitative performance of these methods, the number of quantifiable proteins were also recorded. Samples prepared with the Buffer-TFE-SP3 method yielded the

[3]It was not possible to combine the "in-cap" method with RIPA buffer as this pipeline did not have a detergent elimination step and would disrupt mass spectrometry output.

highest number of quantifiable proteins at 961 (mean), followed by Buffer-RIPA-SP3, Buffer-RIPA-ISD, and Cap-RIPA-SP3 with 790, 759 and 731 (mean). The samples processed solely with the "in-cap" method performed worst.

This series of experiments showed that TFE overall performed the best as a protein extraction/lysis buffer. However, it is important to note that TFE has a high vapour pressure that causes steady evaporation during sample collection and subsequent protein purification steps – which may lead to an overall decrease in protein IDs in LCM experiments requiring longer capturing time. Furthermore, additional care needs to be taken when storing and handling TFE as it poses a higher health risk than the second-best performing buffer (RIPA). Moreover, results using RIPA buffer were highly correlated with experiments that used TFE - see Figure C.2. As such, it was decided that catapulting directly into RIPA buffer and processing with the SP3 method would be the best approach moving forward[4]; this combination of parameters was among the best performing in identification and quantification and the reduced evaporation rate of RIPA buffer should pose fewer difficulties in sample preparation and blend well into future studies requiring high-throughput automated workflows. The performance of the RWS approach was also investigated on FFPE material - see Appendix C.4.

4.3.5 Effect of Histology Staining on LCM Proteomics

To explore how the variance in different staining methodologies could affect downstream LC-MS/MS, cerebellar Granule cells were stained with a variety of dyes (H&E, Hematoxylin and Nissl), microdissected (260,000µm^2; triplicate), and processed using the RWS method. Granule cells were selected as their large density and prominent feature in the cerebellum and also allowed them to be microdissected without any histological stains, providing the ideal baseline to assess how staining may alter the results of LCM proteomics. Moreover, the dense nature of the Granule cell layer meant that the whole area could be microdissected rather than manually annotating thousands of cells.

[4]Referred herein, as the RWS method (RIPA Buffer, Wet-cap, SP3).

Impressively, using the RWS approach, it was possible to identify just under 3,500 proteins in many of the samples and replicates. No significant difference was found in the number of proteins identified between the unstained samples and samples stained with Haematoxylin and Nissl (p value: 0.8496 & 0.3327, respectively). There was, however, a noticeable difference between the unstained samples and the H&E sample (p value: 0.022). All cell fractions were highly reproducible with the majority of proteins being present in each replicate: 91.7%, 84.9%, 94.2% and 92.1% overlap for the unstained, H&E, Hematoxylin and Nissl, respectively - See Figure C.11. Furthermore, when comparing proteins across all staining techniques including unstained tissue, there was also a large overlap in identified proteins, with 3,340 (94.7%) of proteins being present in 4 fractions - Figure 4.11.

As expected, within-sample replicates the level of correlation was very high (Pearson correlation coefficient (PCC): 0.98)) and while there was still an overall high level of correlation across samples (lowest value: 0.94 PCC) there was a clear reduction in H&E samples compared to the others (PCC: 0.94 vs PCC 0.97) correlation between the unstained samples and Nissl and Hematoxylin samples. Nissl samples had the strongest correlation between replicates - see Figure C.11. PCA showed moderate separation between the unstained, Haematoxylin and Nissl staining; whereas frozen Granule cells stained by H&E was clearly delimited - see Figure C.12.

A Welch t-test (FDR threshold: 0.05, 250 randomisations) was then performed to investigate if the different staining methods affected the detection of any particular protein/protein groups. Of the 3,535 identified in this data-set 2,503 were suitable for comparative quantification, the expression data of the x3 staining methods were compared to the unstained microdissected samples. When comparing the H&E sample to the unstained Granule cells 292 proteins were recognised as being significantly lower and 89 proteins were significantly higher in the H&E sample - see Figure 4.12. The Nissl sample had 25 proteins found to be significantly lower compared to the unstained sample and 7 proteins were higher. No significant difference in protein abundance was found between the unstained and haematoxylin samples. ORA showed that among the proteins significantly higher in the H&E

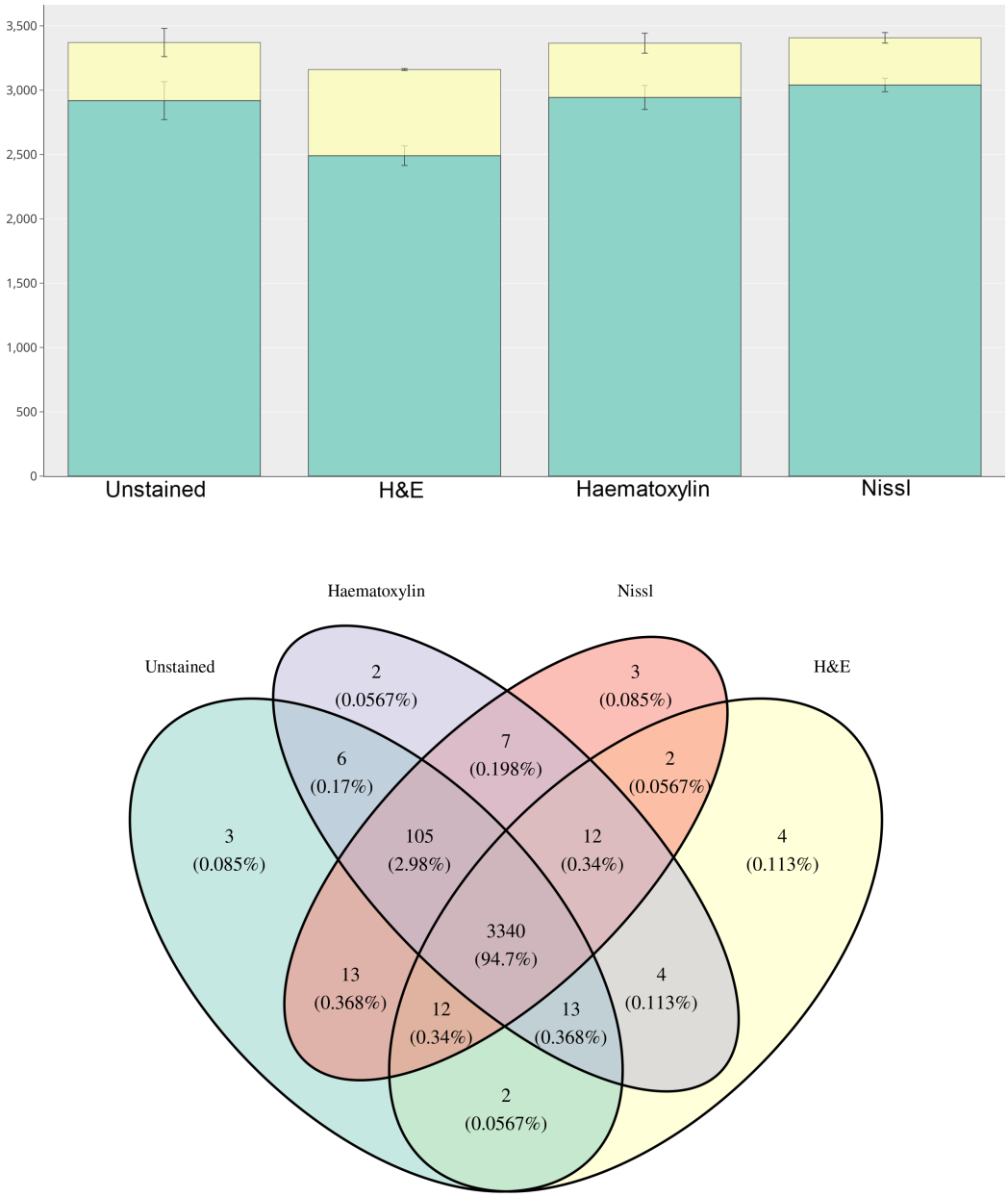


Figure 4.11: Effect of Different Granule Cell Staining on Number of Identified Proteins. Top: Overall number of proteins identified was uniform across all samples. Area microdissected: $\approx$ 260,000µm^2. Turquoise bars: by MS/MS, Cream Bars: by MBR. Vertical bars denote standard deviation. y-Axis: number of proteins identified. Bottom: Venn diagram of proteins identified in each sample.

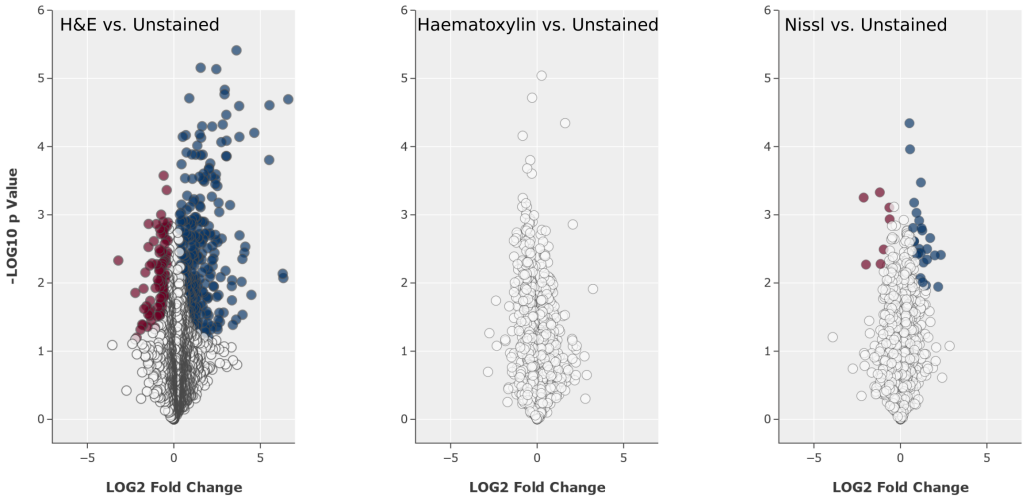


Figure 4.12: Effect of Different Granule Cell Staining on Protein Abundance. Volcano plot showing the effect of H&E vs. unstained (left panel), Haematoxylin vs. unstained (middle panel), Nissl vs. unstained (right panel). Each point represents one protein. Coloured points passed predetermined alpha thresholds (0.05). Red: proteins whose expression was lower in the stain tested. Blue: proteins whose expression was higher in the stains tested.

sample there was a large enrichment in proteins (CDSN, POF1B, DSP, JUP) which localise to the cornified envelope (enrichment: 10.21, FDR q-value: 5.59e-01). Additionally, TGM3 and S100A7 along with the above were also found to contribute to an enrichment of terms indicating skin contamination – keratinocyte differentiation (enrichment: 10.8, FDR q-value: 4.92e-01) and epidermal cell differentiation (enrichment: 9.36, FDR q-value: 5.42e-0). When exploring the proteins found to be significantly reduced in the H&E sample there was an apparent reduction in proteins found in a variety of subcellular components, mainly small nuclear ribonucleoproteins, histone proteins, and ribosomal proteins along with proteins found in the methylosome, nucleosome, and small ribosomal subunit (data not shown). Despite identifying 32 DEPs, no protein pathways were found to be enriched in the Nissl sample, although interestingly – there was a reduction in a collection of proteins known to be highly relevant in neurodegeneration: APOE, SNCA, PARK7 and SUMO2. No DEPs were found between the Haematoxylin and unstained samples.

Hierarchical clustering was used to better compare the global expression of the x3 staining methods and determine similarities in an unbiased fashion. Results showed that replicates from individual samples clustered around each other, except the unstained Granule cells in which the Nissl replicates were found to be more similar to unstained samples than one of the unstained replicates - Figure C.13. As expected from prior analysis, the H&E sample was the most dissimilar to the other samples, with the Nissl stain being the most similar to the unstained Granule cells, even though Haematoxylin did not have any significantly differentiated proteins.

4.3.6 Effect of immunohistochemistry on LCM Proteomics

Histological stains are useful for providing a global view of the cytoarchitecture of tissue, however, identifying cells solely by morphology may mask subpopulations with similar morphology. Additionally, histological stains may provide too much information, highlighting most/all cells in a targeted region. As such, there is increasing pressure to molecularly target cells using nucleic acid probes or complementary antibodies to facilitate cell-type-specific investigations. However, in the context of LCM proteomics, such probes/antibodies require considerable sample preparation and a variety of chemicals, any of which may affect the underlying proteome of cell population/tissues. To test how immunohistochemistry affects LCM Proteomics, an experiment was performed on 100 microdissected Purkinje cells. It is almost impossible to visualise Purkinje cells on unstained sections with a standard bright-field microscope. The previous subsection highlighted that neurons microdissected with Haematoxylin or Nissl staining were similar to microdissected unstained cells. As such, Nissl was selected as a substitute for unstained tissue as Nissl better highlights neuronal morphology and overall, was more closely related to unstained samples.

Purkinje cells were either stained with Nissl (quadruplicate) or stained with an anti-CALB1 antibody (triplicate)[5]. To minimise potential large proteomic changes, antigen retrieval and exogenous peroxidase quenching were omitted. Despite the

[5]One sample was lost during sample preparation.

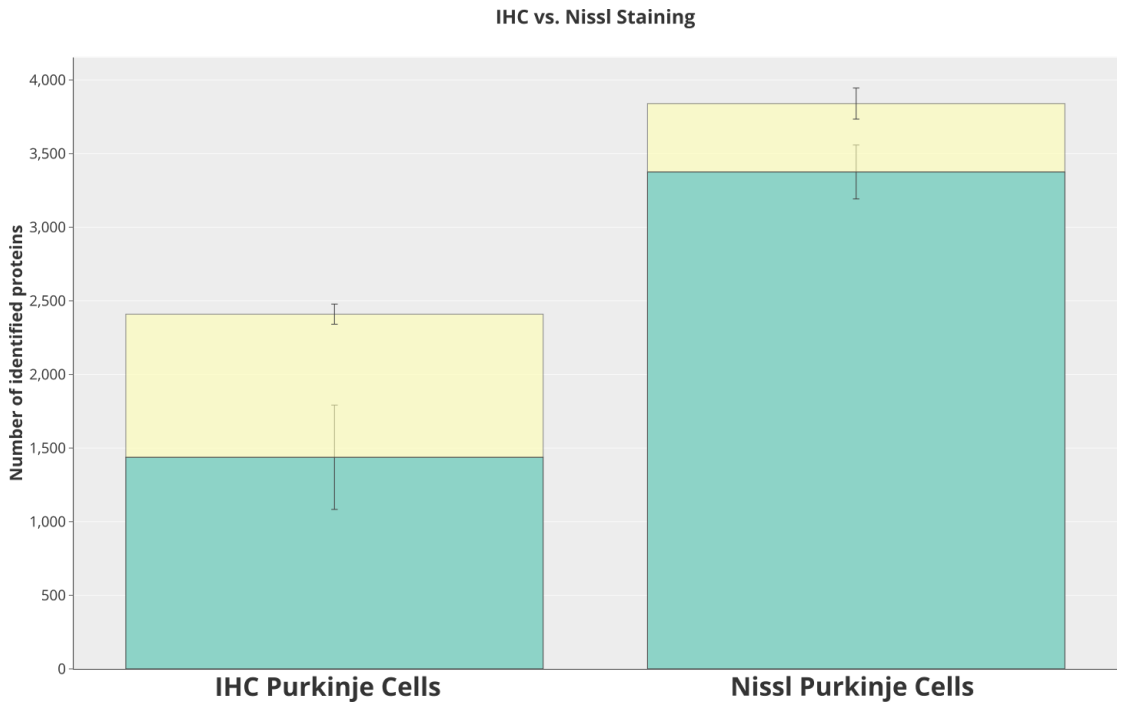


Figure 4.13: Effect of IHC on No. of Identified Proteins.
It was possible to identify a larger number of proteins from Nissl-stained Purkinje cells. However, CALB1 immuno-stained Purkinje cells saw a marked reduction in the number of identifiable proteins. Area microdissected: $\approx$ 60,000µm^2. Turquoise bars: by MS/MS, Cream Bars: by MBR. Vertical bars denote standard deviation.

primary/secondary antibody-HRP chromogen complexes, enzyme digestion efficacy was not affected within IHC samples, with 79.21% of peptides having zero missed cleavages, compared to the 77.89% of peptides with no missed cleavages in Nissl stained Purkinje cells. An impressive amount of proteins was resolved from the samples containing just 100 Purkinje cells (60,000µm^2) slices stained with Nissl ($\bar{x}$ Protein ID: 3,739) - see Figure 4.13; likely due to the optimised RWS workflow and discovery that Nissl staining has a better performance than H&E staining. For reasons not entirely clear, IHC samples suffered from a significant reduction in protein IDs (Welch t-test, p value: 0.0166, $\bar{x}$ Protein IDs: 2,409), although this number of identified proteins is still significantly better than the first attempts of microdissecting 100 Purkinje cells (73 proteins in the first titration experiment, and 1,497 in the second titration experiment).

Among replicates, correlation was high; the IHC samples had a PCC of 0.932

and 0.916 within Nissl samples, however, the correlation was only moderate between the two samples ($\bar{x}$ PCC: 0.61) - Figure C.14. There was also a strong overlap in protein IDs between all Nissl stained replicates (82.8%) but a lower overlap within the IHC samples (57.6%), which can be attributed to one replicate containing 638 not identified in the other replicates. Overall, 69.9% of proteins (3,033) were shared between the two staining methods, with 1,258 (29%) found solely in the Nissl microdissected Purkinje cells and only 46 (1.06%) found solely in the Anti-CALB1 stained Purkinje cells - Figure C.15.

There was an over-representation of several GOCC terms within the 46 proteins solely found in the IHC samples. A 4.07 enrichment in proteins within the nucleolus (FDR: 1.61e-03), a 4.72 enrichment in proteins relating to chromatin FDR: 4.77e-02) and finally a 21.53 enrichment in proteins found within the 90S preribosome (FDR: 4.77e-02). There was also a 7.59 (FDR: 2.55e-02) and 3.56 (FDR: 1.85e-05) enrichment in GOBP terms relating to rRNA processing and regulation of transcription, DNA-templated, respectively. Finally, there was a 14 fold enrichment in GOMF terms ubiquitin-protein ligase activity (FDR: 1.05e-02), 5.27 fold enrichment zinc ion binding (FDR: 1.12e-02), 5.12 fold enrichment in sequence-specific DNA binding (FDR: 2.78e-02) and a 5.08 fold enrichment in transcription regulator activity (FDR: 6.97e-03). When comparing enriched GO terms in the 1,246 proteins only found within the Nissl stained Purkinje cells, there was only a slight over-representation of hydrolases (Enrichment: 1.83, FDR: 7.41e-03).

Of the 3,033 proteins shared between the two staining methods 61.06% (1,852) were eligible for quantitative analysis. A large difference was found in protein expression between the two staining methods (Welch t-test, FDR threshold: 0.05, 250 randomisations), of the 1,852 proteins included in this comparison 1,056 were regarded as significant DEPs. Of these 1,056, 322 DEPs had higher expression within the anti-CALB1 stained samples and 724 DEPs had a lowered expression in the anti-CALB1 stained samples - Figure 4.14. Surprisingly, a large fold change

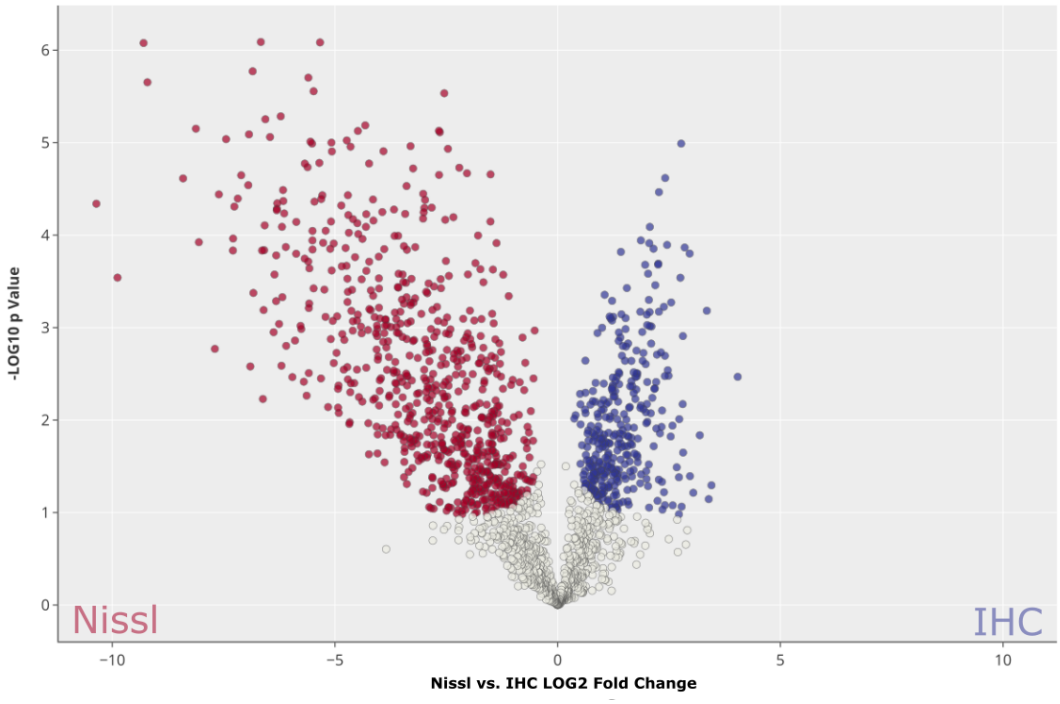


Figure 4.14: Volcano Plot - IHC vs. Nissl Stained Purkinje Cells. Large differences were observed in the underlying proteome of Purkinje cells stained with either Nissl or anti-CALB1. A marked reduction in IHC stained Purkinje cells was seen when comparing the MaxLFQ values. Each point represents one protein. Coloured points passed predetermined alpha thresholds (0.05). Red: proteins whose expression was lower in IHC samples. Blue: proteins whose expression was higher in the IHC samples.

was seen in this data-set - more so than previous experiments - with 169 proteins having a Log2 Fold Change (LogFC >4).

Among the 724 DEPs, there was an over-representation of proteins found in the synapse (enrichment: 1.32, FDR: 1.62e02), plasma membrane (enrichment: 1.33, FDR: 5.14e-04), and secretory vesicles (enrichment: 1.56, FDR: 1.38e-02). Over-representation tests were also performed on the DEPs with a LogFC of >4; this showed a 6.41 and 7.54 enrichment in protein relation to gluconeogenesis (FDR: 2.47e02) and glycolysis (FDR: 3.93e-03). A 3.84 enrichment for dehydrogenases (FDR: 9.33e03) was also seen. Several proteins involved in the regulation of exocytosis were also found in the DEPs with a LogFC of >4 (Enrichment: 2.28, FDR: 2.38e-03), finally, Alpha-Synuclein, Gamma-Synuclein, Synaptotagmin-1,

Synaptotagmin-2, 4-aminobutyrate aminotransferase and Ras-related protein Rab-3A were all found to have a LogFC of >4 and were responsible for a 9.34 enrichment (FDR: 4.96e-02) in the regulation of dopamine secretion. In an attempt to find a relationship with the proteins that were increased in anti-CALB1 stained samples, ORA was performed and several terms relating to mRNA processing was enriched: mRNA Splicing - Major Pathway (FDR: 5.78e-06), Spliceosome (FDR: 8.05e-05), and RNA-binding domain superfamily (FDR: 0.00033).

4.4 Discussion

Understanding the biochemical composition of individual cell populations is of paramount importance in understanding the organization of the human brain and the role different cells have in disease. *'omic'* technologies have rapidly evolved in recent years and have dramatically improved our understanding of these topics. However, these studies either rely on well-characterized biomarkers, use large amounts of starting material or implement complicated clustering algorithms to retrospectively classify and define cell-types. Few studies attempt to provide spatial cell resolution data. While useful in biomedical research, several pitfalls still exist with these approaches, especially when working with cells that are poorly understood and lack specific biomarkers. This chapter describes an optimised LCM LC-MS/MS workflow that has been tested against various upstream histological considerations and generates unprecedented proteomic depth from post-mortem input material of little as 600,000µm^3.

The first attempt to isolate and identify proteins from post-mortem neurons yielded promising results; while only a small amount of protein was identifiable in the lower cell fractions (10 to 200 Purkinje cells) this dramatically improved when microdissecting 400 and 800 Purkinje cells. Looking at this data set, it was encouraging to see canonical Purkinje cell biomarkers being the most abundant proteins, demonstrating the enrichment for our cell-type of interest and highlighting that this technique may be a worthwhile approach for identifying potential biomarkers

from future exploratory experiments. Furthermore, it was possible to validate the proteins identified in our data set by selecting Gene Ontology annotations involved in Purkinje cell function and cross-referencing them to the Human Protein Atlas. While identifying 1,517 and 2,051 proteins is competitive with other LCM proteomic methods, the non-linear increase in Protein IDs across samples and the sharp increase in protein IDs from the 200 and 400 cell fractions indicated that the reduction in performance was most likely due to the poor efficacy of the workflow. As discussed previously, when working with minute amounts of material, a likely contributor to reduced data output is the loss of material by the various pipetting steps in sample preparation. Moreover, analysis of the remnant material left in the cap after sample preparation demonstrated that a large amount of protein was unrecovered with standard LCM sample preparations. Our results show that although the adhesive caps typically used in LCM can successfully capture material and keep it from falling back down onto the tissue section; the microdissected material can often become lodged in the glue, making it difficult to retrieve post-LCM even with prolonged centrifugation. When the remnant material was digested directly, without any pipette transferring a dramatic increase in protein IDs was revealed, with 1,497 proteins being identified from the 100-cell sample, compared to only 73 proteins identified in the first attempt. A similar increase was seen in the 200-cell sample with 2,224 protein IDs being found compared to 107 in the first Purkinje cell titration. A more modest increase was seen in the 400 and 800-cell samples, which is most likely because the larger starting material is less affected by sample loss via pipetting/plastic loss. Notably, there was a clear plateau in protein IDs between the 200, 400, and 800-cell samples. This could be due to the fact that although the adhesive caps are proficient in capturing material the overall adhesive area of the cap could slowly become saturated as more cells are captured into it. Once saturated, microdissected cells may fail to adhere to the cap and fall back onto the slide – this phenomenon can be seen on the underlying histology slide when cutting a large amount of material (data not shown). Although untested, we believe that if

cells were pooled together from several 200-cell fractions or equivalent sized areas then the plateau effect could be reduced and more protein IDs would be identified.

Knowing that pipetting and transferring microdissected cells from the cap to subsequent tubes for sample preparation were the biggest contributors to protein ID loss, the next objective was to optimise the sample preparation workflow by testing various permutations of the sample preparation process. To align these permutations with standard practices in mass spectrometry, it was decided to keep the routinely used denaturing and alkylating steps. Different lysis buffers were used to test which was most effective in lysing cell fragments and releasing proteins. The use of trifluoroethanol (50%) as a lysis buffer produced the most protein IDs, however, considerable health concerns are working with it, it is expected to be a carcinogen and can form explosive oxides [286]. RIPA buffer also performed consistently well across permutations and does not pose significant health risks, and both the performance and protein identified was similar to TFE-based sample preparations. As such, it was decided to use RIPA buffer as the final lysis buffer in our final optimised LCM proteomic approach. Capturing cells directly into a lysis buffer greatly improved performance compared to capturing cells into adhesive caps. Rationally, capturing directly into lysis buffer may minimise the potential issue of cells being adhered too strongly to be recovered and may also reduce the potential cap saturation effect previously mentioned. While catapulting into a lysis buffer yielded better results, a considerable downside of this method is that it is not possible to inspect the cap for successful cellular capture after LCM, which makes may be critical for some experiments. To minimise sample loss by pipetting, a comparison between the standard in-solution digest and a novel approach that utilises magnetic beads to bind proteins/peptides in place was performed. The SP3 method performed best, a key feature of the SP3 approach is that all mass spectrometry sample preparation steps can be performed within the same tube originally used for LCM, dramatically reducing sample loss by unnecessary pipetting to transfer material to subsequent tubes.

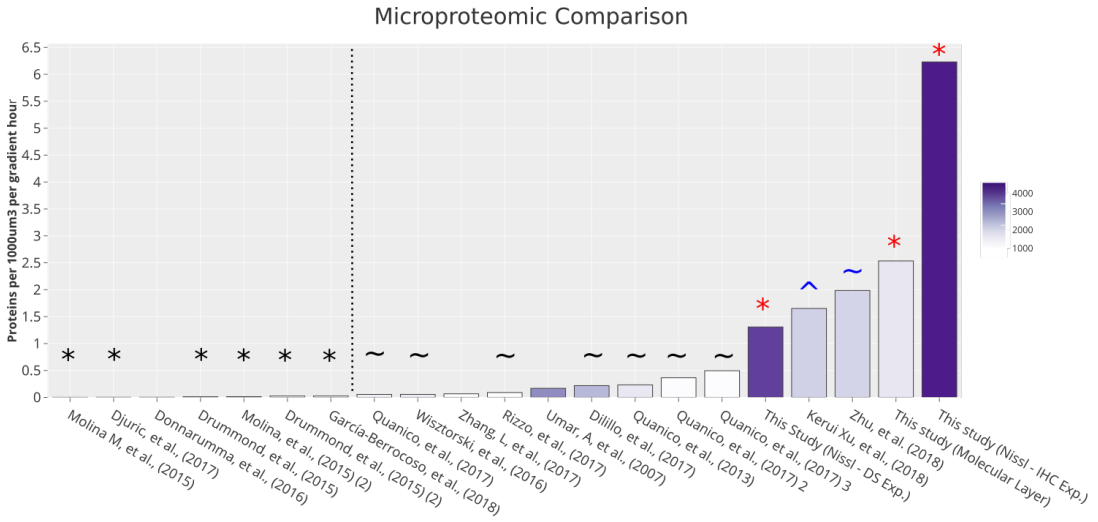


Figure 4.15: Comparison of LCM Proteomic Approaches Reported in the Literature. A comparison of LCM proteomic approaches in the reported in the literature. Our optimised RWS workflow and NanoPOTs perform best at maximising proteomic depth from minute amounts of material. $*$ = studies using human brain material, $\sim$ = studies using rodent brain material, $\wedge$ = studies using HeLa cells. Red: in-house experiments. Blue: NanoPOTs based experiments. *Comparison done 2020.*

Although other groups have combined LCM and mass spectrometry to gain insights into the proteome of enriched cells, their output often lacks high proteomic depth relative to the starting material, or their approaches use long chromatography times making experiments lengthy and cumbersome. To see how our approach compares to previously published approaches a literature review of LCM studies using proteomics was performed. To normalise outputs from studies with different experimental procedures the following parameters were recorded: number of proteins identified, amount of material microdissect, and chromatography time. Collectively, this *Number of proteins identified per 1000µm^3 per one gradient hour* measurement provides an opportunity to compare the RWS method with others in the literature. The RWS approach identified 2.535 proteins per 1000µm^3 per one gradient hour in the sample preparation optimisation experiment. The RWS approach also identified 1.30 proteins per 1000µm^3 per one gradient hour when microdissecting the Nissl stained Granule cells (260,000µm^2) and an impressive 6.23 proteins per 1000µm^3 per one gradient hour when only 100 Nissl Stained Purkinje cells were microdissected in the Nissl/IHC comparison.

Recently, the NanoPOTS technique has outperformed other methods in proteomics of small samples [86, 132, 133, 135, 287, 288]. This approach drastically reduced sample loss by reducing the working volume to <200nL by using automated liquid handling systems to transfer material. Furthermore, they capture directly into Dimethyl sulfoxide (DMSO) which has a low evaporation rate, therefore, avoiding the issue of lysis buffer evaporating during LCM. With these considerations in place, the NanoPOTS technique manages to achieve an impressive proteomic depth of >3000 from as little as 10 HeLa cells [86], and more recently has been used to perform proteomics on a single human cell isolated by FACS [289]. It is important to note, however, that working with HeLa or mammalian cell lines allows for strict regulation of protocols to be implemented to maximise protein quality and efficient extraction. This remains a significant challenge for human tissue, especially tissues of post-mortem origin. Surgical tissue provides a realistic intermediary between post-mortem tissue and animal tissue. When the NanoPOTS approach was applied to human surgical biopsies containing pancreatic islets, the results were still impressive albeit, slightly lowered ($\approx$ 2,400 proteins) [290]. It has yet to be seen how NanoPOTs compares with human post-mortem nervous tissue. As Figure 4.15 suggests, the RWS approach manages to compete with the NanoPOTS method despite the difficulties with working with human post-mortem tissue, additionally, our approach can be implemented in any laboratory with access to a laser microdissection machine, mass spectrometer, and standard laboratory reagents and does not require nanowell chips to be 3D printed and installed, or the use of liquid handling robots. However, the performance of the RWS approach is particularly highlighted when compared with the next best performing LCM proteomic investigation on microdissected human brain tissue [148] - see dotted line in Figure 4.15. While our approach does perform well, it is important to note that many of the studies that produce the best results – including ours - benefit from using cutting-edge mass spectrometers that are capable of identifying increasing numbers of peptides from decreasing amounts of starting material, this makes it difficult to directly compare to older studies whose instruments are not as sensitive.

Nevertheless, since the RWS approach was optimised to increase efficiency and minimise material loss, it is expected that as mass spectrometers become more sensitive in the future that this approach will scale with technological innovation.

There are a variety of methods available to visualise cells of interest for LCM. In the context of neuronal samples, this can be narrowed down to generic stains to visualise many cell-types (H&E or Hematoxylin) or more specialised dyes highlighting features unique to neuronal morphology i.e. Nissl staining or Luxol Fast Blue. Each staining technique contains differing chemicals and protocols to optimise neuronal identification, and while common practice for many histology laboratories, it is not clear how this variance in protocols can potentially affect downstream applications such as mass spectrometry. Unless studying pigment-containing neurons, e.g. dopaminergic neurons of the substantia nigra, it is near impossible to identify cell-types without staining tissue sections. To study the changes of different staining protocols on proteomic analysis it would be useful to use unstained material as a baseline control. To test for these differences, cerebellar Granule layer areas were microdissected from tissue stained with either H&E, Haematoxylin, or Nissl prior to LCM. The unique feature of the Granule cell layer, allowed for unstained Granule cells to also be microdissected and processed in parallel. An impressive amount of proteins was identified using the optimised RWS workflow across all staining techniques. Only the H&E staining produced a slight reduction in proteins identified compared to unstained Granule cells. This reduction in IDs due to H&E staining has been previously reported in the literature [291, 292]. Overall, there was a 94.7% overlap in protein IDs across all samples, highlighting that there is little concern in the reproducibility or identity of protein IDs across staining methods tested here.

One of the most popular techniques to identify cells of interest is to use antibody-based detection systems to visualise proteins that are unique to the cells of interest or that better highlight their unique morphology. Immunohistochemistry, while ideal for staining tissue involves several experimental steps and a variety of different chemicals and although not uncommon in its use in conjunction with LC-MS/MS,

there have not been many studies investigating how these additional steps may affect the underlying proteome of the cells of interest. To test how IHC performs with the RIPA-Buffer-SP3 method, an experiment was performed using Purkinje cells. While biomarkers for Granule cells of the cerebellum do exist, microdissecting areas rather than individual neurons are not the ideal benchmark for judging how the RIPA-Buffer-SP3 method will perform on future individually pooled cell samples. Choosing Purkinje cells for this experiment increases the labour involved, and as such, it was decided to only dissect 100 Purkinje cells for each sample. It is virtually impossible to use standard light microscopy to identify Purkinje cells from unstained sections, so an alternative approach was needed. As Nissl stain is known to highlight neuronal morphology and the previous experiment showed that Nissl stained neurons had little difference to unstained counterparts; it was decided to use this to stain Purkinje cells. As a well-studied marker for Purkinje cells, Calbindin-1 (CALB1) was selected as the protein target. Combining Nissl staining with isolated neurons rather than previous experiments using H&E yielded impressive results. On average, 3,375 proteins were identified from a sample containing just 100 Purkinje cell segments ($\approx$ 60,000µm^2). A potential concern with the IHC approach was that the established antibody-HRP complexes would be difficult to denature and linearize, resulting in poor enzyme digestion and lowered protein IDs, potentially restricting access to other neighbouring proteins. While enzyme digestion remained unchanged there was a significant decrease in the number of proteins identified from immuno-labelled neurons ($\approx$ 1,400 vs. $\approx$ 3,300 proteins). There was only a modest correlation between the anti-CALB1 samples and Nissl stained Purkinje cells (PCC: 0.566), similarly only 70.2% of proteins were shared between the two stained samples as opposed to the >94% in the previous comparison between histological stains. Considering that IHC remains a popular method in visualising cell-types it is concerning that a large quantitative difference can be seen between anti-CALB1 stained Purkinje cells and Nissl stained material. Further studies will need to be done to validate if this reduction in metabolic proteins is systematic across immune-stained samples; if true, careful consideration is needed

with using immuno-labelled material for LCM proteomics. The substantial reduction in identifiable proteins seen in the IHC samples, while adequate for answering a variety of research questions, may pose experimental design issues for future work where immunostaining must be used. Other groups who have attempted to compare LCM enriched immuno-stained tissue to unstained/unstained surrogate material found a reduction in RNA quality but no reduction in DNA or protein quality and did not report a decrease in protein IDs [146, 293–295]; moreover, few have quantitatively compared the expression of proteins between immuno-labelled and regular histology staining on the same cell-type. It is important to note, that the performance of immuno-labelled cells for LCM proteomics still resolved more protein IDs than many of the above studies; as such it may be appropriate to use IHC with LCM proteomics if alternative stains are limited.

IHC-directed LCM LC-MS/MS has the potential to be a very powerful technique. However, several important questions remain unanswered with immuno-stained LCM LC-MS/MS; (1) are proteins relating to metabolism consistently reduced in immuno-stained tissues? (2) Does the type of antibody used affect downstream IDs and protein abundances? (3) In this case, the protein of interest (Calbindin1) was one of the most diminished proteins in the proteomic data for CALB1 stained samples, is this coincident, or does the immunostaining protocol increase the difficulty of resolving the targeted protein? Despite these questions, it was still possible to generate a relatively large proteomic depth from immuno-stained microdissected post-mortem neurons. While future work is needed to gain a better understanding of the potential biases, it is also important to realise that often IHC remains to only way to study specific populations of cells and in a scenario where IHC is the only suitable approach to visualise cells it may be justifiable to proceed with caution, as others have done in the literature.

The work presented in this chapter shows the successful development of a micro-proteomic workflow that allows for robust comparative quantitative proteomic analysis from as little as 100 pooled LCM cell sections. This approach can be adapted to target almost any cell-type in the central nervous system that can be

identified histologically and, although untested, may work well with cells from other organ systems. The minimal input material and high quantitative output from this workflow should allow for previously understudied cell-types to be enriched and large amounts of data generated on their underlying proteome. Part of the data presented in this chapter has resulted in an original publication that describes the narrative of optimising mass spectrometry workflow from frozen human LCM enriched material and settling on the RIPA-Buffer-SP3 approach [217]. At the time of writing, this approach outperformed several others used by research groups and showed a proteomic depth:input material ratio previously unseen in post-mortem human brain studies. Further tests have since been performed to establish this workflow further and to test possible beneficial alterations - see E.1, discussed in a future chapter. Now established, the next chapter aims to utilise this RWS workflow to gain biological insights into the proteome of the enigmatic Betz cell, which previously has mostly been characterised using histological or stereological techniques. More generally, the next chapter will highlight the sensitivity of this LCM proteomic workflow to study neurons that differ substantially, as well as test if this approach can resolve differences in neurons that are physiologically very similar.

4.4.1 Main Conclusions

- A LCM Proteomic workflow (referred to RWS (RIPA-Wet Cap-SP3), herein) has been optimised and established for individually pooled neurons.
- This chapter showed it is possible to identify 3,000 - 4,000 proteins from as little as 100 Purkinje cell slices (60,000µm^2) using the above method.
- The RWS workflow outperforms other LCM Proteomic approaches reported in the literature, especially those performed on post-mortem human brain.
- The performance of the RWS workflow has been tested on different neurons, different staining techniques, and on FFPE and frozen material - see Appendix C.4.

5

Explorative LCM Proteomics of the Human Betz cell

Contents

5.1 Aims of this Chapter 119
5.2 Introduction 120
5.2.1 Breakdown of this Chapter 121
5.3 Results 122
5.3.1 Building a Suitable Betz Cell Research Cohort 122
5.3.2 Betz Cells and Purkinje Cells - Initial Data 125
5.3.3 Relative Proteomics - Betz Cells vs. Purkinje Cells . . . 125
5.3.4 Digital Pathology 132
5.3.5 LCM Proteomic-Directed Immunohistochemistry 137
5.3.6 Comparative LCM Proteomics of the Betz Cell, L5 Motor Neuron, Purkinje Cell, and Cerebellar DNN 151
5.3.7 The Search for Preliminary Betz Cell Biomarkers 165
5.4 Discussion 180
5.4.1 Main Conclusions 191

5.1 Aims of this Chapter

- Generate a tissue cohort that will generate high-quality frozen/FFPE material for future comparative Betz cell investigations.
- Investigate if the optimised RWS proteomic workflow can discern cell-type-specific differences within distantly and closely related neurons.

- Gain novel insights into the proteome of the Betz cell and investigate how this compares to other neurons.
- Validate LCM proteomic findings with immunohistochemistry and try to identify novel potential Betz cell biomarkers.

5.2 Introduction

As discussed in Section 2.3.6, little is known about the biocomposition of Betz cells. Since their discovery in 1874, there has been limited work performed to investigate the molecular composition of this motor neuron, with most work focusing on its unique morphology and how this morphology is altered in ALS. While several statements on its function and role have been proposed, these ideas have been based for the most part on stereological studies and have not been validated with *'omics'* technology. The degeneration of Betz cells in ALS must reflect aberrant cellular mechanisms in these neurons; understanding these pathways may advance knowledge into the mechanism of ALS and lead a path to potential therapeutics. However, before such investigations can be performed to determine how ALS biochemically affects Betz cells, the molecular profile/cellular pathways of the healthy Betz cell needs to be determined. Revealing the molecular nature of the Betz cell and discovering Betz cell biomarkers would accelerate research by removing the need to comprehend subtle morphological changes between Betz cells and neighbouring L5 neurons, aiding histological examination and making Betz cells compatible with immuno-based isolation techniques or allow Betz cell-specific information to be teased out from unbiased snRNA-seq/bulk analysis.

The previous chapter has shown the ability to investigate the proteome of laser microdissected neurons at unprecedented depths; in this chapter, the same approach will be used to study the biomolecular composition of human Betz cells and gain insights and validate its biological activity. To validate and ensure confidence in the data generated, Purkinje cells will also be studied in the same manner. There are several reasons for this logic: both neurons are topographically,

morphologically and functionally distinct with different operating neurotransmitters; (gamma-aminobutyric acid in the Purkinje cell and presumed glutamate in the Betz cell) and originate from different germinal zones in the human developing brain. However, they still share several similarities; (1) they are both used to define dissimilar cytoarchitectonic regions of the brain, (2) both are among the largest cells in the human CNS, (3) both are involved in the motor system of the CNS, and (4) both are the principal neuron in their respective neuroanatomical regions. Furthermore, unlike the Betz cell, there is a plethora of data in the literature regarding Purkinje cells, allowing data generated to be validated by *a priori* knowledge. Therefore, these neurons provide a perfect testbed to demonstrate if the RWS approach could accurately distinguish between the two neuronal cell-types and provide insights into their respective proteomes.

5.2.1 Breakdown of this Chapter

1. The rationale behind prospectively sampling fresh post-mortem brains for this thesis will be outlined.
2. The RWS workflow will be used to compare the proteomes of the human Betz and Purkinje cells.
3. An image analysis approach will be described to allow proteomic data to be quantitatively validated, using an orthogonal approach.
4. Immunohistochemistry will be used to validate proteomic findings from the initial Betz vs. Purkinje cells experiment.
5. Another LCM proteomic experiment will be performed comparing Betz cells to Purkinje cells, Cerebellar Dentate Nucleus neurons, and L5 motor neurons. The cellular pathways enriched in Betz cells will be discussed.
6. The accumulative data generated on the proteome of the Betz cell will be used to screen for potential Betz cell biomarkers.

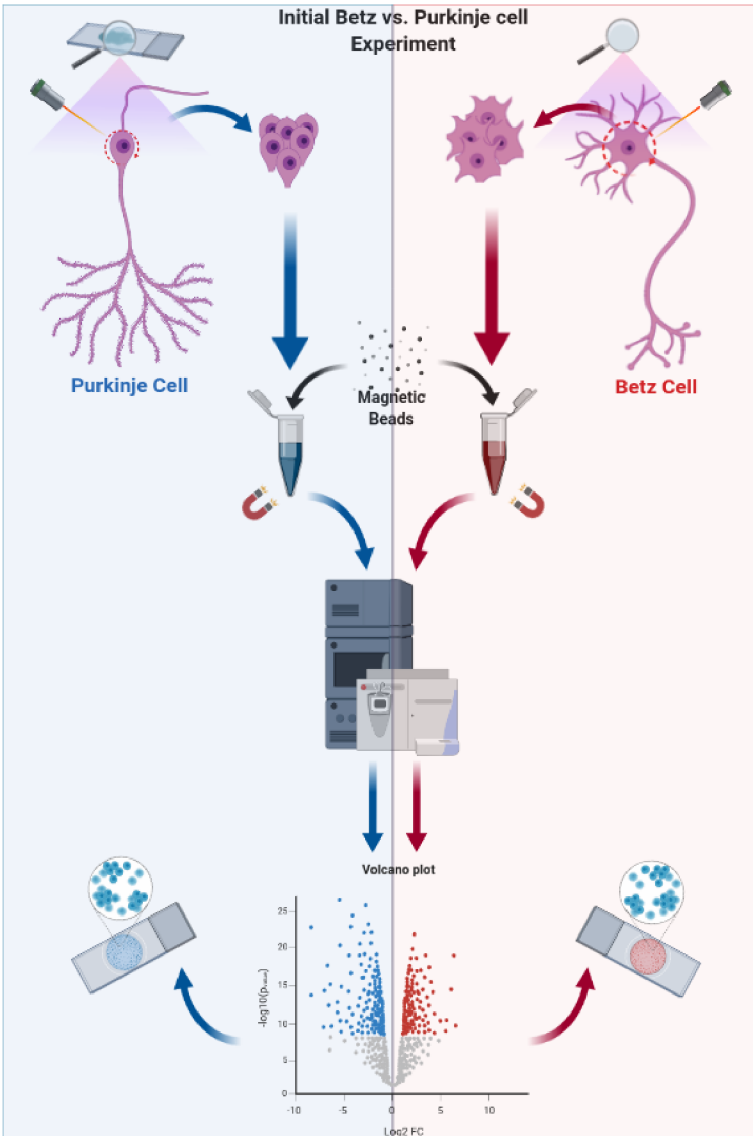


Figure 5.1: Proposed LCM Approach for this Chapter.
This chapter aims to microdissect the Betz and Purkinje cells using the RWS workflow optimised in the previous chapter. Cell-type resolved data will be analysed to gain novel insights into the molecular biology of the Betz cell. Promising findings will be validated and tested by using IHC. Finally, other cell-types will be studied in the same way to better investigate the proteome of the Betz cell.

5.3 Results

5.3.1 Building a Suitable Betz Cell Research Cohort

As repeated throughout this thesis, a significant challenge in studying Betz cells is the difficulties associated with gaining access to suitable human tissue. Any

dedicated analysis of Betz cells would require abundant amounts of the primary motor cortex, however, this is not always sampled by biobanks - and if it is, then it is typically restricted to a limited number of biopsies. Moreover, biobanks preferentially use FFPE tissue as their method of choice when archiving tissue. Although some biobanks freeze coronal sections to preserve neuroanatomy, it can be difficult to identify the primary motor cortex as surrounding neuroanatomical structures can become deformed by the freezing process. Furthermore, extreme care must be taken when sampling from archival frozen tissue as excising the tissue can result in either mechanical or heat damage.

Due to the above limitations, inspired by Rivara et al., (2003) it was decided the best action would be to dissect fresh post-mortem brain donations prospectively, dissecting the entirety of the motor cortex, and prepare the tissue to be compatible for sensitive downstream applications. For FFPE tissues, this meant applying a short-fixation protocol to minimise the deleterious effects of formalin on RNA and protein. For frozen tissue, fresh tissue was immediately frozen using liquid nitrogen vapour, which in-house experiments (data not shown) have been shown to reduce ice crystal formation and preserve cytoarchitecture. Importantly, the primary motor cortex from these fresh post mortem cases was sampled in its entirety, providing access to the total population of Betz cells from both hemispheres - See 3.1.1 for methodology and Figure 5.2. The original experimental design to study the Betz cell with LCM proteomics was planned on both hemispheres to accumulate proteomic information on both preservation counterparts. However, since the RWS approach has not been optimised for FFPE material and standard experimentation yielded unsatisfactory results - see Section C.4, it was decided to solely perform LCM Proteomics on frozen sections. FFPE primary motor cortex blocks were still generated as IHC validation could be used on the cytoarchitectural preserved FFPE blocks. In total, 31 brains (x9 Control, 18x ALS, and x4 FTD[1]) were prepared in this way. A list of all cases part of this cohort can be seen in Appendix F.

[1]FTD post-mortem brains were taken to investigate the ALS/FTD spectrum of disease and how Betz cells may contribute to this. However, the results of Betz cell FTD analysis will not be discussed in this thesis.

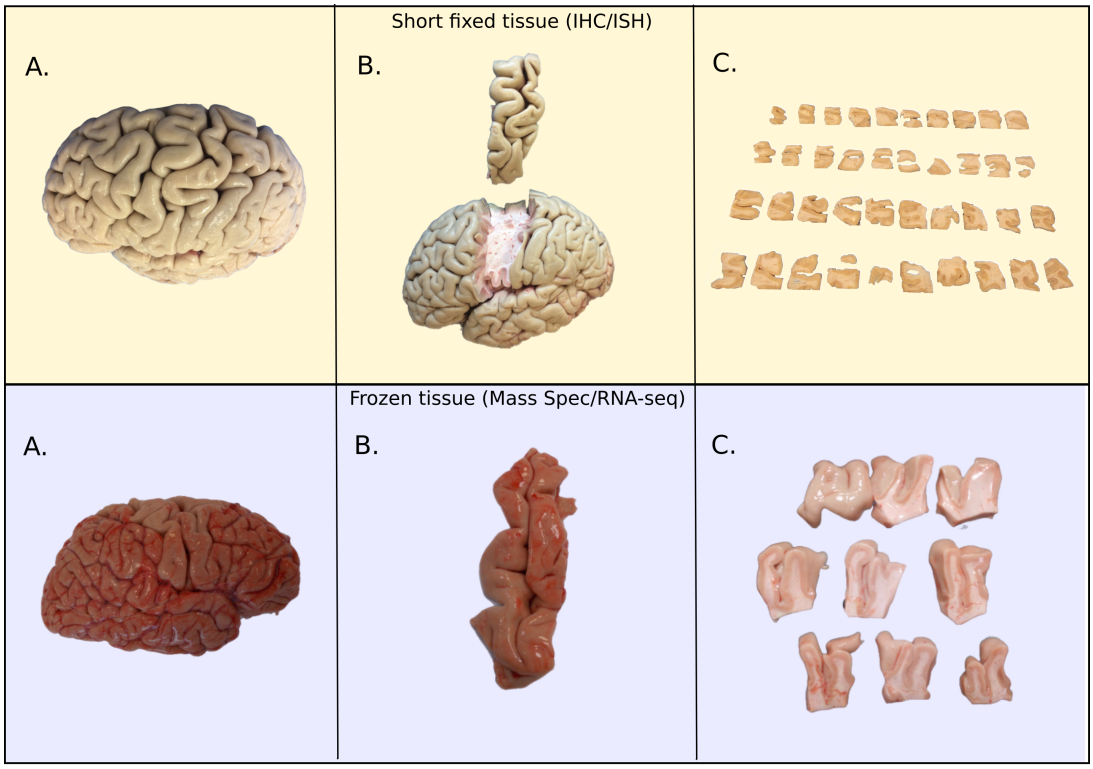


Figure 5.2: Preparing Fresh Post-Mortem Brains to Access Entire Betz Cell Population. A. Whole brains were split along the longitudinal fissure and each half of the brain was processed determining on downstream applications. One hemisphere was frozen for LC-MS/MS and RNA-seq analysis. A short-fixation protocol was applied on the other hemisphere - optimised for minimal RNA/protein degradation. B: The primary motor cortex was identified based on neighbouring neuroanatomical landmarks (central sulcus, lateral fissure, paracentral lobule) and the prominent hand-knob region of the primary motor cortex. The M1 strip was removed, along with the adjacent somatosensory cortex. C: The M1 strip was divided into several tissue blocks vaguely representing the motor homunculus.

5.3.2 Betz Cells and Purkinje Cells - Initial Data

Human Betz cells and Purkinje cells were stained with cresyl violet, microdissected in triplicate, and processed with the RWS workflow. Figure 5.3 shows a representative example of Betz cell microdissection. Previous work has shown that sufficient proteomic depth can be resolved from as little as 60,000µm^2, however, considering the rarity of Betz cells it was decided to microdissect a larger amount of material to ensure any single experiment was well above the threshold for identifying a high number of proteins IDs and would be resistant to potential mishaps in downstream experiment steps. As few as 150 Betz cell segments were isolated ($\approx$ 260,000µm^2) along with $\approx$ 400 Purkinje cells. The difference in cell number reflects the size difference between Betz and Purkinje cells. To maximise proteomic depth, library samples comprising of a cryosection of the motor cortex and cerebellum (unstained) was taken and analysed in parallel.

Using the RWS approach, it was possible to resolve a high number of proteins from the Betz cell sections ($\bar{x}$: 3,174 proteins by MS/MS), which was significantly higher (p value: 0.0008) than the numbers of proteins identified in the microdissected Purkinje cells ($\bar{x}$: 2,020 proteins by MS/MS). Despite the higher amount of material microdissected in this experiment, it was still possible to achieve an impressive 1.4 proteins per 1000µm^3 per one gradient hour and a total proteomic depth of 3,924 was found within Betz cell samples. There was a high correlation in MaxLFQ expression among both Betz cells ($\bar{x}$ PCC: 0.926) and Purkinje cell replicates ($\bar{x}$ PCC: 0.958). As expected, the majority of proteins found in this data-set were shared between the two cell-types (3,304, 79.9%), with 5.08% (210 proteins) found solely within the Purkinje cell fraction and the remaining 15% (620 proteins) found solely in the Betz cell fraction - See Figure D.1.

5.3.3 Relative Proteomics - Betz Cells vs. Purkinje Cells

To investigate the specific proteomic differences between the Betz and Purkinje cells, the maxLFQ values of all shared proteins between each cell-type were

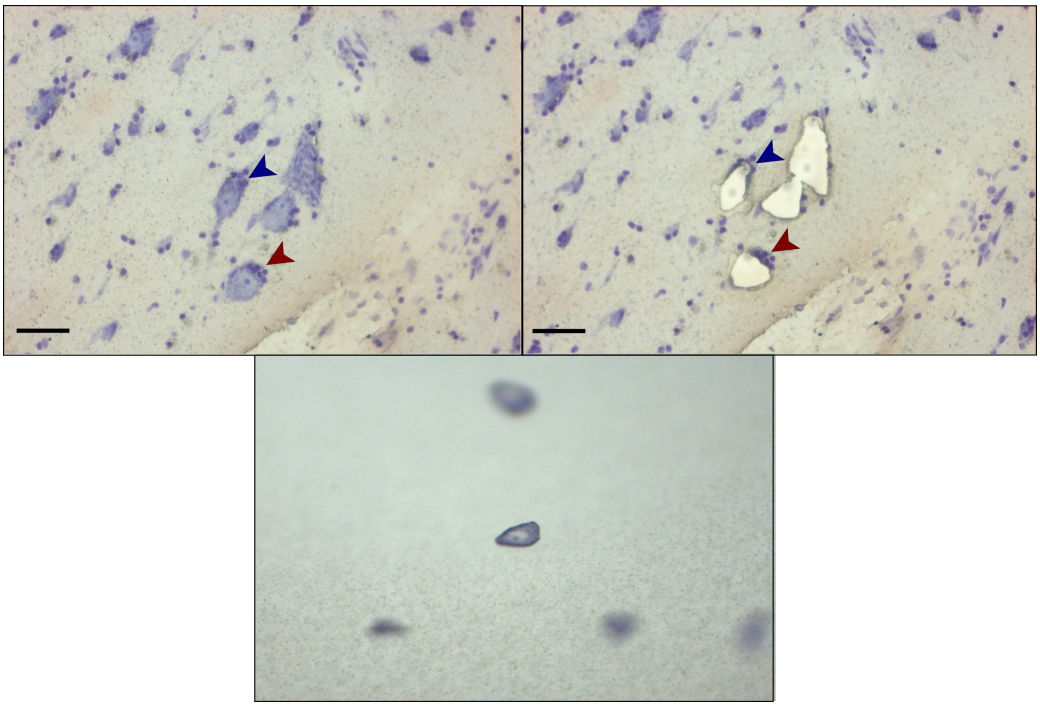


Figure 5.3: Representation of Microdissecting Betz cells - Before/After/In-cap. First panel: A nest of Betz cells identified by size, location, abundance of Nissl bodies, and pronounced nucleolus/nucleus. Second Panel: Successful microdissection of the previous nest of Betz. Bottom panel: Microdissection can be confirmed by inspecting the collection device and confirming captured cells contain Betz cell morphology. Red Arrow shows how Betz cells can be microdissection without capturing surrounding glial cells. However, some glial cells may be accidentally captured (Blue arrow) - potentially reflected in downstream data. Scale bar: 70µm^2.

compared. After filtering for only cases which contained at least 70% of valid values (2,622 proteins), data were submitted to NAguideR [268] which tested 23 different imputation methods to find the most appropriate for the data. The robust sequential imputation of missing values (impseqrob) was found to be the optimum method. Next, a standard t-test (Randomisations: 250, FDR Threshold: 0.05, S0: 0.1) was performed to compare variance; 1,124 (42.87% of proteins analysed) were found to be significantly differentially expressed. Of these 1,124 DEPs, 483 were enriched within the Purkinje cell samples and 641 were enriched within the Betz cells.

Promisingly, several GO pathways were enriched in Betz cells relating to cytoskeleton organisation (FDR value: 7.2e-03) - a known hallmark of Betz cell

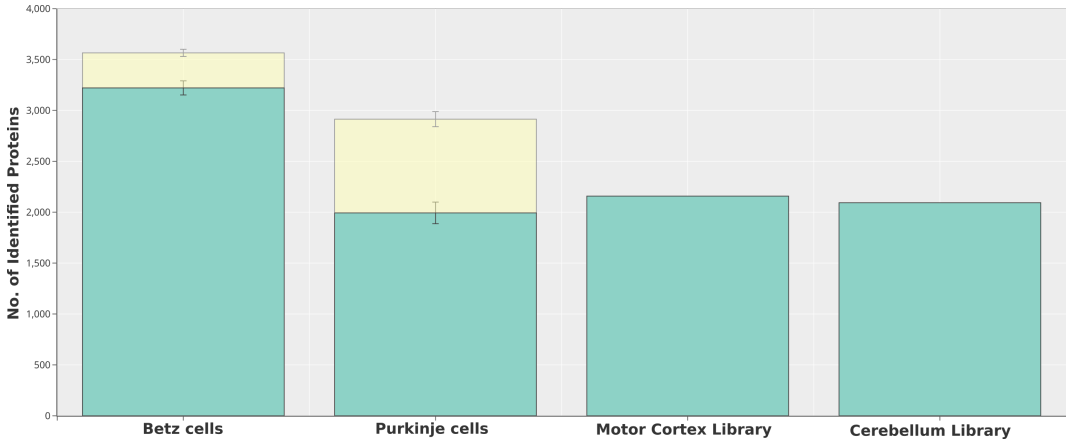


Figure 5.4: Betz Cells vs. Purkinje Cells - No. of Identifiable Proteins. A large number of proteins were identified in Betz cell samples. Samples were matched to library cryosections from the same neuroanatomical area for MBRs. Area microdissected: $\approx$ 260,000µm^2; Turquoise bars: by MS/MS, Cream Bars: by MBR. Vertical bars denote standard deviation.

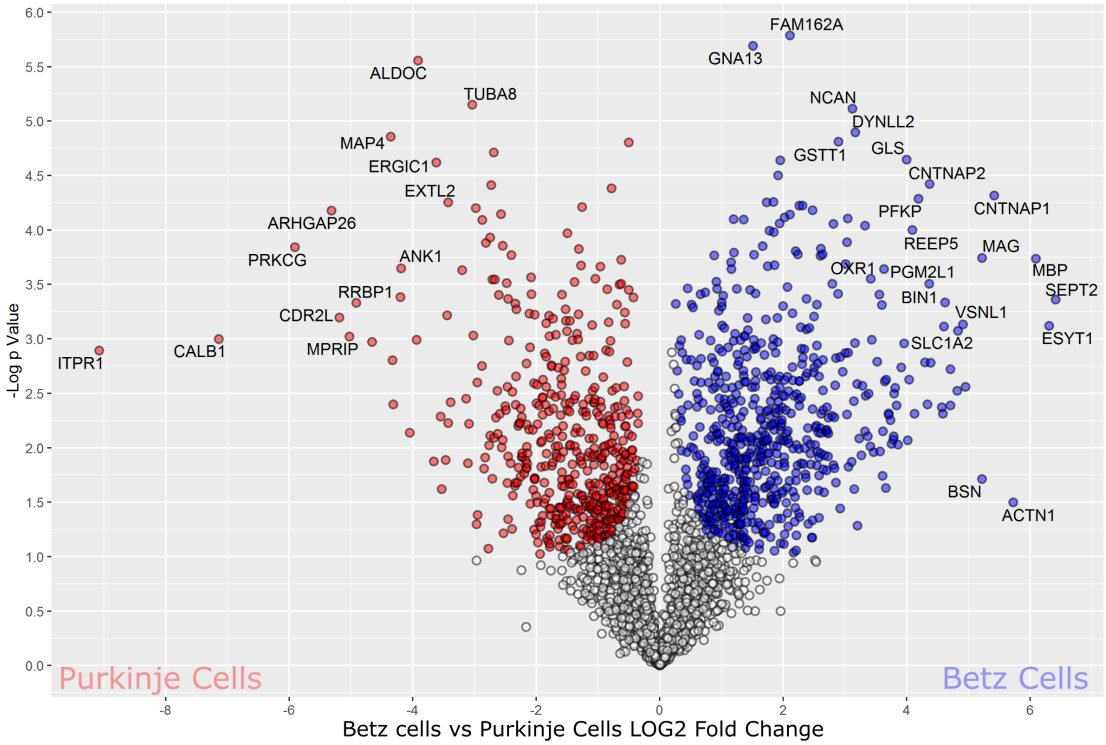


Figure 5.5: Volcano plot - Betz Cells vs. Purkinje Cells. A large number of deferentially expressed proteins (DEPs) were found when comparing Betz and Purkinje cells. Each point represents one protein. Coloured points are DEPs that passed predetermined alpha thresholds (0.05). Red: proteins whose expression was higher in Purkinje cells. Blue: proteins whose expression was higher in Betz cell samples.

physiology. Moreover, proteins known to be involved in upper motor neuron dysfunction (FDR value: 8.2e-03) were enriched within the Betz cell samples.

Although likely true, the Betz cell is commonly stated to be a glutamatergic neuron despite the lack of evidence in the literature. When ORA was performed on the above 641 DEPs, several terms relating to glutamatergic neurons were found: glutamate secretion (FDR value: 5.0e-03), glutamatergic synapse (FDR value: 2.9e-03), and glutamate neurotransmitter release cycle (FDR value: 4.0e-03) were all enriched in the Betz cell samples. Considering that the Purkinje cell is a known GABAergic neuron, it is reasonable to assume from these data that the Betz cell is glutamatergic. However, more work is needed to determine if glutamate is the sole neurotransmitter in Betz cells, or if the Betz cell has co-transmitter properties. Although, absolute quantification between these cell-types found an abundance of energy-generated proteins in both neurons - Figure D.2. Relative quantification suggests that Betz cell still contained an enrichment of energy-generated pathways compared to the Purkinje cell: the generation of precursor metabolites and energy (FDR value: 4.6e-05), cellular respiration (FDR value: 6.1e-04), ATP metabolic process (FDR value: 1.4e-03), oxidative phosphorylation (FDR value: 2.3e-03), electron transport chain (FDR value: 5.0e-03), NADH metabolic process (FDR value: 9.2e-03), and glycolysis/gluconeogenesis (FDR value: 1.9e-03). Furthermore, proteins that were localised to mitochondria (FDR value: 7.4e-06) were more enriched within Betz cells. This aligns with previous assumptions of the Betz cell based on their large soma and axon, in that they are highly metabolically active. Interestingly, several terms were enriched relating to neurotransmitter release: synaptic signaling (FDR value: 9.2e-03), synaptic vesicle priming (FDR value: 1.3e-02), synaptic vesicle membrane (FDR value: 1.3e-03), SNARE complex (FDR value: 1.1e-03), and neurotransmitter release cycle (FDR value: 5.3e-03).

1D enrichment [296] was then performed to rank proteins based on their expression, rather than relying on predetermined alpha thresholds; 404 GO pathways were found to be significantly enriched (213 within the Betz cell and 191 in the Purkinje cell samples). The 213 GO terms were processed using REVIGO to remove

redundancies, resulting in 61 terms, as shown in Figure 5.6. Unsurprisingly, many of the same terms from the above ORA were highlighted as shown in Figure 5.7, however, the 1D approach highlights proteins whose overall shift may be too subtle to detect on an individual level. For example, the Betz cell samples had a higher abundance of proteins relating to the cytoskeleton and axon guidance, however, with a 1D score of 0.172 and 0.207, this would be otherwise difficult to observe.

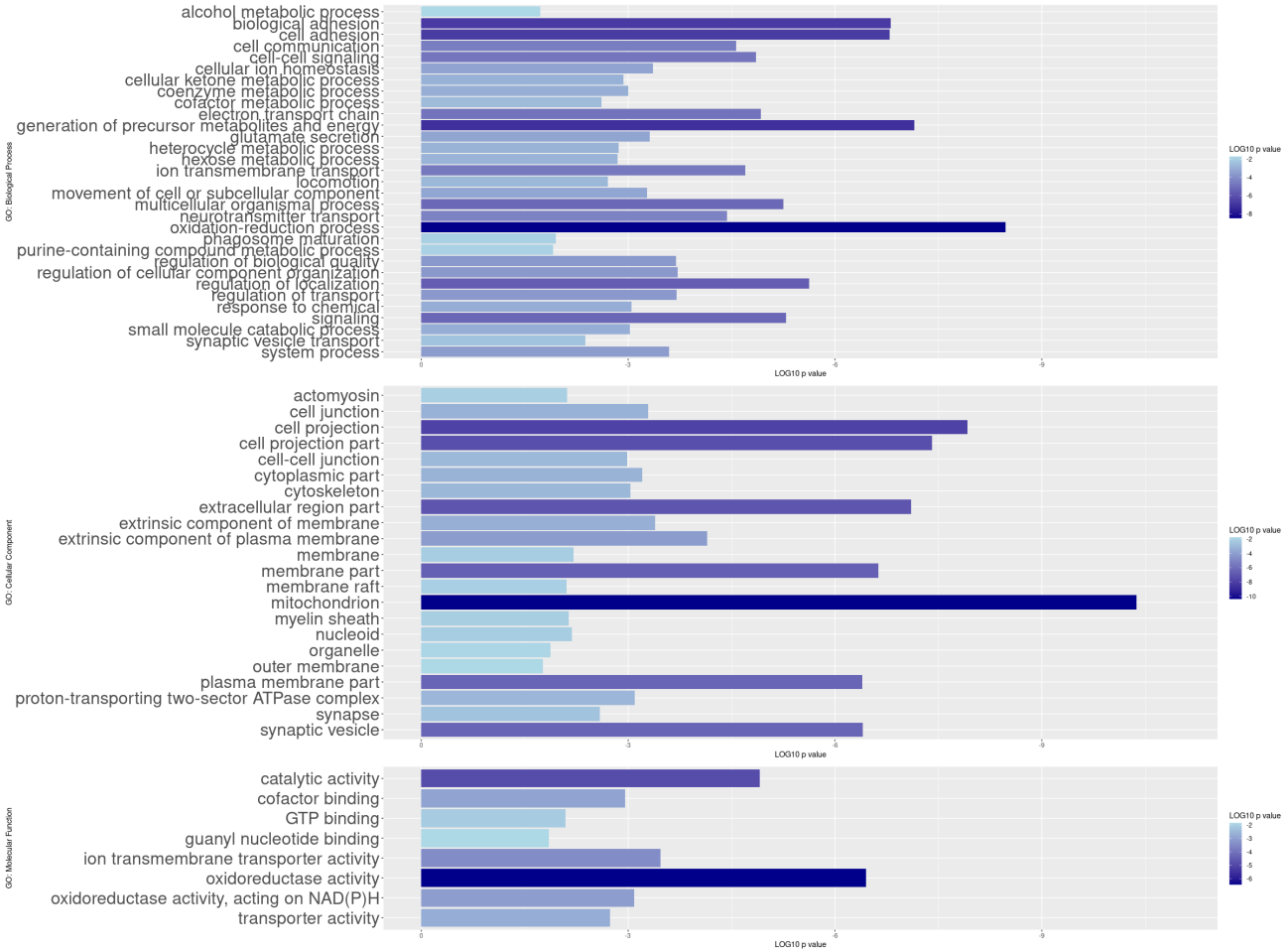


Figure 5.6: Enriched Gene Ontology Terms in Betz cells.
A list of GO terms that were enriched within Betz cellswhen compared to Purkinje cells. GO terms included: cellular adhesion, metabolism, neurotransmitter transport, energy generation, mitochondria, and cell projection.

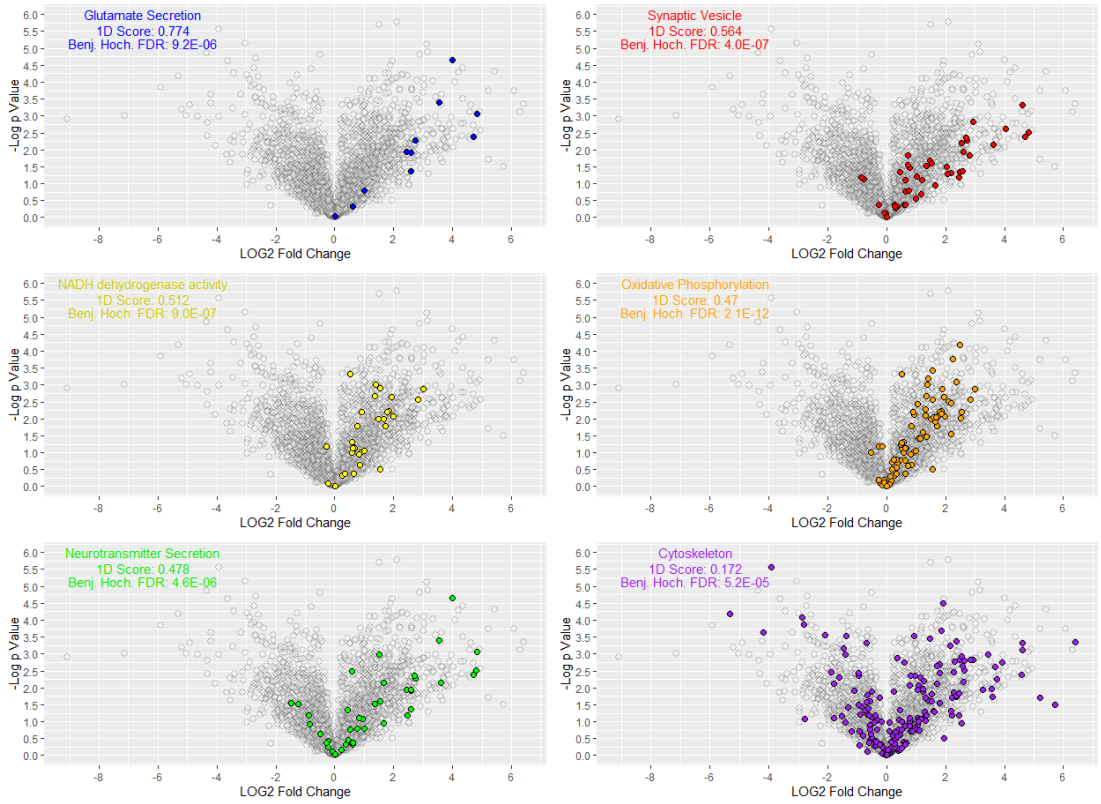


Figure 5.7: 1D Annotated Volcano plot - Betz Cells vs. Purkinje Cells. Volcano plots highlighted several GO terms that were enriched from 1D Enrichment. Betz cells contain an enrichment in neurotransmitter and energy-generating pathways, along with an enrichment in cytoskeletal proteins. Coloured points show proteins that belong to the named pathway. Proteins with a LogFC > 0 are more abundant in Betz cells. 1D Score (-1 to 1) reflects how enriched that GO term is within that sample. A 1D score > 0 reflects an enrichment of that pathway in Betz cells.

5.3.4 Digital Pathology

Promisingly, the proteins that differed most from Betz cells were known biomarkers for the Purkinje cell e.g. ITPR1, CALB1, PRKCG, ALDOC. It was decided to perform IHC on ITPR1 to confirm Purkinje cell localisation - see Figure 5.8. To validate LC-MS/MS findings, proteins not immediately known as Purkinje cells markers were also targeted by IHC. One of the highest-ranking DEPs was ARHGAP26, a GTPase activator for Rho-type GTPases. Staining for ARHGAP26 showed strong immunogenicity for Purkinje cells - similar to ITPR1 - seemingly highlighting their dense dendritic arbour, cell soma, descending axons, and synaptic boutons on the innervated cerebellar dentate nucleus neuron (DNNs) - See Figure 5.9. A similar staining pattern for the Betz cell would represent a "gold standard" biomarker. To confirm LCM proteomic observations that Betz cells contained comparatively reduced amounts of ARHGAP26, the primary motor cortex was also stained, see Appendix D.6. As shown, Betz cells were immunoreactive for ARHGAP26, however, the staining was not as pronounced as in the Purkinje cell; as with most immunochemical stains, it was also difficult to identify Betz cells as surrounding cell-types were also immunoreactive.

Back-mapping promising proteomic hits from LCM LC-MS/MS to tissue allows immunohistochemistry to be used as an orthogonal validation method. Moreover, results from immuno-labelling can highlight what regions or cell-types are enriched for the protein of interest, along with its subcellular localisation - increasing confidence in LC-MS/MS results. It can be used to identify other cell-types that are immunoreactive, either because they share cellular features or due to significant contamination during the LCM process. However, IHC interpretation is often subjective. As such, a quantitative measurement of immuno-stained material would be beneficial to future work. Similar to ITPR1, CALB1 was stained on both Betz and Purkinje cells - see Figure 5.10. As expected Purkinje cells expressed a high amount of CALB1 and in accordance with the LC-MS/MS data, Betz cells showed minimal immunoreactivity with CALB1, with only the synaptic boutons of presynaptic

neurons appearing to be positive[2]. In this example, the results of IHC were clearly identifiable by visual inspection, however, this may not always be the case. To solve this, an automated IHC image analysis pipeline was designed to provide a quantitative measurement of IHC. First, stained tissue sections were digitised into virtual microscopy slides. Next, a multilayer perceptron (feed-forward artificial neural network) was trained to recognise cell-types of interest (e.g. Betz cells or Purkinje cells). Once recognised, annotations were automatically drawn around the cell's perikaryon and the resulting DAB intensity was measured to get a quantitative measurement[3] - see Figure 3.8.2. This pipeline was applied to FFPE material rather than frozen tissue, due to its superiority in preserving tissue architecture. To ensure the workflow could identify cells accurately, 4 cases (with 5 tissue blocks each) were stained with SMI311 to visualise neurons, including Betz cells, in the primary motor cortex. The above image analysis workflow was then applied and the results blinded. Manually counting the same histology sections was used as a benchmark. Overall, the results of the image analysis and from manually counting Betz cells had a PCC of 0.96. In total, this image analysis approach allows large numbers of cells to be comparatively analysed without laboriously annotating hundreds of individual neurons. An example of this output can be seen in Figure 5.10.

[2]It is important to note that despite the enrichment performed by LCM there will always be a level of contamination due to the interwoven network of neural synapses, and must be considered when drawing biological conclusions.

[3]Optical density was used to measure the level of chromogen staining which has been shown to be related to protein expression [297].

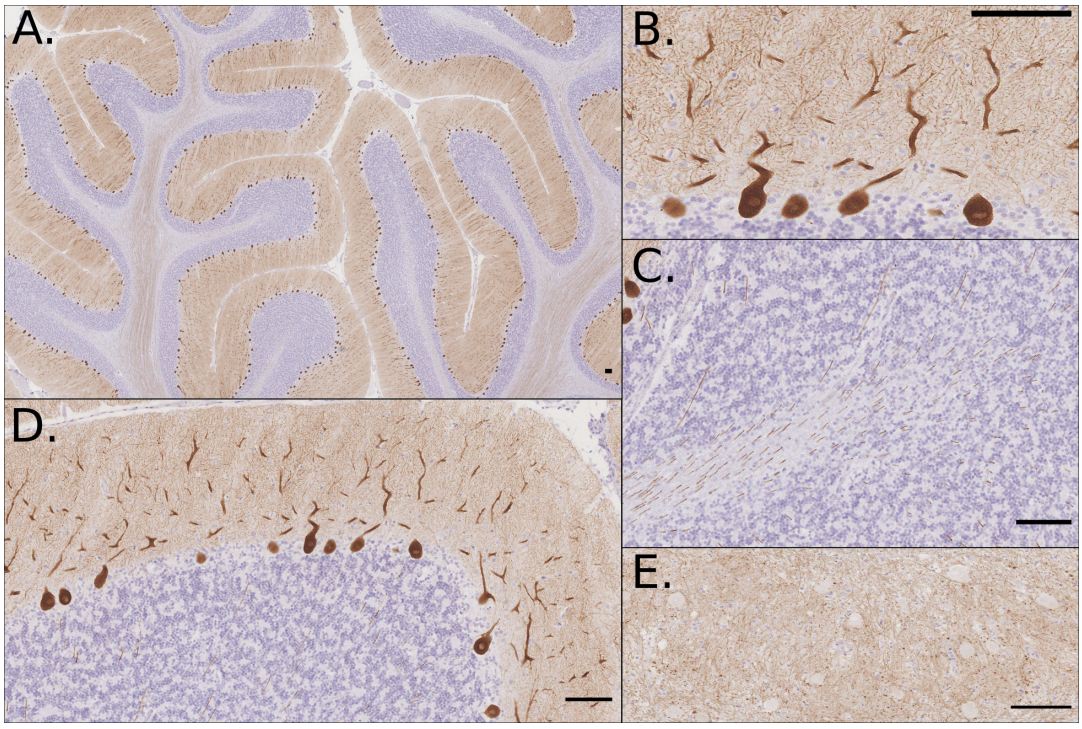


Figure 5.8: Validating LC-MS/MS data - Immunostaining for ITPR1.
A: Low magnification shows intense staining of ITPR1 within Purkinje cells. The Purkinje cell soma (B) dendritic arbour (D) were strongly positive for ITPR1. C: Staining was unique to Purkinje cells with no observable immunoreactivity in the granule layer of the cerebellum. E: No staining was seen in the cerebellar dentate nucleus neuron, although the innervating axons and synaptic boutons of Purkinje cells were immunoreactive. ITPR1 highlights the totality of the Purkinje cell, with minimal staining to other cell-types representing a "gold standard" biomarker. Scalebar: 100µm.

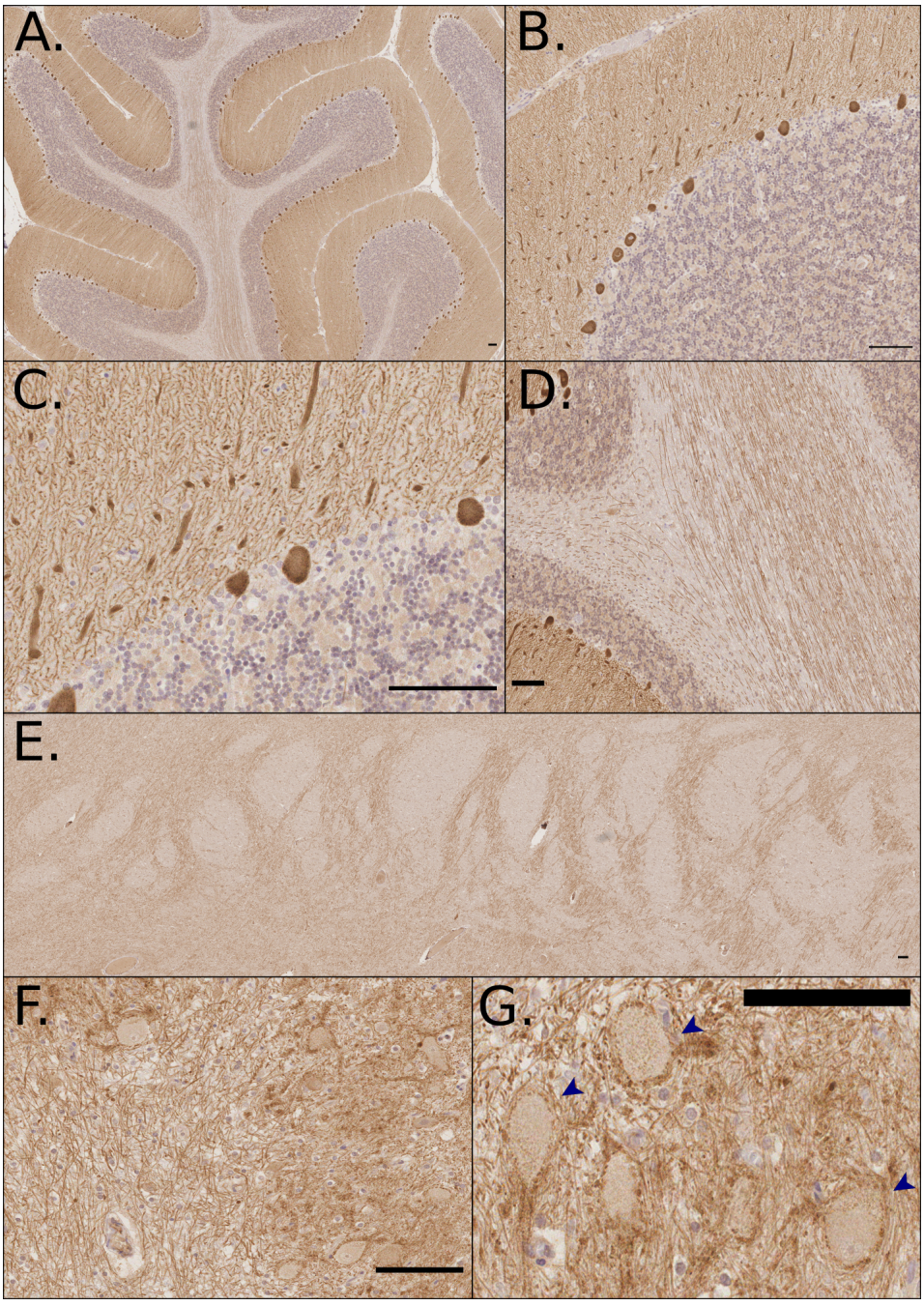


Figure 5.9: Validating LC-MS/MS data - Immunostaining for ARHGAP26. ARHGAP26 mimics the Purkinje cell specificity seen in ITPR1 IHC. A: Low magnification showed ARHGAP26 staining reveals Purkinje cell morphology in the molecular layer and white matter tract. B & C: ARHGAP26 highlights the dense dendritic arbour of Purkinje cells. Some immunoreactivity is seen in the granule layer. Axons of Purkinje cells are immuno-reactive with ARHGAP26 (D), and can reveal the striated pattern of white matter tracts. As with ITPR1 ARHGAP26 also does not stain the cerebellar dentate nucleus neuron (F & G), however, the encompassing innervating synaptic boutons on the Dentate Nucleus can be seen - Blue arrows. Scalebar: 100µm.

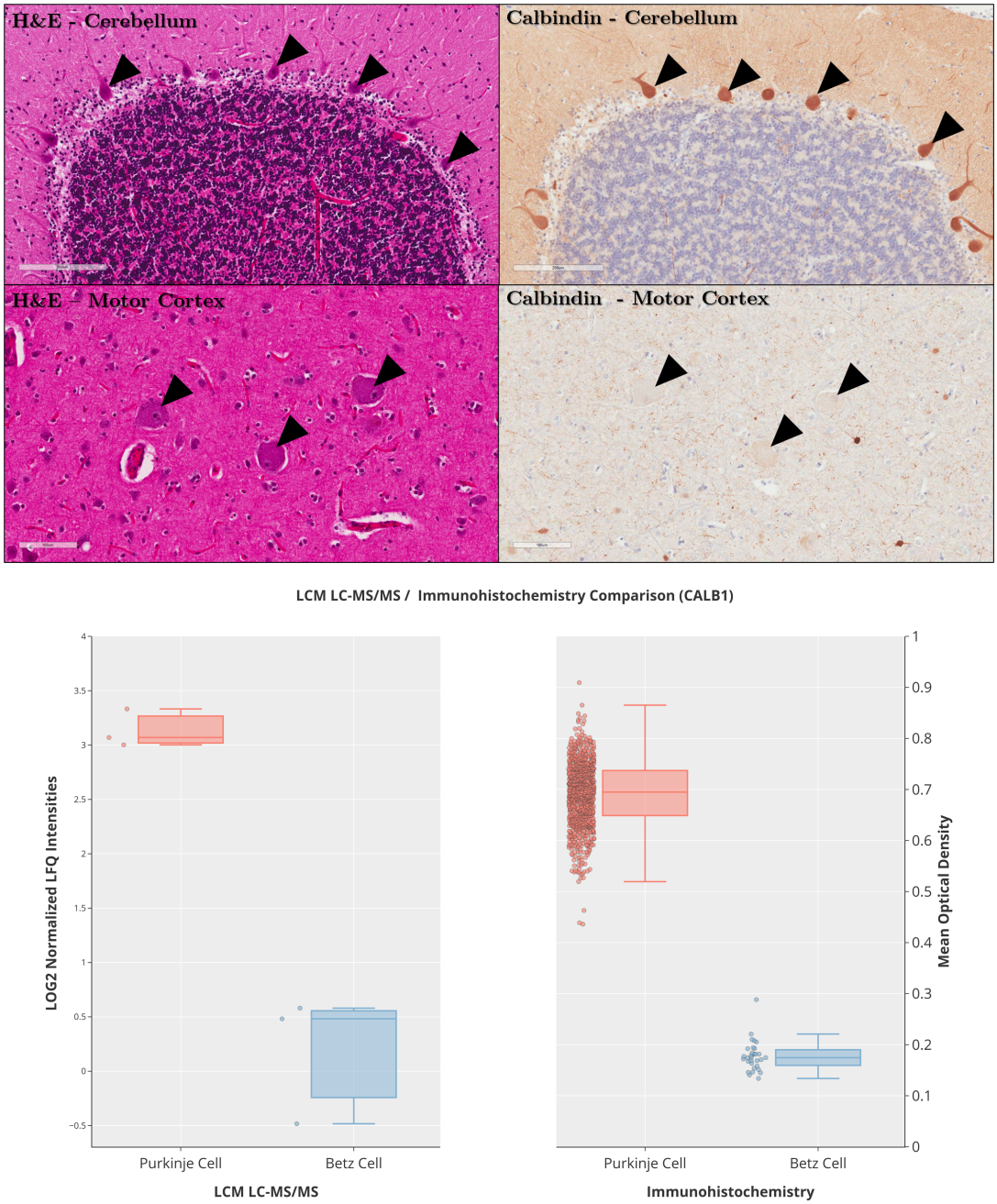


Figure 5.10: Quantitative IHC can Validate LC-MS/MS Findings. Proof-of-principle on one cerebellar and one motor cortex section. Using digital image analysis (Positive pixel count) it is possible to recapitulate protein expression independent to LC-MS/MS.

5.3.5 LCM Proteomic-Directed Immunohistochemistry

Considering that the highest DEPs in the Purkinje cell were known Purkinje cell markers, it was decided to target Betz DEPs to confirm LCM LC-MS/MS observations and potentially identify proteins that may ease the task of identifying Betz cells in the future, without having to rely on pre-existing knowledge of their morphology.

Generic Mitochondrial Marker -

As proteomic investigations suggested an abundance of mitochondria and mitochondrial activity in Betz cells it was decided to first attempt to use a generic mitochondrial marker (Clone: Ae-1, see Figure 5.11). As expected, immunoreactivity to this antibody was not Betz cell-specific, nevertheless, IHC of Betz cells confirmed that they contain abundant amounts of mitochondria. The characteristic L5 band of Betz cells was the most obvious cellular feature at low magnifications and was not seen in the adjacent primary somatosensory cortex - Panel D, Figure 5.11. Accumulations of mitochondria were localised to both Betz cell soma and dendrites, highlighting some of the dense dendritic commissures of the Betz cell. Although other cell-types were also immunoreactive for Ae-1, the high expression within Betz cells make them easily identifiable; which is a similar IHC pattern seen with neurofilament - the current best example of an antibody that can highlight Betz cells from the surrounding L5 neurons.

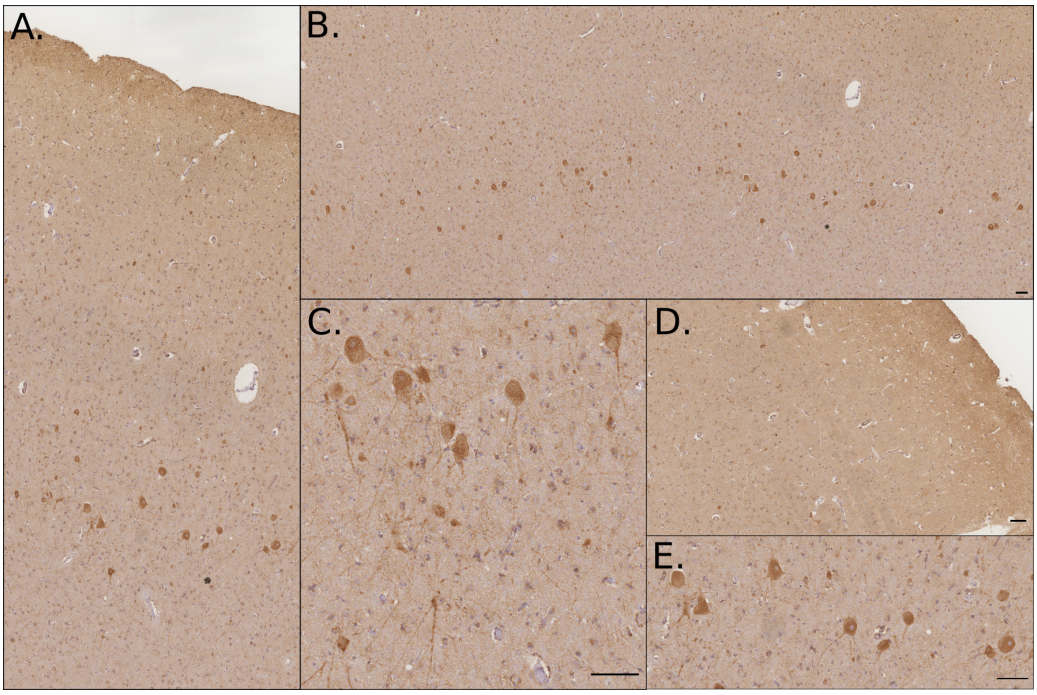


Figure 5.11: AE-1 Staining of the Primary Motor Cortex.
A: At low magnification, it was possible to visualise Betz cell & it possible to identify the L5 band of Betz cells (B). C: At higher magnifications, it was possible to identify Betz cells, as immunoreactivity was strongest in the Betz cell soma and some basilar dendrites. D: No remarkable cytological features were seen at low magnifications when Ae-1 staining was applied to the somatosensory cortex. While Ae-1 staining was enriched within Betz cells, the neuropil and surrounding cell-types were also immunoreactive (E), potentially limiting Ae-1 as a useful marker for Betz cells in the future. Scalebar: 100µm.

ATPB -

Considering the abundance of ATP synthase related proteins in the Betz vs. Purkinje data-set it was decided to immunochemically target ATP synthase, in particular, ATPB (ATP synthase subunit beta, mitochondria; LogFC: 1.42) to confirm LCM LC-MS/MS findings. As expected, ATPB was located primarily in the grey matter, where the vast majority of neuronal bodies are located. At low magnification it was possible to distinguish the L5 band of Betz cells - see B, C, and D of Figure 5.12. At higher magnification, ATPB was seen localised to the cytoplasm, including dendritic processes, allowing for the circumferential dendritic tree of the Betz cell to be visualised. The abundance of lipofuscin within Betz cells could also be visualised, increasing the confidence that these L5 neurons were indeed Betz cells. As a critical protein in cellular respiration, ATPB would be expected in the vast majority of cell-types and at higher power, it is seen with smaller neurons. Further experiments would need to be performed to determine if Betz cells contain larger amounts of ATPB than neighbouring neurons or if their prominent staining is simply due to their large soma. Nevertheless, with anti-ATPB staining, it was possible to distinguish Betz cells from surrounding neuropil and cellular environment, although the fact that other motor neurons may also be strongly immunoreactive may still lead to errors in Betz cell identification. Echoing the proteomic data, DAB optical density measurements of individual cells highlighted a significant increase in ATPB in the Betz cells compared to the Purkinje cell (p value: <0.001). Similar to the neocortex, the white matter of the cerebellum contained little ATPB immunoreactivity, with the molecular layer also containing low amounts - likely as a result of the lowered density of neurons in these two locations. The distinct staining pattern of ATPB in the Granule layer is characteristic of the complex cerebellar glomerulus, one of the most complex synapses in the CNS [298]. Interestingly, when anti-ATPB IHC was applied to the cerebellum the large neurons of the dentate nucleus also produced high immunoreactivity, comparable to that of Betz cells - see Figure 5.14. ATPB may be used as a surrogate for a cell's energy output, as such, these data confirm previous assumptions about the Betz cell; in that, it is a highly active

neuron with large energy demands, potentially more so than the principal neuron of the cerebellum - the Purkinje cell, likewise, DDNs appear to also have large energy demands, comparable to that of the Betz cell.

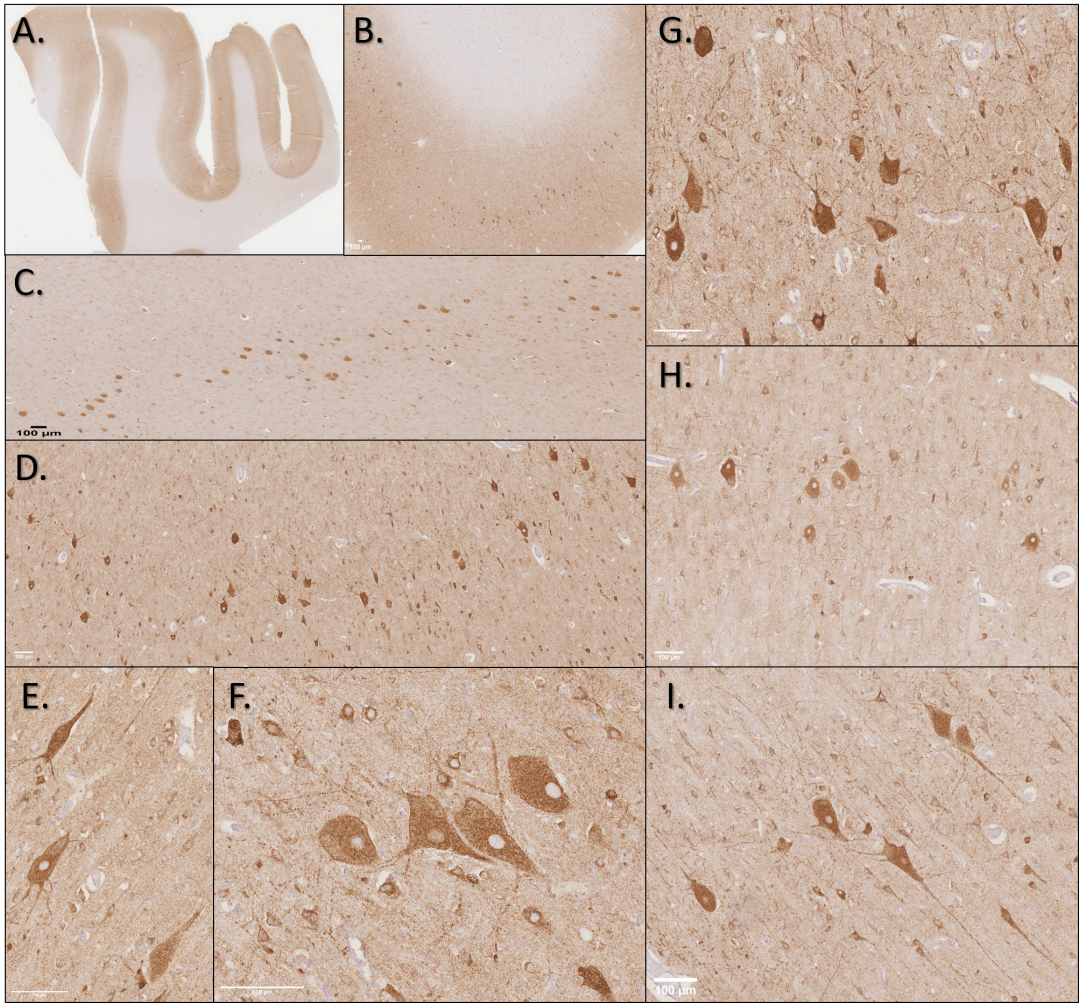


Figure 5.12: ATPB - Validating Proteomic Data.
ATPB IHC showed selective immunoreactivity to the grey matter of the primary motor cortex (A). Despite the diffuse staining, Betz cells contained intense immunoreactivity, visible at low power (B-D) to higher magnification (e-I). As ATPB also stained some dendrites, it was possible to identify Betz cells based on their unique somatodendritic pattern. Some surrounding pyramidal neurons also contained strong immunoreactivity (E, F, G), however, the intense immunoreactivity among Betz cells still made them easily identifiable (H and I). Scalebar: 100µm.

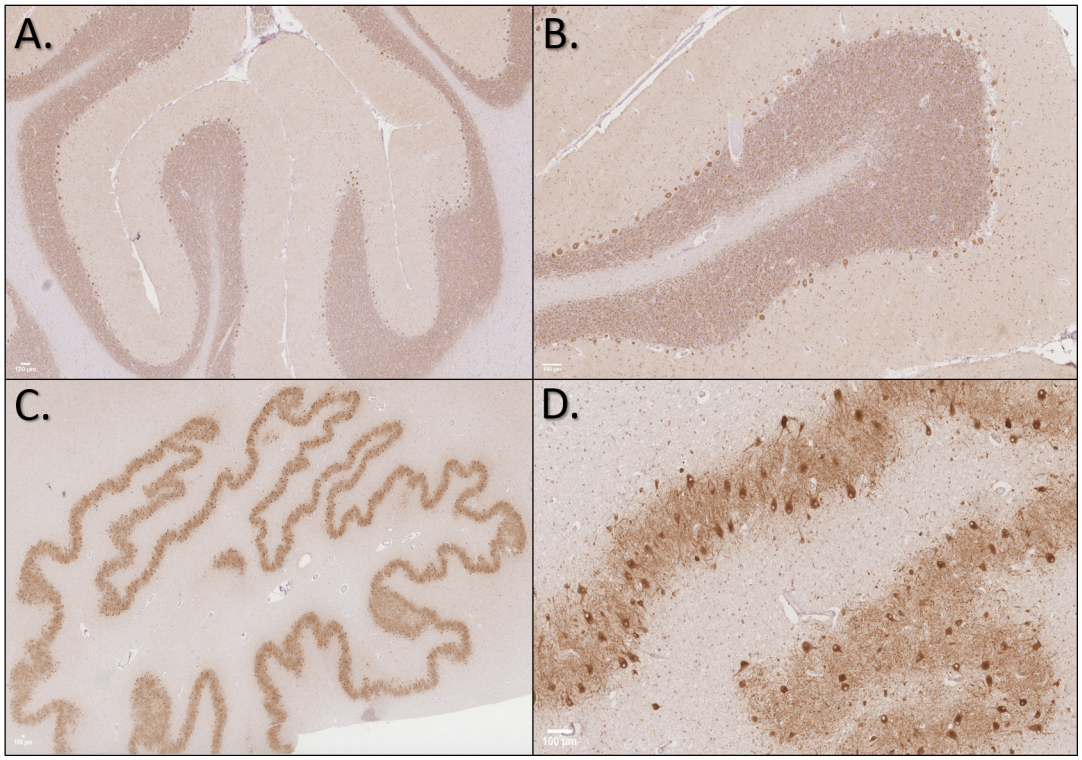


Figure 5.13: ATPB IHC - Cerebellum.
ATPB staining revealed moderate staining in Purkinje cells (A & B), and intense immunoreactivity within the cerebellar Dentate Nucleus (C & D). Scalebar: 100µm.

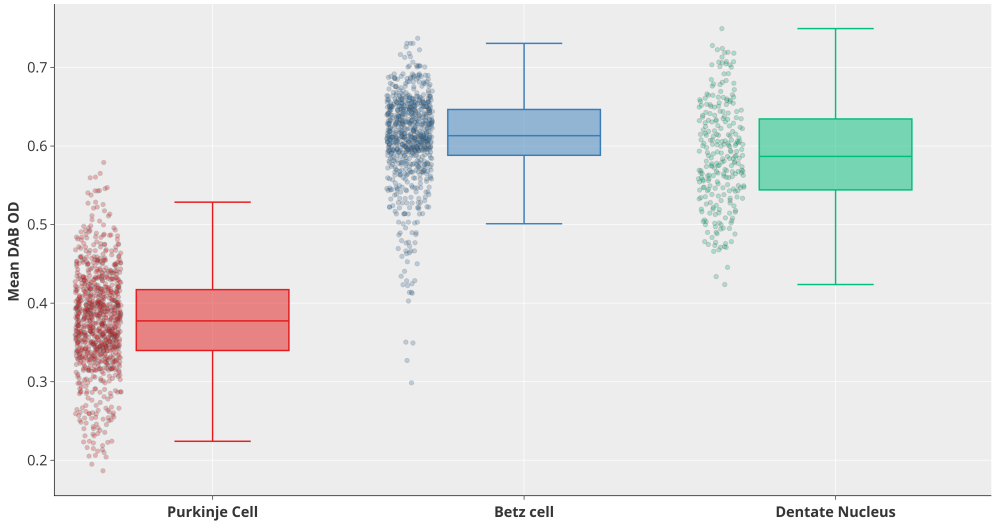


Figure 5.14: ATPB - Quantitative IHC.
Image Analysis results of ATPB staining. The mean optical density of DAB was calculated via positive pixel counts and compared for each cell. Betz cell vs. Purkinje: p value: <0.001; Betz cell vs. Dentate Nucleus p value: <0.001.

CYCS -

Cytochrome c oxidase (CYCS) is a large transmembrane protein and the last enzyme in the electron transport chain, localised in-between the inner and outer mitochondrial membranes. In this proteomic data-set, CYCS had a 3.71 LogFC in Betz cells. Anti-CYCS staining revealed ubiquitous immunoreactivity. The strong immunoreactivity in Betz cells made them easily distinguishable from the surrounding neuropil. As with ATPB, non-Betz L5 neurons also had strong immunoreactivity, which limits its use as a useful biomarker for the Betz cell - see Figure 5.15 (D). Despite the large difference in the MaxLFQ values, Purkinje cells appeared to have similar amounts of CYCS as Betz cells. However, when analysed with the automated image analysis workflow, a slight but significant difference (p value: <0.001) in anti-CYCS staining between the two cell-types. Similar to ATPB staining, DNNs also had strong immunoreactivity, and of the three cell-types had the largest abundance of CYCS (p value: <0.001).

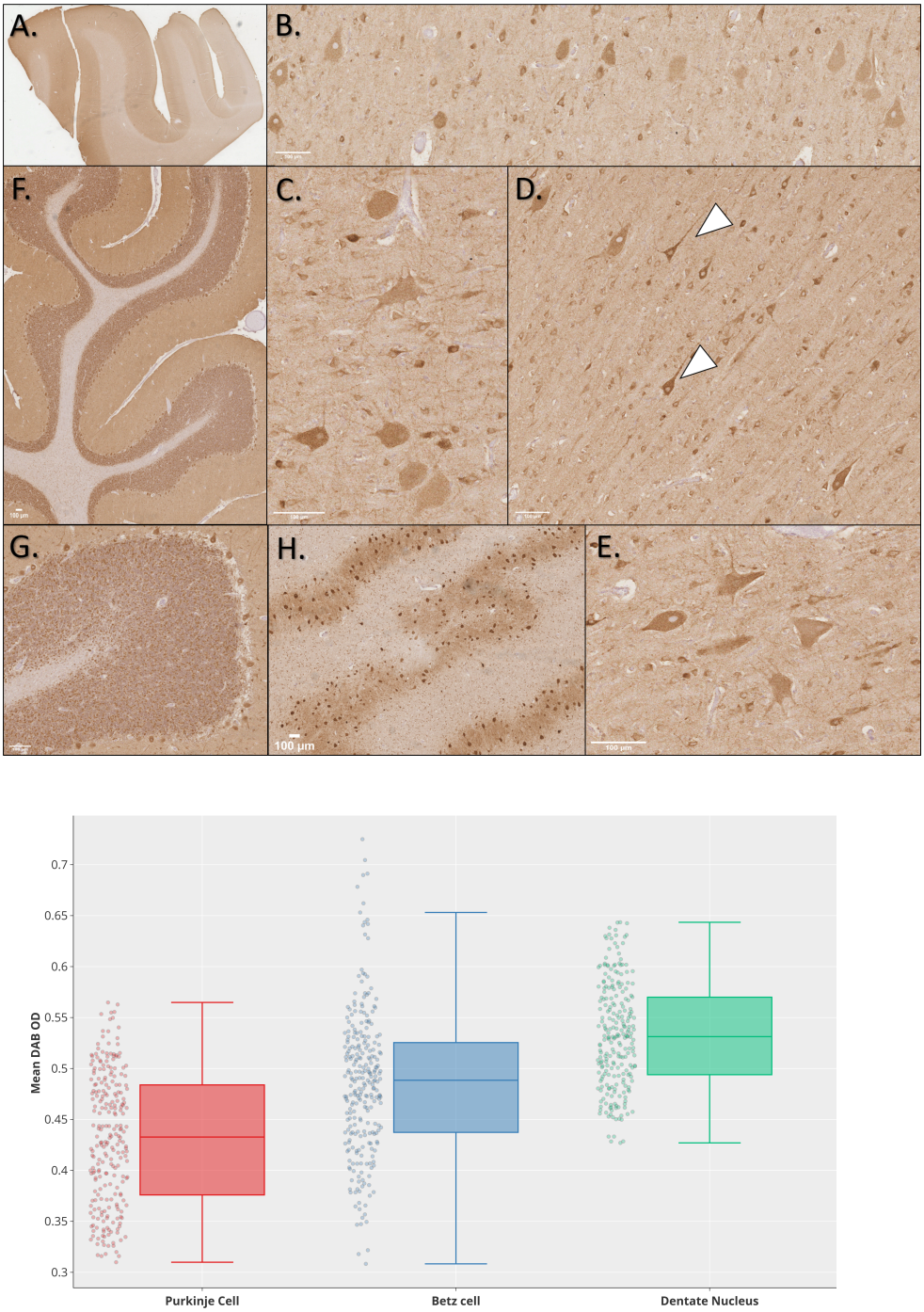


Figure 5.15: CYCS - Validating Proteomic Data and Quantitative IHC. Diffuse CYCS staining was seen throughout neocortical (A) and cerebellar tissue sections (F). Betz cells were immunoreactive for CYCS (B, C, E), however, at times they were difficult to confidently identify due to neuropil staining (B & C). Non-Betz cell L5 neurons were also strongly immunoreactive (white arrows) making CYCS, not a suitable biomarker for selectively studying the Betz cell. As with ATPB, Purkinje cells showed moderate immunoreactivity (G), while the Dentate Nucleus contained a high expression of CYCS (H). Scalebar: 100µm.

VDAC1 -

Next, to investigate the robustness of our approach, a protein was selected that had failed to pass the significant threshold for our proteomic data. Voltage-dependent anion-selective channel 1 (VDAC1), a channel that allows the diffusion of small molecules in the mitochondrial outer membrane and plasma membrane; had a small LogFC difference (0.417) and fell short of our alpha threshold p value: 0.104. Betz cells showed strong immunoreactivity for VDAC1. It was possible to identify the characteristic finger-like dendritic projections within some Betz cells - panel E of Figure 5.16. Unlike the previous proteins targeted, VDAC1 have very little diffuse staining in the neuropil, making the task of identifying Betz cells simpler than previous antibodies. Non-Betz cell L5 neurons were also positive for VDAC1. Immunostaining showed similar immunoreactivity between Betz cells and Purkinje cells, although, despite the small fold change in the LC-MS/MS data, unbiased image analysis showed a relatively large difference in VDAC1 expression between these neurons (LogFC: 2.43). It is unclear if this discrepancy is due to bias in the image analysis pipeline, the non-linear relationship with DAB and protein expression, or the limited number of replicates in the proteomic analysis.

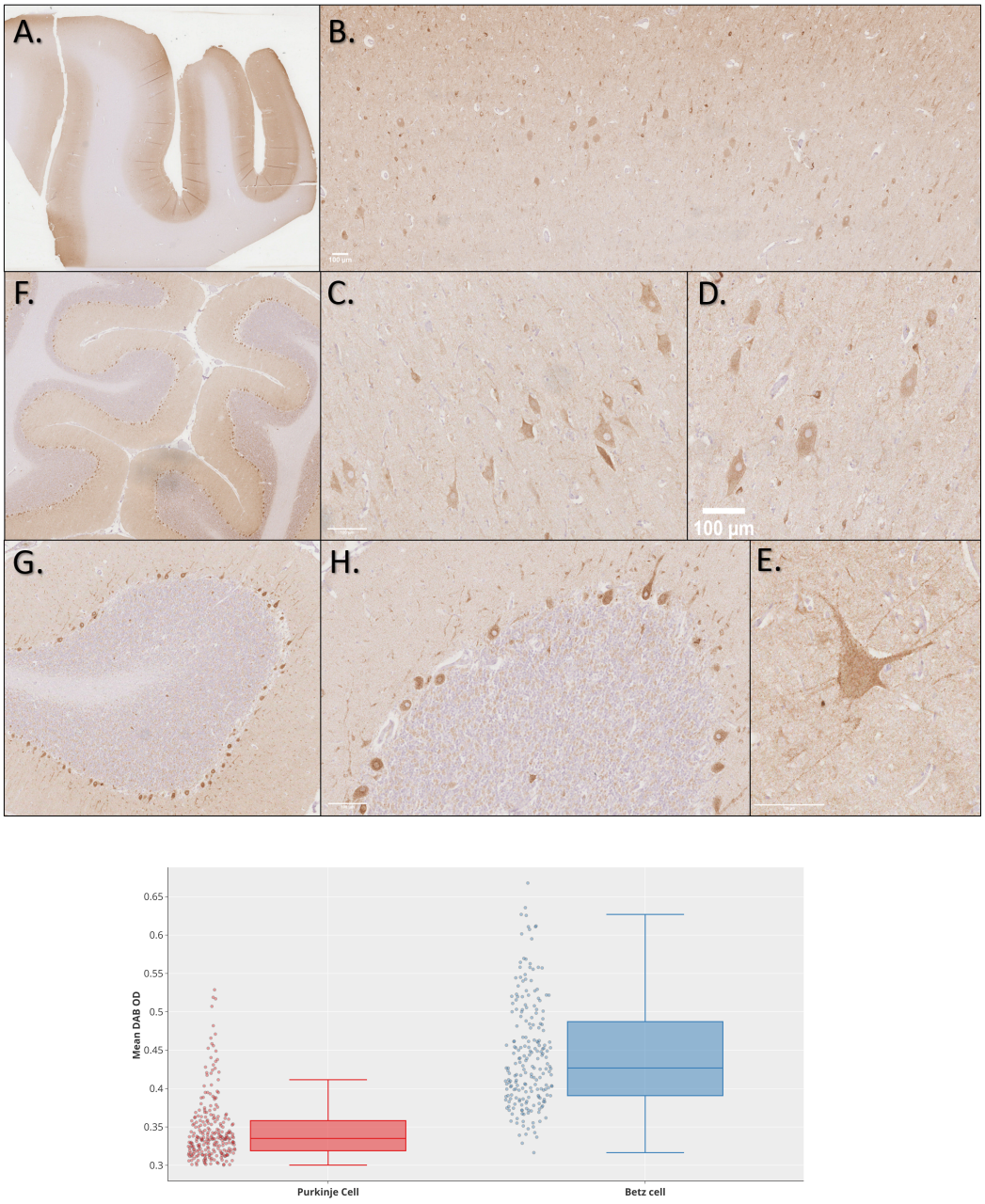


Figure 5.16: VDAC1 - Validating Proteomic Data and Quantitative IHC. Similar to ATPB, VDAC1 showed high expression in the grey matter but minimal staining in the white matter (A & F). Betz soma was enriched in VDAC1 and was clearly identifiable (B, C, D). Purkinje cells also had increased expression of VDAC1 (G & H). At higher magnification, VDAC1 IHC stained the dendrites and the characteristic finger-like dendritic projections of the Betz cell (E). VDAC1 was not unique to Betz cells, with other L5 neurons being immunoreactive (B). Only a subset of L5 neurons was immunoreactive, unlike SMI32 antibody staining, where most neurons are stained. This, plus the fact that VDAC1 IHC highlighting some Betz cell dendritic morphology, may make VDAC1 a useful marker for studying the Betz cell in the future. Scalebar: 100µm.

DST -

Dystonin (DST) is a cytoskeletal linker protein that anchors intermediate filaments to the actin cytoskeleton in neural, muscular, and epithelial cells. In sensory neurons, it is known to regulate the organisation and stability of the microtubule network to facilitate axonal transport [299]. In particular, it docks with the dynein/dynactin motor complex to vesicle cargos for retrograde axonal transport. This protein was selected for IHC as Betz cells are known to express an abundance of cytoskeletal proteins. Moreover, their long axons and the fact that they are the principal neurons of the primary motor cortex would suggest that they would undergo a large amount of axonal transport. Furthermore, DST was one of the top 50 most differentially expressed proteins in this data-set (LogFC: 3.43; adj. p value: 0.002). The anti-DST antibody selected (PA5-56396) targeted a multitude of DST isoforms, including isoform 6 which is required for bundling actin filaments around the nucleus, and isoform 7 which has been shown to facilitate axonal transport in sensory neurons. When stained, Betz cells were highly immunoreactive, with only some grey matter diffuse positivity - Figure 5.19. As with the previous proteins, it was possible to identify and track the L5 "Betz cell band" across the motor cortex, even at low magnification - Figure 5.17. While other neurons in the grey matter also stained positive for DST - some staining was seen in L3 - it was abundant in the gigantopyramidal neurons. At higher magnification, it was possible to view DST immunoreactivity along with the presence of excessive lipofuscin. Anti-DST staining of the cerebellum showed a pronounced reduction in DST expression among Purkinje cells, confirming LCM proteomic findings (p value: <0.001). While DNNs contained a higher expression of DST than Purkinje cells, Betz cells contained the most abundant of DST in all three neurons. While, neurofilament heavy (identified by SMI32/SMI311 antibodies) better highlights the full somatodendritic structure of the Betz cell (see Appendix A, DST appears to be better than teasing out Betz cells from the cellular diversity of M1, and may represent a potential robust biomarker for the Betz cell.

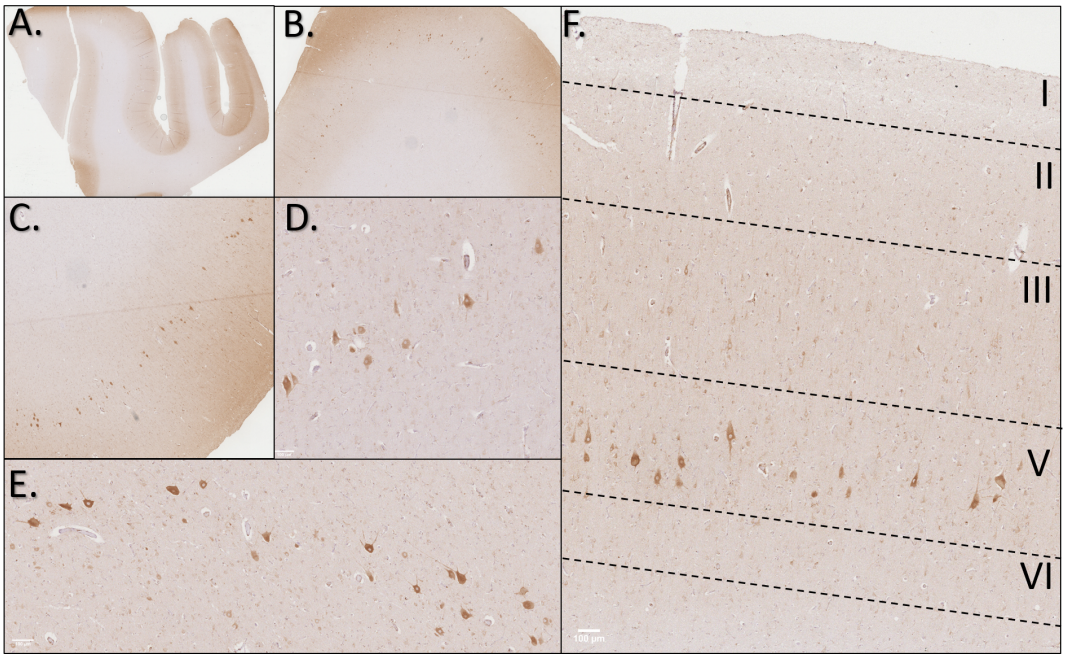


Figure 5.17: DST (Low Magnification - Motor Cortex) - Validating Proteomic Data. DST expression was up-regulated in Betz cells. DST IHC showed some diffuse staining in the grey matter (A), and the characteristic L5 band of Betz cells was easily identifiable with minimal staining of other cell-types (B-E). DST expression was enriched within L5, with faint staining in L3 pyramid neurons. Scalebar: 100µm.

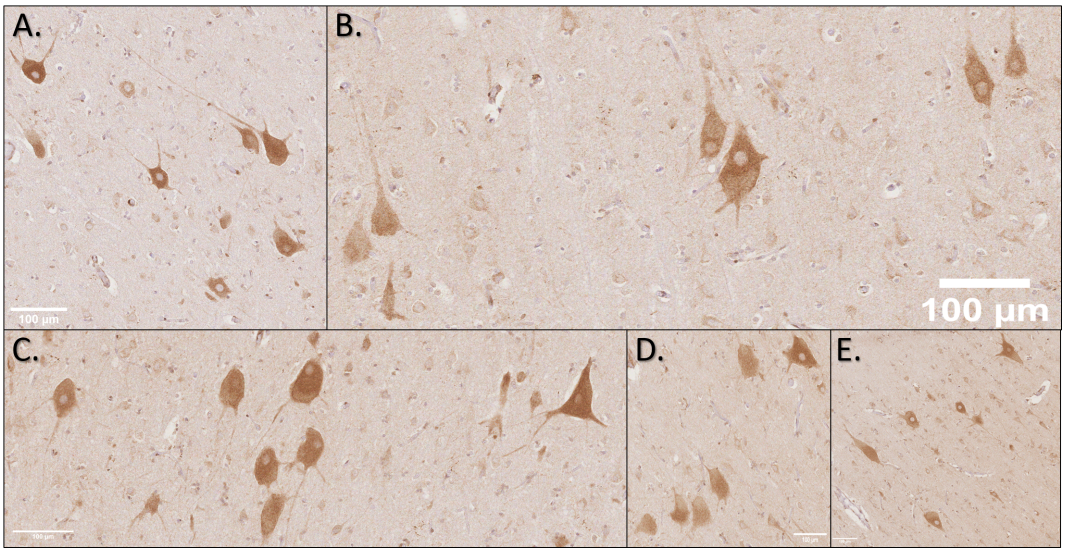


Figure 5.18: DST (High Magnification - Motor Cortex) - Validating Proteomic Data. At higher magnifications, Betz cells were the most intensely stained cell. DST was mostly localised to the Betz cell soma, with some staining of basilar dendrites. The increased intensity of DST in Betz cells and only the faint expression of other L5 neurons makes it a promising Betz cell marker. Scalebar: 100µm.

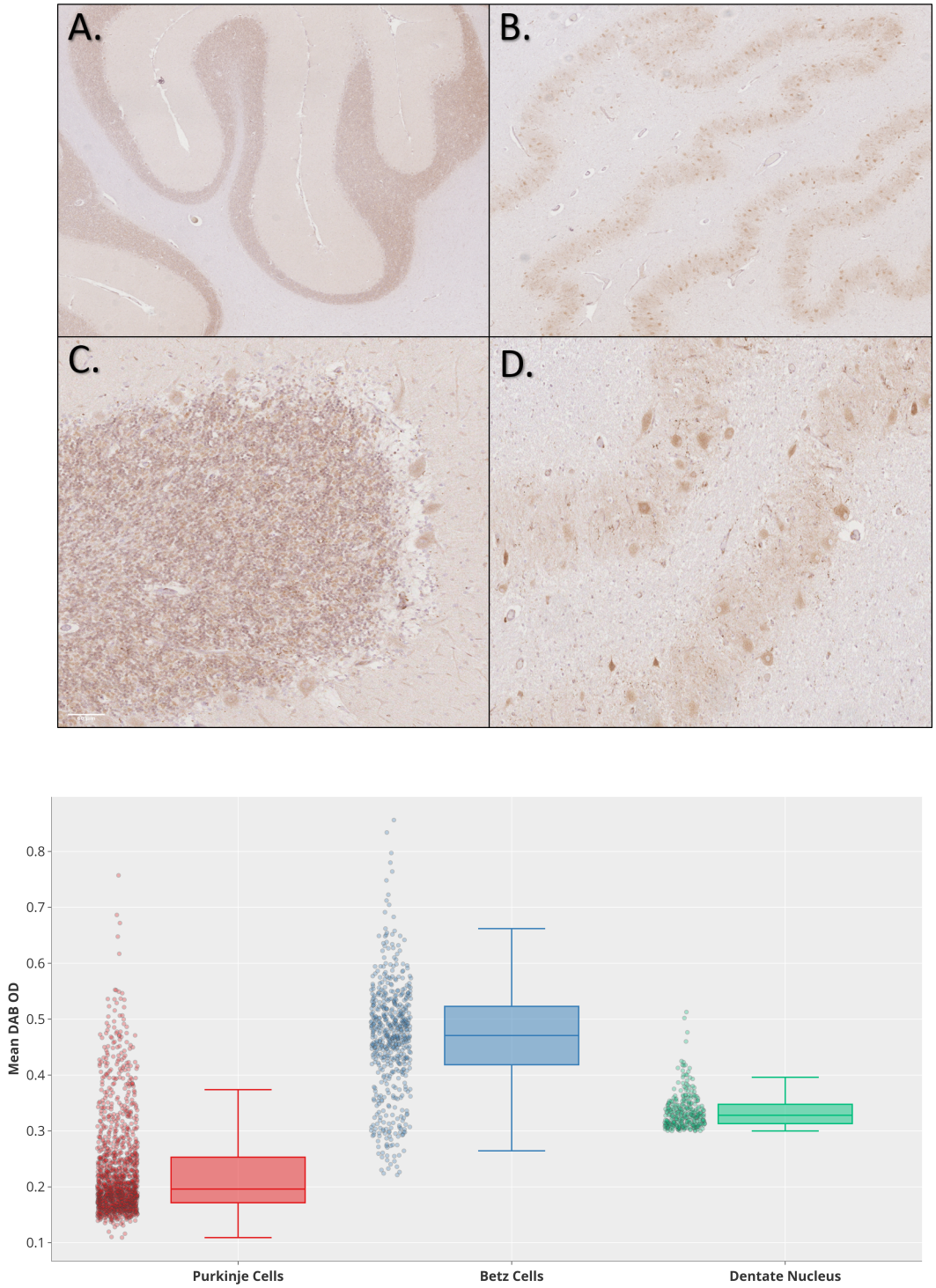


Figure 5.19: DST (Cerebellum) - Validating Proteomic Data and Quantitative IHC. Both Purkinje cells and Dentate Nucleus were faintly positive for DST measured via positive pixel counts while Betz cells contained a higher abundance of DST - reflecting the generated LC-MS/MS data. Scalebar: 100µm.

ESYT1 -

Extended Synaptotagmin 1 (ESYT1) is an endoplasmic reticulum-anchored protein that mediates the distance and connection of the endoplasmic reticulum and plasma membrane in response to cytosolic Ca^{2+} [300, 301]. These ER-PM junctions are the facilitators of store-operated Ca^{2+}, a calcium entry mechanism that activates when ER stores of calcium are reduced. The full role of ESYTs in ER-PM contact connections is unclear, in neurons, there is evidence to suggest that they play a role in lipid transfer between the two membranes, along with contributing to protein trafficking, synaptic function, and neuronal morphogenesis [302]. Interestingly, in *Drosophila* the overexpression of Esyt in motor neurons was associated with increased synaptic growth and enhanced resistance to synaptic depression, thereby promoting neurotransmission [303]. ESYT1 was the 2ndhighest-ranked DEP in Betz cells (LogFC: 6.06), with SEPT2 being 1st. ESYT1 staining of the primary motor cortex showed little immunoreactivity, with the only positive staining being pyramidal neurons in L3 and the Betz cells in L5 - Figure 5.20. Although staining was faint, within L5, the cytoplasmic staining of ESYT1 made Betz cells easily identifiable; in some cases, ESYT1 extended to dendrites where Betz cell projections could be seen. Within L3, pyramidal numbers were also immunoreactive, albeit with lowered expression. The presumed apical dendrites of the Betz cells could be seen in the superficial layers, but further work is needed to confirm their origin. Per the LCM proteomic data, Purkinje cells showed very little expression of ESYT1 - Figure D.19. However, surprisingly, DNNs contained a striking abundance of ESYT1. The high expression of ESYT1 in DNNs show that this is not a "Betz cell-specific" biomarker, however, in the context of studies investigating the primary motor cortex, ESYT1 may be used (either solely, or in conjunction with other proteins) as a robust biomarker, for immuno-labelling the Betz cell.

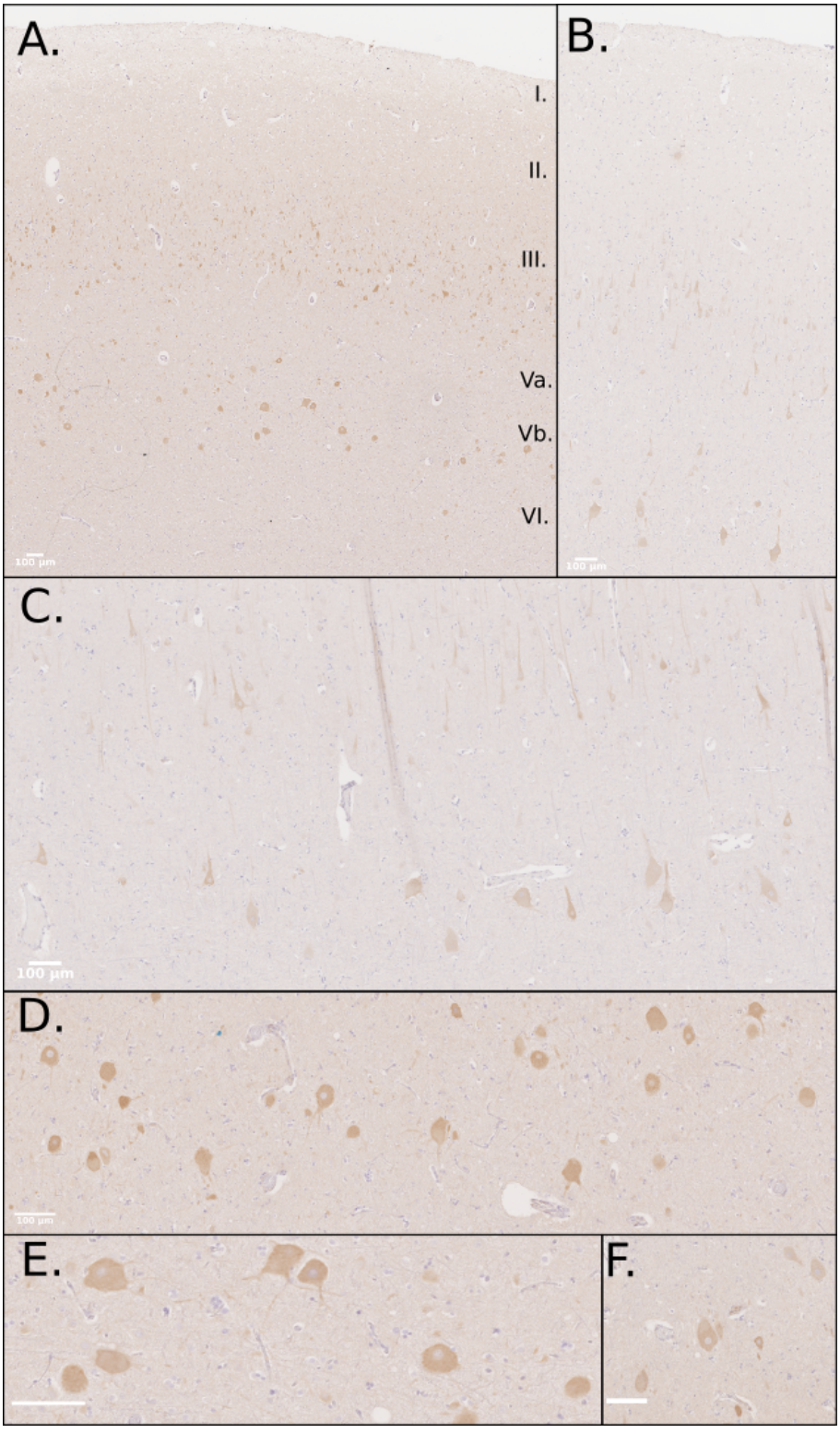


Figure 5.20: ESYT1 - Validating Proteomic Data and Quantitative IHC. Anti-ESYT1 IHC of the primary motor cortex showed that ESYT1 was localised to Betz cell soma and L3 pyramidal neurons. Minimal immunoreactivity was seen in non-Betz L5 neurons, with ESYT1 showing high specificity to Betz cells. Scalebar: 100µm.

5.3.6 Comparative LCM Proteomics of the Betz Cell, L5 Motor Neuron, Purkinje Cell, and Cerebellar DNN

The above immunohistochemical analysis highlighted the similarity between the Betz cell and giant neurons of the cerebellar Dentate Nucleus. Conceptional, this is unsurprising as both neurons represent a class of large excitatory glutamatergic neurons. While the Purkinje cell provides a robust benchmark for LCM proteomics; to test if the RWS approach was capable of discerning proteomic differences between neurons of similar biochemical properties the Betz cell was compared to the large cerebellar neurons of the Dentate Nucleus (DNNs). Moreover, comparing neurons with similar neurophysiology may result in an enrichment of Betz-specific findings rather than differences most attributed to glutamatergic/GABAergic differences. As the non-Betz cell L5 neurons are often too small for individual microdissection, it was decided to microdissect areas of the L5 microenvironment containing neurons that did not contain typical Betz cell morphology, in an attempt to try and find potential differences between these two types of motor neurons.

As shown in Figure 5.21, on average, 3,547 proteins were identified in the Betz cell samples, 3,628 proteins in the Purkinje cell samples, 3,526 proteins in the DNNs, and a slight decrease in non-Betz L5 neurons with an average of 2,827 being identified. Overall, there was a high correlation between all samples (Figure D.8). The lowest correlation with the Betz cell samples was the non-Betz cell L5 neuron samples (0.752 PCC) whereas, Purkinje cell samples and microdissected DNNs showed a high correlation with a Pearson Correlation Coefficient of 0.837 and 0.889, respectively. As expected, the vast majority of proteins were shared within the 3 replicates of each cell-type (Betz: 3,280 proteins, 87.6%; non-Betz L5: 2,374, 74.3%; Purkinje: 3,350, 87.5%; Dentate Nucleus: 3,192, 84.8%). In total, 3,009 (75.9%) proteins were shared across all samples, or 3,554 (89.6%) when excluding non-Betz motor neurons - see Figure 5.22.

Promisingly, Principal Component Analysis was able to separate the individual samples based on the protein expression (Figure D.9); highlighting the robust

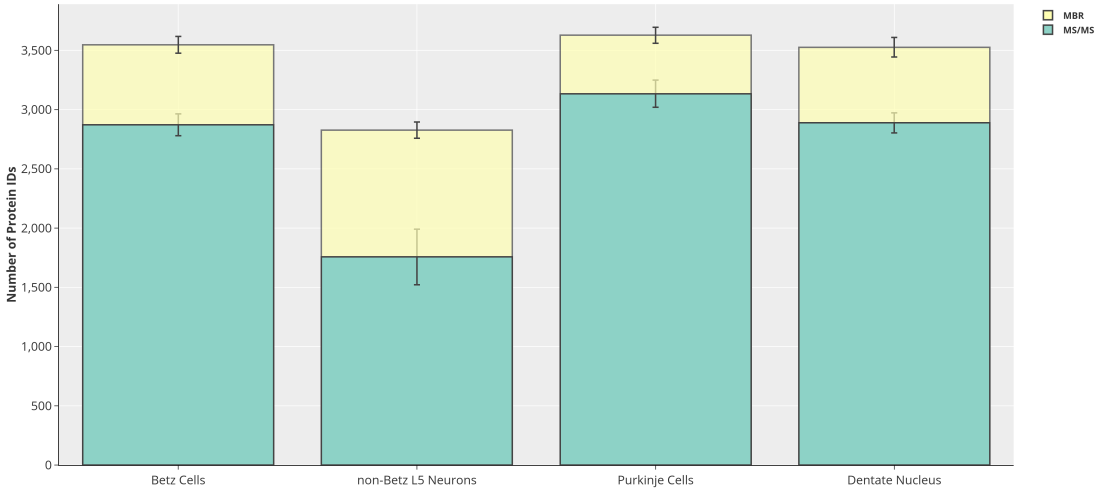


Figure 5.21: Betz, Purkinje, L5, and DNNs - No. of Identifiable Proteins. Area microdissected: $\approx$ 150,000µm^2. Turquoise bars: by MS/MS, Cream Bars: by MBR. Vertical bars denote standard deviation.

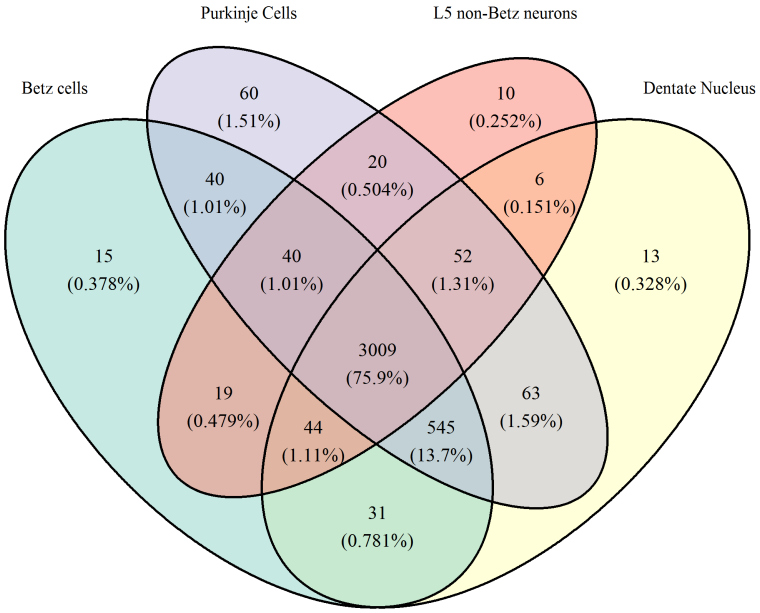


Figure 5.22: Venn Diagram of Proteins Identified in the 4-Cell Type Comparison. All cell-types (Betz cell, L5 neurons, DNNs, and Purkinje cells) contained a large overlap of identifiable proteins. Betz cells, DNNs, and Purkinje cells contained the highest overlap (89.6%).

nature of the LCM-proteomic approach to achieve the depth of proteins needed to separate both dissimilar and closely related neurons. PCA also highlighted that the L5 samples substantially differed in their proteomic expression compared to the other 3 samples types. As with previously generated data, the proteins most responsible for Purkinje cell differentiation were known Purkinje cell biomarkers, i.e. ITPR1, CALB1, PPKCG. ARHGAP26, and DGKZ. Several proteins were also found to be disproportionately responsible for distinguishing Betz cells from the other samples. Auspiciously, ESYT1 was one of these proteins, whose previously expression in Betz cells increases confidence in the result of this PCA. Strangely, MX1 (Interferon-induced GTP-binding protein, Mx1), a GTP-metabolizing protein traditionally involved in the cellular antiviral response, outperformed ESYT1 in the dataset; further work needs to be performed to validate this finding. Limited studies have indicated that MX1 may regulate calcium channel activity of transient receptor potential cation channels (TRPCs), these channels are heavily expressed in corticolimbic brain areas which open when intracellular calcium stores are depleted [304, 305]. Figure D.9 also highlights several proteins that contribute most to the delimitation of the L5 samples from the Betz, Purkinje, and DNNs. These proteins (MOG, PLP1, MBP, BCAS1, CLDN11, and GAP43) are either known components of myelinated sheaths and/or are known to be abundant in the neuropil of the cortex. These data indicate that the L5 neuron samples are most likely heavily contaminated with neuropil protein, likely due to the alternative approach used to microdissect L5 neurons.

Hierarchical clustering using the Ward2 method was used to compare the proteomic landscape of each cell. This resulted in a dendrogram that clustered the 4 cell samples into their respective sample-specific clusters - see Figure D.10. This dendrogram showed that the overall protein expression between Betz and DNNs were closely related, echoing the results seen in the previous immunohistochemical staining experiments. Purkinje cells were the next closely aligned to Betz cell samples, with the L5 samples being the least related to the 3 other cell-types - likely reflecting the excess neuropil contamination seen in the above PCA. The resulting

hierarchical clustering heatmap could be simplified into 10 discrete clusters[4]. The 9 clusters can be seen in Figure 5.23 and showed interesting protein profiles. Cluster 1 highlighted proteins that were highly expressed in the DNNs, moderately expressed in the L5 samples, and low expression in Purkinje and Betz cells. Cluster 2 consisted of proteins highly expressed in Purkinje cells and low expression in the remaining samples. Cluster 3 contained proteins high in Betz and DNNs, and lower in the other 2 sample types. Cluster 4 highlighted proteins that were solely abundant in Betz cells, while Cluster 5 contained proteins low in Purkinje cells but moderately high in all other samples studied. Cluster 6 contained proteins selectively higher in DNNs, while cluster 7 showed proteins selectively lower in Betz cells. Finally, clusters 8 and 9 highlighted proteins that were uniform among Betz, Purkinje, and Dentate Nucleus cells, but differentially expressed in L5 samples.

As cluster 4 contained the proteins which were highest among Betz cell samples it was selected for further analysis. Cluster 4 contained 116 DEPs; protein-protein interactions (PPI) between the 116 proteins identified via STRING [306] found that 73.27% of DEPs interacted with at least one other protein in cluster 4, suggesting underlying related biological functions between them - see Figure D.11. STRING Functional network analysis showed an enrichment of proteins localised to the recycling endosome (FDR: 0.0084) and enrichment of proteins with the associated keywords: microtubule (FDR: 0.0311) and Coiled-coil (FDR: 0.0221) within cluster 4.

Promisingly, when performing ORA on cluster 4, neurofilament proteins (NEFL, NEFM, NEFH) were among the 116 proteins in this cluster. A ranked list of all proteins found in cluster 4 can be found in Appendix D.4. Other enriched GO terms include Organelle Localisation, Cytosolic Transport, Regulation of Intracellular Transport, and Response to Peptide Hormones. Additionally, there was an abundance of proteins localised to the Endoplasmic Reticulum Membrane, Microtubule, and Recycling Endosome - Figure 5.24.

[4]One cluster only contained 2 proteins, and was omitted from further analysis.

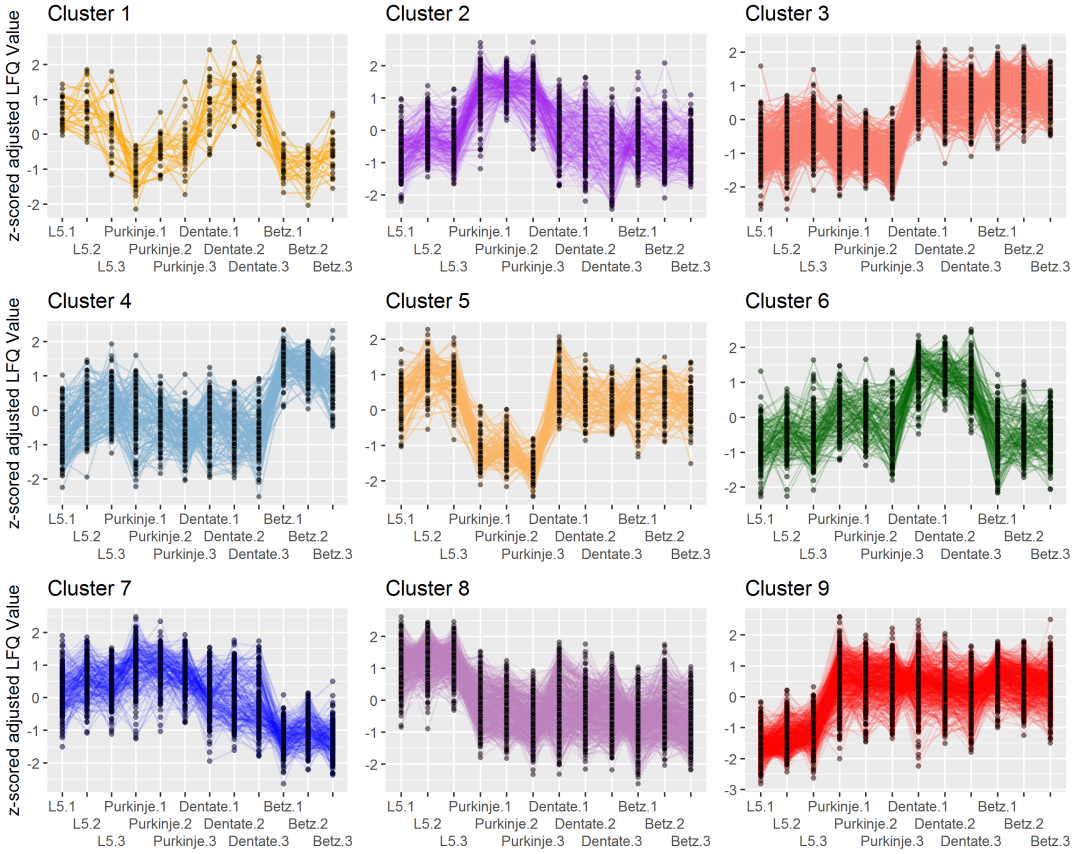


Figure 5.23: Hierarchical Clustering Revealed Ten Unique Clusters.
Ten unique clusters were identified from unbiased hierarchical clustering. Cluster 1: highlighted proteins that were highly expressed in the DNNs, moderately expressed in the L5 samples, and had low expression in Purkinje and Betz cells. Cluster 2: highly expressed in Purkinje cells and low expression in the remaining samples. Cluster 3: contained proteins high in Betz and DNNs, and lower in the other 2 sample types. Cluster 4: highlighted proteins abundant in Betz cells. Cluster 5: contained proteins low in Purkinje cells but moderately high in all other samples studied. Cluster 6: contained proteins selectively higher in DNNs, while cluster 7 showed proteins selectively lower in Betz cells. Clusters 8 and 9 highlighted proteins that were uniform among Betz, Purkinje, and Dentate Nucleus cells, but differentially expressed in L5 samples. Cluster 10 was omitted as it contained only 2 proteins.

To investigate the underlying differences that caused the L5 samples to differ substantially from the other samples in this data-set, clusters 8 and 9 were subjected to ORA. From ORA, it was clear that Cluster 8 had a plethora of GO terms relating to extracellular properties. These included terms relating to cortical cytoskeleton, cell adhesion, plasma membrane, cell migration, tissue homeostasis, regulation of angiogenesis, and immune system function. The proteins in this cluster also had

enriched GOCC terms indicating that were localised, at least in part, to extracellular organelles, extracellular matrix, and extracellular vesicles - see figure D.21. Moreover, there was an increase in proteins found in the myelin sheath, paranode regions of an axon, synaptic membrane, and glutamatergic synapses. When compared against the Human Protein Atlas there was a large enrichment with proteins that are typically expressed highly (adj. p value: 1.16e-11), or moderately (adj. p value: 5.8e-5) in the cerebral cortex neuropil. Whereby, proteins in cluster 9 had strong associations with molecular functions that relate to intercellular processes, i.e. gene expression, RNA processing, peptide and macromolecule synthesis, with these proteins typically found within ribonucleoprotein complexes, nucleoplasm, and cytosolic ribosomes. When compared to the HPA, cluster 9 also showed enrichment in proteins that the HPA have reported are know to be moderately expressed in neuronal cells.

Next, proteins in Cluster 3 were investigated to look for unique similarities between the Betz cell and DNN. It was found that this cluster contained a large enrichment of proteins involved in mitochondrial protein import (Fold enrichment: 4.54, FDR: 1.02e-02) and the processing of SMDT1 (Fold enrichment: 8.07, FDR: 2.09e-02), along with an increase in proteins found within mitochondrial protein complexes (Fold enrichment: 3.20, FDR: 5.40e-03). Cluster 5 did not contain any over-represented GO terms and Clathrin-mediated endocytosis (Fold enrichment: 5.18, FDR: 2.74e-02) was the only enriched term in Cluster 2. Interestingly, Cluster 6 (Dentate Nucleus cluster) contained a large enrichment of oxidoreductases (Fold enrichment: 3.71, FDR: 7.14e-06), and large fold increase in proteins involved in ATP synthesis (Fold enrichment: 12.5, FDR: 1.60e-04), and a large increase in proteins found in the mitochondrial respiratory chain complex (Fold enrichment: 7.23, FDR: 1.09e-02).

Finally, to look for specific and potentially subtle differences between the Betz cells and the other samples, the individual abundance of each protein/protein group were compared. Although the revealed proteome of the DNNs was found to be highly similar to Betz cells, there remained some striking differences. A Welch

t-test (FDR threshold: 0.05, S0: 0.1, 250 randomizations) revealed that 770 DEPs between these two cell-types, 330 DEPs were enriched in Betz cells and 440 DEPs were enriched in DNNs, however, only 59 DEPs (7.99%) had a LogFC of more than 2. CRYM, PRUNE2, GAP43, CAPRIN2, and MAP2 were the most enriched among Betz cells. While, CALB1, PRKCG, ITPR1, PVALB, and SPTB had the highest fold changes in the DNNs. 1D annotation enrichment was performed to find trends in the expression values of individual proteins. In total, 243 GO terms were enriched; and reduced to 116 GO terms using REVIGO, a list of which can be seen in figure D.14 & figure D.15. The results of the 1D annotation enrichment showed that proteins intrinsic to or proteins targeting the endoplasmic reticulum were more abundant in the Betz cell. Additionally, the vast majority of proteins found in the large ribosomal subunit (89.8%) in this comparison were more abundant in the Betz cell than in DNNs. A large difference was also found for proteins involved in Protein Kinase A binding in addition to proteins involved in the tRNA metabolic process and/or aminoacyl-tRNA ligase activity - see Figure 5.27.

Next, the Purkinje cell samples were compared to the Betz cells in the same manner; 1,248 proteins were significantly differentially between these two cell-types with 134 proteins (11%) having a LogFC of more than 2. There were 597 DEPs enriched in Betz cells and 651 DEPs enriched in Purkinje cells. The aforementioned MX1 and ESYT1 were the highest-ranked Betz DEPs when compared to Purkinje cells, followed by PRUNE2, CPT1A, NACAD, and CAMK2G. The proteins with the largest differential expression in Purkinje cells mimicked those of our previous Betz-Purkinje comparison, namely, PRKCG, ITPR1, CALB1, DGKZ, and ARHGAP26. 1D annotation enrichment identified 302 significantly enriched pathways between the two cell-types, which were reduced to 52 terms with REVIGO and can be found in Figure D.12 & Figure D.13. As when compared to the Dentate Nucleus, these data showed that there was a pronounced increase in protein abundance in Betz cells relating to protein kinase A and C, aminoacyl-tRNA biosynthesis, and the large ribosomal subunit. Similar to the previous work comparing Betz cells to Purkinje cells, there was also a prevalence of mitochondrial proteins in the

Betz cell. Additionally, this comparison showed a strong tendency for proteins localised to the lysosomal lumen to be expressed in higher amounts in Betz cells. A smaller, yet still significant tendency was also seen in Betz cells concerning proteins involved in lipid transport - see Figure 5.28. Despite the large differences from other cell-types and the accumulation of neuropil proteins, the L5 samples were also compared to Betz cells with the same methodology. As expected, there were large differences between the two samples with 1,628 (71.95%) DEPs and 356 (21.93%) of these having a LogFC of more than 2. There were 958 Betz DEPs and 670 L5 DEPs. CAPRIN2, MX1, PRUNE2, GOLGB1, and CLUH were the highest-ranked DEPs in the Betz cell sample, with PRKCH, CLDN11, CADM2, SYNGR3, and PLP1 being the highest DEPs in the L5 samples. A large number of GO terms (724) were found and reduced to 165 with REVIGO, which can be found in figure D.16 & figure D.17. Similar to the above comparisons, Betz cells contained an abundance of proteins relating to aminoacyl-tRNA synthetase, cytosolic ribosomes, lipid transport, and proteins targeting the endoplasmic reticulum; in addition to proteins found in the COPI vesicle coat - see Figure 5.29.

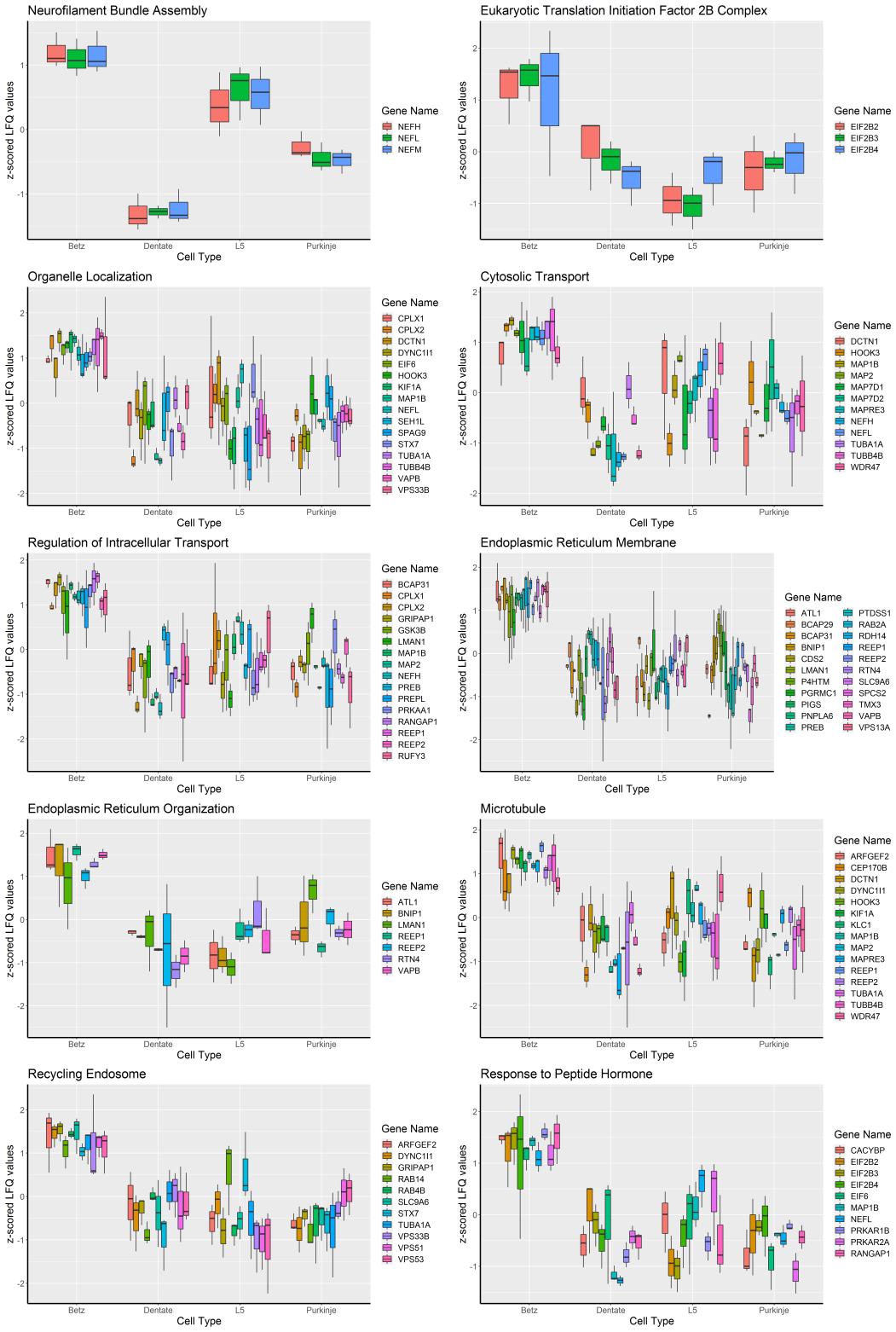


Figure 5.24: Enriched GO Terms in the Betz Cell Cluster (Cluster 4). Neurofilament, endoplasmic reticulum membrane and organisation, recycling endosome, and macromolecule transport were all terms significantly enriched in Betz cells compared to other cell-types investigated.

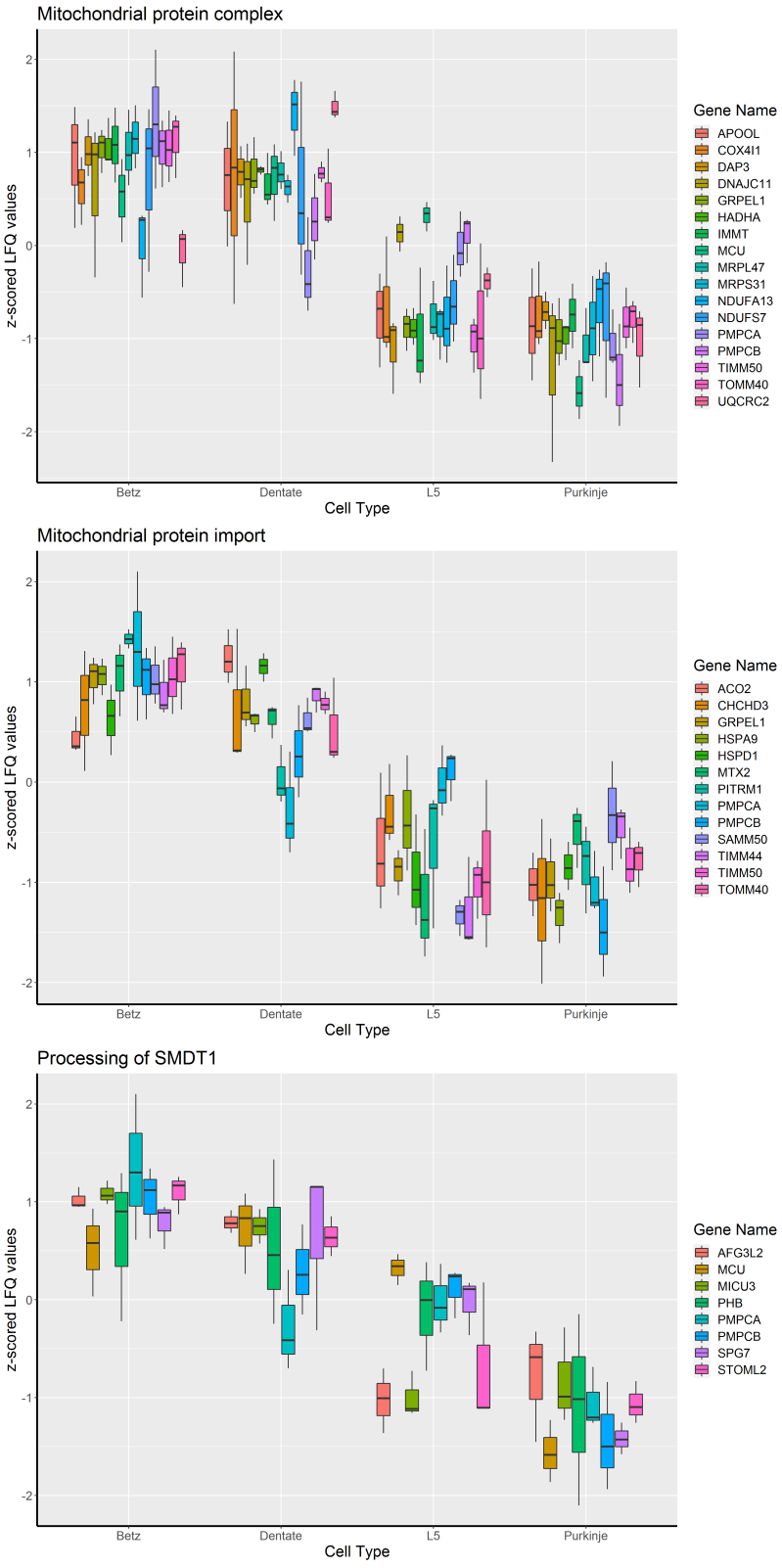


Figure 5.25: Enriched GO Terms in the Betz Cell/DNNs Cluster (Cluster 3). The only significant terms in this cluster related to mitochondrial protein import and protein complexes within mitochondria.

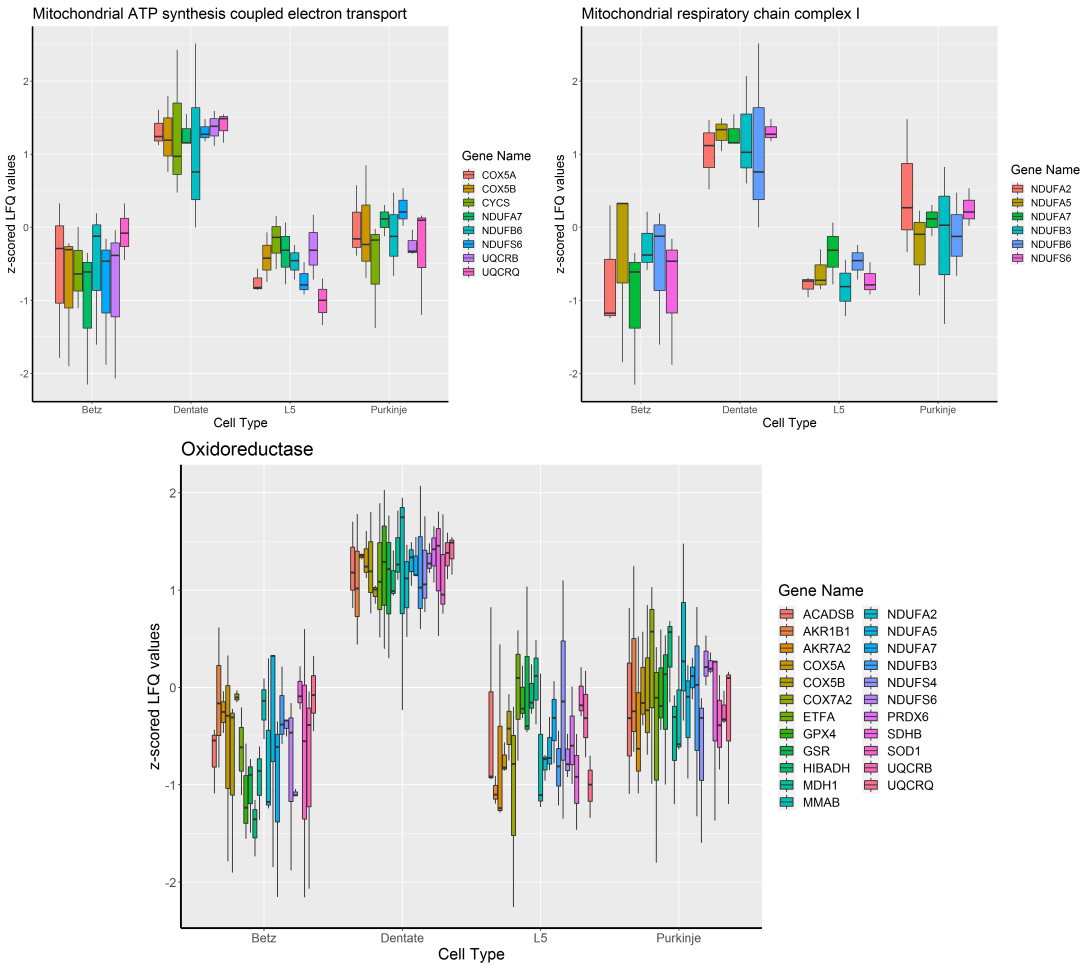


Figure 5.26: Enriched GO Terms in the DNNs Cluster (Cluster 6). The only significant terms in this cluster related to ATP synthesis, the mitochondrial electron transport, and oxidoreductase reactions.

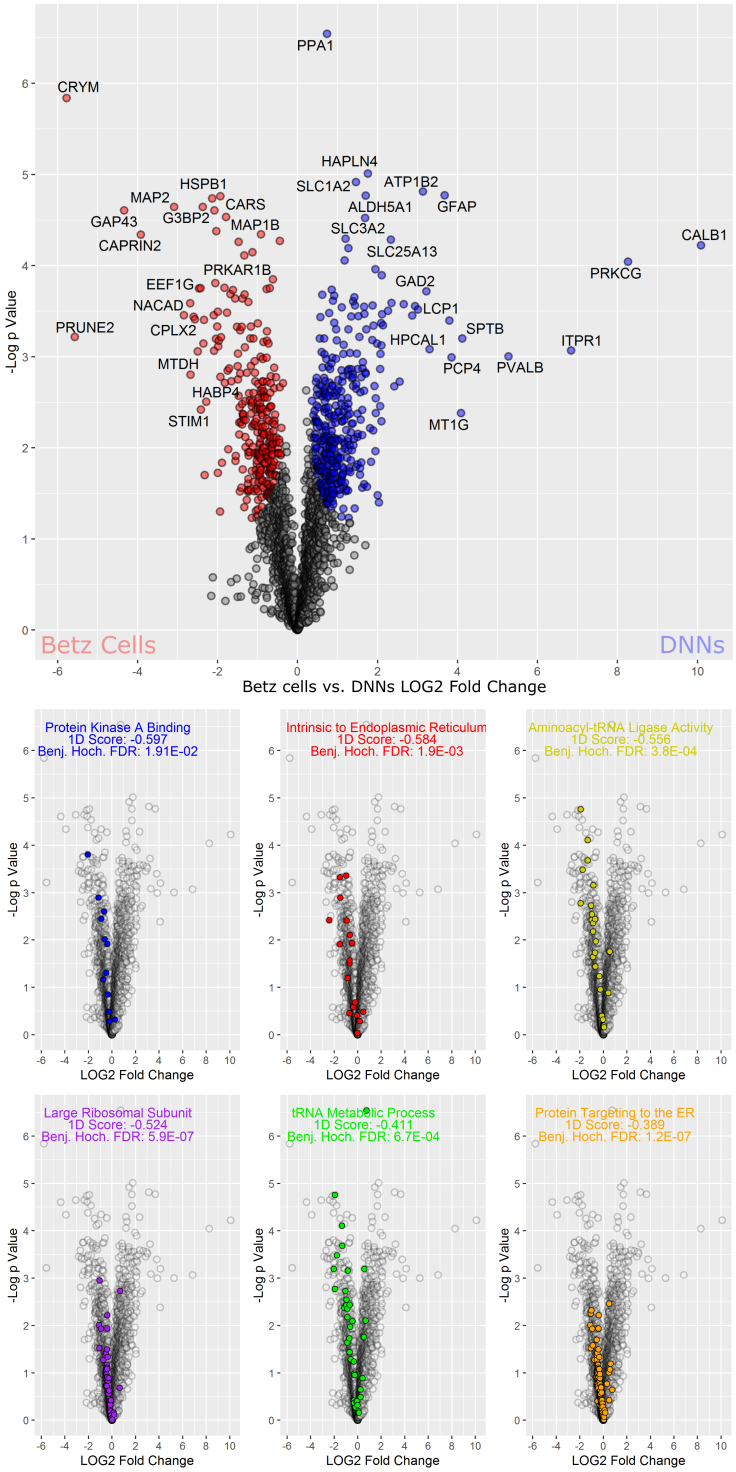


Figure 5.27: 1D Annotated Volcano plot - Betz Cells vs. DNNs.
Several proteins and associated pathways were found to be enriched in Betz cells when compared to DNNs. Among the most enriched pathways were tRNA metabolism and activity, protein kinase A binding, and transport of proteins to the endoplasmic reticulum. Each point represents one protein. Coloured points are DEPs that passed predetermined alpha thresholds (0.05). Red: Betz DEPs; Blue: DNN DEPs.

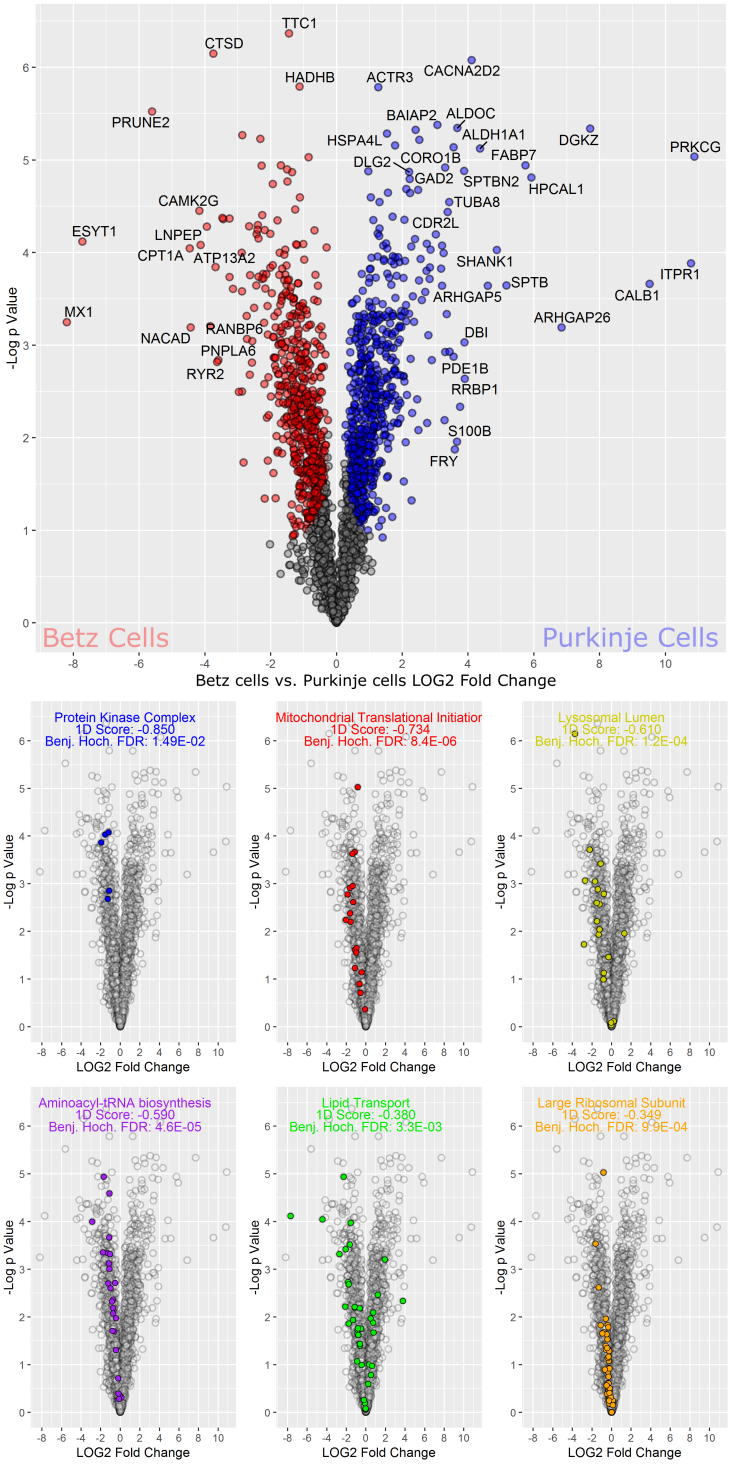


Figure 5.28: 1D Annotated Volcano Plot - Betz Cells vs. Purkinje Cells. Several proteins and associated pathways were found to be enriched in Betz cells when compared to Purkinje cells. Among the most enriched pathways were aminoacyl-tRNA synthesis, protein kinase activity, lysosomal lumen, lipid transport, and echoing the Betz vs. DNNs experiment: large ribosomal subunit. Each point represents one protein. Coloured points are DEPs that passed predetermined alpha thresholds (0.05). Red: Betz DEPs; Blue: Purkinje DEPs.

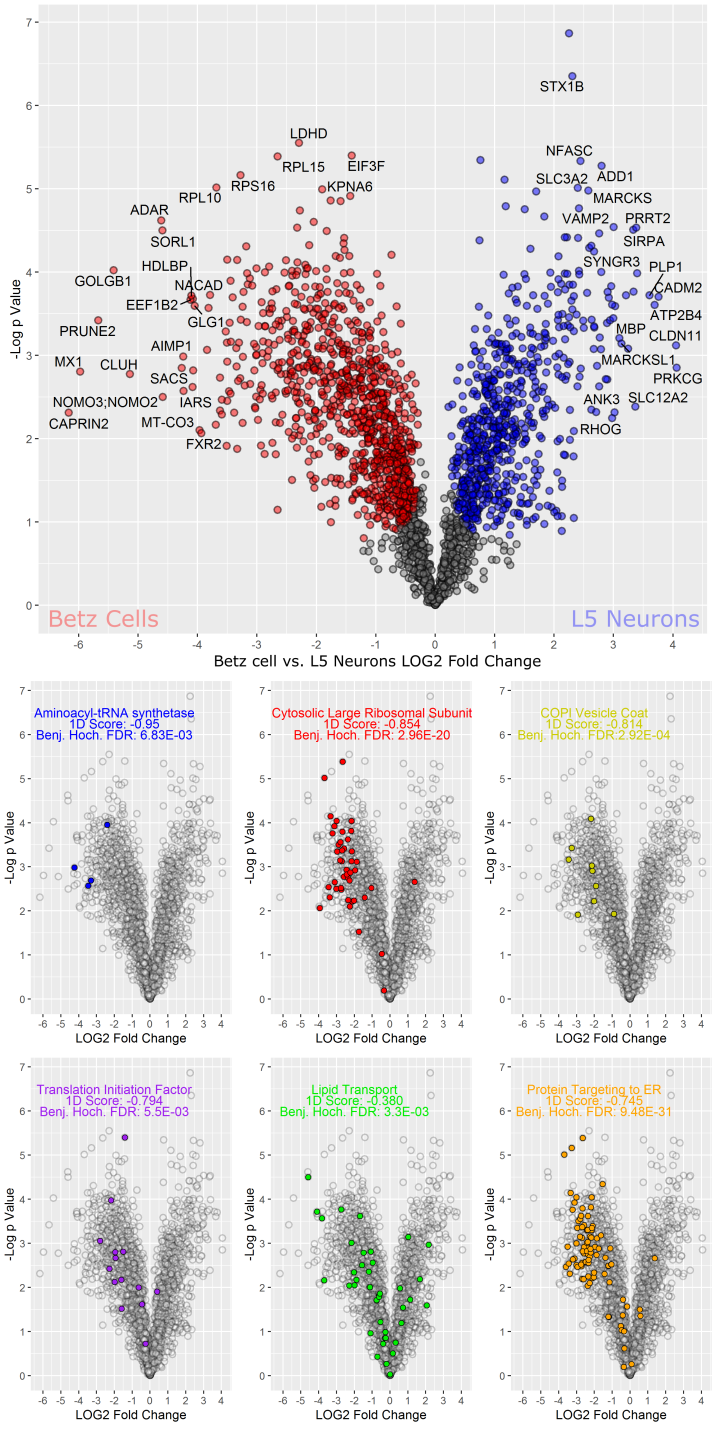


Figure 5.29: 1D Annotated Volcano plot - Betz Cells vs. L5 Neurons. Several proteins and associated pathways were found to be enriched in Betz cells when compared to L5 neurons. Among the most enriched pathways were aminoacyl-tRNA synthesis and tRNA acitivty, lipid transport, COPI vesicle coat, transport of proteins to the endoplasmic reticulum, and large ribosomal subunit. Each point represents one protein. Coloured points are DEPs that passed predetermined alpha thresholds (0.05). Red: Betz DEPs; Blue: L5 neurons DEPs.

5.3.7 The Search for Preliminary Betz Cell Biomarkers

For the first time, rich proteomic data has now been generated on the human Betz cell; moreover, the proteome of these cells has been comparatively analysed to both distant and closely related neurons; highlighting both the large and subtle differences between them. Similar to 5.3.5, promising DEPs were targeted by IHC, in an attempt to identify potential Betz cell biomarkers, to either aid histological investigations or help Betz cells to be compatible with *'omics* technologies.

CRYM -

µ-crystallin (CRYM) was highlighted as a DEP in both DNNs and Purkinje cell comparisons. CRYM is a protein believed to be a key regulator of the thyroid hormone T3, although its exact function is still unknown. It is thought to play a role in cell survival, lysine metabolism, neurotransmission, and neuroprotection [307]. CRYM was selected because the literature reports that CRYM may be selectively expressed by large neurons across the human neocortex. However, contradictory to the above proteomic data, previous observations have suggested that Betz cells are only weakly positive for CRYM [307]. Hommyo et al., (2018) describe a mosaic pattern of positive and negative neurons within L5 of the primary motor cortex. Curiously, this pattern was also seen in our IHC data, with a subset of Betz cells seemingly positive for CRYM, and others negative - see Figures 5.30 and 5.31. CRYM was highly expressed in some small motor neurons within L5, along with moderate immunoreactivity in the neuropil - likely explaining why CRYM was not being highlighted as a DEP in the L5 - Betz cell comparison. Although IHC results show that CRYM would not be a suitable Betz cell biomarker, the interesting mosaic staining of assumed Betz cells[5] warrants further research into this matter.

[5]CRYM staining alone does not highlight dendritic staining, as such, it is difficult to have absolute confidence that the large motor neurons that are strongly positive for CRYM are Betz cells rather than large L5 neurons.

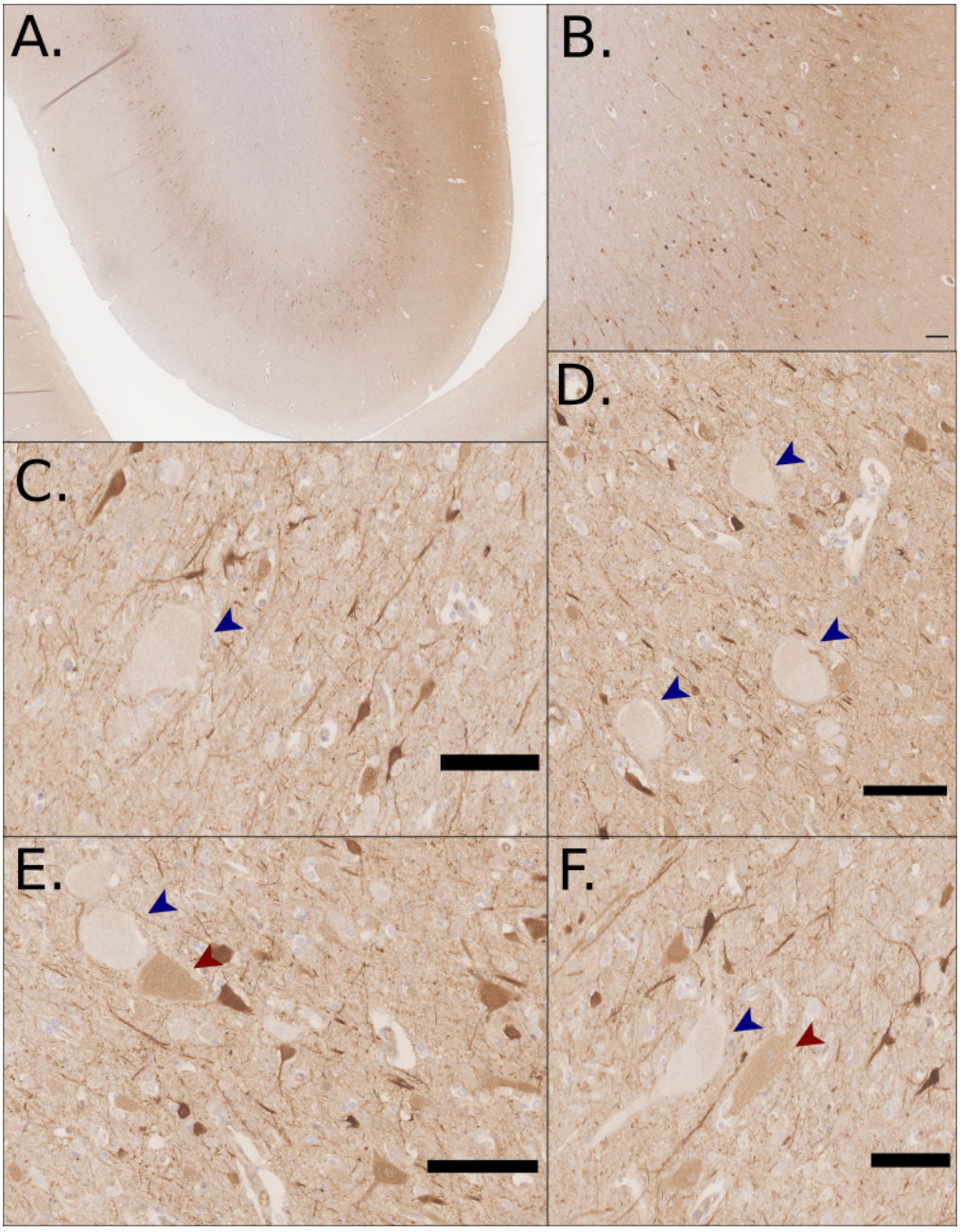


Figure 5.30: Mosaic Pattern of CRYM in Betz Cells.
A & B: CRYM IHC was enriched within L5 of the primary motor cortex. Several L5 neurons contained strong expression, however, Betz cells only faintly expressed CRYM (Blue arrows). There existed several large neurons that were immuno-reactive for CRYM (Red arrows) but the lack of morphological features makes it difficult to confirm Betz cell identity. Scalebar: 100µm.

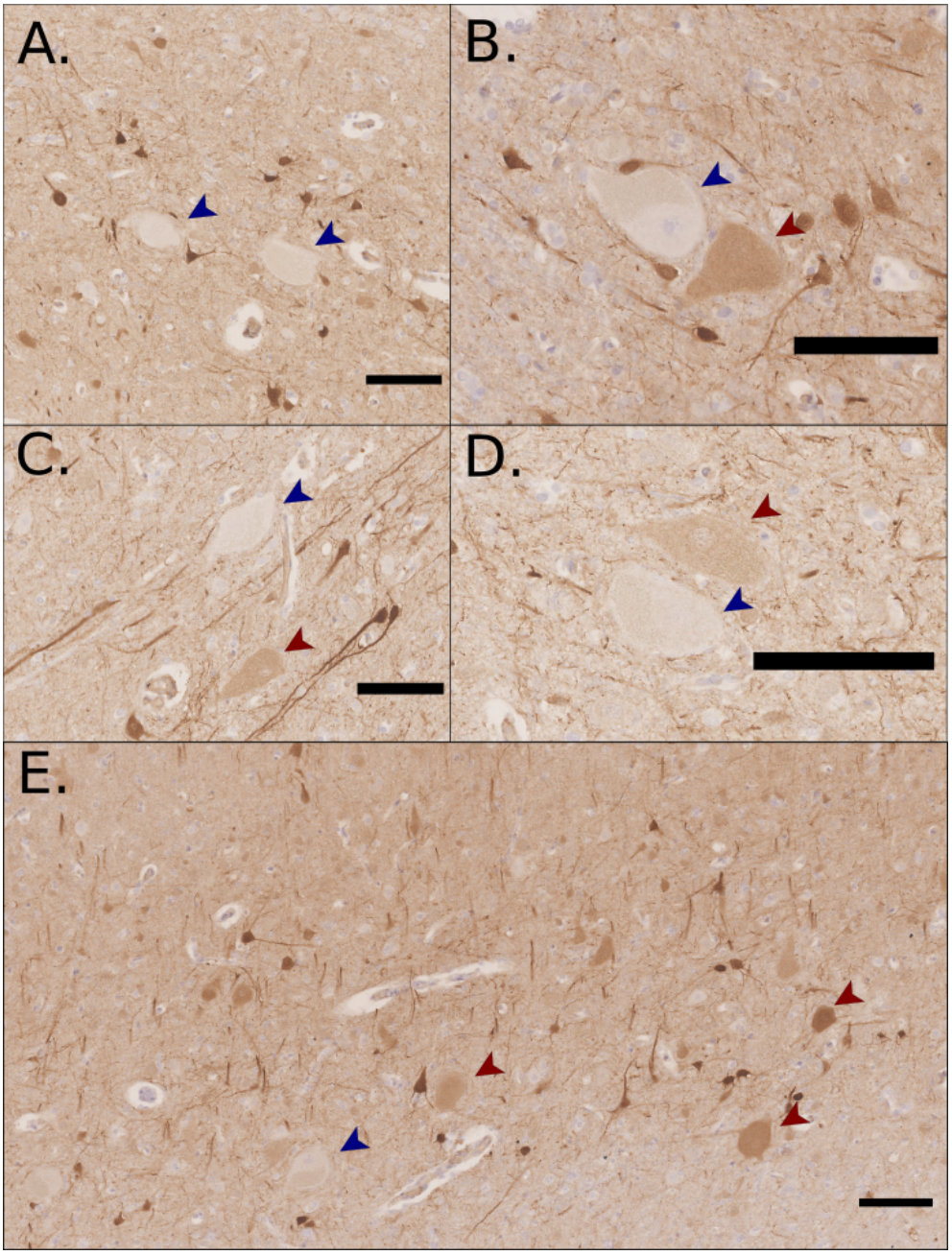


Figure 5.31: Mosaic Pattern of CRYM in Betz Cells - More Examples. As with Figure 5.30, Several non-Betz L5 neurons contained strong expression of CRYM. The majority of Betz cells were negative (Blue arrow), however, several assumed Betz cells (B, C, D, E) were positive for CRYM (Red arrow). Scalebar: 100µm.

REEP1 -

Receptor Expression-Enhancing Protein 1 (REEP1) was the 8thhighest-ranked Betz DEPs in the Betz cell cluster. REEP1 is found in both mitochondria and the ER and is believed to be critical for ER network formation, shaping, and remodelling [308–310]. It has also been suggested that REEP1 may attach the cytoskeleton to ER tubules and may play a role in the long-term maintenance of axons. REEP1 was selected as earlier data has shown a huge dependency of Betz cells to rely on ER function, along with the dual role REEP1 may have in axon and cytoskeletal maintenance. Immunoreactivity to REEP1 was mostly selective to Betz cells, see Figure 5.32 and 5.33. Macroscopically, REEP1 was restrictive to the grey matter with only some axons immunopositive in the white matter. Moreover, the Betz cell layer is the only cytology feature distinguishable at lower magnifications. At higher magnification, Betz cells were strongly positive for REEP1, which was localised mostly to the cytoplasm, with limited staining in non-Betz neurons. When in-plane, the dendrite architecture was visible, including basal and apical dendrites - see panel H 5.32. Examination of the white matter tract showed obvious selective immunoreactivity for REEP1 in some axons. It is tempting to assume that the positive axons belong to immunopositive Betz cells in the associated grey matter, but further work is required to confirm the true source of these REEP1 positive axons. However, such a protein that selectively targets Betz cell axons would be of great interest to neuroanatomy and ALS communities. While previous IHC has shown results that help in the identification of Betz cells, the preferable signal:noise ratio of REEP1 staining, the strong immunoreactivity of gigantopyramidal neurons in the primary motor cortex, and the ability to visualise both apical and somatic dendrites may make REEP1 a robust biomarker for the Betz cell marker.

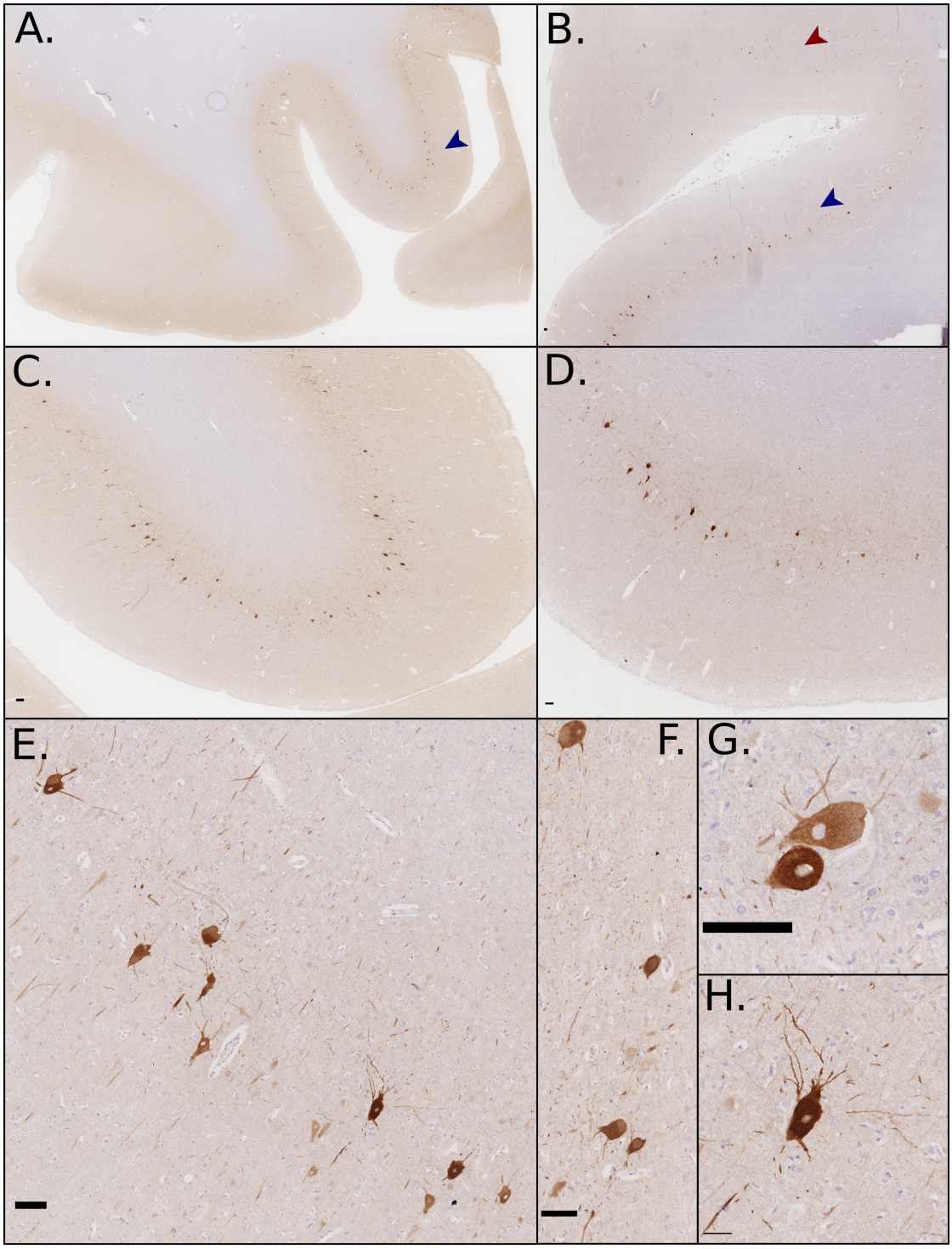


Figure 5.32: REEP1 Selectively Stains for Betz Cells.
A: At low magnification, REEP1 IHC only stains Betz cells (Blue arrow). B: The boundary of M1 and S1 (Red arrow) can be clearly delimited. The Betz cell layer can be easily identified with REEP1 (C, D, E). F-H: Betz cells have a high expression of REEP1. H: Importantly, REEP1 IHC can identify the apical dendrite and basilar array of Betz cells, making identification easy to validate. Scalebar: 100µm.

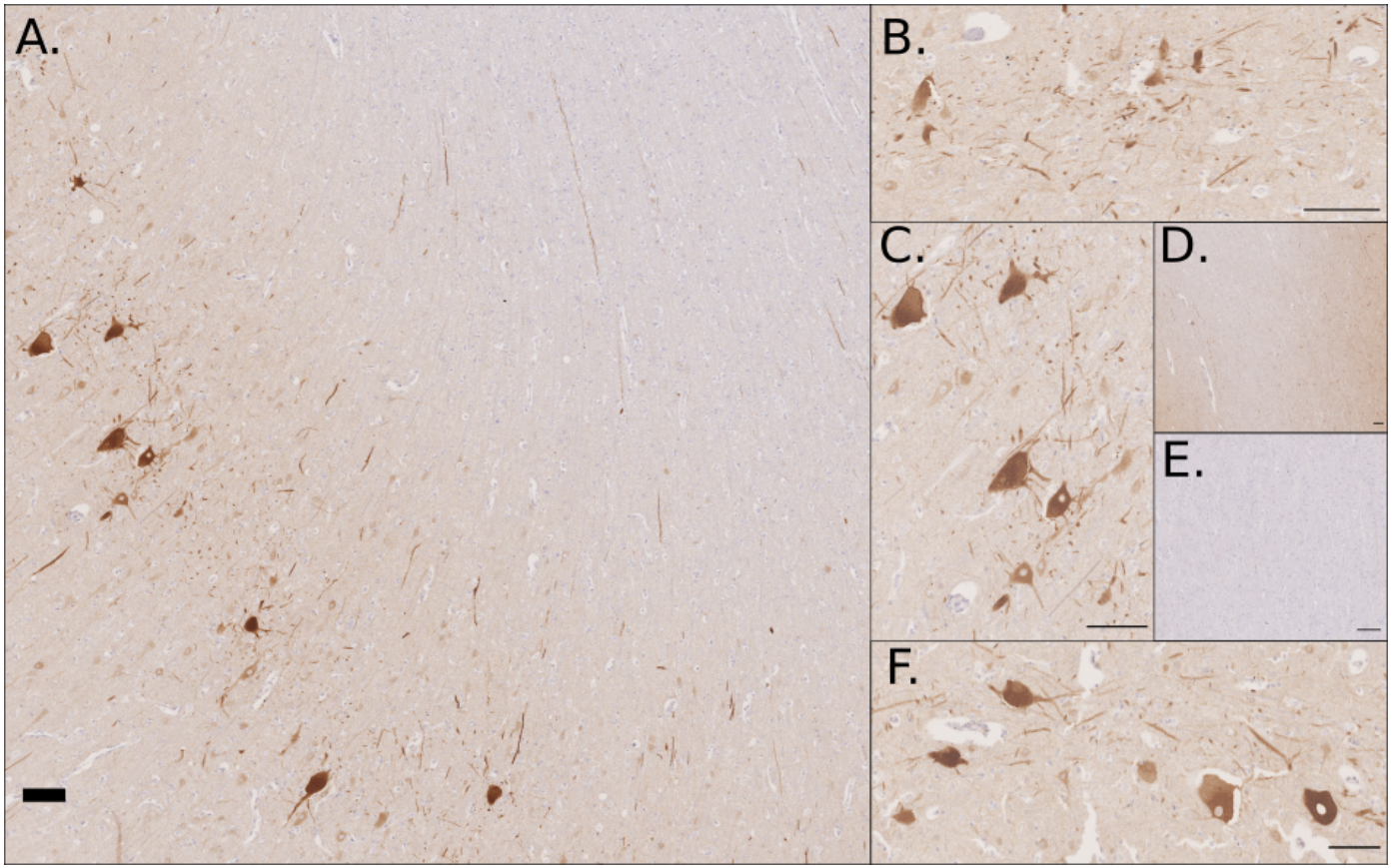


Figure 5.33: REEP1 Selectively Stains for Betz Cells Cont.
A: REEP1 stains Betz cell soma, the dense network of Betz cell basilar dendrites, and immuno-reactivity of axons - presumed to belong to Betz cells - throughout the white matter of the primary motor cortex were observed. B: Another example of Betz cell dendritic staining throughout L5. C & F: More examples of Betz cell soma being selectively targeted by REEP1. The white matter of the primary motor cortex (D) showed a small subset of axons being positive for REEP1, while no staining was seen in the white matter tract of the adjacent somatosensory cortex (E). Scalebar: 100µm.

KIF1A -

Kinesin Family Member 1A (KIF1A) is an anterograde motor protein that transports organelles along axonal microtubules in addition to transporting synaptic vesicles [311]. Building upon the conclusions drawn from the proteomics of the Betz cell, it is assumed that the long axon of the Betz cell and high neurotransmitter/energy demands would have a large dependency on cellular motor proteins. In addition to this, KIF1A was among the top 116 DEPs in the Betz cell cluster. Immunostaining of KIF1A showed diffuse immunoreactivity in the grey matter. Although KIF1A is involved in axonal transport, it was the soma of Betz cells that were immunopositive - Figure 5.34. Betz cell immunoreactivity was strong enough to allow clear identification despite the diffuse neuropil staining. Some non-Betz cell neurons were also positive, including neurons of the somatosensory cortex, however, the intensity of Betz cells was a clear cytology feature that could be seen both macroscopically and microscopically. Moreover, it was possible to visualise the somatodendritic morphology of the Betz cells. Despite the diffuse staining, the intensity of KIF1A within the gigantopyramidal neurons may make it a potential biomarker in Betz cells, at least in conjunction with some of the others identified.

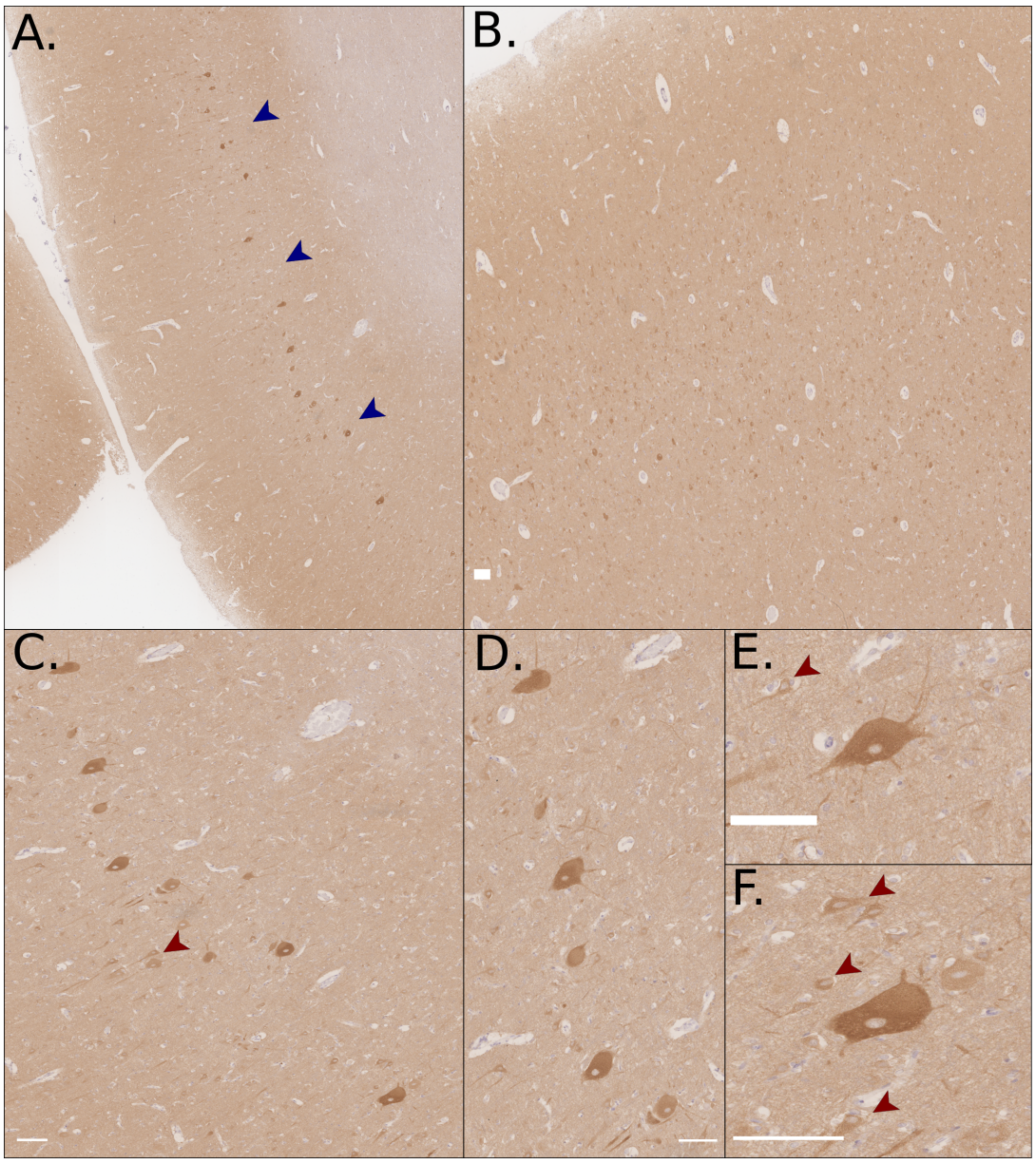


Figure 5.34: KIF1A Highlights Betz Cell Morphology.
Betz cells have high expression of KIF2A, staining the primary motor cortex (A) and Primary Somatosensory Cortex (B) revealed strong Betz cell immunoreactivity in M1; Betz cells could be seen at low magnifications (Blue arrows). Only faint cytoplasmic staining was seen in S1. C-F: Betz cells was the most immunoreactive neurons in the primary motor cortex, with only a small subset of surrounding neurons staining positively (Red arrows). Scalebar: 100µm.

CAPRIN2 -

In all cell-type comparisons; CAPRIN2 (cytoplasmic activation proliferation associated protein 2) was significantly upregulated in Betz cells. CAPRIN2 has been reported to promote the activation of the canonical Wnt signalling pathway by regulating LRP5/6 phosphorylation [312]. Wnt signalling in neurons has been shown to regulate neuronal maturation by governing dendrite formation and dendritic complexity [313]. Moreover, it has been stated that neurons with excess Wnt signalling develop longer dendrites and more complex dendritic branches [313]. As such, CAPRIN2 was selected as a potential marker for Betz cells. While not unique to Betz cells, the strong immunoreactivity of Betz cells for CAPRIN2 made their presence obvious - Figure 5.35. Furthermore, the circumferential dendrites were prominent. Immuno-staining of L5 also revealed a dense network of thick dendrites near Betz cells and may represent the thick Betz cell basilar dendritic arbor that was described in [191].

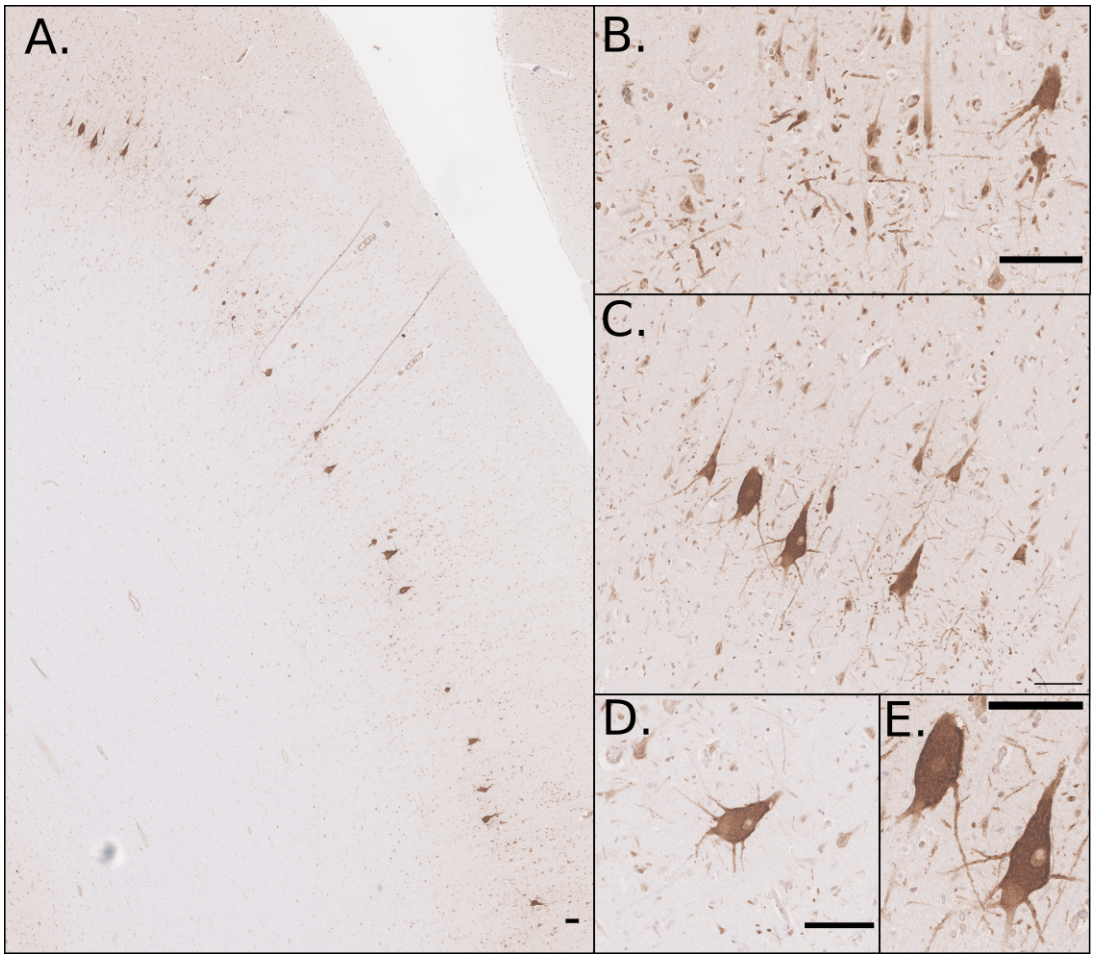


Figure 5.35: CAPRIN2 is Enriched within Betz Cells.
A: The characteristic L5 band of Betz cells could be seen at low magnification. B: Dense dendritic networks were visible in L5. C: Betz cells contained abundant amounts of CAPRIN2; non-Betz L5 neurons were also immunoreactive. Importantly, CAPRIN2 also revealed the circumferential dendrites of the Betz cells (D, E). The strong intensity of CAPRIN2 in Betz cells may make CAPRIN2 a useful marker to immunochemically enrich/target Betz cells. Scalebar: 100µm.

Transcription factors (OTX1 & HDAC1) -

Orthodenticle Homeobox 1 (OTX1) is a transcription factor involved in brain development. In the above data-set, OTX1 was enriched in all comparisons to Betz cells (Purkinje cells, DNNs, and L5 neurons). Interestingly, OTX1 has already been noted as being present in a subpopulation of L5 pyramidal neurons - see Figure 2.8. OTX1 is found in corticotectal, corticobulbar, and corticospinal neurons, and is critical in the refinement of L5 axons being guided to their appropriate subcortical targets. Some have noted that OTX1 is required for the differentiation and development of neurons expressing calcium-binding proteins [314]. Similar to KIF1A, OTX1 staining produced a diffuse staining pattern across the cortex in both the primary motor cortex and primary somatosensory cortex. Despite this, Betz cells could be identified at low magnifications; at higher magnifications the Betz cell cytoplasm and a few somatodendrites were visible - see Figure 5.36.

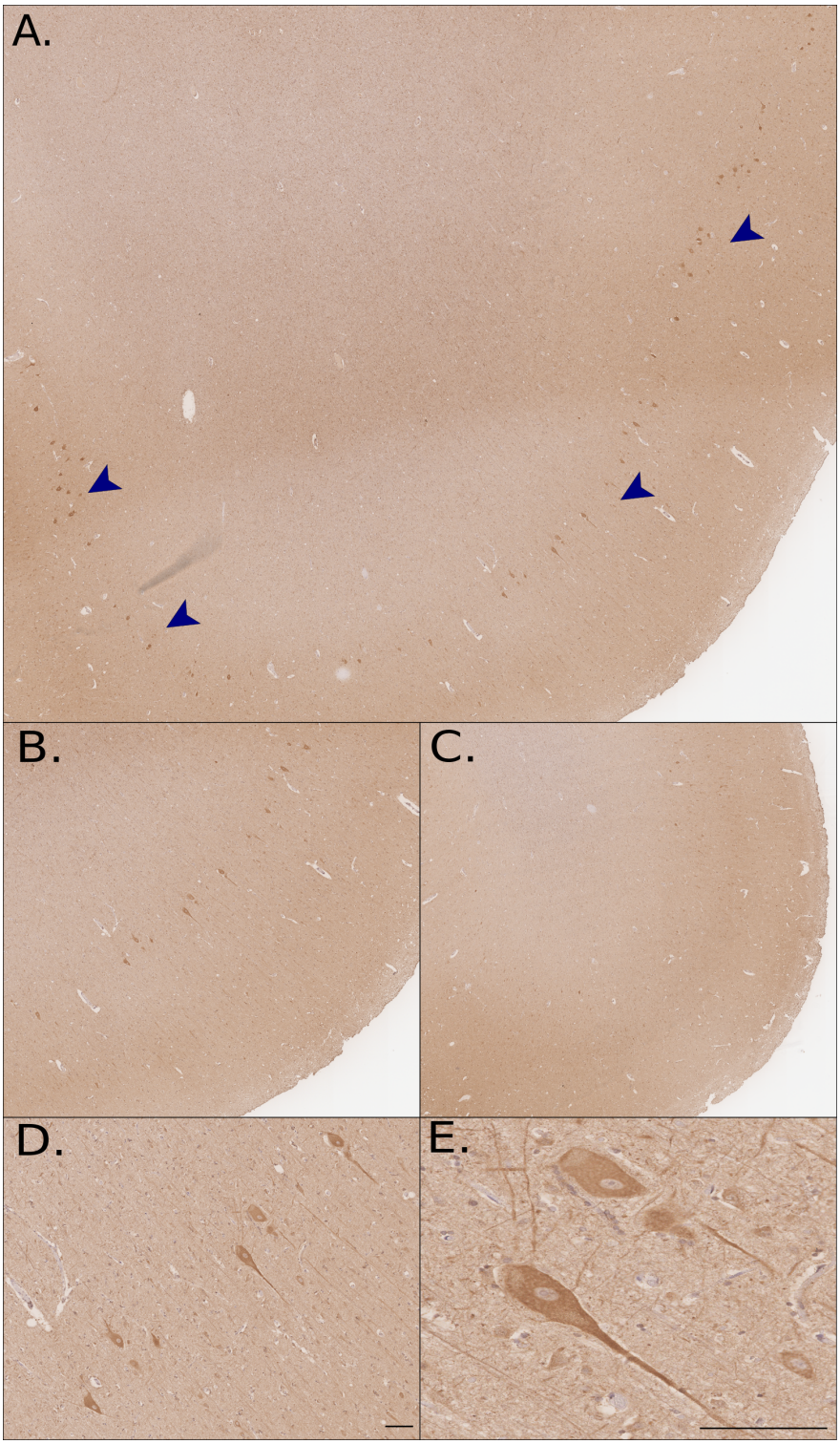


Figure 5.36: Transcripton Factor OTX1 Selectively Targets Betz Cells. OTX1 staining reveals diffuse staining across the primary motor cortex (A & B) and Primary Somatosensory Cortex (C). Blue arrows show the boundary of L5 and the associated Betz cells within them, no obvious cytology features were seen in the Primary Somatosensory Cortex. Scalebar: 100µm.

HDAC1 -

As current proteomic technologies can only identify the most abundant proteins in a cell rather than all proteins, BART (Binding Analysis for Regulation of Transcription) [315] was used in an attempt to identify potential transcription factors based on the 116 DEPs within the Betz cell cluster. In total, 919 potential TFs were identified, with the highest-ranking being Histone Deacetylase 1 (HDAC1) (p value: 5.227e-05). Immunostaining for HDAC1 did highlight Betz cells within the primary motor cortex. Only at high magnifications was it possible to discern immunoreactivity in other cell-types. HDAC1 appears to be localised to the nuclear membrane, cytoplasm, and plasma membrane of Betz cells. Although immunoreactivity of HDAC1 was also found in surrounding cell-types, this staining was mostly nuclear, allowing the Betz cells to be identified with ease. Similar to REEP1 staining, HDAC1 IHC of the motor cortex also revealed selective axonal staining and more work will be needed to confirm if these axons belong to Betz cells and if it is possible to track their projection through the cerebrum in thicker tissue sections.

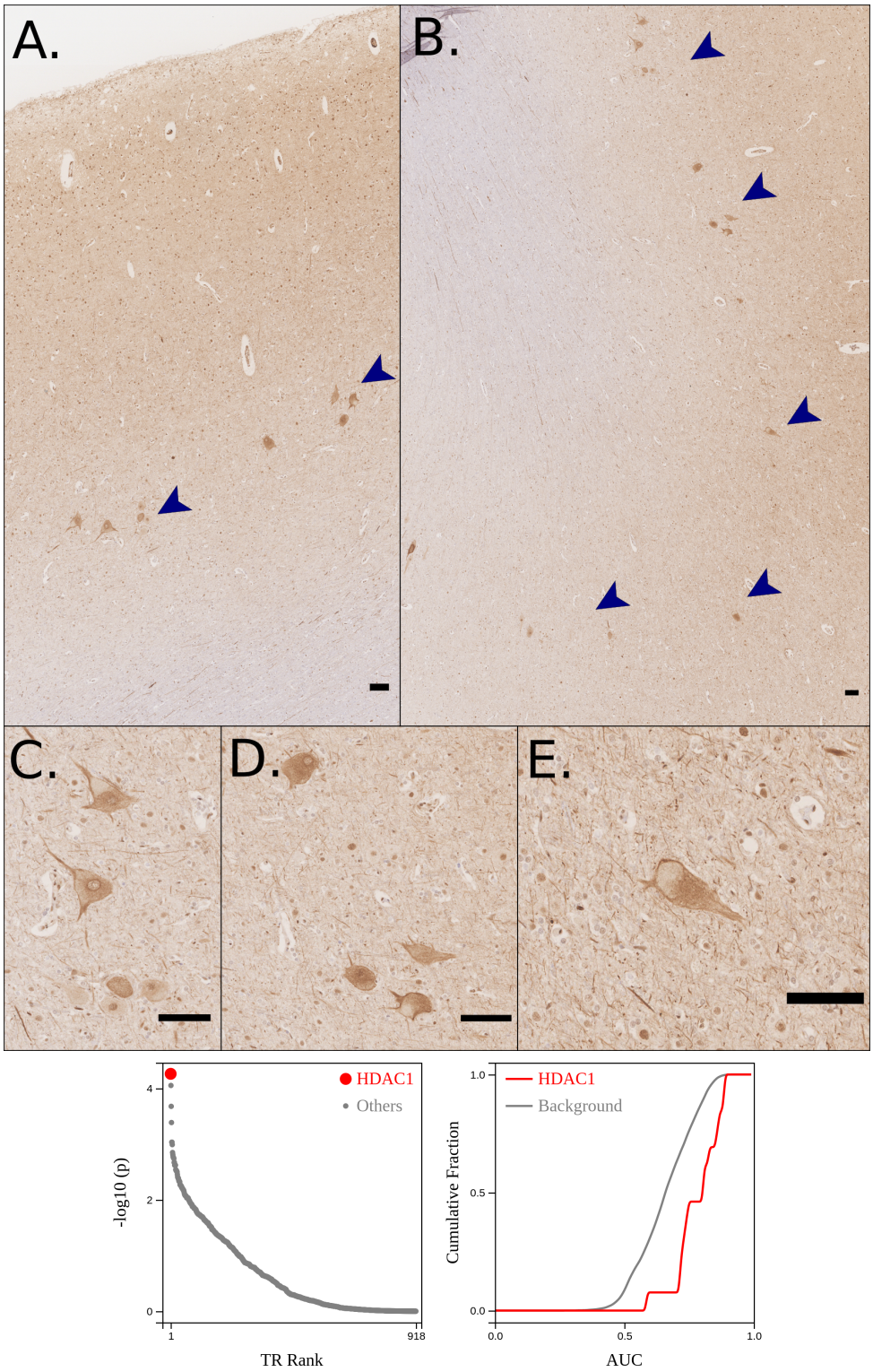


Figure 5.37: Transcripton Factor HDAC1 Highlights Betz Cell Morphology. Curiously, HDAC1 did stain other neurons, but was mostly localised to the nucleus. In Betz cells, HDAC1 expression was high, particularly in the cytoplasm. This staining pattern allowed Betz cells to be pronounced both at low (Blue arrows) and high magnification (C-E). Scalebar: 100µm.

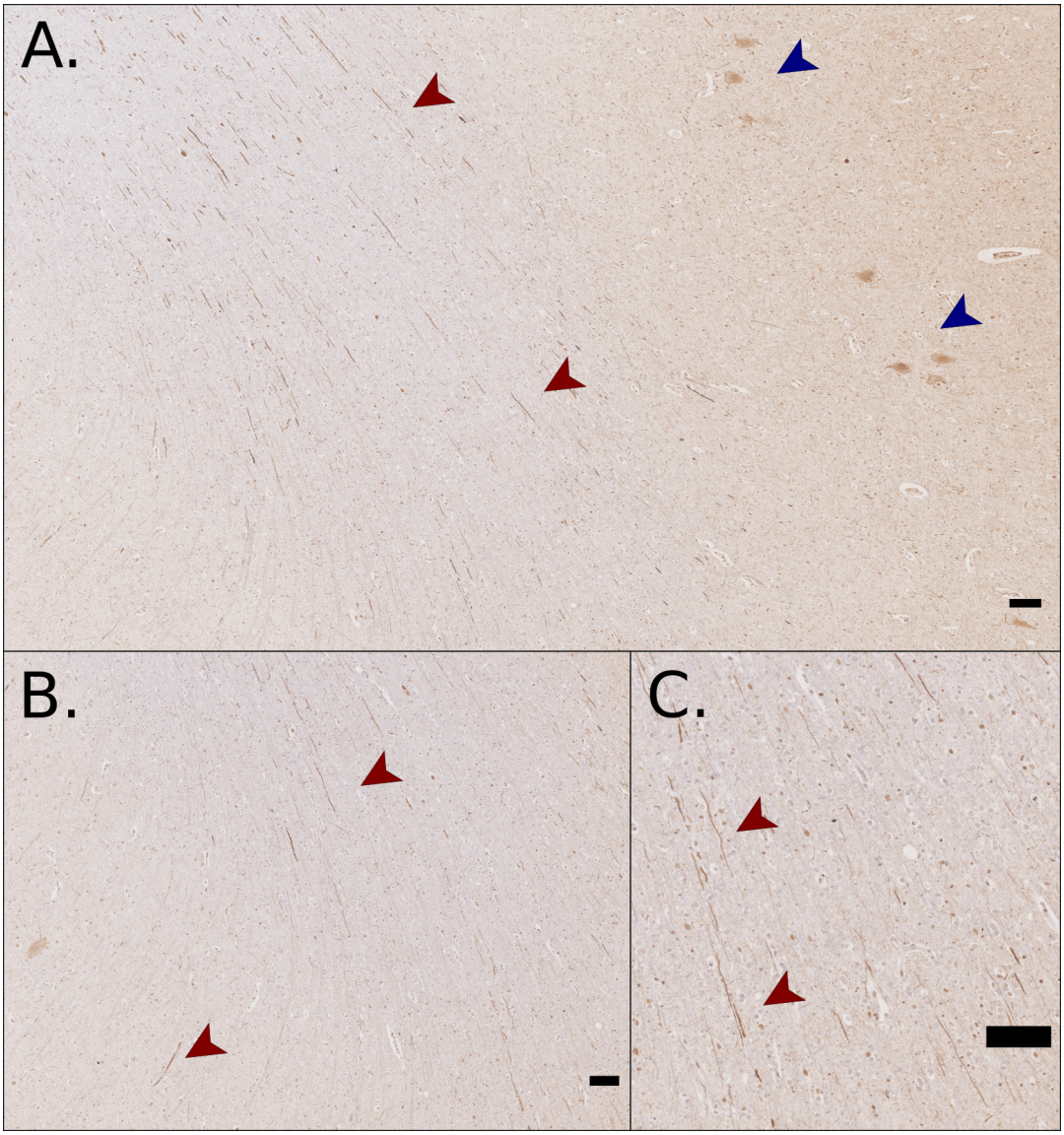


Figure 5.38: Transcripton Factor HDAC1 Highlights Betz Cell Morphology Cont. Similar to REEP1 IHC, HDAC1 staining revealed enriched staining in Betz cell soma (Blue arrows) and positive immuno-reactivity in a subset of axons in the white matter of the primary motor cortex. More work is needed to confirm that these axons originate from Betz cells; the ability to track Betz cell axons through the neocortex would be particularly useful for future research into Betz cell projection sites or studying ALS axonal degeneration in Betz cells. Scalebar: 100µm.

5.4 Discussion

The work in this chapter has - to the best of our knowledge - investigated for the first time, the proteomic composition of the Betz cell; and has also compared the Betz cell proteome to that of other CNS cell-types. To study the Betz cell and gain a better understanding of their biochemical composition we utilised our robust LCM proteomic approach which was optimised in the previous chapter. For this and future work, a substantial research cohort had to be generated that provided access to abundant numbers of high-quality Betz cell populations; providing access to the entire Betz cell pool and allowing tissue processing steps to be taken with *'omics* technologies in mind i.e. short fixation, minimal exposure to RNases, and immediate freezing. Having access to both FFPE and frozen material from the same cases meant that back-mapping validation could be achieved by discovering proteomic hits from 'gold standard' frozen tissue and applying immunohistochemical assays on morphologically superior FFPE tissue.

To gain novel preliminary data on the Betz cell and to see if it was possible to discern cell-type-specific differences using our approach, an experiment was performed in which Purkinje neurons were microdissected in parallel to Betz cells, and the underlying protein expression of both was assessed. Absolute quantification (Appendices D.2 / D.3) comparing the distribution of individual proteins *within* each cell-type showed several pathways were significantly enriched. The most abundant proteins in the Betz cell-related primarily to energy generation, namely oxidative phosphorylation and ATP synthase activity. While high in Betz cells, protein involving ATP activity was also highly abundant within the Purkinje cells. This may be intuitive as both cells are the principal neurons of their respective neuroanatomical areas and their large size has been previously attributed to high energy demands. Although, without the presence of a "typical" neuron sample, it is impossible to say if proteins relating to energy generation are specific to large neurons like the Betz and Purkinje cell or if it is a feature of all neurons. Proteins

that localise to the ribosome, nucleosome, and cytoskeleton were among the most abundant in the Betz cell, which reflects observed Betz cell morphology.

Large differences were seen when comparing the relative protein abundances between these two cell-types, with 1,124 marked as being differentially expressed. Confidence in our data increased since well-characterised canonical Purkinje cell biomarkers were found to be among the highest-ranked DEPs in the Purkinje cell samples. It was noted that many of the most differentially expressed proteins in the Betz cell were known to be immuno-reactive with the complex neuropil which comprises the grey matter of the neocortex. The reasoning behind this abundance of neuropil proteins is unclear, however, there are several possibilities. (1) The process of LCM while useful for enriching Betz cells does not result in the total isolation of them from their surrounding micro-environment, likely due to the dense interconnection between the two. (2) The micro-environment of the Betz cell is more protein-rich/complex than the micro-environment of the Purkinje cell. (3) Although laser microdissection can isolate neuronal soma, the incredibly high abundance of proteins found in the neuropil such as GFAP, BSN, MBP, MAG, and MOG can still result in their identification in the data-set. (4) The aforementioned proximity of Betz cells to glial satellite cells makes pure isolation particularly difficult, as such contaminating glial-specific proteins can be observed in the final data, unlike in the Purkinje cell samples. The truth is most likely a combination of the above. Despite these shortcomings, it was still possible to perform informative proteomic analysis. While absolute quantification showed that both Betz and Purkinje cells represented highly metabolically active neurons, the direct comparison of protein abundances highlighted that despite their similarities, the Betz cells had an overall higher abundance in proteins involved in oxidative phosphorylation and cellular respiration, and proteins localised to mitochondria; strong indicators that Betz cells are the more active neuron.

Recent electrophysiology of the Betz cell has suggested their ability to perform prolonged firing, which when compared to rodent ET neurons, is dramatic and distinctive of primate/human ET neurons and may be a result of Kv1-containing

voltage-gated K+ channels [163]. KCNC1 was not identified in this data-set, however, KCNA1 (L5/Purkinje), KCNA2 (DNNs), and KCNAB2 (L5) were found to be enriched within Betz cells when compared to the other neurons studied. Along with their large size and characteristically long axon, this prolonged firing may in part explain the substantial energy demands of the Betz cell as seen in this data. This may also explain the enrichment of protein pathways involved in neurotransmitter release and synaptic vesicle priming, in particular, a strong enrichment in proteins involved in the SNARE complex, and thereby neurotransmitter release. Finally, Figure 5.6 highlights the protein pathways which are relatively enriched in the Betz cell when compared to Purkinje cells. Moreover, the data from this initial investigation highlight that the human Betz cell is a highly metabolically active neuron with high energy demand and an increase in synaptic vesicle priming and transport. This preliminary data has also shown that Betz cells have an increase in proteins related to the regulation, localisation, and transport of macromolecules.

LCM LC-MS/MS validation was performed via immunohistochemistry by back-mapping promising hits to the FFPE counterparts of the primary motor cortex and cerebellum. CALB1 and ITPR1 were first selected to be back-mapped, although known biomarkers and confirmed via the HPA data-set to be highly expressed in Purkinje cells these were still selected to confirm findings independently and to investigate if Betz cells truly expressed these markers at lower amounts. While there was some CALB1 immunoreactivity localised to Betz cells, on closer inspection this appears to be instead, localised to synaptic boutons, most likely originating from local inhibitory interneurons - a feature previously reported in Betz cells [316]. To make the immunohistochemical results quantitative and therefore comparable to LC-MS/MS data, a machine learning approach was employed. Using this workflow, it is possible to validate LCM LC-MS/MS data with an automated workflow. From this initial study, several proteins were targeted to confirm proteomic findings. Except for one (VDAC1), all antibodies correlated with mass spectrometry data. It is important to note, however, as VDAC1 was selected primarily because it failed to reach our alpha thresholds. Several proteins, particularly, ESYT1 and DST showed

highlight selective immunoreactivity for neurons with L5. While these proteins do not highlight all aspects of Betz cell morphology they performed well in selectively staining the soma of Betz cells, which may allow them to be used in future studies as potential Betz cells markers. All of the proteins targeted by IHC also contained strong immunoreactivity within DNNs. DNNs are large glutamatergic neurons responsible for processing a plethora of afferent/efferent signalling, some of which are from the pre-motor and supplementary motor cortices. Their physiology, therefore, may be more similar to Betz cells than Betz cells to the inhibitory Purkinje cells - a feature reflected in the proteomic data. This may mean, that the enriched Betz cell proteins identified are likely to contain a large overlap of proteins enriched in glutamatergic vs. GABAergic neurons; reflecting different neurophysiology between Betz and Purkinje cells, rather than cell-type-specific changes.

Despite this large difference in the fundamental biology of Betz cells to Purkinje cells, they were still microdissected in subsequent experiments, as the wealth of literature on them meant that the results of future LCM experiments could at least be validated, in part, by the presence of known Purkinje cell markers or pathways known to be enriched within Purkinje cells. To identify Betz-specific protein expression, the ideal neuron to compare it to would be the neighbouring neurons in L5 of the motor cortex, this may also help find molecular differences between these two neuronal populations; which was the rationale for the next LCM experiment. Moreover, neurons from the Dentate Nucleus were also microdissected as they appeared to have similar staining patterns to Betz cells in the previous IHC experiments, and comparing them to Betz cells might reveal subtle Betz specific protein expression.

Promisingly, the RWS approach continued to achieve high proteomic depth with $\approx$ 3,500 proteins being identified across the Betz, Purkinje, and DNNs, although a slight reduction was seen in the non-Betz L5 neuron samples. A clear bias was seen in this experiment, as L5 samples were the most dissimilar from the other neuronal populations. PCA revealed that neuropil-rich proteins were driving this separation (PLP1, MBP, MOG, BCAS1, GAP43, and CLDN11). The alternative microdissecting strategy of cutting areas rather than individual

neurons appears to have been the reason for this. ORA of L5 samples also found enrichment in extracellular/neuropil proteins. While large structures such as neuronal layers, giant neurons, and blood vessels can be microdissected with relative ease, smaller structures (small neurons and individual glial cells) require the surrounding microenvironment to also be microdissected; and as seen here, can lead to pronounced differences in downstream results. This may pose a significant issue in the previously mentioned LCM literature that microdissect whole areas rather than individual cells. To combat this, it may be possible to microdissect a representative sample of neuropil and use that as a background filter for cell-specific applications where cell bodies cannot be isolated individually. Nevertheless, despite these shortcomings, neuronal-specific proteins were still present in the non-Betz L5 neurons samples and it was still possible to discern differences between all 4 sample types.

In total, 116 DEPs were found to be consistently enriched within the Betz cells when compared to all other samples, with ABHD16A being the highest ranked. ABHD16A is an endoplasmic reticulum protein that regulates cellular lysophospholipid levels in neurons. In particular, it is thought to generate lysophosphatidylserine lipids which are critical in homeostatic cell signalling [317]. Interestingly, the second-highest ranked protein is the vesicle transport protein BNIP1, another endoplasmic reticular protein whose role is to facilitate transport to the endoplasmic reticulum. BNIP1 is also part of the SNARE complex and has been described as a mitophagy-associated, proapoptotic protein [318]. Reduced levels of BNIP1 have been reported in spinal motor neurons in ALS patients, and BNIP1 has also been classified as an ALS risk gene [319]. Similarly, several proteins regarding endoplasmic reticulum organisation or signalling were identified within these 116 enriched proteins including RAB2A and ARFGEF2 and indicate the increased priority of protein synthesis, calcium storage, and lipid metabolism within Betz cells. Similar to the previous Betz-Purkinje experiment, another Cytochrome c-associated protein (COX7A1) was also one of the top-ranking proteins. Promisingly, light, medium, and heavy chains of neurofilament, which are known to be highly expressed in Betz cells and are currently used as a non-specific Betz cell marker were also found in this Betz cell cluster - a

full list is found in Appendix D.4. Analysis of the 116 proteins highlighted several GO terms that were enriched in these samples, including intracellular transport, endoplasmic reticulum organisation, and response to peptide hormone. Moreover, there was also enrichment in proteins found in the endoplasmic reticulum membrane, microtubules, and recycling endosomes, which suggests that Betz cells have a high demand for protein/lipid synthesis, transport, and recycling - and may provide clues to the characteristic abundance of lipofuscin in these neurons. It is important to note that the abundance of mitochondrial protein that was seen in the previous Betz-Purkinje data-set was not seen in Betz-specific cluster 4, however, this was seen in cluster 3 which comprised of proteins that were the most abundant within both Betz and DNNs. In addition to this, both Betz cells and DNNs contained an enrichment of proteins involved in mitochondrial protein import and mitochondrial calcium regulation, specifically, in the processing of SMDT1, an essential regulatory subunit of the mitochondrial calcium uniporter complex. These data indicate that both these neurons have a high dependency on energy production mechanisms - which mirrors the previous similar staining pattern seen between these two neuronal cell-types.

Other pathway enrichments were apparent when comparing the proteome of the Betz cells directly to the other cell-types; in the case of DNNs, the two cell-types had remarkably similar profiles with only 59 proteins having a LogFC greater than 2. Previous experiments have shown that Betz cells contain little or no expression of CALB1, ITPR1, and PRKCG, while our in-house CALB1 and ITPR1 immunohistochemistry have shown that these are expressed at moderate amounts in the Dentate Nucleus neuron, likely from Purkinje synaptic boutons; it was expected to see these proteins as the most differentially expressed between these neurons. These results also showed that Betz cells contained a larger abundance of proteins localised to ribosomal subunits and endoplasmic reticulum. Conversely, the DNN possessed a higher expression of proteins localised or involved in the NADH dehydrogenase complex or mitochondrial activity. Considering that earlier work has suggested these features are especially abundant in Betz cells, these data suggest that the Dentate Nucleus neuron may have an even greater demand for

energy generation than the Betz cell. Interestingly, tRNA activity, protein kinase activity, lipid transport, and protein targeting to the endoplasmic reticulum were all terms that were enriched in Betz cells when they were directly compared to DNNs, Purkinje cells and non-L5 neurons.

PRUNE2 was seen to be a major protein that appeared to be significantly expressed in Betz cells compared to DNNs (LogFC: 5.58). PRUNE2 is considered responsible for many cellular processing including synaptic function, apoptosis, cell transformation, and post-endocytic trafficking [320]. PRUNE2 has also been identified as a susceptibility gene for Alzheimer's disease [321], however its exact localisation in the brain is unclear, although an investigation into the Prune2 mRNA in the mouse brain has shown it to be highly expressed in cranial motor nuclei and spinal motor neurons [320]. Interestingly, this study did not identify Prune2 mRNA expression in the motor cortex. Within our experiment, PRUNE2 also showed strong enrichment in Betz cells in direct Purkinje cell and L5 comparisons (LogFC: 5.6 and 5.67, respectively). Follow up investigations should be performed to confirm the localisation of PRUNE2 in the motor cortex.

The comparative analysis of the Betz cell to other neurons has identified several proteins which are particularly enriched within them, providing insights into the cellular mechanisms that are predominant within these giant neurons. Importantly, immunostaining for a subset of these proteins has identified a handful of proteins that may act as potential Betz cell biomarkers, by either selectively or abundantly staining Betz cells. Initial immunostaining of energy-associated proteins (ATPB, CYCS, VDAC1, Ae-1) did reveal increased immunoreactivity in Betz cells, confirming that these neurons are highly active. Although some of these proteins do stain the unique somatodendritic architecture of the Betz cell, the ubiquitous nature of proteins involved in cellular respiration renders these proteins less than ideal for selectively studying the Betz cells from surrounding motor neurons. Conversely, DST staining revealed strong immunoreactivity in Betz cells. As DST is essential for neuronal cytoskeleton organization [322] it is unsurprising that Betz cells have an enrichment of DST. While diffuse staining existed in the superficial layers

of the motor cortex and some non-Betz cell L5 neurons were also positive; the intense immunoreactivity of DST in Betz cells, along with positive staining of dendrites makes Betz cells easily identifiable. While neurofilament heavy cortical IHC expression does highlight the full range of Betz cell morphology it also intensely stains surrounding neurons, making DST and other-Betz "soma"-enriched markers potential alternative biomarkers. ESYT1 IHC also revealed selectively staining of Betz cells, unlike DST, non-Betz cell L5 staining was rarer. Interestingly, ESYT1 staining of the motor cortex also stained the neurons in L3, potentially representing the similarities of L3 and L5 neurons - both of which are the primary drivers of action potentials in the motor cortex. The full role of ESYT1 is unclear, however, increased amounts may be indicative of increased neurotransmission, synaptic growth, and in response to depressed cytosolic Ca^{2+} levels. It may be that ESYT1 is an indirect marker for neurons that have large excitatory neurotransmission demands, which could explain the selective pyramidal cortical staining in the motor cortex along with the abundance of ESYT1 in DNNs but not Purkinje cells. The clear Betz cell staining of ESYT1 in L5 makes ESYT1 a tantalising protein to be used for future work that wishes to target neurons with large excitatory energy demands.

Observations in the literature have noted that Betz cells are weakly positive for CRYM; however, this data-set highlights CRYM as along the most highest-ranked DEPs when comparing Betz cells to Purkinje cells, DNNs, and L5 samples. Confusingly, when CRYM IHC was performed, Betz cells were indeed weakly negative to CRYM. As reported in the literature, CRYM did selectively stain L5 of the motor cortex. However, in our tissue, there were several large cells whose cell soma was positively stained with CRYM, these cells are presumed as Betz cells, however, some Betz defining morphological features - Appendix A - were not immediately apparent. If CRYM is only expressed highly in a subset of Betz cells, this raises interesting questions both to its function and the wider physiology of Betz cells. As CRYM is believed to be a crucial determinant of neuronal death/survival in neurodegenerative disease [323], perhaps the obvious Betz cell degeneration seen in ALS is reflective of the Betz cells with only low levels of CRYM, with CRYM

positive Betz cells being resistant to degeneration? However, more work is needed to reveal the role of CRYM in the healthy and diseased neuron, before such claims can be validated. The existence of subtypes of Betz cells have been reported previously from electrophysiology investigations of the Betz cell - see B.2 - and may justify dedicated experiments to reveal the molecular profile of these potential subtypes.

There has been a significant amount of work aimed at molecularly defining the subpopulations of L5 neurons. Curiously, LC-MS/MS data validated by IHC revealed that Betz cells were immuno-positive for both ER81 and OTX1, a staining pattern that some have argued is very rare or non-existent as it is believed that ER81 and OTX1 stain different subpopulations of neurons in L5 [172, 324–326] [6]. Additionally, the same authors have suggested that ER81-positive neurons never co-localise with neurofilament (N200)-positive neurons, which as already discussed are enriched in Betz cells [324]. Much of the work on molecularly classifying L5 neurons has been performed on the rodent or non-human primate brain, and many do not use the primary motor cortex. The immunoreactivity of neurofilament heavy, OTX1, and ER81 make it difficult to place Betz cells in any preexisting neuronal niche proposed by these models. Further work is needed to stain human Betz cells with established L5 markers to investigate where they fit into the global context of L5 pyramidal neurons. Moreover, co-localisation experiments are needed to confirm if it is the same population of Betz cells that are stained with ER81 and OTX1, along with the other markers identified in this thesis, or if sub-populations of Betz cells exist.

Immunostaining of REEP1 revealed intense immunoreactivity in Betz cells; with anti-REEP1 it was also possible to stain some of the somatodendrites around the Betz cell aiding in identification. REEP1 may be involved in the transport of specific cargo, regulation of ER membrane, and vesicular transport [308]. Critically, REEP1 has been suggested to link ER tubules to the cytoskeleton [310] in addition to being involved in long axon maintenance [309]; all of which are cellular features

[6] It is important to note that in our hands ER81 was not restricted to just L5 neurons, as seen in these studies; a different antibody clone was used, which may explain the dichotomy between our data and those reported in the literature.

that are either known or have been shown to be abundant in Betz cells from earlier proteomic data. If the immunoreactive axons in the white matter do belong to Betz cells, it would make REEP1 an interesting molecule to target in the future to confirm the proportion of Betz cell axons in the CST and beyond. Some immunoreactivity was seen in smaller neurons within L5, however, some cellular features may be of Betz cells not currently in the histological plane. In light of this, REEP1 represents a strong candidate for a marker to immunochemically target Betz cells; as it allows the unique morphology of Betz cells to be studied without the distracting neuropil or surrounding motor neurons. Finally, binding analysis for regulation of transcription (BART) revealed HDAC1 as a potential transcription factor that is abundant within Betz cells. Although it does have a clear role as a TF, HDAC1 is critical for mature motor neurons - neuronal progenitor cells without HDAC1 are unable to differentiate into mature neurons and eventually die, resulting in abnormal cortical layering [327]. Other studies have shown that HDAC1 represses NOTCH target genes and promotes the generation of motor neurons in zebrafish [328]. Relevantly, the HDAC family have also been found to be critical in neuronal development and axon regeneration, with HDAC1 having a role in dendritic wiring [329]. Interestingly, HDAC1 inhibition has been reported to play a role in ALS [330, 331]. As with REEP1, HDAC1 represents a strong candidate for a marker to best distinguish Betz cells from other neurons in the context of the motor cortex. Moreover, if the HDAC1 immunoreactive axons in the white matter originate from Betz cells, it may help answers the remaining questions around the projection of Betz cells, i.e. are all Betz cells corticospinal? Do some project to other targets? Or to the contralateral motor areas?

Overall, this initial comparative analysis of the proteome of the Betz cell confirms assumptions on the nature of the Betz cells along with new clues on their cellular mechanisms. For example, these data suggest that proteins relating to protein kinase A binding, cellular component disassembly/recycling, intracellular transport, protein transport, aminoacyl-tRNA ligase activity, translational initiation and elongation were all enriched within human Betz cell. Collectively, this confirms that Betz cells

are highly dynamic neurons, with an increased priority on protein/lipid synthesis and cellular transport, more so than the other large neurons studied. Although these cellular functions may be assumed with neurons with Betz cell's unique morphology, i.e. large fast-firing glutamatergic principal neurons, it cannot be understated the importance of developing approaches to confirm basic assumptions. The increase in microtubule proteins, intracellular transport, and emphasis in biomolecule synthesis can also be expected due to the added demands of the Betz cell's signature elongated axon. Furthermore, the work outlined in this chapter showcases the robustness of our LCM proteomic method; with it, it was possible to generate a rich proteomic data-set from a minimal amount of material, allowing proteomic differences between both very similar and very different neurons in the human brain to be compared. This work in this chapter also highlights several proteins which may be used as surrogate markers for the Betz cell (REEP1, DST, ESYT1, HDAC1, CAPRIN2, OTX1) to increase the ease of correctly identifying Betz cells without needing intimate knowledge of their morphology. It is important to note that the proteins outlined here are not 'true' Betz cell markers as other cell-types also showed immunoreactivity, and other cells in the CNS or elsewhere may also be positive for these proteins. However, in the context of the primary motor cortex, these proteins provide an immunohistochemical way to selectively target/study Betz cells and may allow Betz cells to be compatible with immuno-based enrichment techniques e.g. immunoprecipitation or FACS in the future, allowing them to be studied to the same degree as other neurons. Similarly, data generated could aid other research, either in helping find a Betz cell signature in bulk tissue or by guiding/validating unbiased scRNA-seq experiments.

While this chapter provides novel insights into the biochemical composition of Betz cells, along with providing a rich proteomic data-set of microdissected Purkinje cells, DNN, and L5 microenvironment; more work is needed to get a strong grasp on the biological nature of the Betz cell. Although new insights have been generated, the work in this chapter has been performed on relatively few samples. In the next chapter, the RWS approach will be used to investigate the function and

composition of the Betz cell in a larger number of post-mortem cases. In addition to this, investigations will be made to see how the proteome of the Betz cell and other neurons are affected in patients with ALS. Collectively, by combining the novel insights gained in this chapter, with the additional validation of large-scale LCM proteomics and the introduction of LCM transcriptomic data in the next chapter, the goal of this thesis is to lay the foundations of biochemically defining a molecular profile for the Betz cell and begin to pave the way for future studies that wish to selectively research this neuron in the healthy and diseased brain.

5.4.1 Main Conclusions

- It is possible to discern proteomic differences between neurons of similar physiology using the RWS workflow.
- Proteomic data regarding the human Betz cell has been generated for this first time.
- LC-MS/MS analysis using the RWS approach reflects known/assumed biology of the Purkinje and Betz cell and reveals individual proteins that can be investigated/validated using standard IHC.
- The generated data confirms previous assumptions about the Betz cell; they are highly metabolically active glutamatergic neurons, with enrichment of cytoskeletal and mitochondrial proteins.
- The generated proteomic data reveals the enrichment of ER proteins, tRNA acetylation proteins, cytosolic transport proteins, and recycling endosomal proteins within Betz cells, highlighting their dependency on protein synthesis, cellular recycling, and macromolecule transport.
- LCM Proteomic analysis revealed $\approx$ 1,000 proteins that were differentially expressed in Betz cells compared to other neuronal subtypes; which was narrowed down to 116 DEPs that were consistently enriched in Betz cell samples across comparisons.

- IHC validation has resulted in several proteins that can be used to selectively visualise the Betz cell in the context of the primary motor cortex. These proteins may have the ability to act as future Betz cell biomarkers.

6

LCM Proteomics and Transcriptomics of the Human Betz cell in the healthy and ALS brain.

Contents

6.1 Aims of this Chapter 193
6.2 Introduction 194
6.2.1 Breakdown of this Chapter 194
6.3 Results 197
6.3.1 Developing a High Throughput Micro-Proteomic Workflow 197
6.3.2 Case Selection for LCM Proteomics and LCM-seq . . . 197
6.3.3 LCM Proteomics of the Human Betz cell 198
6.3.4 LCM-seq 203
6.3.5 LCM-seq Analysis of the Human Betz cell 208
6.3.6 LCM-seq Analysis of ALS Neurons 219
6.4 Discussion 225
6.4.1 The Betz Cell in the Healthy Brain. 225
6.4.2 The Betz Cell in the ALS Brain. 227
6.4.3 Main Conclusions 234

6.1 Aims of this Chapter

- Build upon the proteomic findings found in the previous chapter.
- Adjust our LCM proteomic approach to allow high-throughput mass spec-

trometry.

- Investigate how the proteome of motor neurons, including Betz cells, is affected in ALS.
- Optimise an LCM transcriptomic protocol to gain novel insights into the transcriptome of Betz cells.
- Investigate how the transcriptome of motor neurons is affected in ALS.

6.2 Introduction

As previously mentioned, despite the clear degeneration of Betz cells in Amyotrophic Lateral Sclerosis there have been limited studies that (molecularly) investigate the exact mechanisms which become aberrant in these neurons. Identifying the biomolecules involved in these affected pathways is fundamental to understanding the full impact of this disease. Thus far, this thesis has described, optimised, and implemented an LCM proteomic workflow that produces cutting-edge proteomic performance from small amounts of pooled cells. This approach has provided novel insights into the proteinous nature of the Betz cell and has highlighted several proteins which can be targeted with immunohistochemistry to better visualise the Betz cell in the cellular environment of the primary motor cortex. Next, this chapter aims to maximise the utility of this technique; using it to build upon the Betz cell insights seen in the previous chapter, but additionally, to explore how the proteome of the Betz cell changes in individuals with ALS. This requires isolating Betz cells from a larger number of post-mortem cases, which poses significant challenges. To complement the proteomic workflow, the experience gained from the RWS approach will be used to optimise an LCM-seq protocol to analyse the expression of mRNA in Betz cells and investigate how the abundance of these transcripts change across different cell-types and in ALS.

6.2.1 Breakdown of this Chapter

1. The RWS workflow will be adapted and optimised to allow for high-throughput proteomics. Developing such an approach will accommodate LCM proteomic

studies wishing to analyse a large number of samples.

2. Proteomic data will be generated on a large number of Betz cells, L5 motor neurons, and Purkinje cells microdissected from the healthy and ALS brain.
3. LCM-seq, a complementary technique to LCM proteomics, will be optimised to maximise performance on in-house post-mortem microdissected neurons.
4. LCM-seq will be used to directly study the transcriptome and potential changes of the Betz cell, L5 motor neurons and Purkinje cells from the healthy and ALS brain.

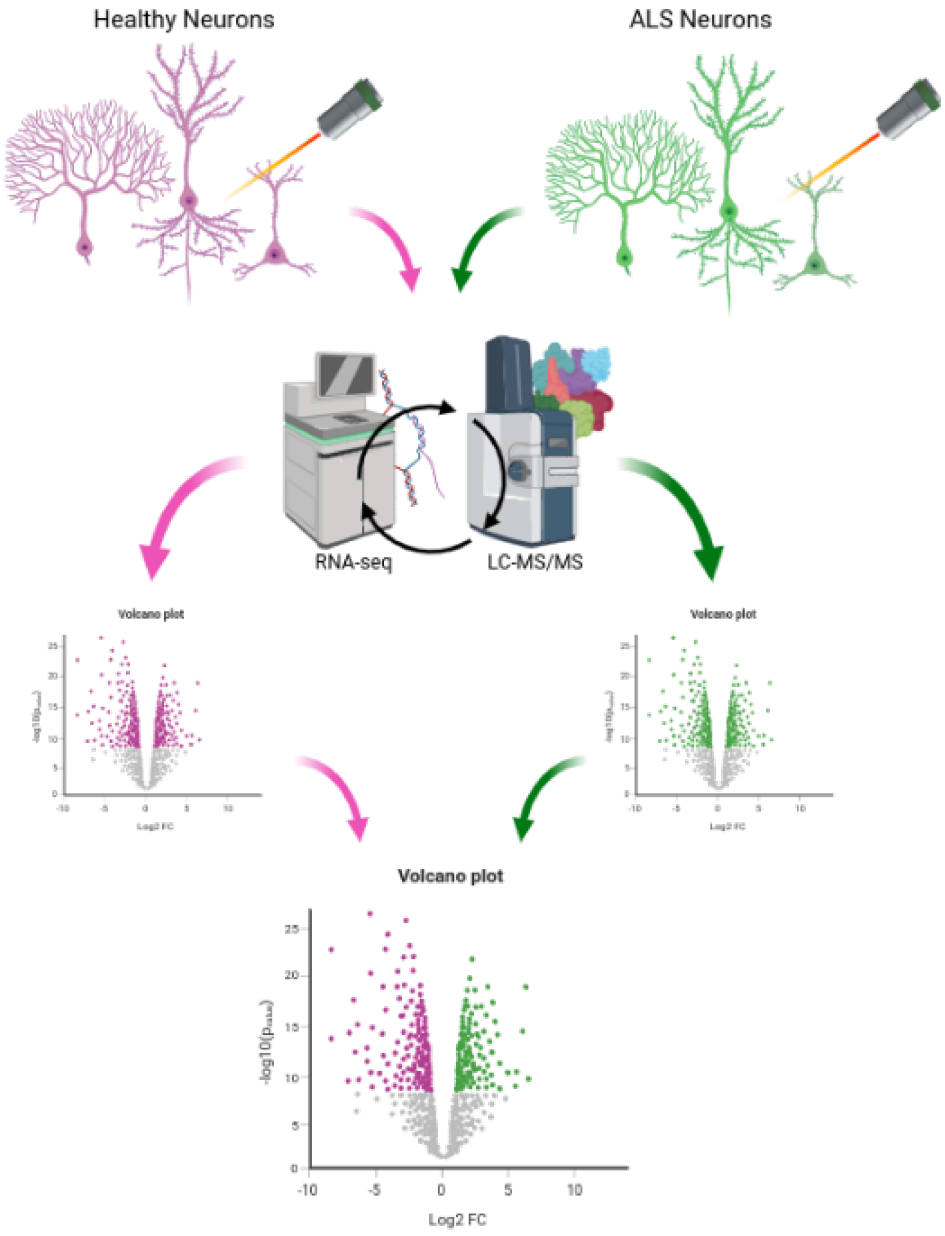


Figure 6.1: LCM *"-omics"* of Healthy and Diseased Neurons.
In this chapter, the proteome and transcriptome of Betz cells, L5 motor neurons, and Purkinje cells will be investigated in multiple post-mortem brains, with the goal of gaining more information on the cellular pathways enriched in Betz cells. Next, the same neuronal cell-types will be microdissected from multiple ALS post-mortem brains to investigate how their proteome and transcriptome is affected by disease.

6.3 Results

6.3.1 Developing a High Throughput Micro-Proteomic Workflow

To perform an exploratory analysis of the Betz cell in the healthy and diseased brain, both biological and technical replicates are needed to increase confidence in the generated data. However, the fact that the solution containing peptides from microdissected cells are injected entirely into the mass spectrometer makes technical replicates difficult. Instead, replicates of microdissected cells from the same brain were processed and injected into the mass spectrometer in parallel. In addition to this, cells were microdissected from multiple post-mortem brains. This strategy meant, that the number of samples to be analysed far exceeds that of previous experiments and poses a significant challenge for the current RWS approach. Currently, each sample uses a 1-hour gradient making analysis lengthy and prone to data drift in large-sized projects. To address this, an alternative high-throughput workflow was optimised and implemented to minimise the resources needed (time, cost, labour) and to better adjust for batch bias - see Appendix E.1 for details. In summary, the optimised high-throughput method still uses the RWS workflow, however, an alternative mass spectrometer was used; utilising trapped ion mobility to resolve peptides with similar properties, in conjunction with using the same nanoflow liquid chromatography system used in the previous experiments - boosting sensitivity, despite samples being run on only a 17-minute gradient.

6.3.2 Case Selection for LCM Proteomics and LCM-seq

Before performing LCM Proteomics/LCM-seq on our research cohort, several quality checks were performed to select only the best cases within the cohort. For each post-mortem case, the individual tissue blocks representing the upper third of the frozen motor strip (i.e. leg to hand areas) was sectioned and stained with Nissl and the degree of freezing artefact was assessed. Tissue blocks with severe artefacts were excluded from future analysis. Next, the number of Betz cells was counted per tissue section to determine which part of the motor strip should be used for

microdissection. As time spent microdissecting is critical for preserving mRNA quality, only tissue blocks that contained abundant Betz cells with excellent histology and proper orientation were selected for LCM-seq[1]. Overall, ALS cases had fewer Betz cells per section, likely attributed to their known atrophy in ALS, although it is unclear if they had degenerated completely, or just enough that it was difficult to histologically distinguish them from surrounding neurons. The final cohort selected for LCM-seq included x7 Control and x6 ALS cases, see Appendix F for details.

6.3.3 LCM Proteomics of the Human Betz cell

To maximise statistical power, 156 samples were prepared for LCM proteomics using the optimised high-throughput approach: healthy Betz cells (Replicates per case: 4, Total microdissected: 1,404), ALS Betz cells (Replicates per case: 4, Total microdissected: 1,707), healthy Purkinje cells (Replicates per case: 4, Total microdissected: 4,549), ALS Purkinje cells (Replicates per case: 4, Total microdissected: 4,184), healthy L5 neurons (Replicates per case: 4), and ALS L5 neurons (Replicates per case: 4). Each sample contained 100,000um^2 worth of cellular material[2]. Despite the shorter gradient time, the high-throughput approach was comparable to earlier low-throughput experiments - Figure 6.3, identifying $\approx 3,000$ proteins per sample. Overall, 6,545 proteins were identified in the Betz cell samples, 6,449 in Purkinje cells, 5,618 in L5 neurons; and 7,224 and 6,683 proteins from the cerebellum and motor cortex library samples, respectively.

Cell Type	Disease	No. Microdissected	Soma Size	Sum
Betz Cell	ALS	1,707	1,406	2,400,152
	Control	1,404	2,019	2,834,000
Purkinje Cell	ALS	4,184	601	2,513,765
	Control	4,549	614	2,819,741

Table 6.1: Neurons from 13 post-mortem brains (7 control, 6 ALS) were microdissected in quadruplicate. Each sample consisted of 100,000µm^2 pooled material. Soma size (µm^2): mean of all neurons microdissected; mean Betz cell was in accordance to the literature. Sum (µm^2): total amount of all cell microdissected.

[1]The tissue block representing the assumed hand region was often selected as the best block.
[2]This corresponds to approximately 60 Betz cells slices.

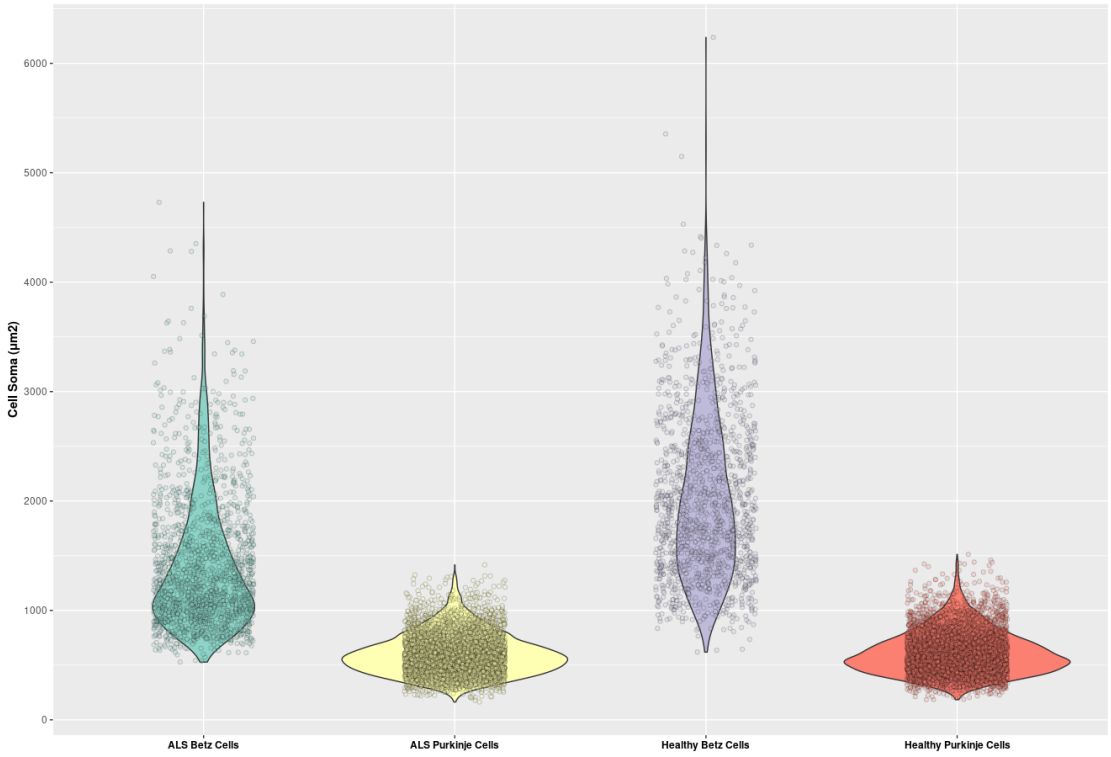


Figure 6.2: Summary of Cells Microdissected for Large-Scale LCM Proteomics. Median: healthy Betz cell size: 2,019µm^2. Median ALS Betz cell size: 1,247µm^2. Median healthy Purkinje cell size: 614µm^2. Median ALS Purkinje cell size: 601µm^2. For L5 neuronal samples, areas were taken so individual soma sizes were not recorded.

Relative proteomics of Betz cells against Purkinje cells and L5 neurons mimicked the results of small-scale experiments in the previous chapter. In total, 544 proteins were enriched in Betz cell samples compared to Purkinje cells, and 531 proteins were enriched in Betz cell samples compared to L5 neurons. Of these, 324 proteins were found to be consistently upregulated in Betz cells. ORA of these 324 DEPs showed enrichment of mitochondrial, cytosolic ribosomal, and endoplasmic reticulum proteins; similar to earlier experiments - see Figure 6.5 for a summary of enriched terms.

When ALS Betz cells were compared to healthy Betz cells, 517 DEPs were revealed; 246 upregulated in ALS Betz cells and 271 downregulated in ALS Betz cells. 1D Enrichment Analysis of healthy and ALS Betz cells revealed 273 pathways that were dysregulated. Notably, an increase in gliogenesis regulation was seen

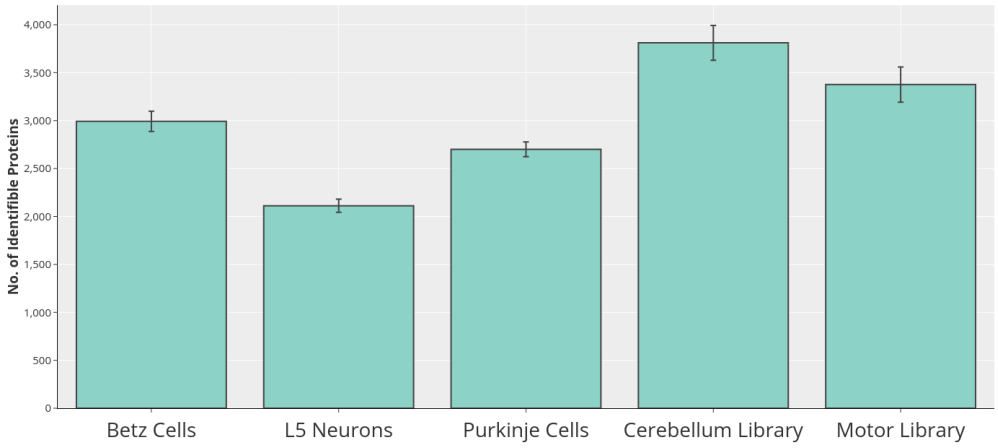


Figure 6.3: Large-Scale LCM Proteomics - No. of Identifiable Proteins per Sample. Turquoise bars denote the number of proteins identified per sample. Vertical bars: standard deviation. In total, 6,545 proteins were identified in the Betz cell samples, 6,449 in Purkinje cells, and 5,618 in L5 neurons. Library samples were large microdissected areas (x8, 1,000,000 µm^2) of the representative neuroanatomical areas, which were processed using the RWS method, pooled, and fractionated to generate more protein IDs. In total, 7,224 and 6,683 proteins IDs were found from the cerebellum and primary motor cortex library samples. Samples were matched to library samples via the MBRs algorithm.

(FDR: 0.01), potentially reflecting the reactive gliosis known to affect Betz cells in ALS. Moreover, glutamate signalling (FDR: 0.02), long-term synaptic depression (FDR: 0.0001), positive regulation of synaptic transmission (1D score: 0.622, FDR: 0.004), and the regulation of calcium ion transport (FDR: 0.009) were all enriched with ALS Betz cells, echoing the same pathways known to be affected in the excitotoxicity hypothesis of ALS. Several pathways were downregulated in ALS Betz cells, the most depleted being terms relating to tRNA aminoacylation (FDR: 1.05e-05) and tRNA biosynthesis (FDR: 1.5e-03). There was also a reduction of oxidative phosphorylation (FDR: 3.21e-04), cellular adhesion (FDR: 1.45e-03), and regulation of cytoskeleton (1.94e-04) in ALS Betz cells. Of the proteins upregulated in ALS Betz cells, they were an enrichment of those typically localised to the apical dendrite (FDR: 8.25e-03), spectrin cytoskeleton (FDR: 1.72e-02), methyltransferase complex (FDR: 1.23e-02), synaptic vesicle membrane (FDR: 1.11e-04), and chromatin (FDR: 1.58e-02). Conversely, of the proteins downregulated in

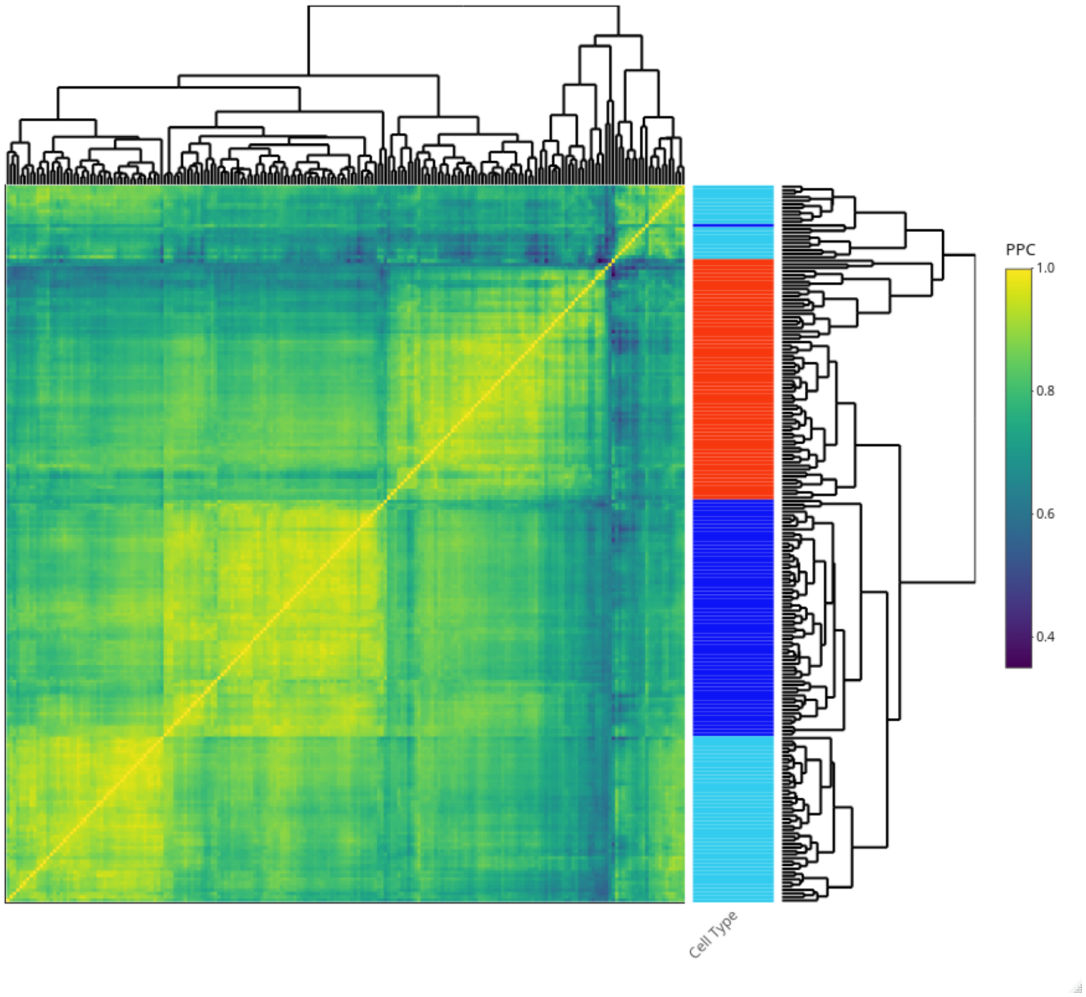
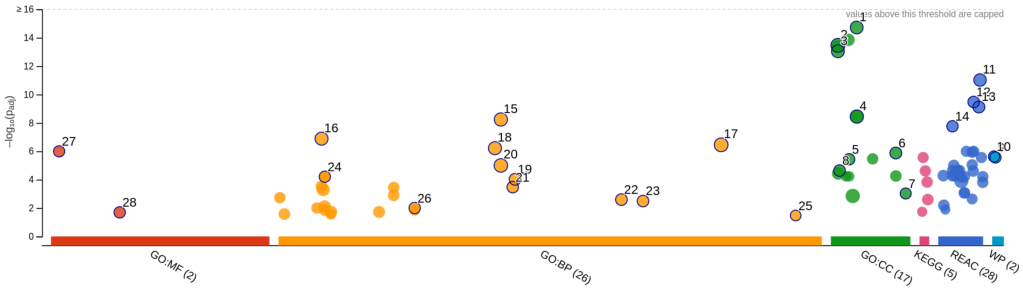


Figure 6.4: Pearson Correlation Coefficient Heatmap - Large LCM Proteomic Study. Overall, all samples correlated strongly with each other. Unbiased clustering showed that each cell-type cluster among their respective groups. L5 neurons (Teal) were similar to Betz cells (Blue), with Purkinje cells (Red) being the most dissimilar. A subset of samples clusters, mostly L5 neurons and 1 Betz cell sample, clustered away from the remaining samples. These samples had an abundance of glial markers, although it is not clear if this reflects disease mechanism or significant contamination from neighbouring cell-types.

ALS Betz cells, there was an enrichment of those localised to the polysome (FDR: 1.40e-02), endoplasmic reticulum (FDR: 1.34e-03), mitochondria (FDR: 8.52e-06), and aminoacyl-tRNA synthetase multienzyme complex (1.76e-02). Within L5 samples, 92 DEPs were found between healthy and ALS neurons. Eighteen DEPs were upregulated in ALS L5 neurons while 94 were downregulated. In total, 1D Enrichment highlighted 29 pathways that were downregulated and 1 pathway



ID	Source	Term ID	Term Name	p_{adj} (query_1)
1	GO:CC	GO:0031966	mitochondrial membrane	1.916×10^{-15}
2	GO:CC	GO:0005739	mitochondrion	3.439×10^{-14}
3	GO:CC	GO:0005740	mitochondrial envelope	9.006×10^{-14}
4	GO:CC	GO:0031975	envelope	3.605×10^{-9}
5	GO:CC	GO:0022626	cytosolic ribosome	3.673×10^{-6}
6	GO:CC	GO:0098798	mitochondrial protein-containing complex	1.336×10^{-6}
7	GO:CC	GO:1990204	oxidoreductase complex	9.448×10^{-4}
8	GO:CC	GO:0005840	ribosome	2.292×10^{-5}
9	WP	WP:WP477	Cytoplasmic Ribosomal Proteins	2.754×10^{-6}
10	WP	WP:WP111	Electron Transport Chain (OXPHOS system in mito...	2.340×10^{-6}
11	REAC	REAC:R-HSA-72...	Translation	9.567×10^{-12}
12	REAC	REAC:R-HSA-24...	Selenoamino acid metabolism	3.366×10^{-10}
13	REAC	REAC:R-HSA-14...	The citric acid (TCA) cycle and respiratory electro...	7.621×10^{-10}
14	REAC	REAC:R-HSA-15...	Eukaryotic Translation Elongation	1.737×10^{-8}
15	GO:BP	GO:0043604	amide biosynthetic process	5.770×10^{-9}
16	GO:BP	GO:0006412	translation	1.297×10^{-7}
17	GO:BP	GO:1901566	organonitrogen compound biosynthetic process	3.525×10^{-7}
18	GO:BP	GO:0043043	peptide biosynthetic process	6.099×10^{-7}
19	GO:BP	GO:0045333	cellular respiration	9.512×10^{-5}
20	GO:BP	GO:0043603	cellular amide metabolic process	1.001×10^{-5}
21	GO:BP	GO:0045047	protein targeting to ER	3.384×10^{-4}
22	GO:BP	GO:0070972	protein localization to endoplasmic reticulum	2.546×10^{-3}
23	GO:BP	GO:0072599	establishment of protein localization to endoplas...	3.227×10^{-3}
24	GO:BP	GO:0006614	SRP-dependent cotranslational protein targeting t...	6.195×10^{-5}
25	GO:BP	GO:1990542	mitochondrial transmembrane transport	3.365×10^{-2}
26	GO:BP	GO:0022900	electron transport chain	9.983×10^{-3}
27	GO:MF	GO:0003735	structural constituent of ribosome	1.016×10^{-6}
28	GO:MF	GO:0016874	ligase activity	2.030×10^{-2}

Figure 6.5: Manhattan Plot: Summary of DEPs in Betz Cells.
A summary of all the enriched GO terms within Betz cells when compared to both Purkinje cells and L5 neurons - from the large-proteomic experiment. As with previous proteomic experiments on smaller case numbers, Betz cells appear to contain an enrichment of mitochondrial and ribosomal proteins, along with proteins involved in cellular respiration and the transport of proteins to the endoplasmic reticulum. All points represent significantly enriched pathways and are colour-coded by their representative pathway type.

(complement and coagulation cascades; FDR: 1.2e-02) that was upregulated in ALS L5 neurons. Downregulated pathways were all involved in oxidative phosphorylation and NADH dehydrogenase activity.

6.3.4 LCM-seq

Several optimisation steps were performed to ensure high-quality material could be generated from post-mortem microdissected material - see Appendix E.2. The number of Betz cells per section was used to calculate the number of LCM slides needed to pooled sufficient amounts of material. Each sample contained 100,000um^2 worth of cellular material, as this was found to be sufficient for cDNA synthesis and matched the amount of material used in the large complementary proteomic study. In addition to Betz cells, L5 neurons[3], and Purkinje cells from the same brain also microdissected. In total, 420 healthy and 337 ALS Betz cells were microdissected and 1,325 healthy and 1,245 ALS Purkinje cells microdissected - see Figure 6.6. On average, Betz cell size was slightly smaller in ALS cases (1,703 µm^2 vs. 1,280µm^2; p value: <0.001); the same was seen in Purkinje cells (535 µm^2 vs. 484µm^2; p value: <0.001), although it may be flawed to see this as a direct reflection of atrophy in ALS.

Sequencing data reported that ≈ 60% of reads were duplicates, with ≈ 30 million reads being unique across all samples. The saturation of read depth suggests that the full RNA diversity of within our samples have been captured and identified by the LCM-seq experiment. Early quality control steps identified a higher than expected abundance of rRNA transcripts across samples, which were later removed *in silico*. Once processed, the cleaned data were normalised using DESeq2's variance stabilizing transformation (VST) and PCA was applied to look for apparent batch artefacts/outlier samples. The results of the PCA highlighted a subset of samples that differed substantially from all samples within the LCM-seq experiment - see Figure E.8. Initial attempts to resolve these outlier samples failed (data not shown). Instead, removing these samples resulted in the remaining samples being tightly clustered together delimited strongly by cell-type (Figure 6.7). To avoid potential bias' downstream it was decided that the potential outlier samples would be removed permanently from downstream analysis rather than attempting to computationally resolve the underlying issues behind their altered mRNA expression. The differential

[3]Care was taken to ensure L5 neurons were microdissected with minimal neuropil contamination.

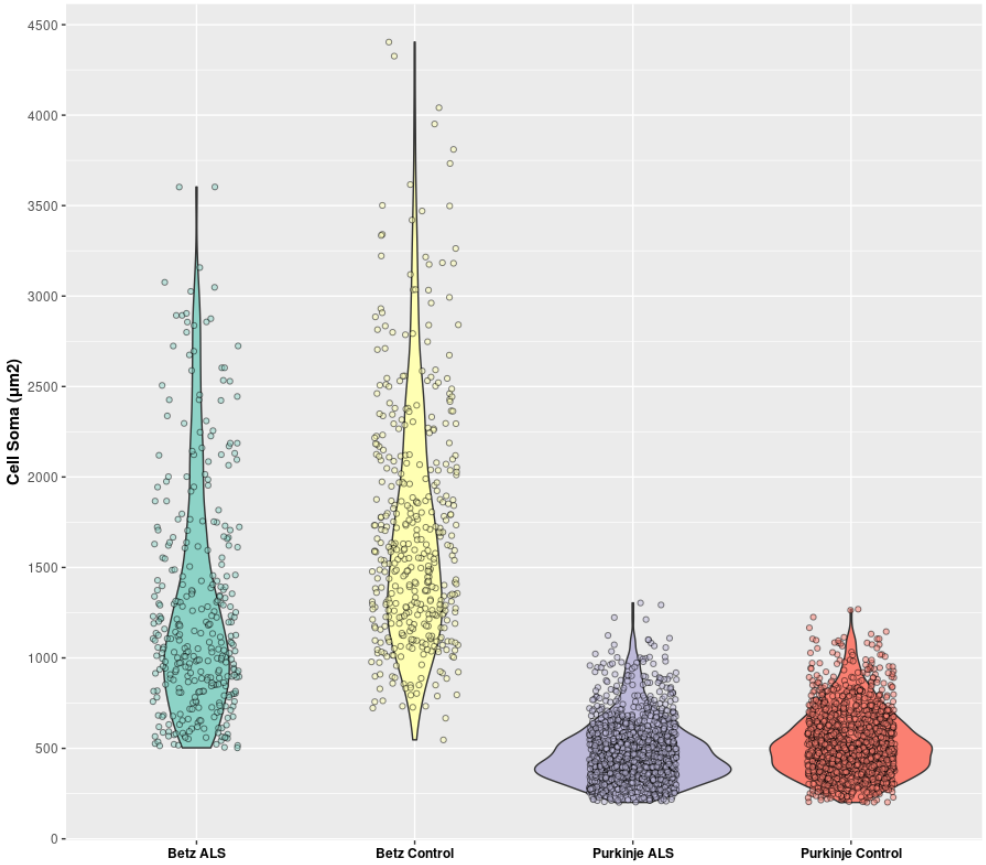


Figure 6.6: Summary of Cells Microdissected for LCM-seq Analysis. Mean healthy Betz cell size: 1,703µm^2. Mean ALS Betz cell size: 1,193µm^2. Mean healthy Purkinje cell size: 614µm^2. Mean ALS Purkinje cell size: 601µm^2. For L5 neuronal samples, areas were taken so individual soma sizes were not recorded.

mRNA expression of Purkinje cells - which are known to be substantially different from the excitatory neurons of the motor cortex - was responsible for 77.4% of the variance in this data set. While the variance between Betz cells and L5 samples was small (7.88%), it was still possible to separate these samples in an unbiased manner based on their underlying transcriptome. Interestingly, among the top 10 transcripts within Betz cells that best differentiated them from L5 neurons was neurofilament proteins (NEFH/NEFM) - see Figure E.13. Nine of the top 10 proteins in Purkinje cells driving separation between themselves and motor neurons had IHC data available as part of the HPA; and of these 9, 8 showed strong immunoreactivity in Purkinje cells. It was not possible to clearly separate individual cell-types by disease state.

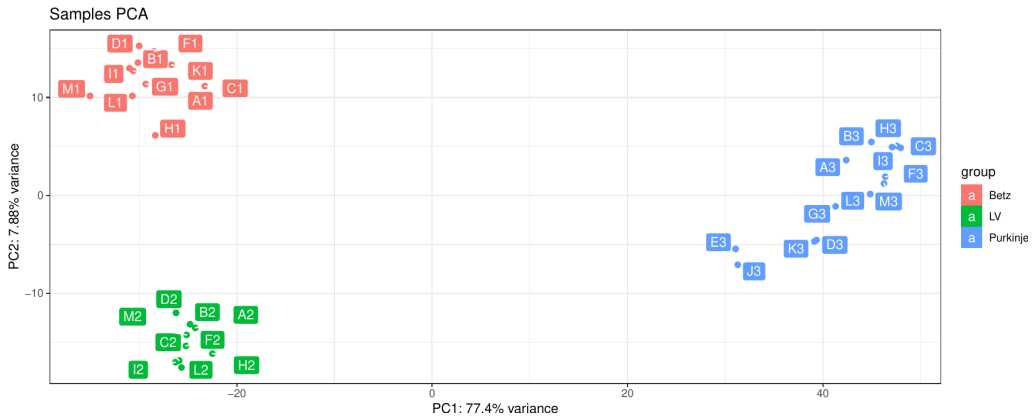


Figure 6.7: Principal Component Analysis - LCM-seq.
Once irregular samples were removed, the remaining samples clustered within their own cell-type groups.

To confirm that microdissection successfully enriched for the desired CNS cell-type, multi-subject single-cell deconvolution (MuSiC) [332] was performed on our samples using the Darmanis et al., (2015) single-cell RNA-seq data set as a reference [51]. The results of this showed that samples containing Betz cells were highly enriched with neuronal-specific markers ($\bar{X}$: 99.2%), with minimal contamination with astrocytes, microglia, oligodendrocytes, or endothelial cells - see Figure 6.8. Samples containing microdissected L5 neurons were also enriched for neuronal-specific transcripts ($\bar{X}$: 91.7%), however, there was a noticeable increase in oligodendrocyte specific transcripts ($\bar{X}$: 6.35%). Repeating MuSiC analysis on the handful of samples that were removed from the analysis showed pronounced differences in cell-type proportions, with the removed samples containing a marked increase in endothelial, astrocytic, and microglial markers - see Figure E.10, which may explain why they differed from the rest of the samples in this experiment.

Pairwise comparisons between samples found high levels of correlation across all samples ($\bar{X}$: 0.835) and within biological replicates (Betz, $\bar{X}$: 0.913, L5 neurons: $\bar{X}$: 0.885, Purkinje: $\bar{X}$: 0.876) - Figure E.9. When unbiased clustering was performed to investigate similarity between samples, all 3 sample types clustered among their respective cell-type, with Betz cells being more similarly related to L5 neurons than Purkinje cells, as expected (Figure 6.9). In total, 29,184 RNA species were

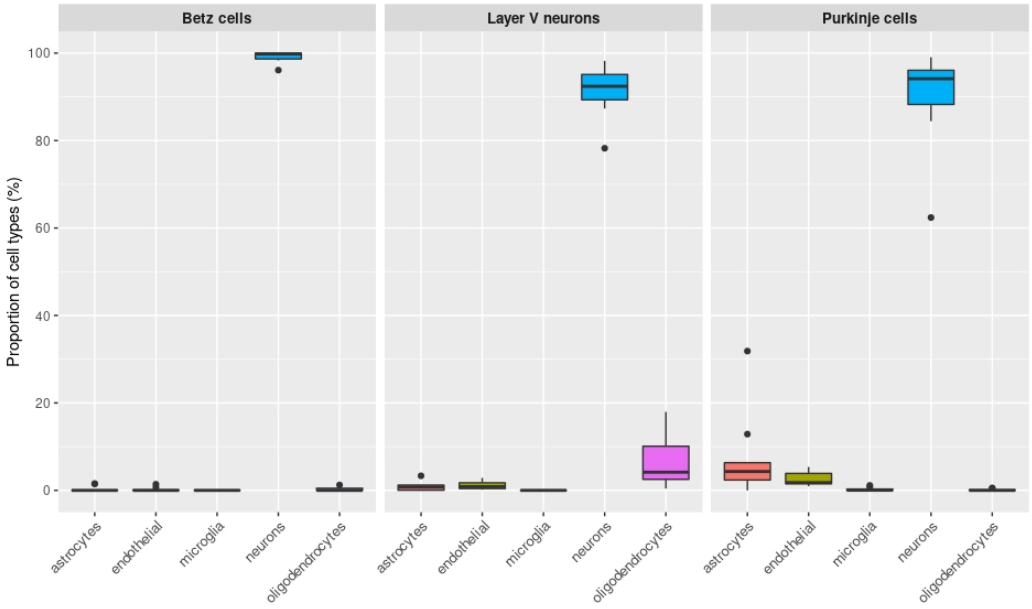


Figure 6.8: MuSiC Results of LCM-seq Cohort - Outliers Removed. Betz cells samples were highly enriched for neuronal-specific genes, with minimal contamination from other cell-types. L5 neurons were also enriched for neuronal genes, but an increase in oligodendrocytes was found, likely due to neuropil being microdissected. Purkinje cells samples were also enriched with neuronal genes, with increased contamination from astrocytes and endothelial cells.

identified in this data set; enrichment of mRNA was confirmed by the abundance of protein-coding RNA (21,712 IDs, 74.4% of RNA identified). Long non-coding RNA was the second most abundant type of RNA (3,185 IDs, 10.9% of RNA identified), followed by microRNA (1,899 IDs, 6.5% of RNA identified). The remainder of RNA consisted of snRNA, snoRNA, scaRNA, and RNA pseudogenes - see Figure E.11. For subsequent analysis, only protein-coding RNA was used; of these, 92.4% (17,443 IDs) were shared among all samples - Figure E.12.

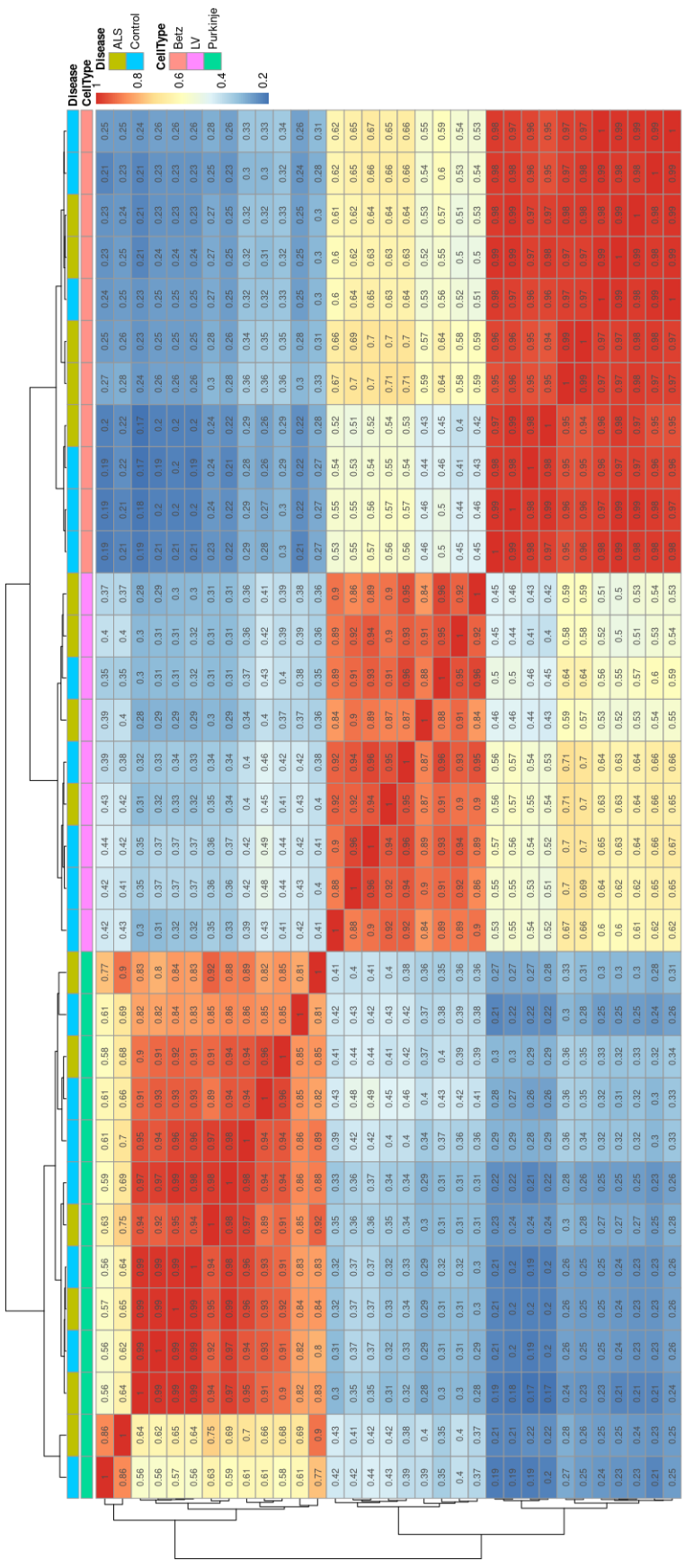


Figure 6.9: LCM-seq - Sample-to-Sample Distance Heatmap.
Unbiased clustering revealed that Betz cell samples were very similar to one another; the same can be said for L5 neurons and the majority of Purkinje cells samples. Two Purkinje cells appear to differ slightly, likely due to the increase in astrocyte/endothelial cells seen in the two samples. Betz cells and L5 neurons were more similar to each other than to Purkinje cells, with Purkinje cells being slightly more similar to L5 neurons than Betz cells. Within cell-types, samples did not cluster by disease type.

6.3.5 LCM-seq Analysis of the Human Betz cell

Differential gene expression analysis was first performed on the Betz and Purkinje cells from healthy individuals. Of the 19,144 transcripts used in this comparison, 18,673 (97.5%) were suitable for analysis. As expected, and echoing previous proteomic work, there were large differences in gene expression between the Betz and Purkinje cells - see Figure 6.10. Overall, 5,700 significantly differentially expressed genes (DEGs) were observed, with 3,444 (18%) being upregulated in Purkinje cells and 2,256 (12%) being upregulated in Betz cells; 13% of mRNA reads were excluded due to low counts and 2% were deemed as outliers and removed. Interestingly, the second-highest-ranking DEG in the Betz cell samples was ESYT1, mimicking previous experiments investigating these two cell-types at the protein level. The L5 specific marker, FEZF2 was also among the highest-ranking DEGs in Betz cells, along with all neurofilament proteins (NEFL, NEFM, NEFH) and the previously studied, KIF1A. Known Purkinje cell markers were also among the highest DEGs in the Purkinje cell samples i.e. ITPR1, CA8, ARHGAP26, and CALB1, again reflecting previous LC-MS/MS observations.

Next, the best performing Betz cells markers in the previous chapter were screened to see if their mRNA expression matched their protein expression - Figure 6.11. ESYT1 had the largest difference to L5 neurons, followed by CAPRIN2, REEP1, and DST. However, the 3 types of neurofilament were the most differentially expressed in the Betz cells. MX1 and PRUNE2, two of the top-ranked DEPs in the last chapter showed strong enrichment in Betz cells, despite IHC being negative (data not shown). FEZF2, ASGR2, and CSN1S1 were also screened as they were the defining molecules in the Betz-containing clusters in the latest taxonomy experiment of the primary motor cortex by the Allen Institute [163][4]. LCM-seq revealed a plethora of transcripts that were upregulated in Betz cell - more so than neurofilament or any of the markers identified in the previous chapter - see Figure E.17.

[4]The transcriptome of microdissected Betz cells were compared to these two clusters from the Allen Institute in Appendix E.5.

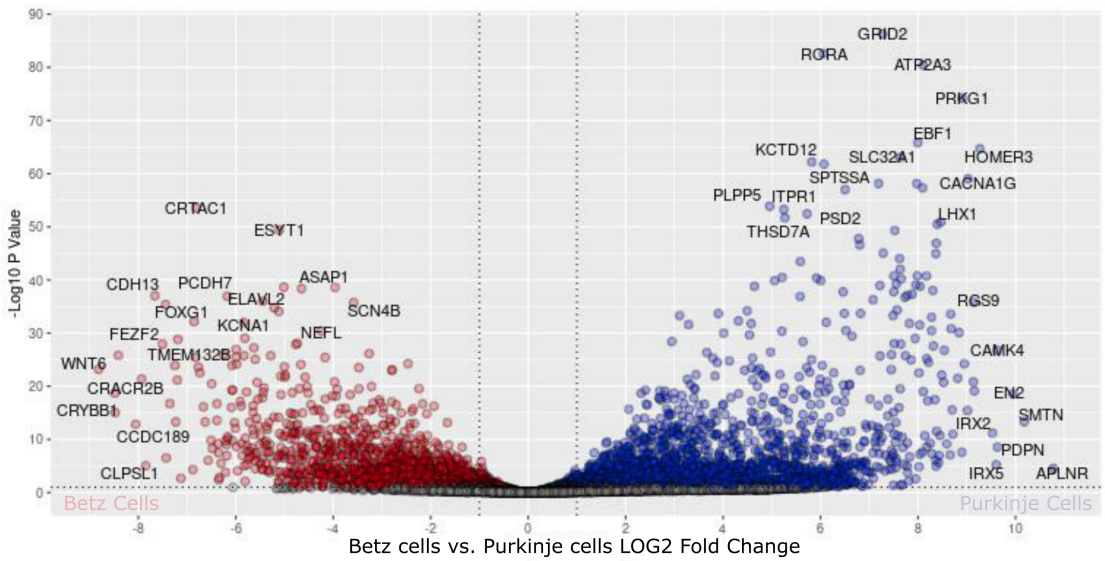


Figure 6.10: Volcano Plot - Betz Cells vs. Purkinje Cells (LCM-seq). Volcano plot - Betz cell vs. Purkinje cell. This volcano plot shows mRNA transcripts that were differentially expressed in Betz cells and Purkinje cells. As with the previous proteomic comparison, there was a pronounced difference between these two cell-types. Additionally, ESYT1, a top-performing protein in proteomic experiments was among the top-ranked mRNA transcripts in this experiment. Each point represents one mRNA transcript. Coloured points are DEGs that passed predetermined alpha thresholds (0.1). Red: mRNA whose expression was higher in Betz cells. Blue: mRNA whose expression was higher in Purkinje cell samples.

ORA on the 2,256 DEGs resulted in 215 GOBP, 44 MF, and 87 CC terms - see Table 6.2 for a summary of the most enriched pathways. The remaining GOBP terms highlighted enriched pathways that have previously been presumed to be particular important in Betz cell physiology; such as, glutamate secretion, calcium ion import, cytoskeleton-dependent intracellular transport, axogenesis, microtubule-mediated mitochondrion localization, modulation of synaptic transmission, regulation of neurotransmitter levels/transport, synaptic vesicle cycle, and synapse organization. Genes relating to the regulation of cellular component/anatomical structure size were also found to be enriched in Betz cells. Similar to previous proteomic investigations, these data show that Betz cells appear to be heavily dependent on cellular respiration, protein kinase A activity, organelle organisation, synapse regulation, and calcium-dependent activities; all of which can be explained by the large and active nature of the Betz cell. When investigating where the 2,256 DEGs tended to be located, several

significant terms were found, including actin cytoskeleton, juxtaparanode region of the axon, apical dendrite, contractile ring, exocytic vesicle membrane, glutamatergic synapse, SNARE complex, and transport vesicles; all of which reflect presumed dominant features in Betz cell physiology. Additionally, a strong enrichment in neuronal dense core vesicles was found (Fold enrichment: 5, adj. p value: 1.37e-03); these are regions of the neuron that are involved in the release of neuropeptides and neurotrophins and may explain the abundance of G protein-coupled receptor activity and binding repetitively seen in previous comparisons in this thesis. These data also indicate that Betz cells may have an enriched amount of serotonin, acetylcholine, catecholamine, and glucagon receptors, shining some light on the type of neurons that innervate them. Strangely, an enrichment (Fold enrichment: 3.21, adj. p value: 7.91e-03) of proteins known to localise to the Calyx of Held Synapse was found in Betz cells. This synapse giant axosomatic synapse is located in the auditory brain stem, and acts as a fast-firing synapse with many release sites, allowing the principal neurons of the medial nucleus of the trapezoid body (MNTB) to fire with high-frequency [333]. This may be an artefact of the hierarchical nature of gene ontology, however, it may also suggest a similar synapse in the motor cortex utilised by Betz cells, or at least reflect similar biology; perhaps the unusually large spines previous reported in Betz cells? [177, 207].

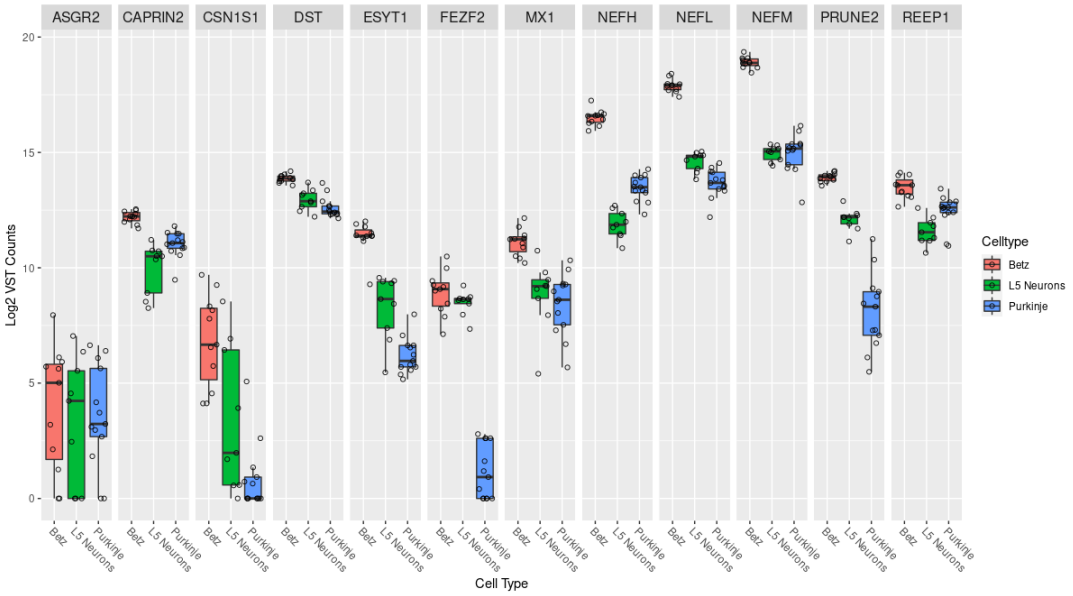


Figure 6.11: mRNA Expression of Previously Identified Biomarkers. Markers known to be enriched in Betz cells (NEFL, NEFM, and NEFH), best performing markers from the previous chapter (DST, REEP1, ESYT1, CAPRIN2), and transcripts shown by the Allen Institute to best characterise Betz cell-containing clusters (FEZF2, ASGR2, and CSN1S1) had their mRNA expression analysed in this data set. Neurofilament transcripts performed the best. DST, REEP1, ESYT1, and CAPRIN2 were all significantly increased in Betz cells compared to L5 neurons and Purkinje cells. PRUNE2 and MX1 - promising hits from the previous chapter were also significantly enriched in Betz cells. FEZF2, CSN1S1, but not ASGR2 were also significant up-regulation in Betz cells. VST: Variance Stabilizing Transformation.

GO Type	GO Term	Fold Enrichment	Adj. p value
GOCC	neuronal dense core vesicle	5.00028768699655	1.37e-03
	contractile ring	4.71455601202532	5.15e-03
	apical dendrite	3.92879746835443	1.66e-03
	juxtaparanode region of axon	3.92879746835443	3.44e-02
	L-type voltage-gated calcium channel complex	3.92879746835443	3.44e-02
	axolemma	3.92879746835443	7.25e-03
	axon initial segment	3.43769778481013	1.75e-02
	calyx of Held	3.21447065592635	7.91e-03
	main axon	2.94659810126582	1.45e-05
	SNARE complex	2.61919831223629	2.06e-03
	exocytic vesicle membrane	2.43018400104398	5.59e-05
	glutamatergic synapse	1.89205543801771	2.44e-07
	actin cytoskeleton	1.33782327586207	2.96e-02
GOBP	synaptic vesicle docking	4.9967890828818	6.58e-03
	spontaneous synaptic transmission	4.48691264585304	6.31e-03
	positive regulation of catecholamine secretion	4.28296207104154	3.36e-02
	neuromuscular process controlling posture	4.22805230089998	1.74e-02
	activation of protein kinase A activity	4.15699259836385	5.31e-03
	axonal transport of mitochondrion	3.92604856512141	2.71e-02
	retrograde axonal transport	3.53344361460927	1.55e-02
	serotonin receptor signaling pathway	3.36518448438978	2.13e-02
	response to acetylcholine	3.08475244402397	1.55e-02
	acetylcholine receptor signaling pathway	3.0200373577857	2.73e-02
	glutamate secretion	2.99127509723536	2.47e-03
	cellular response to glucagon stimulus	2.94453642384106	4.98e-02
	axonal transport	2.7551218000852	1.40e-03
	regulation of catecholamine secretion	2.7311642192149	6.31e-03
	positive regulation of axonogenesis	2.45969307694354	9.69e-04

GO Type	GO Term	Fold Enrichment	Adj. p value
	regulation of neurotransmitter transport	2.141481035552077	2.67e-03
	modulation of chemical synaptic transmission	2.0663413500639	7.86e-11
	cytoskeleton-dependent intracellular transport	1.8714842901O969	1.61e-03
	regulation of neurotransmitter levels	1.7812627749162	2.74e-03
	regulation of anatomical structure size	1.573878499674 22	8.42e-04

Table 6.2: Top Enriched GO Terms from Betz cell - Purkinje cell LCM-seq comparison.

Gene Set Enrichment Analysis (GSEA) analysing all 18,673 genes ranked by the stat variable generated by DESeq2 identified the same GO terms as being enriched in Betz cell samples; with SNARE complex, neuronal dense core vesicle, axolemma, and apical dendrite being most enriched. Dozens of pathways were enriched within Betz cells, as such only a subset of pathways are reported here - an overview of the most enriched pathways can be seen in Figure E.18. Several mechanisms were upregulated in Betz cells: organelle transport along microtubules, synaptic vesicle exocytosis, axonogenesis, and synaptic signalling. When investigating KEGG pathway differences between Betz and Purkinje: steroid biosynthesis and the TCA cycle were among the most enriched pathways, followed by lysosome activity, proteasome activity, axon guidance, endocytosis, cytoskeleton regulation, along with the MAPK signalling pathway, insulin signalling pathway, and terpenoid backbone biosynthesis. Several Reactome pathways also showed strong enrichment within Betz cells, including axon guidance, axon growth, action potential, and myelination: rapid depolarisation, recycling pathway of L1, SEMA3A PAK-dependent axon repulsion, L1CAM and LGI ADAM interactions. Other enriched Reactome pathways included Golgi apparatus to endoplasmic reticulum trafficking, an abundance of kinesins activity, and cholesterol biosynthesis. Interestingly, Betz cells also contained an up-regulation of genes related to selective autophagy and aggrephagy.

The above analysis was repeated for comparing Betz cells to L5 motor neurons in healthy individuals. Large differences were seen in mRNA expression between the two sample types; 1,604 DEGs (8.9%) were upregulated in the Betz cells, and 1,982 DEGs (11%) were upregulated in the L5 samples - Figure 6.12. The top 3 ranking Betz DEGs were the three types of neurofilament (NEFM, NEFL, and NEFH) all known to be abundant in Betz cells, and comparatively lower in surrounding L5 neurons. The 4th ranked gene was RAB3D, a protein transporter that is critical for regulating exocytosis, followed by the neuronal axonal transport gene KIF21A. GSEA revealed more upregulated pathways - see E.4.2 for a summary. Many terms mimicked those of the Betz - Purkinje GSEA, with lipid and carbohydrate processing, lipid transport, organelle transport, fatty acid transport, and vesicle transport being

among the most enriched in Betz cells. Along with terms related to ATP production and energy, there was a particular enrichment of pathways involved in the axonal and axodendritic transport of mitochondria, in addition to fatty acid transport across the mitochondria membrane. Curiously, GSEA also found enrichment in antigen processing and presentation. Comparing GOMF annotations between Betz cell and L5 neurons showed an increase in signalling receptor activity of nuclear import and nucleocytoplasmic carrier activity, MHC I protein binding, binding unfolded ER proteins, ubiquitin-protein transferase, and thioredoxin disulfide reductase activity. Next, differential KEGG and WIKI pathways were investigated; these highlighted that Betz cells had an increase in pyruvate metabolism, TCA cycle, gluconeogenesis, and aminoacyl tRNA biosynthesis - echoing the previous chapter. In addition to enrichment in genes found in the proteasome endopeptidase complex and endolysosome, there was also a strong enrichment in terpenoid backbone biosynthesis as seen in the Betz-Purkinje GSEA.

Next, DEGs from the above Betz-Purkinje and Betz-L5 comparison were combined to identify a subset of transcripts that were consistently upregulated in Betz cells - see Figure 6.13. In total, 754 DEGs were found that met this criterion. Promisingly, several of the previously validated Betz "markers" from the previous chapter were found among these 754 DEGs, including DST, ESYT1, REEP1, and CAPRIN2. Figure 6.14 highlights some of the common pathways shared between these DEGs. Unsurprisingly, several cytoskeleton genes were found to be consistently enriched in Betz cells, including NEFL, NEFM, NEFH, KIF21A, MAP1B and several tubulin genes. Within environmental information processing, a large proportion of these DEGs were associated with ion channels, G protein-coupled receptors, or involved in a variety of signalling pathways including the Wnt, MAPK, Erb8, PI3K-AKT, FoxO, and calcium signalling. There was also a tendency for transcripts related to ubiquitin/proteasome activity to be enriched within Betz cells. Both the proteomap in Figure 6.14 and ORA (see Table E.2) highlighted an enrichment of lipid and steroid metabolism among Betz cells, including several genes involved in low-density lipoprotein particle clearance, in addition to proteins

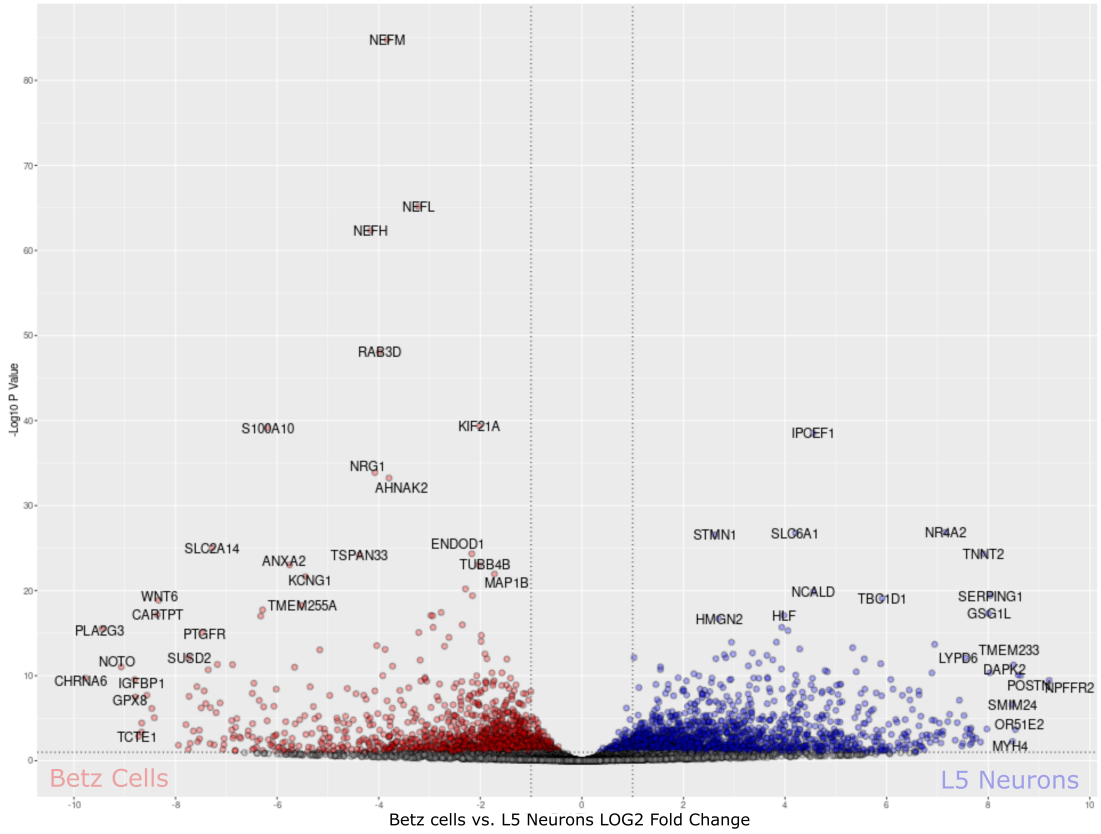


Figure 6.12: Volcano Plot - Betz Cells vs. L5 Neurons (LCM-seq)
This volcano plot shows mRNA transcripts that were differentially expressed in Betz cells and L5 neurons. As with the previous proteomic comparison, there was a pronounced difference between these two cell-types. Interestingly, the top ranking DEGs were the different chains of neurofilament and molecules known to be involved in axonal transport - proteins/pathways shown in this thesis to be enriched within Betz cells. Each point represents one mRNA transcript. Coloured points are DEGs that passed predetermined alpha thresholds (0.1). Red: mRNA whose expression was higher in Betz cells. Blue: mRNA whose expression was higher in L5 samples.

genes localised involved in retrograde axonal transport, dynein heavy chain binding, endolysosomes, and antigen presentation.

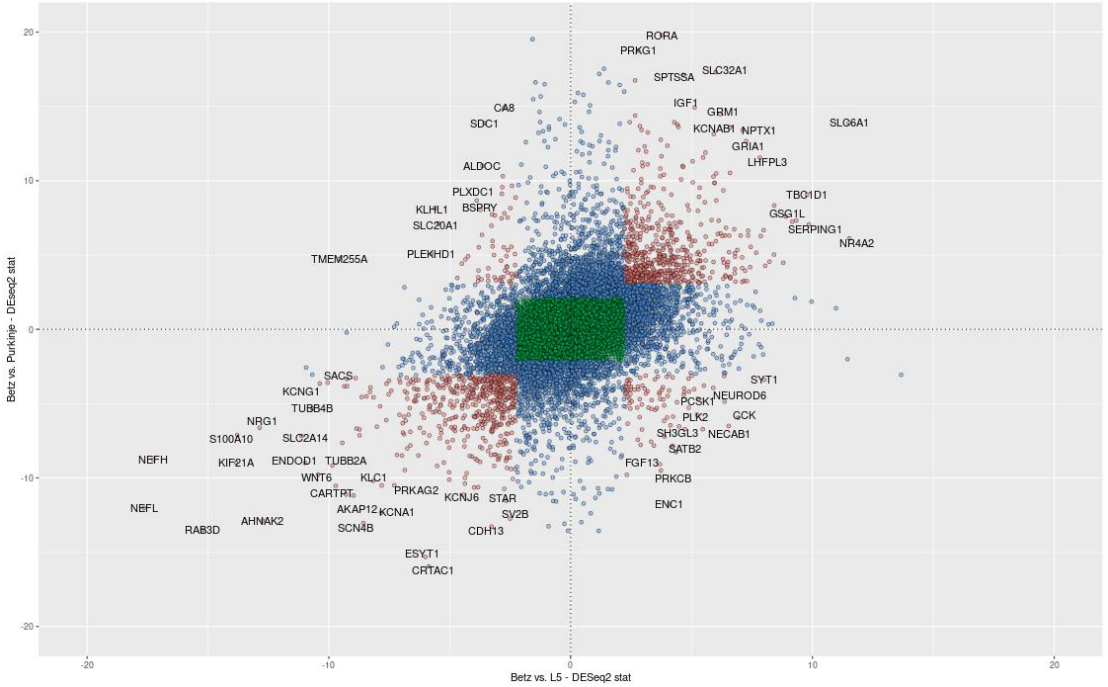


Figure 6.13: Differentially Expressed Genes from All Cell-Type Comparisons. Differentially expressed genes (DEGs) from all healthy cell-type comparisons were plotted. Top Left: Transcripts increased in Betz cells in Betz vs. L5 comparisons but decreased in Betz cells in Betz vs. Purkinje cell comparisons. Top Right: Transcripts down-regulation in Betz cells in both Betz vs. Purkinje and Betz vs. L5 neuron comparisons. Bottom Left: Transcripts enriched in Betz cells in both Betz vs. Purkinje and Betz vs. L5 neuron comparisons (i.e. the previously mentioned 754 DEGs). Bottom Right: Transcripts decreased in Betz cells in Betz vs. L5 comparisons but increased in Betz cells in Betz vs. Purkinje cell comparisons.

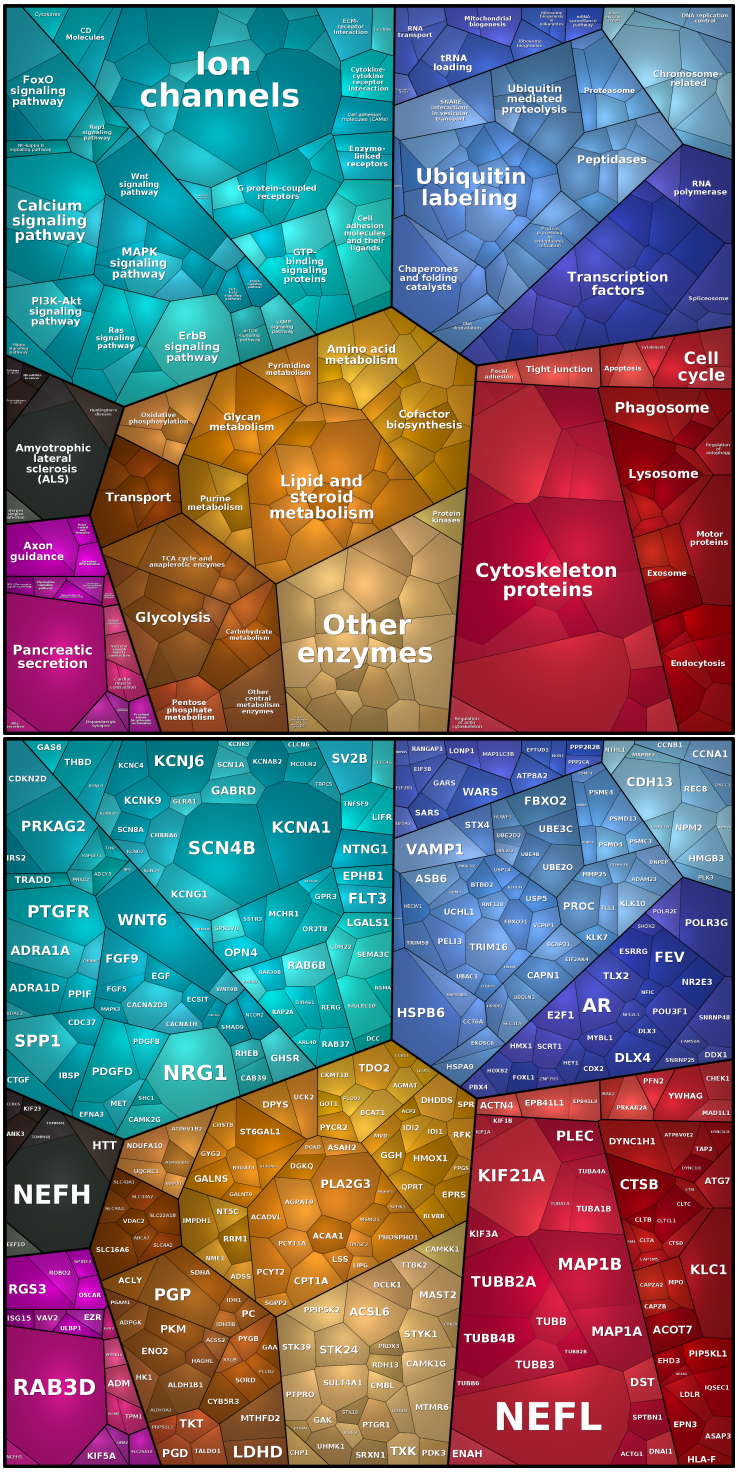


Figure 6.14: Treemap of the 754 Betz Cell-Enriched Transcripts.
This expression-weighted treemap was generated with the Proteomap tool (Liebermeister, et al., 2014) using the 754 differentially expressed genes (DEGs) found to be consistently enriched in Betz cells in LCM-seq cell-type comparisons. Each polygon represents one mRNA transcript and the size reflects its relative expression. Note: Not all 754 Betz cell DEGs were compatible with this approach; only 407 DEGs were supported.

6.3.6 LCM-seq Analysis of ALS Neurons

Next, the transcriptome of Betz cells from healthy tissue was compared to Betz cells from ALS brains. Only a small subset of DEGs was identified. With 88 (0.49%) of genes being upregulated in ALS Betz cells and 69 (0.39%) being downregulated in ALS Betz cells - see Figure 6.15. While only 157 DEGs were found in total, 71.94% had a LogFC >2 or more. Although ORA showed no gene ontology enrichment among these 157 DEGs, several of the DEGs were known to be involved in spliceosome activity (DHX8, PHF5A, SYF2, HSPA2, MAGOHBM), regulation of cell growth (KAZALD1, DDR1, FAM107A, MELTF, NANOS1), response to DNA damage (DDIT3, FANCL, CCDC13, NABP2, and PLK3), and transcription factor activity (DDIT3, FOS, PHF5A, TGIF1, ATF3, RUNX1, TADA2A, ZCAN29, ZNF121, ZNF217, ZNF451, ZNF567, ZNF782).

GSEA on these 17,901 transcripts highlighted a plethora of pathways that were differentially up-regulation in Betz cells with and without ALS, of which the most enriched pathways will be discussed here. Promisingly, GSEA found that ALS Betz cells had a higher expression of genes relating to the Amyotrophic Lateral Sclerosis KEGG pathway. The most enriched pathways involved in ALS Betz cells were ones associated with glutamate synaptic transmission, programmed necrotic cell death, and protein kinase A activity (See Appendix E.4.3 for the summary of the top pathways). Furthermore, glutamate receptor binding, glutamate receptor activity, and glutamate and glutamine metabolism were also found to be enriched in ALS Betz cells. ALS Betz cells also had an increase in both synaptic long-term depression and long-term potentiation. Moreover, genes relating to axon guidance, netrin signalling, and neurogenesis were also enriched in ALS Betz cells and may reflect the aberrant axonal branching that has been previously reported in ALS. Phosphatidylinositol signalling, GNRH signalling, MAPK signalling, and ERBB signalling were all also found to be enriched within ALS Betz cells. Curiously, Vitamin D metabolism was one of the most enriched pathways in the ALS Betz cells, along with protein kinase A activation of glucagon signalling and trafficking of AMPA receptors. Several immune-related pathways were also upregulated in

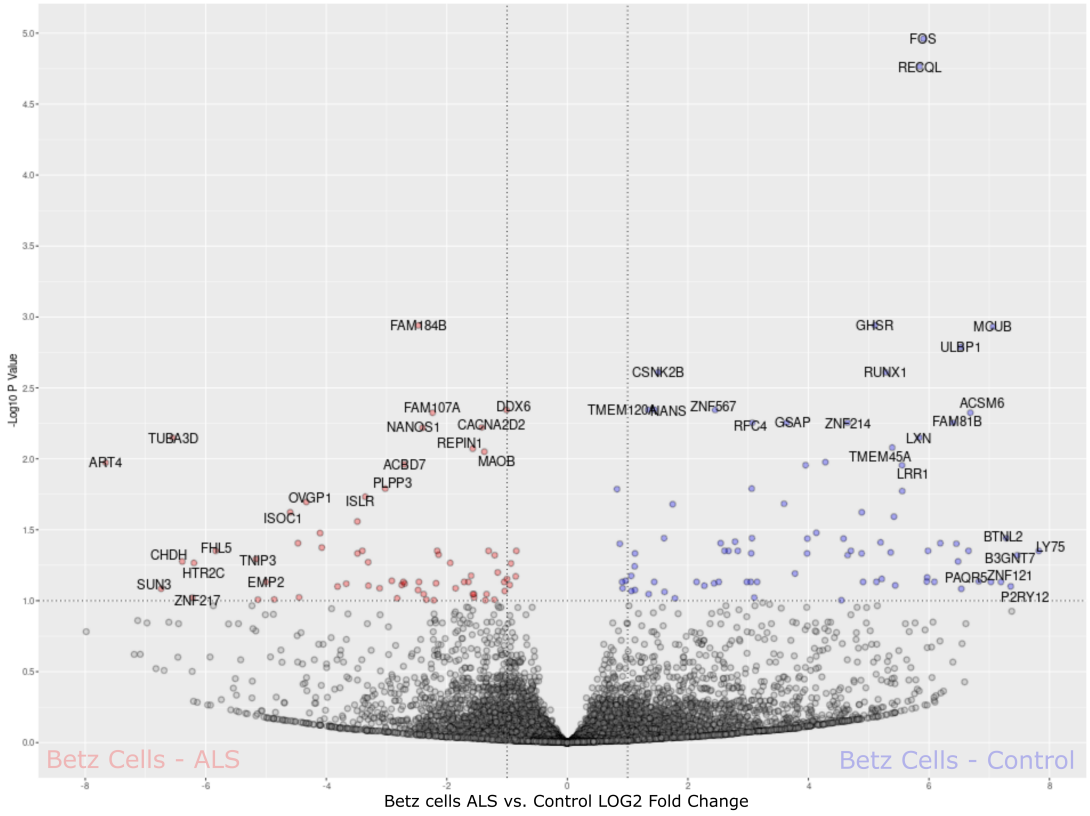


Figure 6.15: Volcano Plot - Healthy Betz Cells vs. ALS Betz Cells. Several mRNA transcripts were found to be differentially expressed between Betz cells in the healthy and ALS brain, with FOS and RECQL being particularly affected. Each point represents one mRNA transcript. Coloured points are DEGs that passed predetermined alpha thresholds (0.1). Red: mRNA whose expression was up-regulation in ALS Betz cells. Blue: mRNA whose expression was down-regulation in ALS Betz cells.

ALS Betz cells, including interferon 1 and gamma activation, which may reflect the pronounced reactive gliosis that is known to surround Betz cell perikaryon in ALS.

Interestingly, several pathways relating to mRNA metabolism were also enriched in ALS Betz cells, including pathways relating to alternative mRNA splicing, pre-mRNA intronic binding, regulatory RNA binding, mRNA 3' UTR binding, mRNA decay by 5' to 3' exoribonuclease, and miRNA binding, pathways known to be affected in ALS.

In contrast, several pathways were found to be comparatively diminished in ALS Betz cells; the majority of which are involved in ATP synthesis. Other diminished metabolic pathways included NADH dehydrogenase activity, carboxyl

lyase activity, FAD-binding, pregnenolone biosynthesis, growth factor activity and several pathways relating to fat metabolism. Finally, several mitochondrial pathways were aberrant in ALS Betz cells; tRNA modification in mitochondria, mitochondrial protein import, mitochondrial fatty acid oxidation of saturated fatty acids, and respiratory electron transport were all reduced in ALS Betz cells.

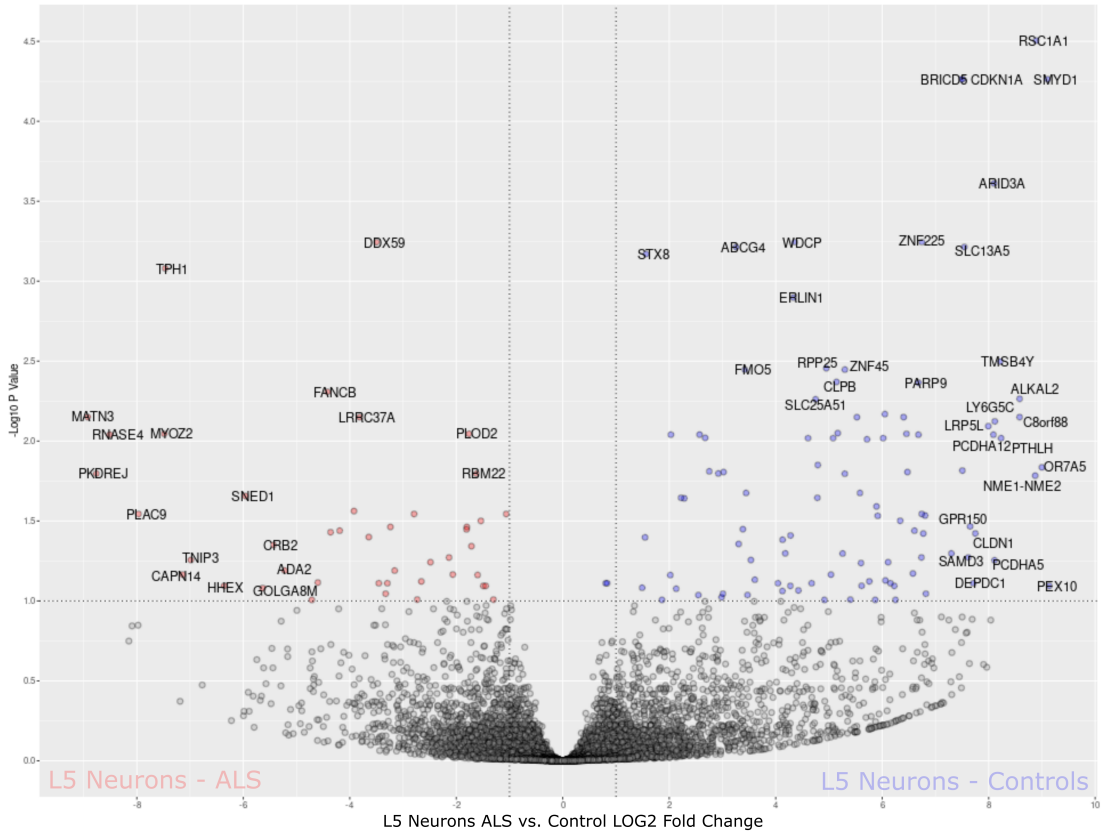


Figure 6.16: Volcano Plot - Healthy L5 Neurons vs. ALS L5 Neurons. Several mRNA transcripts were found to be differentially expressed between L5 neurons in the healthy and ALS brain. In contrast to Betz cells, there appeared to be a large difference between L5 neurons in health and ALS - and may reflect potential pathogenic mechanisms in the microenvironment of L5. Each point represents one mRNA transcript. Coloured points are DEGs that passed predetermined alpha thresholds (0.1). Red: mRNA whose expression was up-regulation in ALS L5 neurons. Blue: mRNA whose expression was down-regulation in ALS L5 neurons.

healthy L5 neurons were compared with their diseased counterparts; which resulted in identifying 148 DEGs, 104 up-regulation in ALS L5 samples, and 44 down-regulation in ALS L5 samples. As with the Betz cell samples, ORA showed no

gene ontology terms were enriched between DEGs. Several of the DEGs identified were known to be involved in ALS: with RNASE4, SNED1, and NXT1 being up-regulation in ALS L5 samples, and ERLIN1, PARP9, and FANCD2 being down-regulation in ALS L5 samples. Unlike Betz cell samples, GSEA of L5 samples highlighted few pathways that were up-regulated in either healthy or diseased L5 neurons. This may be a result of the reduced statistical power of excluding x4 L5 samples during the LCM quality control checks. It is well-established in the literature that L5 motor neurons are affected in ALS, and this is reflected in the large number of DEGs identified in this comparison. However, it was not possible in our data to determine exactly which pathways were aberrantly affected. Although it was not possible to use microdissected L5 samples to determine the molecular cause of the differential ALS pathology seen in L5 neurons and Betz cells; when L5 neurons and Betz cells were merged and their combined transcriptome compared to ALS counterparts, several of the pathways are seen in the Betz healthy vs. ALS comparison were seen e.g. an increase in response to UV response (NES: -1.61; padj: 4.4e-2), chromatin-modifying enzymes (NES: 1.67; padj: 4.5e-02) and downregulation of energy generation - data not shown.

Although the cerebellum is typically thought to be spared in ALS, it was decided to still investigate if the transcriptome of Purkinje cells were affected and to ensure Betz cell findings were not due to bioinformatic error. Only 22 DEGs were identified; with 8 significantly higher in ALS Purkinje cells and 14 significantly lowered in ALS Purkinje cells. GSEA revealed several pathways that were downregulated in ALS Purkinje cells but very few were upregulated. Similar to the Betz cell samples, there was a decrease in pathways related to energy generation, specifically, oxidative phosphorylation, while transcripts involved in the connexin complex were enriched within ALS Purkinje cells. Several nucleic acid binding pathways were also enriched within ALS Purkinje cells; KEGG analysis found ascorbate and aldarate metabolism and RIG-I-like receptor activity was increased among the ALS Purkinje cells. A small reduction in the expression and signalling of ROBO and SLIT receptors

were also found in ALS Purkinje cells. Finally, pathways involved in heat shock activation via HSF1 was decreased in ALS Purkinje cells.

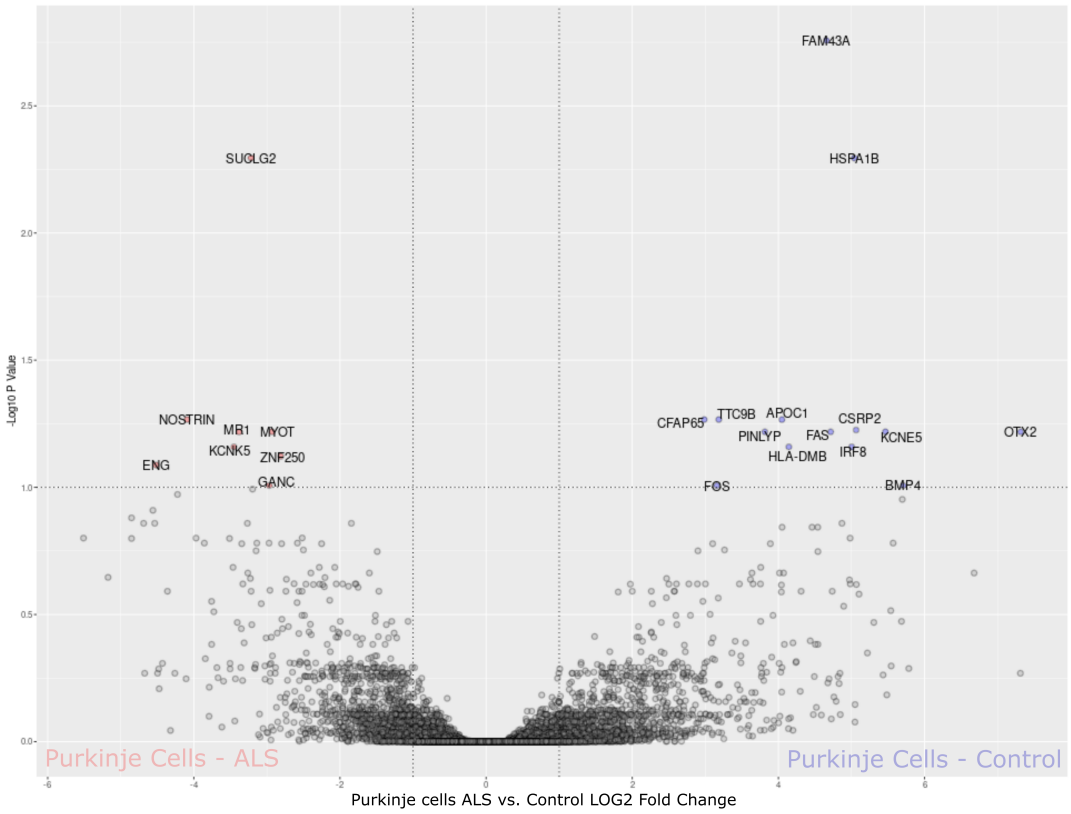


Figure 6.17: Volcano Plot - Healthy Purkinje Cells vs. ALS Purkinje Cells. Although thought to be spared in ALS neuronal degeneration, several mRNA transcripts were found to be differentially expressed between Purkinje cells in the healthy and ALS brain. More work is needed to determine if these changes are a direct result of ALS or instead reflect a compensatory mechanism. Each point represents one mRNA transcript. Coloured points are DEGs that passed predetermined alpha thresholds (0.1). Red: mRNA whose expression was up-regulation in ALS Purkinje cells. Blue: mRNA whose expression was down-regulation in ALS Purkinje cells.

Of all the DEGs identified in this LCM-seq experiment, very little had an overlap across cell-types; potentially reflecting cell-type-specific ALS changes. Of the 3 ALS cell-type comparisons, only 5 DEGs were found in more than one cell-type - Figure 6.18. NOSTRIN was upregulated in Purkinje and L5 ALS samples, FOS was downregulated in Betz and Purkinje cells from ALS patients; L5 samples appeared to also be downregulated but did not pass pre-determined alpha thresholds. TNIP3 was upregulated in ALS Betz and L5 neurons, while differentially regulation was found in Betz and L5 ALS neurons for FANCD2. ALS-associated MAOB was found to be upregulated in both proteomic and transcriptomic comparisons of ALS Betz cells to controls, the same pattern was seen with HSPA2, a heat-shock protein believed to be a key driver of Alzheimer's disease [334]. The opposite was seen with ISCA1, a protein involved in mitochondrial iron-sulfur cluster biogenesis, and mutations in which, are known to cause multiple mitochondrial dysfunction syndromes [335].

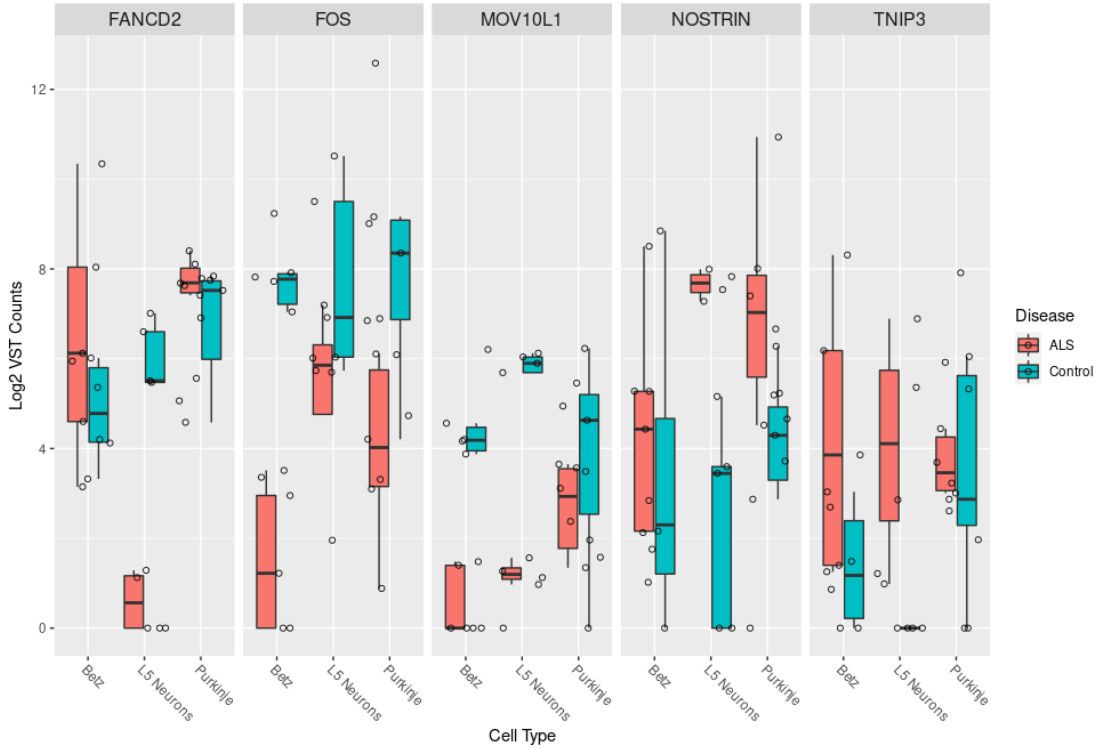


Figure 6.18: Differentially Expressed Genes found in Multiple Cell-Types. A subset of mRNA transcripts that were found to be differentially expressed across multiple cell-types when comparing ALS vs. Control cases.

6.4 Discussion

6.4.1 The Betz Cell in the Healthy Brain.

Using the optimised high-throughput approach it was possible to analyse 54 Betz cell samples, comprised of 3,111 Betz cells, from 13 post-mortem brains. The generated data set has identified 6,390 proteins that provide, to date, the deepest proteomic expression data on the human Betz cell, and the only one to generate data from multiple post-mortem brains. The output from this experiment showcases the ability of the high-throughput approach, as the similar proteomic performance was achieved as in previous experiments, despite the 71.6% decrease in gradient time; reducing the resources needed and the cost associated with large-scale LCM proteomic projects. This chapter revealed and compared the proteome of the Betz cell, L5 Motor neuron, and the Purkinje cell. A complementary approach was also performed to reveal the transcriptome of these neurons and how they differ from other neurons, in health and disease. As the LCM proteomics recapitulated observations seen in earlier experiments, the results of LCM-seq were reported in more depth.

Both LCM proteomics and LCM-seq analysis revealed a magnitude of cellular mechanisms that were enriched within healthy Betz cell samples. LCM-seq, which identified $\approx$ 18,000 genes vs. $\approx$ 6,000 proteins, revealed additional novel insights into the cellular mechanism enriched in Betz cells, e.g. response to glucagon, acetylcholine, serotonin, and catecholamines. These data may reflect a vast receptor diversity within Betz cells, something that has been investigated in previous electrophysiological experiments - see B.2, and provides some insights into the type of signalling the Betz cell receive in the motor cortex. A large (5 Fold enrichment) in neuronal dense-core vesicles (DSVs) was found in Betz cells, these vesicles are used to store and release neuropeptides and neurotrophins, and may in part contribute to the high exocytosis, trans-Golgi network, and G protein-coupled receptor enrichment seen throughout the Betz cell data in this thesis. Typically, DSVs are used to transport neuropeptides, monoamines, and neurotrophic factors away from the soma and are involved in modulating synaptic transmission [336].

It is currently unknown what neuropeptides, if any, are synthesised by the Betz cell, discovering this may provide insights on how Betz cells regulate and maintain their excitatory signalling to faraway targets.

A surprising feature of the LCM-seq data was the large differences seen between Betz cells and L5 neuron samples, although the L5 samples would be expected to contain some contaminating cell-types, earlier MuSiC analysis indicated a strong enrichment in neuronal transcripts, yet when compared to Betz cells a large number of DEGs were identified. When compared to L5 samples, one of the most enriched pathways was the thioredoxin reductase system, which has been shown to repair oxidative damage to neuronal tubulin [337]. It remains to be seen if Betz cells contain additional protection mechanisms to oxidative damage since they are likely prone to oxidative stresses. LCM-seq revealed over a thousand DEGs between Betz cell and L5 samples, any of which may represent a potential "gold standard" Betz cell marker. While neurofilament remains to be one of the most differentially expressed proteins in the Betz cell, both in transcriptomic data, proteomic data, and in IHC investigations, these LCM-seq data highlights 174 transcripts that were more differentially enriched in Betz cells than NEFM. Future work should attempt to screen these transcripts in a high-throughput manner for the best chance of finding a "gold standard" Betz cell biomarker.

Alternative approaches are being developed to study cell-types (in health and disease) without the need for immuno-labelled antibodies. While mass spectrometry-based imaging e.g. MALDI (matrix-assisted laser desorption/ionization) imaging is only starting to achieve single-cell resolution, the gigantic size of the Betz cell places it in good stead to be one of the first to be analysed by such techniques - allowing the molecular profile of the Betz cell to be investigated without the need of histological dyes, immuno-based labelling, or LCM. Preliminary in-house MALDI imaging work has managed to identify Betz-specific profiles from unstained sections via unbiased targeting of the proteome, metabolome, and lipidome of these neurons - data not shown. Although more work is needed to identify what molecules are responsible for this Betz-specific signature. Linking LCM-derived data with whole

section MALDI imaging may provide a worthwhile approach to screening potential biomarkers bioinformatically, rather than labour-intensive IHC.

6.4.2 The Betz Cell in the ALS Brain.

Several transcripts known to be associated with ALS were identified as DEGs in the ALS Betz cell data, most notably being FOS, PCP4, and MAOB. While the overall number of DEGs was low, many cellular pathways were found to be upregulated/downregulated within ALS Betz cells, likely reflecting mechanistic dysfunction. The diversity of such pathways may reflect the complex pathophysiology of ALS. Only the cellular pathways that were the most enriched will be discussed herein.

The most obvious finding in the transcriptomic comparison between healthy and ALS Betz cells was the increase in the expression of transcripts relating to glutamate receptor binding and activity, glutamatergic synaptic transmission, and trafficking of AMPA receptors within ALS neurons. An increase in glutamate signalling was also seen when comparing the proteome of the healthy and ALS Betz cells. This enrichment largely favours the excitotoxicity hypothesis in ALS, in that neurons are progressively damaged by the influx of Ca^{2+} ions from glutamate receptors. These Ca^{2+} ions accumulate, causing cellular processes to accelerate unchecked; and among other damaging mechanisms, cause a significant amount of oxidative stress which eventually results in neuronal degeneration. It is not known if this excitotoxicity occurs due to the overactivation of glutamate receptors or if weakened neurons become more prone to normal synaptic levels of glutamate; however, these data suggest the former. It is believed that glutamatergic neuronal injury occurs via Ca^{2+} influx through either NMDA receptors, AMPA receptors, or voltage-gated Ca^{2+} channels [338]. Motor neurons are believed to be particularly susceptible to the deleterious effects of excess Ca^{2+} ions as they lack the calcium-binding proteins present in other neurons allowing calcium levels to build up with little regulation. This susceptibility is very sensitive and can be triggered just by the consumption of electroactive foods. Previous CALB1 staining has shown that Betz cells do not process this type of calcium-binding protein, and our LCM-seq/proteomics data

also highlights the decreased expression of calcium-binding proteins in Betz cells. As a result of lowered Ca^{2+} buffering proteins, it is believed that the mitochondria of motor neurons play a larger role in calcium buffering [338]. Mitochondrial overloading from excess Ca^{2+} has been shown *in vitro* to lead to excess depolarization and increased generation of damaging reactive oxygen species (ROS) [338]. Curiously, MCUB, a mitochondrial protein that negatively modulates calcium uptake into the mitochondria (via MCU) was the 4^{th}highest-ranked downregulated transcript (LogFC: 7.06) within ALS Betz cells. The exact mechanism behind calcium-mediated neuronal death is not fully understood. Along with the generation of ROS, it is thought that Ca^{2+} ions over-activate several enzymes which lead to degradation, including phospholipases, lipases, proteases, phosphatases, endonucleases, protein kinase C, NO synthase, and xanthine oxidase [338].

The up-regulation of preprogrammed cell death pathways within Betz cells in ALS individuals adds to the wealth of evidence that these neurons are degraded in ALS. Although an established feature of ALS, it remains to be seen if apoptosis is a direct result of ALS pathogenesis or if motor neurons only succumb to apoptosis once they have exhausted functional capacity i.e. they have already undergone excessive atrophy, axonal degeneration, synapse loss, neurite retraction, and cellular stress [339, 340]. Interestingly, blocking the mitochondrial apoptotic pathway halted neuronal loss and prevented axonal degeneration leading to a better pre-clinical outlook; suggesting the former is true [341].

Another highly enriched pathway within ALS Betz cells was vitamin D metabolism. ALS is a multi-faceted disease, believed to cause glutamate excitotoxicity, neuroinflammation, oxidative stress, protein aggregation, axonal transport dysfunction and mitochondrial dysfunction. It is believed that many of these mechanisms are modulated by Vitamin D, leading to the conclusion that Vitamin D exhibits a neuroprotective role in ALS [342]. For example, vitamin D can suppress MHC class II antigens and enhance Toll-like receptors, reducing inflammation. It has also been shown to help increase the expression of calcium-binding proteins, along

with reducing damage caused by oxidation by decreasing TNF-α, IL-1β, and nitric oxide synthase. Vitamin D can also increase microtubule-associated protein, growth-associated protein, and synapsin proteins providing additional neuroprotective effects; and has been shown to protect against the neurotoxicity of excess glutamate exposure. *in vitro* motor neurons receiving vitamin D protected neurons from FAS-induced death [343]. In humans, low levels of Vitamin D has been reported in ALS patients [343]. The increase of Vitamin D metabolism in ALS Betz cells may reflect a compensatory mechanism of these neurons to try and resolve the deleterious mechanisms of excitotoxicity, inflammation, and oxidative stress. Conversely, Vitamin C metabolism was decreased within ALS Betz cells; Vitamin C is a potent antioxidant and has also been shown to prevent excitotoxic damage caused by excessive extracellular glutamate [344], by inhibiting the binding of glutamate to NMDA receptors, thereby, preventing excessive neuron stimulation [345].

Dysregulated energy metabolism has been established as a feature within ALS patients and the notion that ALS is a metabolic disease is beginning to gain traction with increasing evidence that mitochondrial and glycolytic energy metabolism is impaired, along with cellular energy homeostasis disruption within ALS neurons and glia [346]. Both proteomic and transcriptomic data showed that energy supply pathways were consistently downregulated in ALS Betz cells; namely: ATP Synthesis via electron transport, mitochondrial respiratory chain assembly, oxidative phosphorylation, sugar metabolism, and fatty acid metabolism. Moreover, GSEA/1D enrichment of protein localisation showed that among downregulated transcripts, there was an enrichment of those localised to mitochondria.

Transcriptomic/Proteomic analysis also indicated many other pathways that were impaired in ALS Betz cells; axonal guidance transcripts were enriched in diseased Betz cells, likely indicative of the axonal branching see in ALS motor neurons [347]. Transcripts involved with interferon/gliogenesis/complement pathways were also differentially expressed in ALS Betz cells, perhaps reflective of the neuroinflammatory/reactive gliosis that is commonplace in ALS? [202, 348].

There was an increase in the sumoylation of DNA methylation proteins/RNA-binding proteins, along with enrichment in mRNA splicing and binding - both emerging features of ALS [349–352]. Along with miRNA binding, there was also an up-regulation of lipid/phospholipid-binding, chromatin binding, ER retention sequencing binding, LYS63-specific deubiquitinase activity, Pi-bodies, heat-shock protein pathways and an increase of neurotrophic tropomyosin receptor kinases in ALS Betz cells. While some of these pathways echo the same pathways known to be affected in the pathogenesis of ALS, some do not, further work is needed to validate if these pathways are truly affected in ALS, and if so, are they a primary or secondary outcome to ALS, and are they unique to the Betz cell? Finally, pregnenolone metabolism, terpenoid backbone synthesis, mitochondrial protein import, and tRNA mitochondrial synthesis were all reduced in ALS Betz cells.

Despite both being degenerated in ALS there is surprisingly little overlap in the individual transcript and proteins that were aberrantly expressed in the ALS Betz cells and ALS L5 samples. Nevertheless, some overlap was seen. FANCD2 was found to be highly down-regulation in L5 ALS samples but slightly up-regulation in ALS Betz cells, compared to their healthy counterparts. FANCD2 has been shown to maintain genome stability and is involved in the repair of DNA double-strand breaks (DSBs), a potentially devastating molecular event. While there is little information in the literature on the role of FANCD2 in ALS, there is considerable evidence of DSBs in ALS [353]. The increase of FANCD2 in these ALS Betz cells may reflect a compensatory mechanism to repair excess DSBs that have occurred due to the loss of TDP-43. Oxidative stress has also been shown to cause DSBs and so has the over-expression of SOD1. Moreover, proteins involved in UV DNA response were enriched in ALS Betz cells. Additionally, when examining transcripts upregulated in ALS Betz cells, there was also an enrichment of transcripts associated with the mismatch repair complex. NABP2, a critical component of DNA damage repair was also upregulated in ALS Betz cells. Along with repairing DSBs, NABP2 has been shown to repair 8-oxoG lesions (one of the most common DNA lesions caused by ROS) in response to oxidative stress [354]. The increase of NABP2 and FANCD2 in ALS Betz

cells, along with the enrichment of genes in the UV DNA response and mismatch repair pathways provides evidence of excessive DNA damage within ALS Betz cells. Contradictory, several other DNA damage repair proteins were downregulated in ALS Betz cells (CCDC13, DDIT3, FANCL, PLK3) and it remains unclear why a large FANCD2 reduction (LogFC: 6.45) was found within ALS L5 neurons.

MOV10L1 was also found to be down-regulation both in ALS Betz and L5 neurons. The complete role of MOV10L1 is unclear, however, it is believed to be a master regulator of piRNA (Piwi-interacting RNA) biogenesis [355, 356]. The mechanism behind piRNA biogenesis is not known, however, there is some evidence that MOV10L1 binds to piRNA precursors to initiate piRNA processing [355]. piRNAs are the largest class of small non-coding RNA molecules and piRNA complexes are believed to be mostly involved in the epigenetic and post-transcriptional silencing of transposable elements. Little is known about the role of piRNA in the nervous system, even though around 30,000 neuronal piRNAs have been characterized in the mouse brain [357]. Although there is evidence that some piRNAs have been suggested to be involved in axonal regeneration; and may play a role in neurodegeneration [357]. Along with being a master regulator, MOV10L1, also associates with Piwi proteins and RNA-binding proteins. Growing evidence suggests that the dysregulation of piRNA pathways results in genomic instability within neurons leading to the development of various neurodegenerative disorders, such as ALS and tauopathies. In a FUS knockdown Drosophila ALS model, an increase in the number of pre-piRNAs was observed; the data from this study suggests that FUS knockdown causes abnormal pre-piRNAs to translocated and accumulate in the cytoplasm, contributing to the pathogenesis of ALS [358]. This mechanism is similar to the ALS RNA processing seen in mutations of other RNA-binding proteins. The fact that MOV10L1 expression was affected in our ALS samples, is the master regulator of pre-piRNA to piRNA biogenesis, in addition to piRNAs being shown to be aberrantly translocated in affected cells in ALS models - mimicking ALS protein pathology, along with the enrichment of Pi-bodies in ALS Betz cell data makes MOV10L1 a very interesting protein to study in future

experiments. Moreover, there is some evidence that MOV10L1 may have the ability to remodel other ribonucleoproteins [355], which is of interest to the ALS community.

ALS pathology is typically thought to be restricted to upper or lower motor neurons. However, emerging evidence is beginning to investigate how the cerebellum may be affected by this multi-system disease. The cerebellum is believed to be spared in the pathogenesis of ALS, however, some evidence has indicated that the cerebellum is biochemically affected despite the lack of support from imaging studies [359]. It is known that cytoplasmic inclusions maybe represent ALS pathology in the granule layer of the cerebellum in individuals with C9orf72 [360]. Some have observed selective Purkinje cell loss in some genotypes of ALS but not others. Our data suggest Purkinje cells are affected in ALS, albeit, less so than motor neurons; with only 22 DEGs identified between healthy and ALS Purkinje cells. Further work is needed to determine if this is a direct result of ALS or in response to ALS pathology elsewhere in the motor system. NOSTRIN (Nitric Oxide Synthase Trafficking) is a potent mediator of neurotransmission, inflammatory response, and vascular homeostasis [361]. It binds ENOS, the enzyme responsible for nitric oxide (NO) production. While NO is typically used to modulate neuronal behaviour, it has been shown to amplify oxidative damage by reacting with superoxides to produce peroxynitrite. In ALS, physiological levels of NO promote the survival of motor neurons, however, this same concentration has also been shown to promote neuronal apoptosis and glial cell activation [362]. It is unclear if the increase of NOSTRIN in L5 samples reflects neuronal or glial physiology or if it is a marker of oxidative stress. The fact that NOSTRIN was also upregulated in ALS Purkinje cells, raises the possibility that these neurons may be subject to oxidative stress, similar to motor neurons. However, it is not known which molecular pathways would cause these inhibitory neurons to become stressed. Conversely, the biochemically changes seen in ALS may not reflect a disease degenerative process, but rather a compensatory mechanism of the cerebellum responding to the degeneration of the associated motor pathways.

FOS (or c-FOS), was down-regulation in both ALS Betz cell and Purkinje cell samples. FOS is often used as a proxy for neuronal activity as it is often expressed when neurons fire action potentials [363]. Contradictory, it has also been suggested that an increase in FOS expression reflects a neuron in an excitotoxic state and increased FOS expression has been noted in ALS rodent spinal cords cultured with CSF from ALS patients [364]. It is unclear why FOS would be reduced in neurons presumed to be in a hyperexcitability state. Has the firing of these neurons slowed as the burden of ALS has increased? As post-mortem tissue was used, is the data generated reflective of end-stage disease? Conversely, mice lacking c-FOS have been shown to instead have neuronal excitability and neuronal cell death [365], suggesting a complex relationship with FOS and neuronal activity. Alternatively, c-FOS is involved in a large array of functions, especially considering it is one of the components of Activator protein 1 (AP-1), a transcription factor that regulates gene expression in response to cytokines, growth factors, stress, and infections; along with regulating differentiation, proliferation, and apoptosis; as such, the reduction of FOS may be a result of another disrupted mechanism. Some have argued FOS exhibits a neuroprotective effect [366], while some have identified a potential pathogenic role. The multifaceted nature of FOS makes it unclear to interpret why ALS Purkinje cells have lowered expression of FOS.

This chapter has generated a large amount of data on the transcriptome and proteome of the Betz cell in the healthy and ALS brain. It has highlighted a plethora of cellular pathways (not all of which can be described in this thesis) that are enriched within the Betz cells and has discovered many transcripts/proteins that may be of interest to the scientific community. While some ALS associated transcripts/proteins were found to be aberrantly regulated in ALS samples, there remains a considerable amount of DEGs/DEPs whose role in ALS is either unknown or limited. Future work is required to confirm if the differentially expressed molecules or pathways found in this chapter are recapitulated in the ALS brain, and if so, what role they play in disease.

6.4.3 Main Conclusions

- An alternative RWS approach was developed to facilitate LCM proteomic for large-sized projects.
- The high-throughput LCM proteomic approach was used to analyse 3,111 Betz cells and 8,733 Purkinje cells which were microdissected from the healthy and ALS post-mortem brains for LCM Proteomic analysis; identifying $\approx$ 6,000 proteins in each cell-type and providing the largest data-set for microdissected neurons, both in health and in ALS.
- 757 Betz cells and 2,570 Purkinje cells were microdissected from the healthy and ALS brain for LCM-seq Analysis.
- The transcriptome of the human Betz cells was directly investigated for the first time.
- LCM-seq analysis revealed a magnitude of cellular pathways that were enriched in the Betz cell and identified $\approx$ 150 transcripts that may outperform neurofilament as a potential Betz cell marker.
- Both LCM proteomics & LCM-seq revealed a plethora of cellular dysfunction within ALS Betz cells, including, excitotoxicity, oxidative stress, DNA damage, diminished energy production, enriched apoptosis, and increased neuroinflammation pathways.

"Hyrtl said that he regarded the anatomy of the Brain as a book sealed with numerous stamps and that humanity will open it, will read it, and only during the last day of all existence will humanity understand all of the truths comprised within it."

– Vladimir Betz, 1871

7 Conclusions

Contents

7.1 Achieved Objectives 235
7.2 The Human Betz cell 236
7.3 Limitations of this thesis 237
7.4 Future Work 239
7.4.1 Data from this Thesis 239
7.4.2 Future Betz Cell Research 240
7.5 Final Thoughts 242

7.1 Achieved Objectives

This thesis has achieved several of its aims:

- Using cutting-edge techniques (LCM coupled with LC-MS/MS, RNA-seq,) it was possible to generate large amounts of information on the biochemical composition of the Betz cell, and novel insights into their cellular mechanisms.
- Although it is well-documented that Betz cells are degraded in ALS, this thesis adds cell-type resolved molecular evidence for this; and provide insights into the particular proteins or pathways that might be responsible/involved in this neurodegeneration.

- To facilitate future Betz cell research, several molecules were tested to assess their performance as potential Betz cell biomarkers. Several proteins were identified that showed strong enrichment in Betz cells with little immunoreactivity in the surrounding primary motor cortex. While no 'pure' Betz-specific marker was identified, the proteins herein should be suitable for immuno- or *in silico*-based analysis; performing as well, if not better, than the current best non-specific marker for Betz cells – NEFH. LCM-seq analysis also identified >150, yet to be investigated, DEGs that are more differentially expressed than neurofilament in Betz cells.

7.2 The Human Betz cell

The data within the thesis confirms various assumptions about the Betz cell. They are highly metabolically active neurons with large energy demands. Along with oxidative phosphorylation, Betz cells also have an increase in glycolysis, pentose phosphate metabolism, purine/pyrimidine metabolism, glycan metabolism, carbohydrate, lipid, and steroid metabolism. They are glutamatergic neurons, although, there is strong evidence for a yet unknown neuropeptide or alternative neurotransmitter as a means of communication. They are fast-firing and therefore are under enormous pressure to ensure they have the energy, cargo, transporters and transport networks required to deliver prolonged synaptic firing, a task that is made more difficult with their elongated axons. These added pressures are reflected by the abundance of endoplasmic reticulum-related proteins, both in its organisation and maintenance, but also in protein transport to and from this organelle. They have an up-regulation in protein synthesis mechanisms, most likely to meet their physiological needs. Their characteristic lipofuscin accumulations may be due to the additional stresses of having to recycle the macromolecules needed to sustain their demands; the data generated revealed that Betz cells appear to have an up-regulation of endolysosomes, ubiquitin labelling, ubiquitin-mediated proteolysis, and proteasomal pathways.

This thesis has also provided some exposure to the molecular mechanisms that are aberrantly affected in Betz cells of ALS patients. It has not gone unnoticed that many of the mechanisms affected by ALS are the same pathways that Betz cells rely heavily on for their healthy physiology e.g. neurofilament/microtubule structure, axonal transport, rapid-glutamatergic firing, mitochondrial activity, energy metabolism, and protein synthesis. This may explain why these giantopyramidal neurons are particularly vulnerable to ALS. The data generated has also provided clues to where to direct future research into investigating the ALS Betz cell, namely, DNA damage, piRNA metabolism, excitotoxicity, and neuroinflammation. Careful analysis and experimentation are needed to continue to study these gigantopyramidal neurons in health and disease and to validate the observations seen in this thesis.

7.3 Limitations of this thesis

As with all investigations primarily working with post-mortem human tissue, there will always be several limitations. Unlike rodent or cell models the circumstances around the end of life cannot be controlled, as a result, post-mortem tissue often has a higher degree of autolysis making biochemical investigations difficult and raising questions regarding the reliability of *'omics'* interpretation on post-mortem tissue. While significant processing and quality control steps were put in place to minimise degradation of protein and RNA, there will always be a degree of degradation in the underlying samples which may or may not be reflected in the generated data sets. Moreover, post-mortem material from individuals with neurodegenerative diseases almost always represents end-stage disease, as such, the data in this thesis may only report biochemical changes that appear later in the disease process and may not be useful in investigating early pathogenesis.

The convoluted nature of gyri and sulci makes generating tissue sections that best reflect the 3-dimension nature of the primary motor cortex particularly challenging. Unlike Purkinje cells, which have a high degree of uniformity across the cerebellum and are localised to a one-cell thick neuronal layer, Betz cells are more irregular

and can be difficult to section without over/under-sampling their cell soma. While care was taken to ensure each tissue section was perpendicular to the axis of the central sulcus; its serpentine nature meant some degree of off-angle sectioning occurred in our cohort, resulting in some Betz cell perikarya being partially/overly sectioned. This is particularly apparent in Figures 6.2 and 6.6 where the standard deviation of Betz cell soma size is considerably larger than that of Purkinje cells. This off-angle sectioning likely exaggerates the size differences between individual Betz cells and may even result in a depletion of particular organelles, which could affect downstream analysis. As such, these figures should be used to represent the number and size of the individual cellular material used in these experiments, rather than an informative measure of the size difference between these two cell-types.

While the motor strip was sampled in its entirety, the lack of known markers meant that only Betz cells with canonical Betz morphology were microdissected. As this morphology is best seen in larger Betz cells most of the data generated in this thesis is derived from the leg area or hand knob area. As it is likely that there exists regional/metabolic differences or sub-populations of Betz cells within and across the motor strip there may exist significant bias in the data, for example, the potential markers discovered in this thesis by not stain all Betz cells equally. Furthermore, while the Purkinje cells in this thesis are from the inferolateral cerebellar neocortex, the microdissected Purkinje cells may not all be involved in motor control. Rather than defining subtypes of neurons, this thesis aims to discover the common molecular denominator for the 'top level' class of these neurons, so that future experiments can probe the subtle differences between the cell-states within neurons.

Another significant limitation in this thesis revolves around the fact that laser microdissection is an enrichment technique rather than an isolation technique. The extreme heterogeneity of the human brain will always raise questions of contamination in downstream data analysis. The interwoven nature of neural tissue also means there may be noticeable contamination from the innervating synaptic boutons connected to microdissected neurons. The accumulation of synapsing material in pooled cells may result in artefacts in the data. Moreover, only the

soma of neurons was isolated in the aforementioned experiments, limiting data generated to that of the perikaryon – a potential reason why the markers discovered were primarily isolated to the cell soma. As the transcriptome of a neuron has been reported to vary considerably depending on the neuronal component studied, along with the fact that mRNA localisation is emerging as an important feature both in ALS and the healthy brain, the data in this thesis may only reflect one aspect of the total expression of mRNA in the Betz cell.

Finally, the lack of *'omic'* investigations on the Betz cell or limited proteomic & transcriptomic studies on the post-mortem Purkinje cell, L5 neuron, or DNNs makes it difficult to compare and validate findings with existing *'omic'* data sets, as these have often been generated with different approaches, whose underlying technology and pre-processing steps differ considerably. As a result, validation of the above data is reliant on in-house IHC which can be challenging as antibody selection, antibody type, antibody affinity, and IHC protocol can all have adverse effects on the success of an antibody. Promisingly, RNA-seq and LC-MS/MS analysis of microdissected neurons did show complementary findings and were in accordance with the current literature.

7.4 Future Work

7.4.1 Data from this Thesis

While some proteomics & transcriptomics data has now been generated on the Betz cell in the healthy and ALS brain; more work is needed to validate findings. The complex relationship between protein and mRNA expression requires both IHC and ISH to be implemented to spatially resolve if biomolecules outlined in this thesis can either serve as biomarkers or confirm if particular cellular mechanisms are affected in disease.

The LCM-seq observations outlined in this thesis have focused primarily on protein-coding genes. There is increasing evidence of the role of lncRNA and miRNA in neurodegeneration and ALS, and the LCM-seq data generated in this

thesis has revealed 3,185 lncRNAs and miRNAs. Analysis of these RNA molecules identified several differentially expressed lncRNAs and miRNAs between the healthy and diseased Betz cell and L5 neuron (data not shown); further data exploration and validation are needed to understand the role of these two RNA entities in cell-type-specific or ALS-specific processes. The same is true for the pathogenic alternative splicing seen in the ALS transcriptomic data (not shown), which has immediate interest to TDP-43 and how ALS neurons process splicing events. Further, although proteomics has the added benefit of investigating PTMs, this type of analysis has not been performed in this thesis. Future analysis will probe any global/protein-specific PTMs within this data set, as this is also of huge interest to the ALS/FTD community; as PTMs are believed to be the driving force behind pathogenesis and disease spread.

Regarding ALS, additional work also needs to be performed to understand the reasons why pTDP-43 inclusions are seemingly rare in Betz cells despite being a prominent feature of ALS pathology and the fact that Betz cells undergo degeneration despite the lack of inclusions. Are Betz cells resistant to pTDP-43 insoluble inclusions, or are they particularly vulnerable to soluble pathogenic TDP-43? Moreover, it would be important to discover if C-/N- TDP-43 fragments are enriched within ALS Betz cells.

7.4.2 Future Betz Cell Research

Protein and mRNA are the most important molecules for characterising the physiology of a cell. However, to understand the full biology of a cell, its "cellome" needs to be investigated. As single-cell technology advances, and our understanding of the molecular composition of the Betz cell grow; similar workflows should be implemented to begin defining the other parts of the Betz cell's "cellome" i.e. its lipidome, glycome, and methylome.

While there is much to learn about the fundamentals of adult neurobiology, even less is known about the extremely dramatic molecular changes that underpin the mechanisms of the developing brain. The limited knowledge on the developing

motor cortex suggests that Betz cells may be the first mature neurons to develop in the embryonic M1 cortex. However, there is almost no data on the embryonic Betz cell; the extreme difficulty and rarity of accessing embryonic neural material is most likely the cause of this. Targeting these embryonic Betz cells and comparing them to the adult Betz cell is critical in understanding their morphology in the adult brain. Although embryonic cells are dynamic and their molecular profile changes significantly by adulthood, *'omic'* investigations of embryonic Betz cells may reveal biomolecules responsible for their unique morphology and may identify proteins or TFs that remain static throughout the life-span of a Betz cell.

"Betz cell" is the nomenclature given the gigantopyramidal motor neurons within L5 of the primary motor cortex in primates. As discussed in Section B.1, gigantopyramidal L5 motor neurons also exist in many other mammals. While it is known that, unlike some other mammals, primate gigantopyramidal motor neurons directly innervate spinal motor neurons; it is possible that these gigantopyramidal L5 motor neurons represent the same neuronal niche as primate Betz cells - however, this has yet to be confirmed using molecular approaches. The information gained from this thesis could be used to confirm if the Betz cell nomenclature – molecularly speaking – can be used as an umbrella term for all L5 gigantopyramidal motor neurons across species. In particular, staining the non-primate motor cortex with the Betz-enriched proteins highlighted in the previous chapters may yield interesting results; especially if applied to the rodent motor cortex to investigate if they do possess neurons that contain a similar molecular profile to the human Betz cell.

While largely regarded as corticospinal neurons, it remains to be seen if this is true for all Betz cells. The curious staining of CRYM, ER81, and OXT1 suggests a potential for differing projection sites for Betz cells, or perhaps subpopulations of these neurons. Is the coordination of motor control from each hemisphere mediated solely by the communication of corticocortical fibres from surrounding neurons or do a subset of these fibres originate from the Betz cell? Moreover, while some have suggested that axonal collaterals from Betz cells can innervate the

striatum it may be worthwhile to investigate all the potential projection sites of Betz cells. This is a challenging task and requires intimate knowledge of neuroanatomy and a clear biomarker for the Betz cell. If it can be confirmed that the axonal immuno-labelling of REEP1 and HDAC1 are from Betz cells it may be possible to track axon path across the cerebrum.

While the Betz cell of the primary motor cortex remains an understudied neuron, even less is known about the mysterious Betz cells of the cingulate [207]. Now that preliminary next-generation data on the molecular profile of the motor Betz cells has been generated, future work should be performed - potentially using the biomarkers and/or LCM workflows developed in this thesis - to study the differences between Betz cell populations both across the motor cortex and cingulate.

To summarise, although novel insights into the Betz cell have been revealed; many relatively simple questions regarding the nature and physiology of Betz cells remain. (1) Do their axons solely project to the corticospinal tract or does a subpopulation contribute to other tracts, i.e., corticocortical connections? (2) What underlines the regional differences in Betz cell morphology, e.g., are Betz cells from the leg area and hand area molecularly/metabolically distinct? (3) What is the biological significance of Betz cell clusters? (4) How do M1 Betz cells differ from the enigmatic Betz cells of the cingulate? (5) What is their molecular role in the progression of ALS, are they the site of end-stage disease, or the beginning? (6) What underpins their selective vulnerability in ALS? (7) Why is pathological insoluble TDP43 (a hallmark of ALS) rarely found in Betz cells despite their degeneration? (8) Can insights into Betz cell biology be found that lead to future therapies?

7.5 Final Thoughts

The LCM proteomic workflow described in this thesis has potentially vast usage outside of the work in this thesis and could be used to research a magnitude of topics within molecular neuroscience and beyond. This technique along with the complementary LCM-seq has generated *'omics'* data sets for the Betz cell, L5 motor

neuron, Purkinje cell, Cerebellar Dentate Nucleus neuron, Cerebellar Granule layer neuron, and the molecular layer of the cerebellum. While the focus of this thesis has been on the Betz cell, and an equal amount of analysis and interpretation can instead be directed to any of the above cell-types; we hope to make this data publicly available for future researchers to independently validate and explore.

The discovery of several molecules (REEP1, DST, ESYT1, etc.) that are enriched within Betz cells should help remove some of the barriers that have to date, slowed down the progression of Betz cell research e.g. the difficulty in correctly identifying them from surrounding neuronal populations. The comprehensive biochemical data that now exists may pave the way for future non-morphological based molecular experiments; for example, FACS strategies may now be implemented to allow Betz cells to be captured from surrounding neural tissue in a high-throughput manner and spark sc/snRNA-seq investigations. Furthermore, the Betz-enriched proteins identified in this thesis may help future studies which wish to selectively study Betz cells from a neuropathological/stereology perspective without needing extensive knowledge on their unique morphology. Moreover, as attempts to characterise the total "cellome" of the human brain have increasingly adopted unbiased single-cell/nuclei enrichment techniques coupled with DNA/RNA-sequencing; the data produced from this thesis may serve as an additional source of information to validate the results of cellular deconvolution algorithm, to identify specific Betz cell clusters, or tease out Betz cell information from bulk tissue analysis.

The vast majority of biological functions are governed and implemented by proteins. However, proteomics, especially LCM proteomics is still in its infancy compared to RNA/DNA sequencing technology. Directly investigating protein expression (rather than inferring from RNA-seq) and quantifying PTMs remains a difficult challenge for cell-type resolved studies. A full map of Betz cell protein expression may only become available when proteomics has evolved to the point where all proteins, isoforms, and PTMs can be quantitatively measured at a single-cell resolution. Until then, the data generated in this thesis hopes to

provide a foundation for future studies that wish to investigate the proteome and transcriptome of the Betz cell, or for researchers interested in any of the other neuronal cell-types studied herein.

Appendices

A

Accurately identifying a Betz cell

A.1 Accurately Identifying a Betz Cell -

While they are the largest cell in the primary motor cortex and confined to layer 5, other large non-Betz cell neurons also exist in this layer, whose sizes overlap with the smallest Betz cells, making accurate Betz cell identification prone to error [176]. As such, size and location alone cannot be used to accurately identify Betz cells from surrounding pyramidal populations [367–369]. Importantly, these large non-Betz motor neurons lack typical Betz cell morphology which can be used to distinguish between these two cellular populations.

Historically, attempts have been made to classify the Betz cell size spectrum in the Rhesus monkey [370]; in which, Nañagas (1922) argued Betz cells should be grouped into 3 distinct groups representing small, medium, and giant cells. The lack of Betz cell-specific biomarkers has resulted in considerable confusion in distinguishing between Betz cells and large L5 neurons. Unfortunately, Betz morphology can be subtle, especially with non-Golgi based staining approaches, which do not highlight the Betz cell's unique somatodendritic architecture, making accurate identification a deceptively simple task. While Betz cells can be visualised with generic histological stains (e.g. H&E, Nissl or Haematoxylin, see Figure A.1) this requires pre-existing knowledge of Betz cell morphology to accurately identify

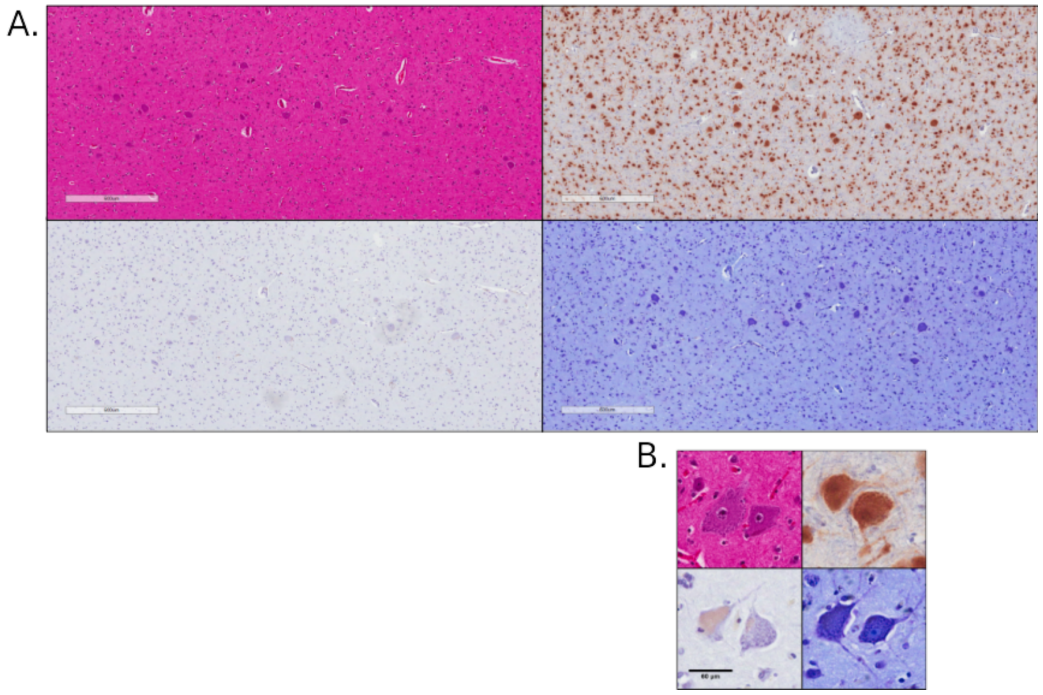


Figure A.1: Basic Stains for Identifying Betz Cells.
It can be difficult to identify Betz cells using generic markers at low magnification (A) and higher magnification (B). A: Shows Betz cells in L5 of the primary motor cortex. B: H&E and Nissl are best at identifying Nissl bodies in Betz cells. Haematoxylin highlights the yellow pigment of lipofuscin, taking up the majority of the cell soma, anti-Neun best highlights Betz cell soma. In this example, finger-like projections can be seen in Nissl and Haematoxylin staining. TL: H&E, TR: Anti-NeuN, BL: Haematoxylin, BR: Nissl staining.

them. Antibody-based approaches are commonly used to visualise cells based on molecular features, however, as this information is limited for the Betz cell, pan-neuronal markers are often used as an alternative (see Figure A.1), resulting in the same problems as generic histological stains. While not Betz-specific, anti-Neurofilament antibodies currently provide the best immunohistochemical approach to label Betz cells (Figure A.3). Both antibody clones, SMI-32 (targeting non-phosphorylated neurofilament heavy) or SMI-311 (targeting pan non-phosphorylated neurofilament) positively stain neuronal cells, with Betz cells being particularly immuno-reactive for NEFH, likely due to their large size requiring additional cytostructural support. However, as neurofilament is a feature for the majority of neurons it cannot be explicitly used to immunochemically label Betz cells from other neurons.

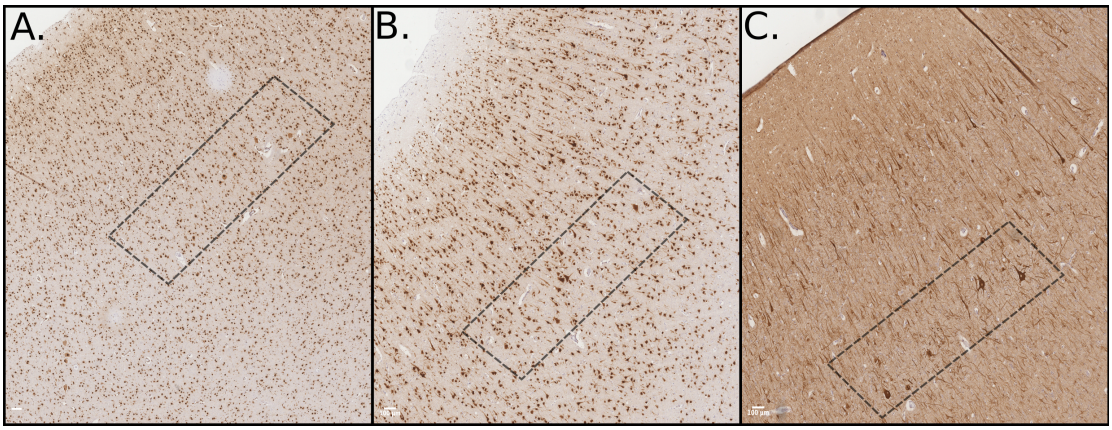


Figure A.2: Immuno-labelling for Identifying Betz Cells.
Pan-neuronal or pan-cellular IHC allows for Betz cells to be identified on their most striking features: size and location. Anti-Neurofilament remains the best method of staining Betz cells. A: Pan-cellular (Anti-TDP-43), B: Pan-neuronal (anti-NeuN), C: Neurofilament (SMI-311 antibody). Dotted box: L5 of the primary motor cortex, containing several Betz cells).

The following criteria/considerations can be used to accurately identify Betz cells in the motor cortex:

Location - Betz cells are found with Layer Vb of the primary motor cortex. The twisting nature of the M1 gyrus, the difficulties with macroscopic dissection, and microtomy can often result in histology sections that do not represent a surface parallel to the M1 gyrus. As such Betz cells may seem to appear in other neuronal layers due to off-angle sectioning. Often Betz cells can be tracked along the L5 boundary of M1. They are particular enriched in the buried M1 of the central sulcus. Considerable forethought is needed when sampling the motor cortex to ensure Betz cells are in-plane, and that the complete laminar pattern of M1 is intact in histology sections.

Size - Betz cells, on average, are the largest neurons in the CNS, their size can be seen in low power magnification, and their existence can be a striking cytology feature. However, as previously shown Betz cell size varies considerably, with some Betz cells being smaller than the surrounding large non-Betz L5 neurons. In light of no-specific markers, if possible, size thresholds must be placed to only

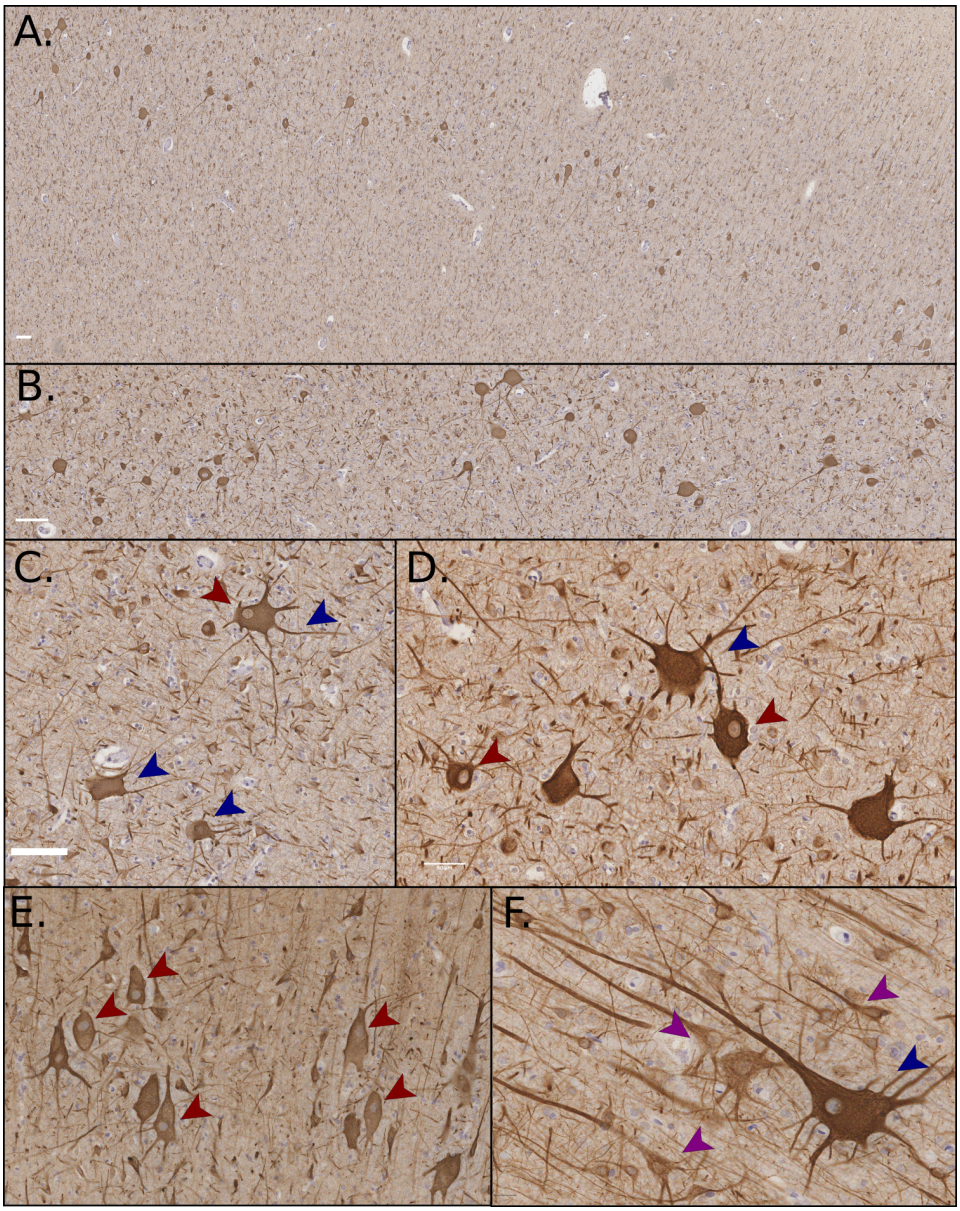


Figure A.3: Neurofilament Heavy Highlights Betz Cell Morphology. A-B: The characteristic band of Betz cells in L5 of the primary motor cortex can be visualised by Neurofilament staining. The strong immuno-reactivity with Betz cells allows Betz cell dendrites to be well defined, highlighting their circumferential dendrites (Blue arrows). The lack of nuclear staining also allows the prominent nucleus of Betz cells to be visualised, aiding in identification (Red arrows). However, it is important to note that non-Betz cells also strongly express neurofilament which could lead to errors (Purple arrows).

study the largest, and therefore, most characteristic Betz cells (e.g. $>1000\mu m^2$). Identifying smaller Betz cells can be achieved by looking for other morphology features that should be used to increase confidence.

Apical dendrite - This can erupt sharply from the Betz soma or gradually depending on the sectioning. Nissl substance is concentrated at the base of the apical dendrite but rapidly diminishes; a feature that can be seen in most basic dye preparations. Bifurcation of the apical dendrite – resulting in two branches of similar size - can also be used to help increase confidence but may be difficult to visualise without immunohistochemistry.

Somatodendritic architecture - Arguably, the most unique feature of the Betz cells is the circumferential dendrites that emerge from the entirety of the Betz perikaryon. In humans, they are mostly oriented towards the white matter track. Nissl substance within Betz cells becomes longitudinally arranged at the sites of their dendrites and can give rise to the finger-like projections seen in light microscopy, which can be a strong identifier of Betz cells.

Nissl bodies - The existence of Nissl bodies can be seen in many standard histological stains; however, they are particularly highlighted using the Nissl stain. These bodies accumulate across the perikaryon of Betz cells, highlighting soma and dendrites.

Nuclear features - Betz cells have a pronounced nucleus with a visible nucleolus. Depending on the plane of the section, this can take up a large proportion of the Betz cell soma.

Lipofuscin - The large accumulation of this auto-fluorescent pigment is almost always present within Betz cells. The lipofuscin can appear anywhere within the Betz soma, however, it is likely posited next to or below the nucleus, typically in the basal part of the soma.

Clustering - In human, Betz cells can be found in neuronal clusters for 2-5 cells, when this striking phenomenon is observed it can be relatively easy to accurately identify them assuming the criteria above is followed. Betz cells found solely across the motor cortex can be more difficult to be resolved.

Nissl staining, methylene blue, H&E, and Haematoxylin can be used to stain the motor cortex so that these features can be visualised. Although no biomarker exists for the Betz cell it is known that they contain an abundance of neurofilament. While antibodies for neurofilament are not Betz specific and stain a large variety of neurons in the motor cortex, it does allow size, location, pronounced nucleus, Nissl bodies, and somatodendritic architecture to be visualised. Pan-neuronal or pan-cellular antibodies may also highlight some of these features, such as NeuN and TDP-43, however, as with neurofilament, the ubiquitous staining of these types of antibodies can cause confusion when studying Betz cells, moreover, they are not suitable in identifying Betz cells immunochemically in non-morphological derived experiments. The "gold standard" would be the discovery of biomarkers specific for Betz cells, allowing them to be identified without extensive knowledge of their morphology, and would allow Betz cells of all sizes and across motor regions to be investigated. This would accelerate research into Betz cells, both in the control and diseased brain, allow them to be studied independently to other L5 neurons; and would also make Betz cells compatible for targeted next-generation isolation techniques such as FACS sc/snRNA-sequencing.

B

Additional information about the Betz cell

Contents

B.1 Comparative Neuroanatomy of Betz Cells 253
B.2 Electrophysiology of Betz Cells 258
B.3 Molecular Composition of Betz Cells 260
B.4 Motor Cortex / Betz Cells in Development 262
B.5 Betz Cells in Ageing and Disease - 265
B.5.1 TDP-43 Pathology in Betz Cells - 268

B.1 Comparative Neuroanatomy of Betz Cells

The limited access to high-quality archival post-mortem human brains has resulted in much of modern neuroscience/neuroanatomy knowledge being extrapolated from animal models. Mice are commonly used when investigating the nature of the neocortex, however, this can be problematic in studying the subpopulations of the mouse motor cortex as there is still an argument on whether mice brains contain Betz cells. If they do possess Betz cells they are not in line with the current defining morphology and cannot be seen with traditional histological investigation. Caution must be taken when extrapolating Betz/M1 data from the mouse brain to the human isocortex. If Betz cells are indeed a molecularly unique subpopulation of L5 motor neurons, there may exist a currently unknown Betz specific biomarker/molecular

profile which if discovered, could be used to validate the presence of Betz cells or closely related Betz cells in rodents.

Many questions remain regarding the evolutionary importance of Betz cells. Despite the fact that Betz cells are prominent features of M1 biology in felines and primates, there is limited overlap in the organisation of their respective corticospinal tract. For example, direct Betz cortico-spinal connections to lower motor neurons are present in primates, whereas non-primates such as cats [371], rats and mice [372, 373] do not possess such connections - although they may exist briefly in the developing rat brain [374]. Primates appear to preferentially use direct cortico-spinal connections over the propriospinal system, resulting in advanced manual dexterity and hand function [166, 375, 376]. In humans, this direct connection has been shown electrophysiologically [377, 378] and anatomically [379], although direct connections appear to be more common in controlling distal muscles and rarer in proximal muscles [377, 378], which may explain the increased size of Betz cells in those representative areas of M1.

Much of the research around Betz cells involve investigating the size, shape, and cellular composition of their soma in health and disease. However, less work has been performed on the morphologically feature that best delimitates them from other L5 neurons, their unique somatodendritic structure - most likely due to the difficulties of wide-scale Golgi stain experiments. Despite the difficulties, some groups have managed to study the somatodendritic architecture in a wide variety of species, including: humans [99, 184], monkeys [380], domestic cat [177, 381], and sheep [381, 382]; quantitative somatodendritic studies have also been performed on the the Panthera genus [190, 203, 383], and giraffe [194, 226] neocortex. Early studies by Brodmann and Betz [224, 384] found that carnivores contained large, wide, and rounded Betz cells, whereas in primates they appear to be large and conically shaped [199], whereas, Betz cells in Artidoactyls are not as large and tend to be elongated [181].

In 2017, Jacobs et al., performed a comprehensive quantitative serological comparison of M1 L5 pyramidal neurons (including targeted Betz cell analysis) across 19 mammalian species from 7 phylogenetic orders [190]. This comprehensive experiment provided new insights into the variation and morphology of the Betz cell across species and highlighted that the somatodendritic architecture is a key criterion for Betz cell identity across species; as such this work along with similar studies will be discussed in depth below. Jacobs et al., (2017) used the Golgi stain to identify Betz cells based on their somatodendritic architecture and dissected the hand/forepaw region of the homunculus in all animals investigated [99, 385, 386]. Critically, the authors confirmed the Primary Motor Cortex was accurately sampled by histologically identifying the agranular nature of M1 and the presence of Betz cells. The authors note that they confidently sampled M1 in all but 4 species (mongoose, rabbit, rat, and wallaby) and stated that although an agranular cortex was found lacking Layer 4 they did not observe the existence of Betz cells in these animals. In total, they studied 181 Betz cells/gigantopyramidal neurons, 203 deep pyramidal neurons (mostly L5, some L6), and 233 superficial pyramidal neurons. Their cohort contained a large range of neocortex sizes, from the smallest (rat, 3g) to the largest brain (human, 1,338g); a 463-fold increase. Overall, the authors measured soma size and depth from the pial surface along with 6 dendritic measurements: dendritic volume, total dendritic length, mean segment length, dendritic segment count, dendritic spine number, and dendritic spine density. As they only had 1-2 brains from each species, the authors decided to avoid comparing individual species, but instead compare individual taxonomical groups, which would allow for robust statistical power. Interestingly, the authors note a positive correlation between somatodendritic measurements and brain mass in superficial pyramidal neurons but not in Betz cells. Furthermore, while superficial pyramidal neurons contained morphological uniformity across species, Betz cells morphology varied the greatest. The largest morphological variation in Betz cells was the average number of primary basilar dendrites; the average across species was 7.01 $\pm$ 3.15 ranging from 3.7 in the clouded leopard to 14.2 in the giraffe. Additionally, they found that

the average number of basilar dendrites was higher than superficial or other deep pyramidal neurons. Interestingly, within feliforms, this value also varied largely, with 3.7 and 3.8 in the clouded leopard and tiger, compared to 9 and 7.3 found in the caracal and lion, respectively. An obvious feature in this data-set was the increased size of gigantopyramidal in carnivores, supporting the same conclusions of Brodmann and Betz. The authors found that with the carnivores, the feliforms contained the largest Betz cell somas: an average of 2,847 μm^2 compared to the average 987 μm^2 found in primates. They noted, however, that feliforms in the Panthera genus possessed the largest Betz cells, with Betz cell soma volumes 12 times larger than surrounding neurons, whereas, in primates, this difference in Betz cell/other L5 neurons was only x2.3.

Jacobs et al., (2017), also found that Betz dendritic morphology was idiosyncratic and varied across different species. For example, in the caracal and lion, dendrites radiated symmetrically from the entire soma, while in the zebra and baboon dendrites were lateralized and asymmetric. Whereas, other species (tiger, dog, mountain zebra, kudu) contained basal dendrites which descended primarily from the bottom of the soma. Supplementary, as previously reported in human Betz cells, these giantopyramidal neurons tended occasional to form long taproot structures. Betz cell apical dendrite morphology also varies across taxonomic groups, bifurcating apical branches was a common observation among Betz cells across species but was a defining feature for perissodactyl and artiodactyls. This bifurcation of the apical dendrite of the Betz cells has also been seen in others studies, including the dolphin [387–389], humpback whale [390] and pygmy hippopotamus [390]. Of note, Baboons possessed more complex dendritic morphology than other species. Interestingly, the authors observed that Betz cells contained lowered numbers of dendritic spines on average than deep or superficial neurons, although this was variable and the approach used by Jacobs et al., has been reported to negatively impact spine estimates [383, 391, 392]. Finally, they found that Betz cell basilar dendrites projected further from their soma than other neuronal types, a feature also reported in humans. Overall, Jacobs et al., (2017) found that the dependent

somal and dendritic measurements recorded were significant differentiators for each taxonomic group and found no correlation between soma size and brain/body mass in primates. They also found no statistically significant relationship between soma measurements in primate species.

Interestingly, the authors found that soma size/volume was a differentiator between Betz cells and other L5 neurons, however, this difference was on average subtle, with Betz cell soma areas being only 1.64 times larger and soma volume 2.3 times larger than other deep pyramidal neurons (in primates). Despite large somatodendritic variations in Betz cell morphology across groups, Betz cell morphology across primates was remarkably uniform and consistent with previous literature. The clustering of Betz cells also varies across species; in the giraffe brain, no more than 4 Betz cells cluster was observed and may be a specific feature of cetartiodactyla [181], while others have shown that cetacean animals have much larger clusters [393–397] e.g. the humpback whale contains Betz cell clusters of up to 15 cells [396]. As mentioned previously, no Betz cells were found with the rabbit, wallaby, or rat. Despite the feliforms containing the most prominent Betz cells, the mongoose contained very small Betz cells, although these were recorded as not having typical Betz cell morphology. In contrast to this study, Brodmann did identify Betz cells in rabbits but noted that they were some of the smallest in his collection and that it was difficult to delimitate them from surrounding neurons [224]. Other groups have also described similar findings in rodents and wallabies [398, 399] and morphological and electrophysiology investigations have not been able to successfully identify them in rodents [214, 215].

While comprehensive comparisons have been made on the evolution of the primate primary corticospinal tract [400], similar observations cannot be made regarding Betz cells; as many species (including closely related primates) have not had their Betz cell biology extensively investigated. However, some conclusions can be drawn between the differences between the human and the well-studied macaque Betz cell. Although counting Betz cells has historically been inaccurate, it had been shown that the macaque brain contains about half as many Betz cells as the human

brain [186]. As seen previously, 75% of all human Betz cells are localised to the upper third of the M1 strip, while in the macaque this upper third only contains 52% of Betz cells. Comparatively, Betz cells in humans have seen an evolutionary medial and upward shift, meaning that in humans more Betz cells are present in the leg region. As such, Old World monkeys appear to contain relatively more Betz cells in the arm region than humans (33% vs. 20%) which is somewhat contradictory with the notion that Betz cells are principally responsible for fine motor control, as humans rank highest in manual dexterity [401]. One explanation given is that in the evolution of the human brain the representation map of the motor cortex has shifted to a higher position on the lateral convexity of the hemisphere. This would result in the centre of the upper body region being shifted in a superior faction in M1 - which is consistent with the M1 evolutionary changes seen in the primate lineage [180, 402]. Another striking difference between the human and macaque brain is the relationship with the number of Betz cells in the central sulcus. In the macaque brain, 31% of Betz cells were found within the anterior wall of the central sulcus while in humans this has increased to 82% [186]. Lassek explained that in the evolution of humans the central sulcus has become deeper and has lengthened pulling the majority of Betz cells into the groove. An MRI experiment studying the Primary Motor Cortex, specifically the hand knob and central sulcus of 131 primates including Old World monkeys, apes, and humans confirmed this observation and reported that great apes and humans do contain a larger central sulcus than Old World monkeys [403].

B.2 Electrophysiology of Betz Cells

Much of the current knowledge on the electrophysiology of Betz cells comes from the extensive work studying the sensorimotor cortex of the domestic feline (*Felis catus*) [404–412], which show feline Betz cells fire repetitively and rapidly, and can quickly repolarise. There is little data on the electrophysiology of primate Betz cells. The Allen Brain Institute has performed PATCH-seq on presumed Betz cells of primate and human origin and recorded their electrophysiology. Within

this data, the authors note that primate Betz/ET neurons displayed a distinctive biphasic-firing pattern. They observed macaque and human ET/Betz neurons being characterised as having an early cessation of firing which was then followed by a sustained and dramatic increase in firing. They also saw that primate ET/Betz cells returned to a steady state quicker than mouse ET neurons.

Early electrophysiology data, suggests that anaesthetic does not render Betz cells inactive, instead, Betz cells discharge characteristically, allowing them to be easily recognised [413]. This feature allowed authors to show that small electrical stimulation that does not provoke movement, still causes Betz cells to fire. The specific functional contribution of Betz cells to motor movement remains mostly unexplored. It has been postulated that the heavily myelinated axons that originate from Betz cells are capable of conducting rapid impulses at 40-100m /s [191, 414, 415]. It has also been suggested that the rapid firing of Betz cells may introduce partial inhibition of extensor muscles, which may result in flexor facilitation before movement, and restoration of extensor tone instantly after movement [229, 415]. It has also been suggested that Betz cells are silent during muscle contraction and briefly fires to relaxes muscles during posture maintenance [191, 192]. Moreover, Betz cell may play a major role in the depression of extensor muscles to control anti-gravity muscles in locomotion and posture [191, 192]. Others have suggested, that for upper limb movement, Betz cells are particularly important for the fine motor control of the wrist and hand [166, 376].

Electrophysiology experiments on the felines have led to the suggestion that Betz cells possess two transient potassium channels (Ca2+ dependent K+), which prep for repolarization and thus allow for repetitive firing, albeit at different speeds [408, 409]. This led to the conclusion by some researchers that there are two sub-types of Betz cells, one fast-firing, and one slow-firing, which is differentially stimulated by serotonin [411, 412]. Betz cells have also been shown to be excited by a variety of compounds: Substance P, acetylcholine, bradykinin, bombesin, eledoisin, DL-

homocysteic acid, serotonin, and angiotensin ii [178, 416–418]; shining a light on the diversity of their receptors, and the type of input they receive from other neurons.

Axon collaterals have been extensively studied in PT neurons of the motor cortex; with Betz cell collaterals noted as having an "all or nothing" response [419]. By staining collaterals, it was possible to visualise that these collaterals surround PT neuron perikaryon and share their basilar dendritic domains [420]. These authors found that small/slow-firing PT neuronal axon collaterals seem to monosynaptically target the dendrites of other PT neurons. Interestingly, it has been proposed that PT neurons, including Betz cells, use axonal collaterals to exert both excitatory and inhibitory effects on neighbouring L5 neurons [419, 421]. The dual stimulating nature of Betz cells has been questionable, although, in 1963, Suzuki and Tukahara proposed various explanations including the presence of a single inhibitory neuron that could inhibit neighbouring Betz cells and are innervated by another Betz's collateral axon [422]. In one ultrastructural study, authors caused lesions within BA4 of the monkey motor cortex and found several degenerating presynaptic synapses attached to Betz cells [423]. These synapses were plentiful, occurring 1-3 times in each section and synapsing with Betz apical dendrite, soma, and proximal processes. The authors believed these synapses belonged to a neuron of inhibitory nature (e.g. a basket cell) and suggested that the inhibitory nature of Betz cell collaterals could be performed via these basket cells/large stellate cells - potentially supporting the work of Suzuki and Tukahara. It is unclear of the consequence of this proposed Betz cell-mediated inhibition, although it has been suggesting that the delicate inhibition of Betz cells and PT neurons via axon collaterals can result in finer motor control [424].

B.3 Molecular Composition of Betz Cells

Most of the molecular information on the Betz cells has come from low-throughput studies using individual immuno-labelling approaches. Adult macaques (*Macaca mulatta & Macaca fuscata*) have Betz cells that strongly express the fast-firing

associated voltage-gated potassium channel Kv3.1b, which are localised to the proximal dendritic cytoplasmic membrane [425]. Curiously, others have shown that Kv3.1b is not expressed in Layer 5 mouse neurons [426]. Supporting the idea that Betz cells are fast-firing neurons, it has also been shown that in macaques Betz cells express high amounts of cytochrome oxidase, a mitochondrial transmembrane-enzyme involved in ATP production [427]. Moreover, as with other fast-firing neurons, parvalbumin has also been shown to be expressed in macaque Betz cells [428], which has recently been observed in human Betz cells [316]. As with other pyramidal neurons, it has been shown that Betz cells are enclosed by an extracellular matrix of chondroitin sulphate proteoglycan (CSPG) immunoreactive perineuronal nets [429]. Another study found that Betz cells have an increased presence of DNAJC30 an auxiliary component of ATP-synthase [430] and regulates mitochondrial respiration. Additionally, acetylcholinesterase has also found to be present in almost all Betz cells and $\approx 80\%$ of L3 neurons [431].

In macaques, different members of the protein kinase C (PKC) family have been shown to selectively stain neurons in M1. βI-PKC stained cells in L1, L3 and L5 neurons, including Betz cells, while α-, βII-, and γ-PKC did not stain Betz cells but did stain neurons in other layers. Intense βI-PKC axonal staining was found in the white matter, but the authors could not confirm if they originated from Betz cells. The presence of βI-PKC could suggest that signal transduction occurs via bradykinin and m1-muscarinic receptors on Betz cells [432]. Betz cells only express a small amount of poly-ADP ribose polymerase (PARP), which has been shown to increase vulnerability to ALS [433]. It is unclear if the enzyme that converts glutamine into glutamate, glutaminase, is present in Betz cells. Some have found phosphate-activated glutaminase is expressed at high levels in human Betz cells [434] suggesting these neurons use glutamate as a neurotransmitter - although highly likely, little work has been performed to prove Betz cells use glutamate as their neurotransmitter, or if they use any secondary neurotransmitters. Contradictory, others have stated that Betz cells are glutaminase negative in IHC studies [435]. Although it should be noted that glutaminase is also used to convert glutamate into

GABA, so may also label inhibitory neurons. Finally, others have suggested that Betz cells contain intermediate levels of GABA, although this could be due to the lack of sensitivity of the assay, or erroneously measuring synapsing interneurons [436].

Several studies have attempted to identify the receptor diversity within Betz cells; Betz cells contain various AMPA glutamate receptor subunits and have previously been shown to be stained positively for GluR2/3 [437, 438]. However, with the invention of an antibody to differentiate between the two receptors, it has now been reported that Betz cells are positive for GluR3 but not GluR2, with staining located in the cell soma and proximal dendrites [439]. This staining pattern suggests the presence of atypical calcium-permeable AMPA receptors on the Betz cell. Such calcium-permeable AMPA receptor without GluR2 has been shown to increase susceptibility to glutamate toxicity, as seen in ALS [440]. Betz cells also show moderate levels of GluR1 – which is localised mostly to dendrites, Betz cells contain high amounts of GluR4 in their dendrites, which was weakly expressed in other L5 neurons. Betz cells have low expression of GluR4/5/6/7 and are also immuno-reactive for NR2A and NR2A/B. Betz cells also contain angiotensin IV receptors [438, 439, 441]. Finally, it has also been reported that acetylcholine receptors are localised mostly to the dendrites of Betz cells [442].

B.4 Motor Cortex / Betz Cells in Development

While the anatomical development of the cerebral cortex is now broadly well delineated, the particular cytoarchitecture of the Primary Motor Cortex suggests that this region, in particular, is governed by distinct transcriptomic controls. The development of the Primary Motor Cortex is complex and has been extensively described in Marin-Padilla's 2011 book [443] where he provides a comprehensive narration of >4500 Golgi preparations performed on the Primary Motor Cortex of 27 prenatal cases ranging from 6th-week gestation to the new-born. Within this, he showed that the embryonic cytoarchitectural organization of the cerebrum begins at approximately the seven-week stage with the development of the lateral and third

ventricle, choroid plexus, and anlage of the hippocampus. Amazingly, at this early stage fibres can be seen passing through the internal capsule and terminating at the subpial zone, which at the time contain Cajal-Retzius (CR) neurons. At seven weeks cortical organization is considered functional active, and deep neuron-like cells from the subplate can be seen with afferent corticofugal fibres coming from their axon terminals. Motor cortex features can be recognised by the 11-week gestational age; the first lamina (adult nomenclature: Layer I) can be seen with CR cells and their horizontal fibres. At this stage, migrating neuroblasts are guided by radial glia fibres and extend to the first lamina where they lose their glia attachment and form an apical dendrite that attaches to the first lamina. Collectively, these neuroblasts will form the pyramidal cell plate (PCP) which eventually will give rise to all the pyramidal neurons in the Primary Motor Cortex. However, at this stage the PCP is not functionally active as it only contains small densely packed neuroblasts lacking dendritic spines, containing a smooth soma, and producing axons that have not yet reached the white matter.

By 15-weeks, it is possible to determine the age of the neurons present based on the length of their apical dendrite, as this reflects their arrival to the first lamina. As the brain develops and thickens older neurons will elongate their apical dendrite to maintain their anchorage to the first lamina, however, their soma location will not change while accommodating the arrival of new neurons [443, 444]. At this stage, there is great variety in the size of PCP cells, ranging from 30µm for the superficial neurons to 275µm for the deeper older neurons. Deeper neurons begin to form basilar dendrites and proximal dendritic spines as they mature and become the first pyramidal cell stratum in the Primary Motor Cortex. This "P1 stratum" corresponds to Layer V in the post-embryonic brain and is the birthplace of Betz cells, which are some of the first neurons to form in this stratum. Marin-Padilla notes that some of the large neurons in P1 grow in size throughout development, and although he does not use typical Betz cell nomenclature, he states that large older P1 neuron gives rise to the giant pyramidal neurons characteristic of the gigantopyramidal Betz cells. This P1 stratum is the first layer to contain an intrinsic anastomotic

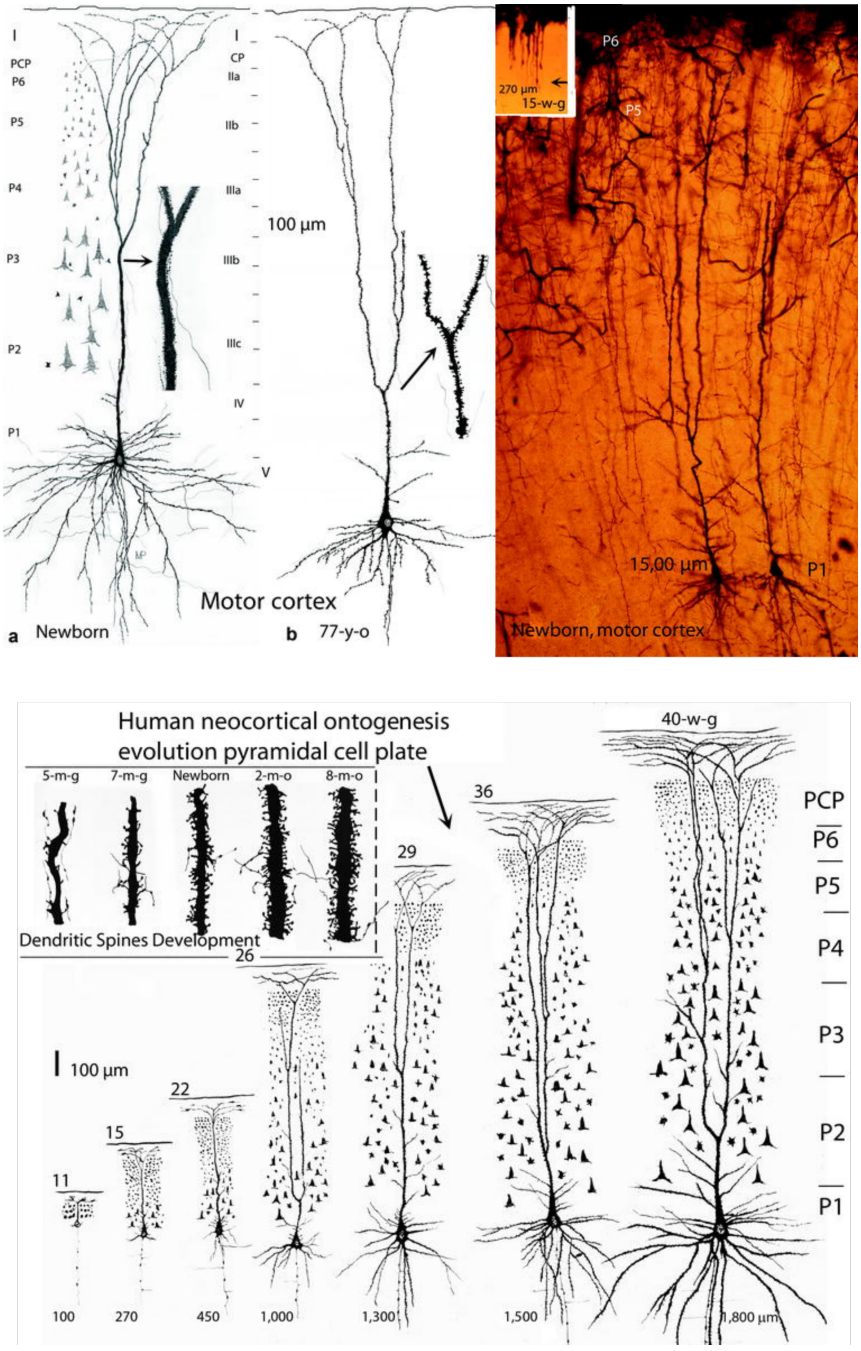


Figure B.1: Large Layer V Neurons in the Developing Primary Motor Cortex. A-B: shows camera lucida drawings of Betz cells in a 77-year old human, and a large L5 motor neuron of the newborn - presumed to be a Betz cell. Golgi image: shows two large pyramidal neurons in P1 of the newborn motor cortex. Note the bifurcating dendrite and dense basilar array. Bottom Image: Shows the development of the Primary Motor Cortex from 11-weeks to 40-weeks, note the prominent presence of the large P1 neuron - likely a Betz cell - with circumferential dendrites. The progressive development of dendritic spines on P1 neurons throughout development is also noted. These large P1 neurons are among the first to develop in this cortex.

capillary plexus. The large neurons of the subplate begin to lose their anchorages with the first laminar, resulting in the primordial CR-subplate system of the motor cortex slowly transition into an organisational system that is now based primarily on pyramidal neurons, CR cells, and afferent thalamic corticipetal fibres. Furthermore, by this stage, the axons of neurons in the P1 stratum have reached the white matter and are now the principal source of corticofugal fibres. The division of the subplate and white matter become unrecognizable as the large subplate neurons retrograde and transform into deep interstitial neurons. By 20-weeks, P2, the next pyramidal stratum is fully formed, and P3 (Layer III) is beginning to form and is completed by 26 weeks, P4 is fully formed at 30 weeks, P5 at 25 weeks, and finally, P6 is fully formed in the new-born. In the 30-week brain, P1 neurons develop more basal dendrites and proximal spines and have synapsed with the ascending corticipetal fibres from the thalamus. By 30-weeks P1-P4 has formed, and P1 neurons have continued to grow in size, containing longer basal dendrites and formed occasional collaterals; there has also been an increase in the number of apical dendritic spines. At the 35-week stage, the intricate connections of the Primary Motor Cortex are overwhelming, the first lamina is fully formed and P1-P4 have matured significantly with extra- and intra- cortical fibres synapsing with the entirety of the motor cortex [443].

B.5 Betz Cells in Ageing and Disease -

It has been speculated that Betz cells are highly metabolically active and high firing neurons, based on their large size, the existence of parvalbumin, Kv3.1b, and above electrophysiological studies, however, this has not been molecularly proven and it remains unclear if these findings are truly reflected in Betz cell physiology. Several morphological changes occur throughout the lifespan of a neuron, with highly metabolically active neurons seemingly being most affected. One of the most obvious features of an ageing neuron is the accumulation of large amounts of lipofuscin, which is also a criterion for Betz cell identification and appear to accumulate at a much earlier stage of life [192, 206]. There exist other age-related degeneration that

affects Betz cells, for example, as with other pyramidal neurons, Betz cells undergo the progressive loss of somatodendrites as they age [192]. In Rhesus monkeys, while somal volume is reduced, axonal axo-somatic apposition size, Betz cell number, and mitochondria per Betz cell remain relatively constant throughout life [380], with the most obvious age-related occurrence being is the extreme accumulation of cytoplasmic lipofuscin, which although seems to be well tolerated by Betz cells [195, 380] may still displace the amount of protein synthesis performed in their endoplasmic reticulum. In macaques, there does not appear to be a preference in the location of accumulated lipofuscin, which starts to appear after 5 years of life, whereas in the cat it tended to be localised close to the apical dendrite [177, 195]. Another prominent feature in the human ageing Betz cell is the progressive loss of circumferential dendrites [192]. Within the Rhesus monkey, Betz cell number decline is only weakly associated with ageing; barely being statistically significant, which is a sharp contrast to the nature of the human Betz cell in ageing. In the human Primary Motor Cortex, Scheibel et al., (1977) note that more than 75% of Betz cells may degenerate by the 8^{th}decade of life and suggested that Betz cells potential role to temporary relax extensor muscle before movement may be progressively lost during ageing, and may result in the pain, stiffness, slowed motor movements seen in the hip/lower extremities of the ageing population. Rhesus monkeys also experience similar mobility reduction with age, however, unlike Scheibel's conclusions, these cannot be associated with loss of Betz cells. Accumulations of Bunina bodies with neurofilaments or cytoplasmic organelles are sometimes seen in older patients without a known neurological disease [198] and in ALS patients [445].

In several age-related neurodegenerative diseases, Betz cells are known to undergo pathophysiological changes. In Parkinson's Disease, Betz cells can be seen with Lewy bodies with insoluble accumulations of α-synuclein. Although, this may only contribute to a small fraction of the overall disease phenotype; one study found that of the 747 Betz cells studied in a patient 8 years from diagnosis only 1.1% (8 Betz cells) contained Lewy Body pathology [446]. Like other L5 neurons,

prominent degeneration can be seen in Betz cells with patients with multiple-systems atrophy [447] and corticobasal degeneration [448]. Others have seen Betz degeneration in Friederich's ataxis [449] and late-infantile galactosidosis [450]. In Pick's Disease, a type of frontotemporal dementia, Betz cells degenerate and clusters of macrophages commonly containing accumulations of Tau can be seen as their replacement [451]. Historical clinical case studies have noted patients with a complete absence of Betz cells with abnormal neuronal lamination [452], although the neurons in these patients may have shrunken and become difficult to identify other morphologically defining features. In lathyrism, a topical disease that results in paraplegia, tremors, and muscle weakness it is currently debated if Betz cells are the cause of observed spastic paraparesis seen in this condition [453]. Betz cells also appear to process a high amount of hnRNP-A1, a heterogeneous nuclear ribonucleoprotein. hnRNP-A1 is a key multipurpose RNA binding protein that can package pre-mRNA into hnRNP particles, and shuttling polyadenylation mRNA from the nucleus into the cytoplasm. This has been shown to act as an autoantigen in patients with human T-lymphotropic virus type 1 (HTLV-1)-associated myelopathy/tropical spastic paraparesis (HAM/TSP) [454], a disease whose clinical features are indistinguishable from multiple sclerosis. Mutations in hnRNP-A1 are also known to cause Amyotrophic Lateral Sclerosis. Marin-Padilla (1976) showed that in a child with Down's syndrome the apical dendrite of Betz cells has an increase in the number of dendritic spines which also contained a spectrum of abnormal morphology with some containing cystic vacuolization. The vacuolated spine-depleted apical dendrites of dying neurons may or may not be part of Down's syndrome neuropathology and may instead be due to cerebral anoxia, or a similar non-trisomy related phenomenon. The author notes that the vast number of elongated dendritic spines dramatically increase the surface area of the already giant Betz cells and struggles to explain how they can maintain their enormous size while being viable and functionally active [455]. The elongated dendritic spine morphology seen in this case is similar to the morphology seen in motor neurons (including Betz cells) in Patau syndrome [455]. Loss of Betz cells have been seen

in two cases of juvenile SMA [456] however, Betz cells are seemingly unaffected in the adult version of the disease [457]. Betz cell loss has also been reported in several types of spinocerebellar ataxia [458, 459].

B.5.1 TDP-43 Pathology in Betz Cells -

TDP-43 is part of a family of heterogeneous nuclear ribonucleoproteins (hnRNP) that can shuttle between the nucleus and cytoplasm and whose multifunctional abilities are critical for nucleic acid metabolism. They can perform nuclear (splicing, transcription, polyadenylation) or cytoplasmic (mRNA transport, mRNA translation, mRNA stability, and mRNA degradation) functions [460]. Several studies have shown that TDP-43 can bind to both RNA and DNA and is involved in axonal transport, cell division, and apoptosis [461, 462]. It is known that TDP-43 can shuttle proteins between the nucleus and cytoplasm, and while TDP-43 is mainly concentrated in the nucleus up to 30% still resides in the cytoplasm [463]. In 2011, an extensive RNA-protein interaction map provides evidence that TDP-43 was critical for regulating the splicing and expression of over 1000 RNA transcripts [464]. In 2006, it was discovered that the insoluble aggregated protein found in ALS motor neurons and glial cells was comprised mostly of TDP-43. This ubiquitinated cytoplasmic TDP–43 inclusions were revealed to highlight a common pathological hallmark within the ALS/FTLD spectrum. In ALS, TDP-43 aggregation is a neuropathological characteristic that is found in 97% of all cases, and 50% of cases with frontotemporal lobar degeneration (FTLD). Pathological TDP-43 burden is distributed across discrete anatomical regions of the CNS in ALS. This distribution follows a progressive that somewhat coincides with the disease progression of ALS. Once the cytoplasmic TDP-43 is ubiquitinated, it is unable to re-enter the nucleus, nor can it be eliminated by autophagy, proteasomal recycling, or metabolized, leading to excess accumulations [184]. Once in the cytoplasm phosphorylated TDP-43 begins to aggregate and form the characteristic TDP-43 inclusions discussed above. It is not known if it is the loss of nuclear TDP-43 or the gain in pathological cytoplasmic TDP-43 that is a potential disease mechanism in ALS [465]. Further, it

is not known if TDP-43 nuclear clearance and subsequently cytoplasmic aggregation is itself neurotoxic or if it is a response to neurodegeneration and may provide a neuroprotective benefit to underlying molecular mechanisms.

TDP-43 pathology – namely, nuclear clearing and the generation of insoluble cytoplasmic inclusions is seen in neurons with long axons i.e. bulbar and spinal α-motoneurons [184], while neurons with short axons are spared in the beginning stages of the disease [466, 467]. Intriguingly, Betz cells, containing the largest axons in the human CNS do not exhibit classical TDP-43 pathology [184]. A pivotal study in 2017 specifically examined the TDP-43 pathological changes in Betz cells in ALS [184]. In this, the authors examined post-mortem brains of 15 individuals with ALS, whose disease progression represented all 4 stages of ALS and contained individuals with arm, leg, bulbar, respiratory and truck disease onset. To minimise sampling errors that could potentially deemphasize TDP-43 pathology the authors sampled sections at 100µm thickness to encompass the entire perikaryon of the Betz cells resulting in >1000 Betz cells being used in their analysis. In this study, the authors observed that Betz cells did exhibit TDP-43 nuclear clearing, however, the ALS hallmark cytoplasmic inclusions were rare, in contrast to bulbar and spinal α-motoneurons in which cytoplasmic inclusions were almost always associated with TDP-43 nuclear clearing. There were no cases in this cohort where Betz cells exhibited abnormal TDP-43 pathology without the classic TDP-43 pathology seen in the bulbar and lower motor neurons. Interestingly, even in early-stage disease both neocortex and spinal cord showed TDP-43 burden. Within controls, $\approx 98\%$ of Betz cells contained a normal cell nucleus whereas across ALS cases $\approx 68\%$ of Betz cells contained a healthy nucleus, with $\approx 8\%$ containing reduced TDP-43 immunoreactivity, and $\approx 14\%$ displaying the complete loss of nuclear TDP-43. The remaining $\approx 10\%$ was marked as Betz cell loss as a result of degeneration as determined by the presence of lipofuscin pigment remnants in the neuropil of L5.

Although severely reduced, a few TDP-43 cytoplasmic inclusions were found within Betz cells in this study. This suggests that TDP-43 may still possess some aggregation properties, albeit considering the subtly of these inclusions and the

gigantic size of the Betz cell nucleus, these inclusions are still disproportionate to the nuclear TDP-43 originally present in the Betz cell. One possible explanation that may explain this TDP-43 Betz cell-specific ALS pathology, is that abnormal TDP-43 does exist in the cytoplasm of these neurons but is retained in a soluble state rather than becoming insoluble and aggregating. This could be due to TDP-43 being more dilute in a large cytoplasm and therefore having less pressure to become insoluble. Alternatively, there may be underlying cellular mechanisms unique to the Betz cell that reduce the insolubility of TDP-43 or modify the protein in such a way that it is not detectable by standard phosphorylation TDP-43 antibodies. It is unclear if this Betz cell TDP-43 phenomenon means Betz cells are selectively resistant to insoluble TDP-43 or if they are selectively vulnerable to pathogenic soluble TDP-43 causing them to degenerate before the formation of cytoplasmic inclusions. It is important to note that with time the rare/subtle inclusions in Betz cells do eventually develop into larger aggregates. Further evidence that soluble TDP-43 in the cytoplasm has a toxic role is apparent when considered that Betz cells with nuclear clearing die prematurely [184, 468].

C

Chapter 4 - Supplementary Data

Contents

C.1 Reproducibility of LCM Proteomics 271
C.2 Comparing Buffers 274
C.3 LCM Proteomics of FFPE Purkinje Cells 274
C.4 LCM Proteomic Comparison: Frozen & FFPE Granule Cells 276
C.5 Effect of Staining on LCM Proteomics 285

C.1 Reproducibility of LCM Proteomics

The time taken to perform LCM is often a critical parameter in RNA-targeted methods, as prolonged-time microdissecting can significantly reduce RNA quality. Most recommendations suggest limiting microdissection to under 1.5 hours as the physical heat from the laser along with natural degradation at room temperature has also been reported to have a deleterious effect on nucleic acids [469]. Unlike RNA, protein is relatively stable at room temperature, however, it is unclear if the process of LCM can negatively affect proteomic output over time. Similarly, as experiments become more intricate; either requiring extra time to microdissect enough rare/small cell types for representative downstream analysis or requiring additional time for complex microdissection strategies; it becomes increasingly

important to understand how the length of time microdissecting material may negatively skew results. To address this, Granule cell areas were microdissected (200,000µm^2) in triplicate over several time periods in an attempt to determine the maximum time that can be spent performing LCM before proteomic data is noticeably affected. Interestingly, there was no observed drop in performance between LCM experiments that took 1 hour to complete compared to experiments that took 4 hours to complete. Promisingly, protein intensity was highly correlated across all samples (minimum PCC: 0.952, maximum PCC: 0.988) with 95.1% (3,519) of all proteins identified being shared across all samples and replicates. Moreover, no pathways or DEPs were found to be significantly enriched across samples. While these data indicate that intricate LCM proteomic experiments requiring additional microdissection time can be performed without noticeable drops in quality, for future experiments in this thesis, all microdissection procedures were performed on a tissue section was within 1.5 hours, with an average time spent microdissecting each sample taking $\approx$ 30 minutes.

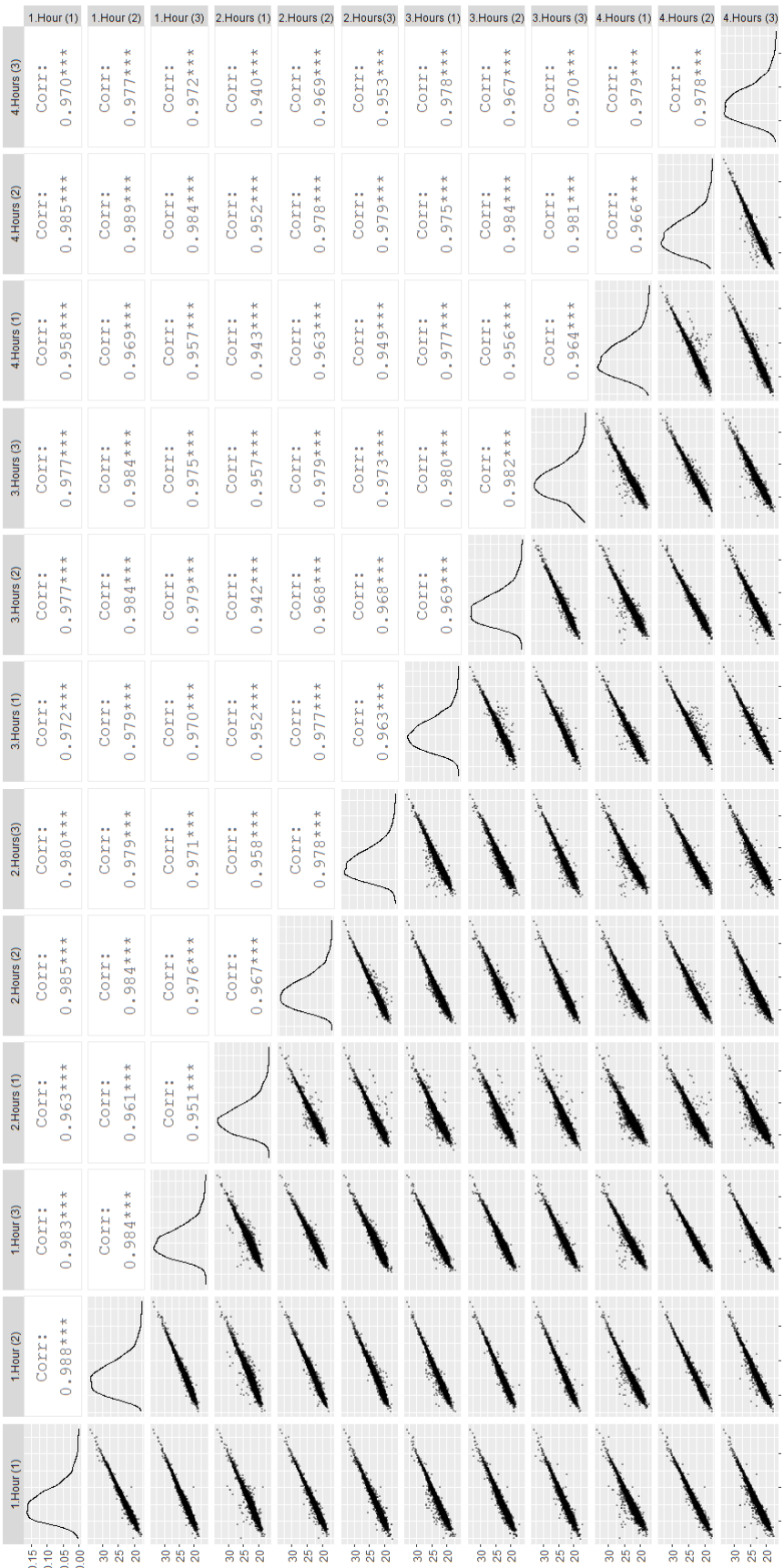


Figure C.1: No difference in LC-MS/MS performance was seen in longer LCM experiments.

C.2 Comparing Buffers

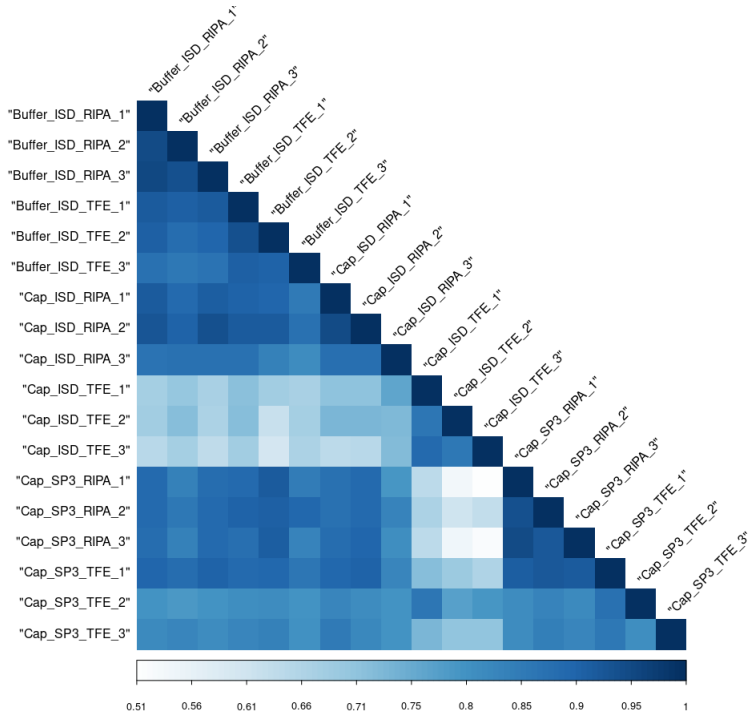


Figure C.2: Pearson Correlation Coefficient of RIPA and TFE.
The overall abundance of individual proteins did not change significantly in preparations using RIPA or TFE as a lysis buffer. ISD: In-solution digestion, indicating the standard sample preparation steps were performed. SP3: Samples were prepared using the magnetic bead approach from Hughes et al., 2014. Cap: Areas were catapulted directly into dry LCM caps, as done in the previous Purkinje cell experiments. Buffer: Areas were catapulted directly into LCM caps prefilled with lysis buffer. Legend: Pearson correlation coefficient.

C.3 LCM Proteomics of FFPE Purkinje Cells

To test whether the RWS method could also be used to gain proteomic insights into tissue fixed in formalin and embedded in paraffin wax another proof-of-principle titration experiment was done using Purkinje cell from FFPE human cerebellum.

While it was possible to resolve proteins from FFPE Purkinje cells using this approach there was minimal difference in the number of proteins identified compared to starting input material - Figure C.4. Despite an 8-fold increase in Purkinje cells,

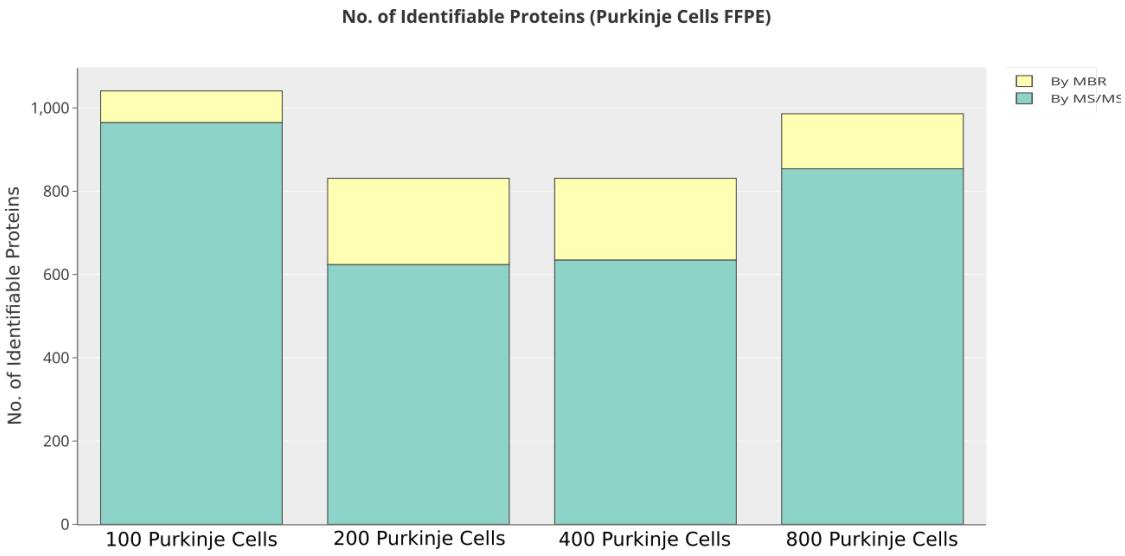


Figure C.3: FFPE Purkinje Cell Titration - No. of Identifiable Proteins. Turquoise bars: Number of proteins identified by MS/MS, Cream Bars: Number of proteins identified by MBR.

the numbers of identifiable proteins were almost identical between the 100 and 800 cell samples, this was most likely caused by proteins not being sufficiently released from the cross-linking effects of formalin, along with the difficulties of working with wax. In addition to this, there was an overall reduction in the number of identifiable proteins compared to the frozen equivalent experiment ($\approx$ 800 vs. $\approx$ 2,000). Despite these shortcomings, it was still possible to resolve proteins relating to Purkinje cell function as with the frozen Purkinje cell titration experiment. The previously mentioned, ATP2B2, HSPA5, SPTBN2, and MYH10 was present in all fractions, with CEND1 found in the 100, 200 and 400 cell sample, and ARCN1 found only in the 100 Purkinje cell sample. Interestingly, the protein UQCRQ (which contained Purkinje cell-specific GO annotations) was present in all FFPE samples but was absent in all previous Purkinje cell frozen counterparts despite their considerable higher number of protein IDs. Promisingly, as with the previous frozen Purkinje cell work, the biomarkers: CALB1, CA8, PRKCG, ALDOC were also found across these FFPE samples.

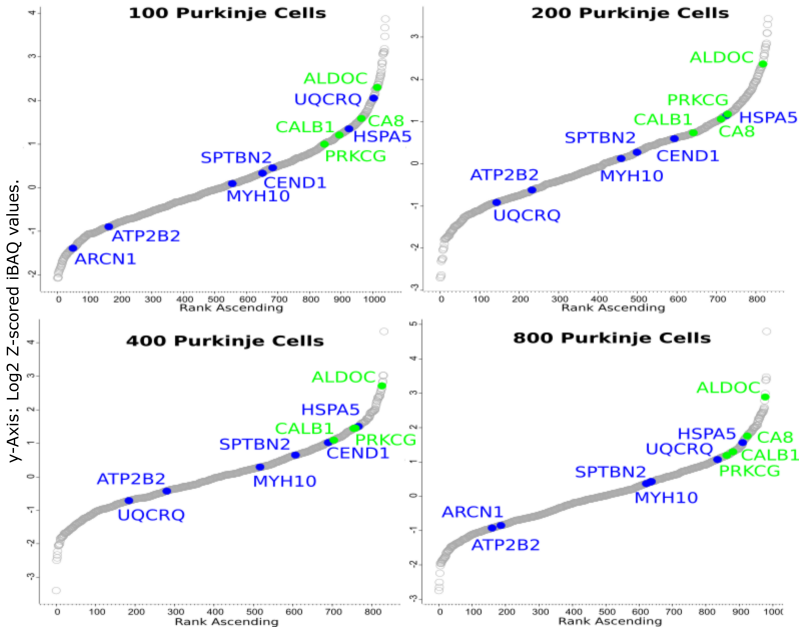


Figure C.4: Purkinje Cell-Related Proteins within FFPE Purkinje Cell Titration. Intensity based absolute quantification (iBAQ) highlights several proteins that are involved in Purkinje cell biology in FFPE Purkinje cells. Blue: Proteins with Purkinje-cell Gene Ontology. Green: Known canonical biomarkers for the Purkinje cell.

C.4 LCM Proteomic Comparison: Frozen & FFPE Granule Cells

Next, it was decided to directly compare the FFPE proteome and frozen proteome as previous evidence has suggested that each preservation method may result in different proteins being identified in their collective fractions. Rather than manually annotating hundreds of Purkinje cells it was decided to find a surrogate material to microdissect. As future work aims to investigate the perikaryon of neurons it was decided to microdissect areas containing the Granule cell of the cerebellum, rather than the molecular layer which is relatively sparse of cellular features. Granule cells originating from either frozen tissue or FFPE tissue were laser microdissected and processed using the Buffer-RIPA-SP3 method. The only difference in upstream processing was that FFPE underwent deparaffination before LCM and contained an additional heating step to reverse the cross-linking of nucleic acids to proteins due to the formalin fixative. Areas equal to 50,000, 100,000, 200,000 and 400,000µm^2

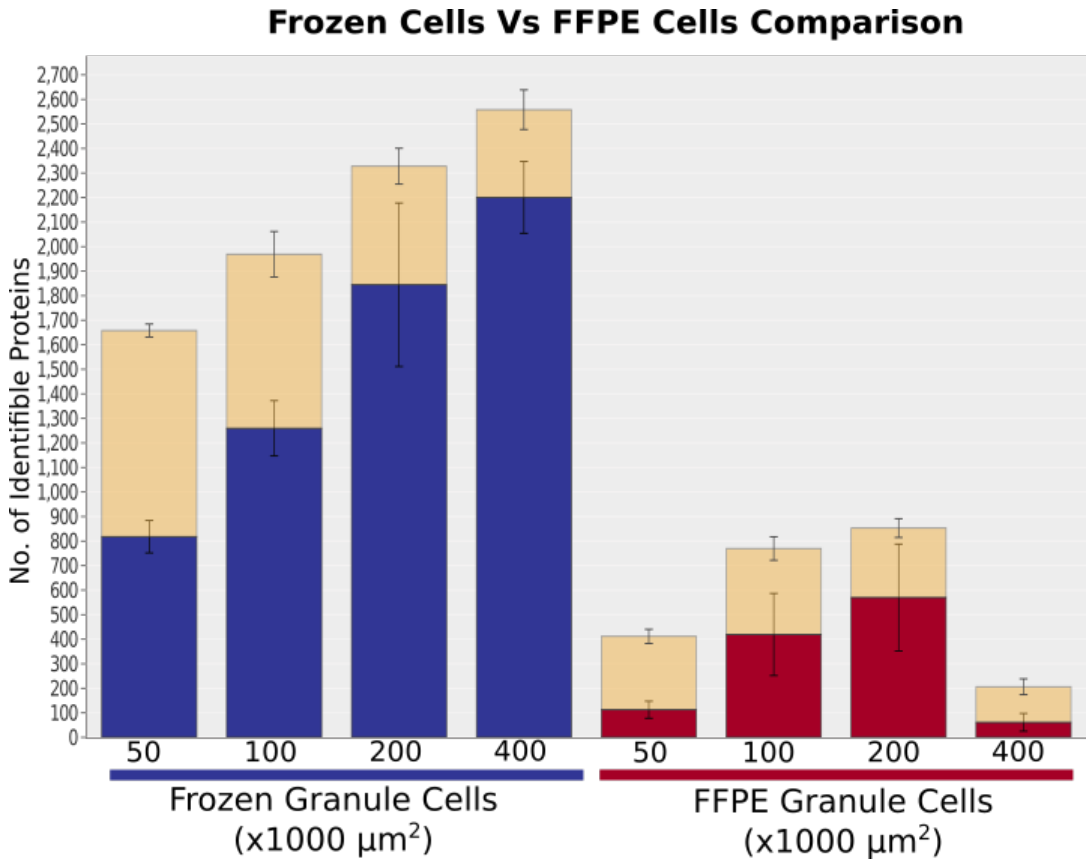


Figure C.5: FFPE vs. Frozen Material - No. of Identified Proteins. Frozen (Blue) or FFPE (Red) Granule cells were microdissected in areas correspondingto 50,000, 100,000, 200,000, and 400,000µm^2. Orange bars: Proteins identified by MBRs. Vertical bars: standard deviation of samples.

were taken in triplicate from the same human cerebellum.

Frozen Granule cells when processed using the optimised RWS protocol yielded high amounts of proteins, likely reflecting the benefit of having a highly efficient workflow with little sample loss. There was a pronounced decrease in total protein identifications by MS/MS in the FFPE counterparts compared to the frozen Granule cells (Mann Whitney Test, p value: <0.0001) - see Figure C.5. Similar to the previous FFPE Purkinje titration experiment there was an unexpected relationship between the amount of input material and the number of identifiable proteins in FFPE samples, unlike the plateau seen previously with frozen tissue, there seemed to instead be a pronounced decreased in identifications in the 400,000µm^2, most likely attributed to the increased difficulty of isolating proteins from larger quantities of FFPE tissue.

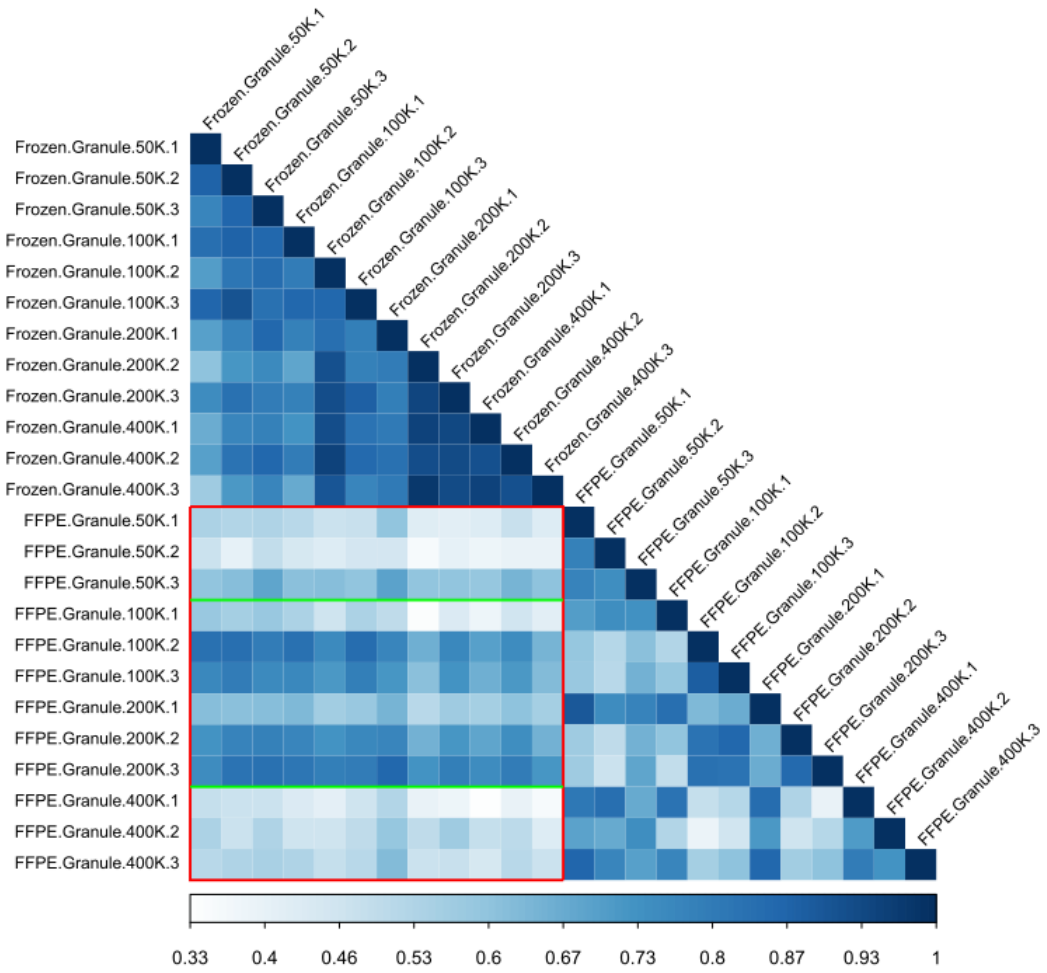


Figure C.6: Pearson Correlation Coefficient of FFPE vs. Frozen material. Red box: PCC of Frozen samples to FFPE samples. Green box: PCC of Frozen samples to FFPE samples that contained higher protein IDs. Frozen Granule cells had better reproducibility among themselves than FFPE. Legend: Pearson correlation coefficient.

Maxquant's MaxLFQ Algorithm was applied to better compare the abundances of individual proteins across samples. There was a high amount of reproducibility within triplicates for frozen Granule cells (PCC: 0.930) which was slightly lowered among FFPE Granule cells (PCC: 0.863) - Figure C.6. Furthermore, there was also a high level of correlation in the label-free intensities across all frozen Granule cell samples (PCC: 0.848) and a moderate level of correlation across all FFPE samples (mean PCC: 0.720). Promisingly, despite their differences in sample preparation and workflow, there was a modest correlation between label-free intensities between frozen cells and FFPE cells (mean Pearson correlation: 0.587) – highlighted by

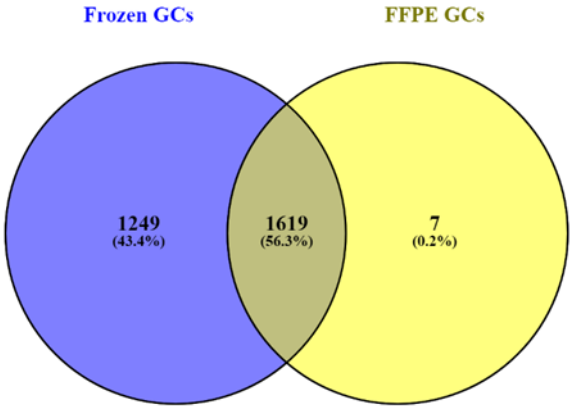


Figure C.7: Overlap in Proteins between FFPE vs. Frozen Samples.

the red box. When comparing LFQ intensities of frozen cells to the subset of FFPE samples that performed best and yielded the most protein identification i.e. the 100K and 200k samples (highlighted by the green box) there was a clear increase in correlation (mean PCC: 0.720) which suggests that if the proteomic depth was increased in FFPE samples there may be a higher correlation between FFPE and frozen sample protein intensities.

Next, individual proteins were investigated to see if they were found mutually within either the frozen or FFPE counterparts or found in both. Previous proteomic studies have noted seeing a marked difference regarding expression and identification of the "frozen proteome" and the "FFPE proteome" suggesting that both compartments need to be analyzed to get the best representation of the total proteome within a sample. However, this is not reflected in this data-set; of the 2,868 proteins were found within the frozen Granule cells and 1,626 proteins were found in the FFPE counterparts (total). Surprisingly only 7 (0.2%) proteins were found solely in FFPE samples with proteins identified in the frozen Granule cells making up 98.7% of the total proteins discovered.

Gene Ontology over-representation analysis (ORA) was used in an attempt to identify why particular proteins were differentially found in samples. The results show that although many ontology pathways were significant within the frozen sample proteome, the overall enrichment of those pathways was modest - data not

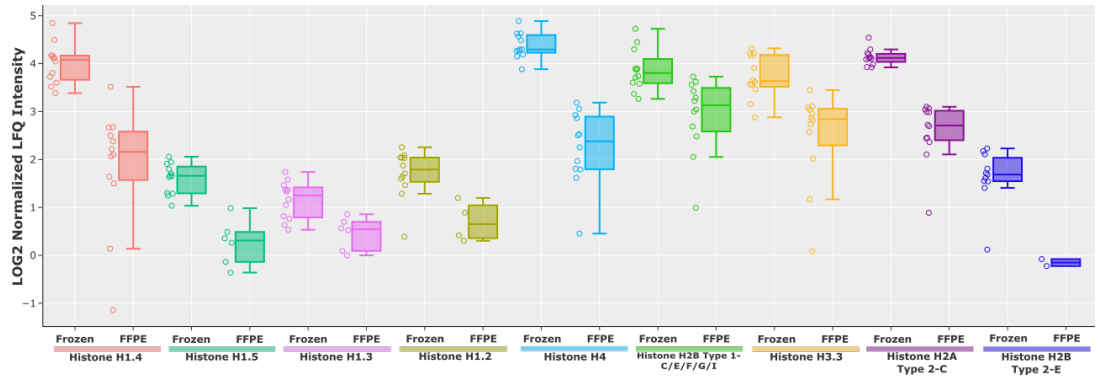


Figure C.8: Histone Proteins were Reduced in FFPE Preparations for LCM Proteomics.

shown - suggesting a non-specific enrichment of all proteins. Interestingly, the FFPE fraction contained a significant enrichment in proteins (PLAG2G4D, PLAG2G4B, ANXA9, DUSP14) relating to enrichment in phospholipid-binding (FDR: 2.99e-02).

When Principal Component Analysis (PCA) was applied to these samples frozen and FFPE samples were divided between the 1st (60.2%) and 2nd (21.1%) component of PCA. Investigating further, the individual proteins attributing to this separation were all histone proteins: Histone H3.3, Histone H2B type 1-C/E/F/G/I, Histone H1.2, Histone H1.4 and Histone H4. Overall, 10 histones were shared between FFPE and frozen samples. Of these 10 histone proteins, 9 were highlight as significant (Welch T-test; FDR threshold: 0.05, 250 randomisation) - Figure C.8. Overall, there was a reduction in histone proteins in FFPE samples, which may reflect the inefficiency of this approach to break DNA-protein links caused by formalin.

Next, the differences in individual protein abundance were accessed between the frozen and FFPE samples. A Welch T-test was performed (FDR threshold: 0.05, 250 randomizations) on the 1,613 quantifiable proteins found within both cell fractions; 987 proteins were suitable for analysis, and 607 were marked as significant. Of all the proteins deemed significant, 96.54% had a bias towards being lowered in FFPE samples - see Figure C.9.

Several GO terms were significantly enriched among the 586 proteins found to be lowered in FFPE samples, although enrichment within pathways was modest, having a fold change of less than 2, with metabolic pathways being the most affected. No

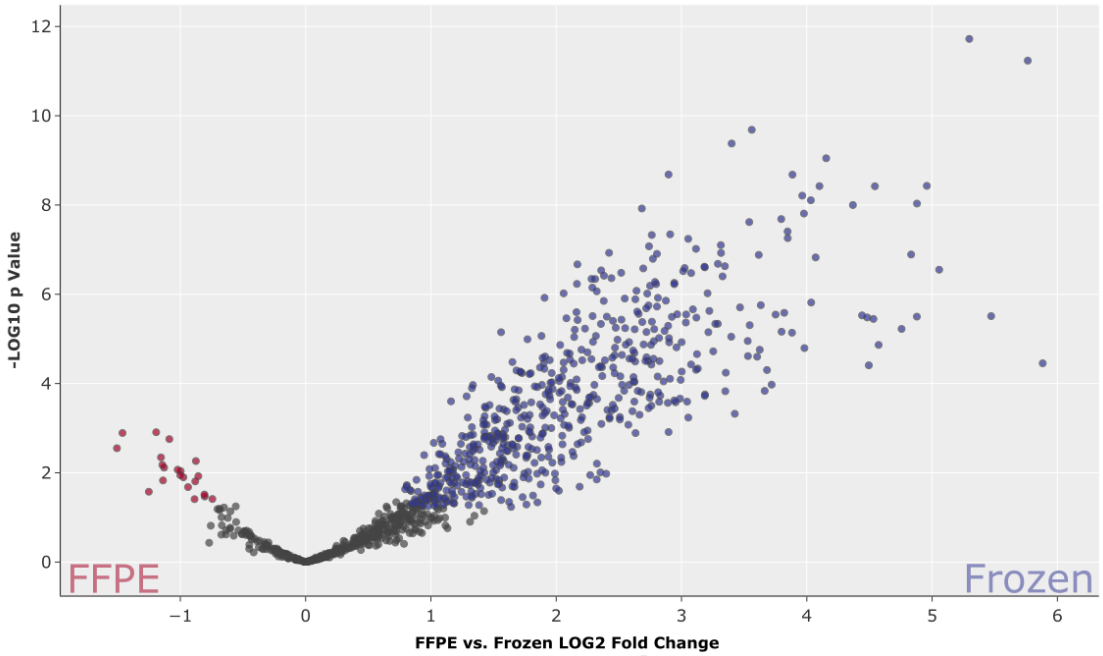


Figure C.9: Volcano Plot of FFPE vs. Frozen Material.
Each point represents one protein. A large reduction in protein abundance was seen in FFPE samples when compared to their frozen counterparts. Coloured points are DEPs that passed predetermined alpha thresholds (0.05). Red: proteins whose expression was higher in FFPE samples. Blue: proteins whose expression was lower in FFPE samples.

GO terms were enriched within the proteins significantly higher in the FFPE samples. The LFQ intensity of all proteins within the frozen samples was compared to the FFPE samples and a global decrease in protein LFQ intensity was noted between the two sample types (Mann Whitney test, p value: <0.0001) - see Figure C.10.

While snap freezing of fresh tissue currently remains the gold standard in preserving biomolecules for downstream *'omics'* applications [470], this approach is rare with the most common method to preserve samples consisting of a formalin-fixed paraffin embedding protocol; this is especially true with biobanks containing clinical samples where storing frozen samples is logistically difficult and expensive. While suitable for histopathology, several groups have identified that formalin has a pronounced effect on the proteins within a sample [471–473]. Although a considerable amount of work has been done to investigate the inter-peptide cross-linking of formalin, complicated models still need to be generated to account for

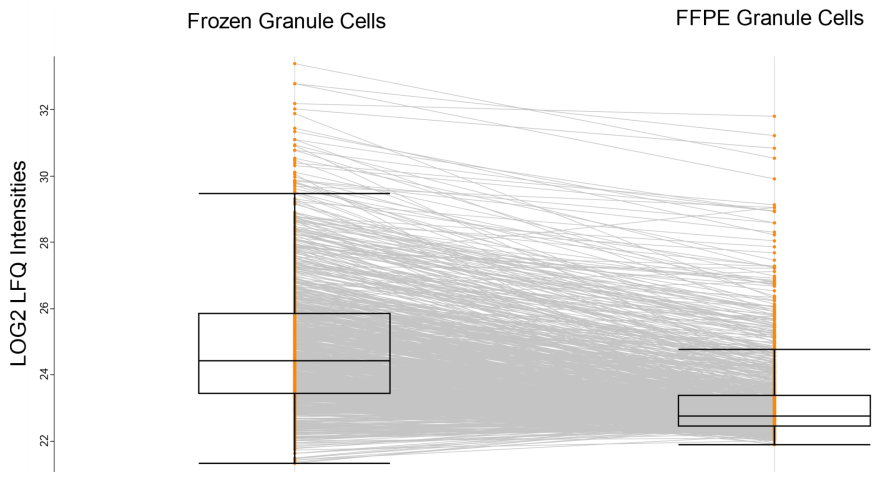


Figure C.10: Overall Abundance of Proteins in FFPE vs. Frozen Granule Cells. As seen in Figure C.9, a global shift in proteins being less abundant in FFPE samples was observed in this comparison.

the vastly complex nature of tissue blocks – this global effect on proteins has led to the historic paradigm that FFPE tissue is too challenging for proteomics. However, there has been a considerable push to gain access to this invaluable source of material for proteomics and it is now possible to get meaningful data due to advances in separation, removal of strong detergents during sample preparation, and mass spectrometry detection [471]. The performance of FFPE proteomics has improved dramatically in recent years; most studies comparing FFPE to frozen material, report only a small reduction in protein IDs with FFPE samples [474–476], while some produce comparable protein IDs compared to their frozen counterpart [477–479], and others find higher protein IDs within the FFPE samples [480]. This discrepancy in studies can most likely be attributed to variation in inter-laboratory FFPE protocols and LCM FFPE proteomic sample preparation. Overall, FFPE proteomic studies suffer from shallow proteomic depth; however, as access to cutting edge mass spectrometers becomes commonplace this dichotomy between tissue preservation methods may reduce. Despite the reduction in protein IDs, it is now possible to use FFPE proteomics to independently study disease [146, 291].

The RWS method uses the maxLFQ algorithm to measure the relative abundance of peptides across different samples without the need for labelling techniques.

MaxLFQ minimises the amount of sample preparation steps and has been shown to rival relative quantification achieved by labelled approaches [260]. However, labelled entities may still be added to increase the sensitivity of mass spectrometers, and in the context of FFPE proteomics may increase performance. Recently, one group has shown that it was possible to achieve >5,600 proteins from FFPE material extracted from post-mortem substantia nigra by using 10-plex tandem mass tag (TMT) labelling [291]. Similar to our frozen tissue work, Griesser et al., (2020) optimised their sample preparation approach by testing a variety of lysis buffers to maximise protein extraction, they also adopted the SP3 method for isolating peptides/proteins and compared how their approach was affected by immuno-staining. Identifying >5,600 is an impressive feat for pooled cells of post-mortem origin – especially from FFPE material. However, to achieve this many procedures were implemented that considerably increase labour and may be unrealistic for laboratories looking to analyse larger sample sizes or without regular access to a mass spectrometry facility. Firstly, the >5,600 protein IDs were resolved from microdissected total areas instead of individual neurons with the authors equating this to approximately 3,000 cells – while this does reduce time spent isolating neurons it does not remove the averaging effect that commonly affects nervous tissue. The authors make use of sample fractionation to maximise protein IDs which is a valid method of increasing the identifications within a sample, albeit, with increased labour. Furthermore, a 4-hour mass spectrometry protocol was employed to maximise time spent on the mass spectrometer, the authors reduced this in one of their TMT batches, however, it resulted in significantly lowered protein IDs, so the 4-hour MS protocol was re-introduced. This increase in MS time coupled with the 24 fractions needing to be analysed per sample dramatically increases the amount of MS running time on an experiment (≈96 hours per sample), and may only be suitable on small experiments. Finally, Griesser et al., (2020) compared their TMT 16 fractions method and their TMT 24 fraction method to a label-free single-shot approach, similar to our own; and found that the TMT approach increased the number of quantified protein groups by ≈220% highlighting the

powerful application of TMT labelling. However, when accounting for the long MS time and large tissue input this equates to 4.88e-05, 5.58e-06, and 1.86e-05 proteins identified per 1000µm^3 per one gradient hour for the TMT 16, TMT 24 and label-free approaches; respectively; highlighting the extreme difficulty and dedicated methods needed to identify large amounts of protein from post-mortem laser dissected FFPE material. Nevertheless, this approach shows the powerful nature of TMT labels, and many groups have adopted to modify their sample preparation to include such tags, namely NanoPOTS, who have established *Nanogram TMT Processing in One Tube* or NanoTPOT [481]. Our approach will undoubtedly benefit from the use of TMT tags and/or fractionation, however, to implement such techniques significant sacrifices will need to be made that will reduce the rapid nature of our approach.

C.5 Effect of Staining on LCM Proteomics

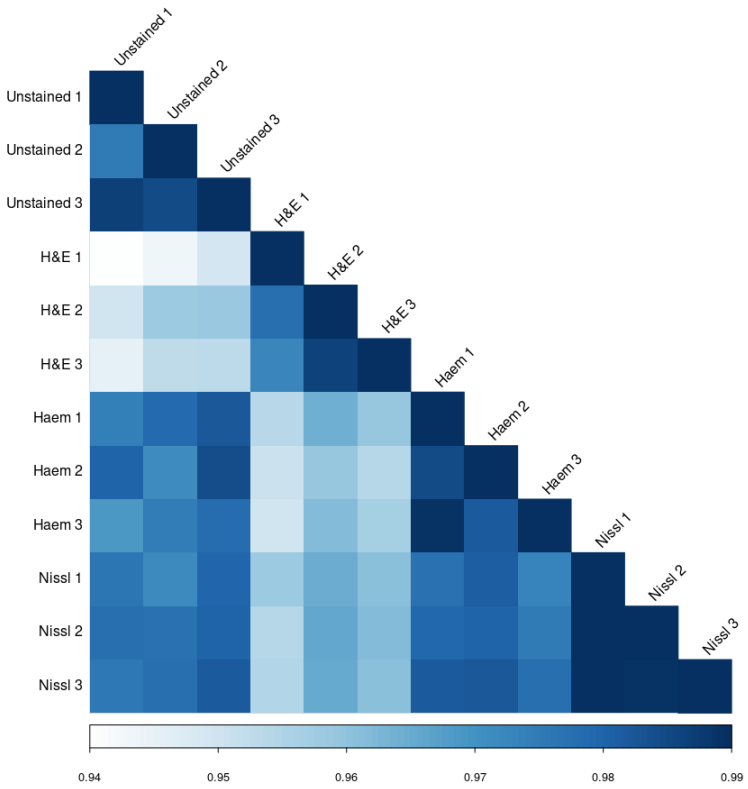


Figure C.11: Pearson Correlation Coefficient Heatmap - Different Staining Methods. All staining methods resulted in a high correlation in protein abundances to unstained material. Nissl was the most reproducible among replicates, and both Nissl and Haematoxylin strongly correlated with unstained cells.

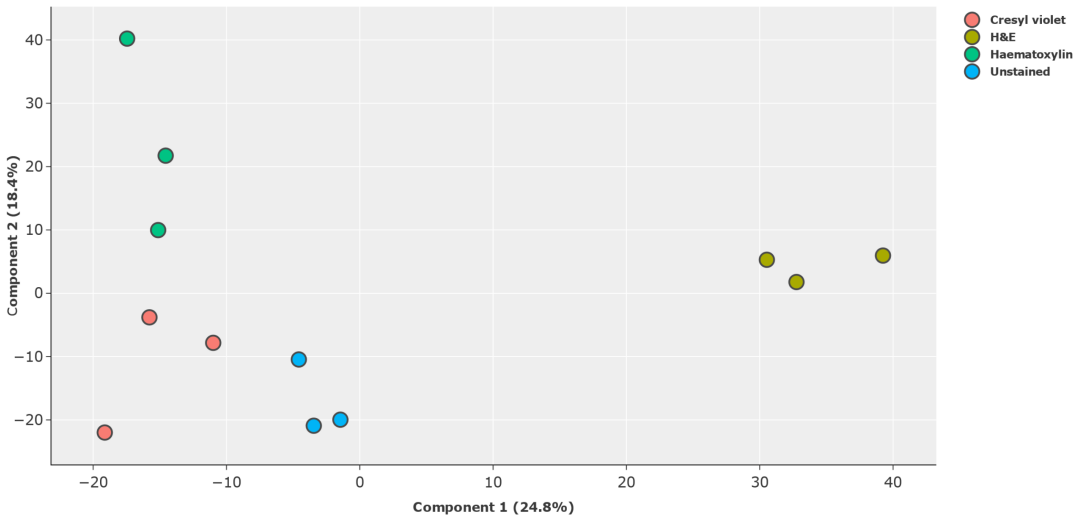


Figure C.12: Principal Component Analysis (PCA) - Different Staining Methods. All staining methodologies grouped with their respective replicates. H&E stained material differed the most from the other stains and the unstained material.

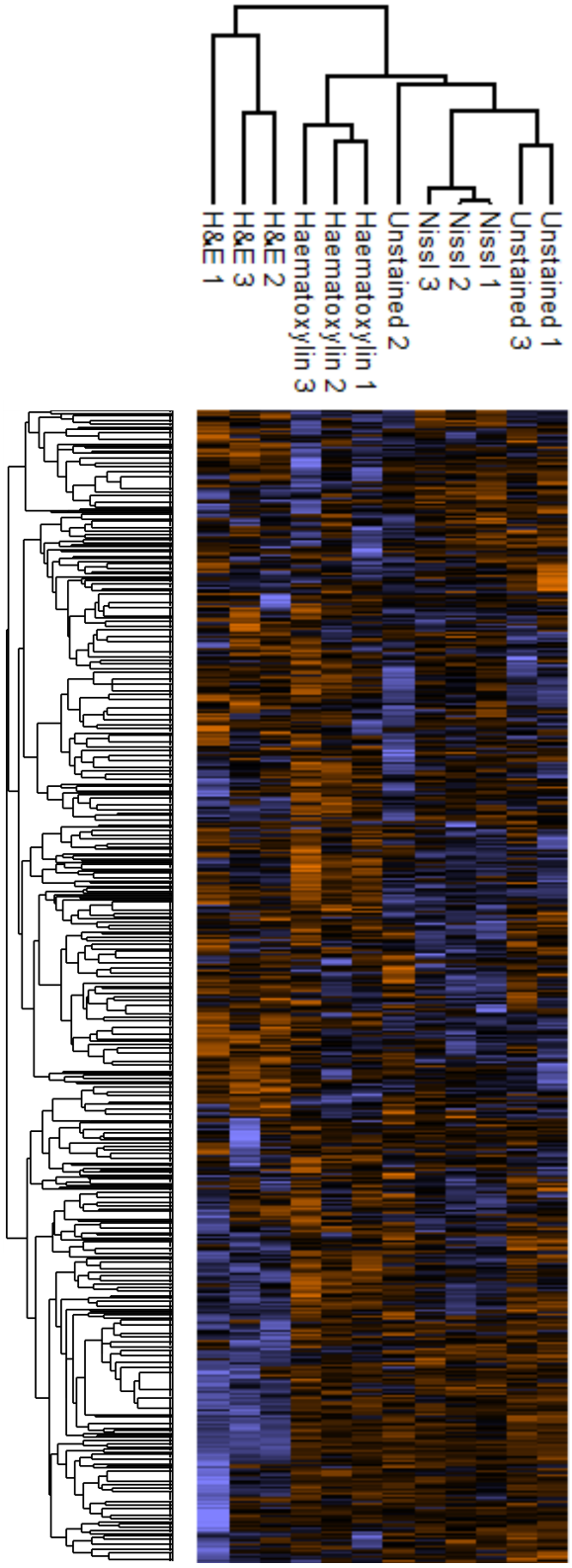


Figure C.13: Hierarchical Clustering - Different Staining Methods. Unbiased clustering confirmed that H&E samples were the most dissimilar to the other samples analysed. Moreover, Nissl staining was found to affect the underlying proteome of granule cells the least.

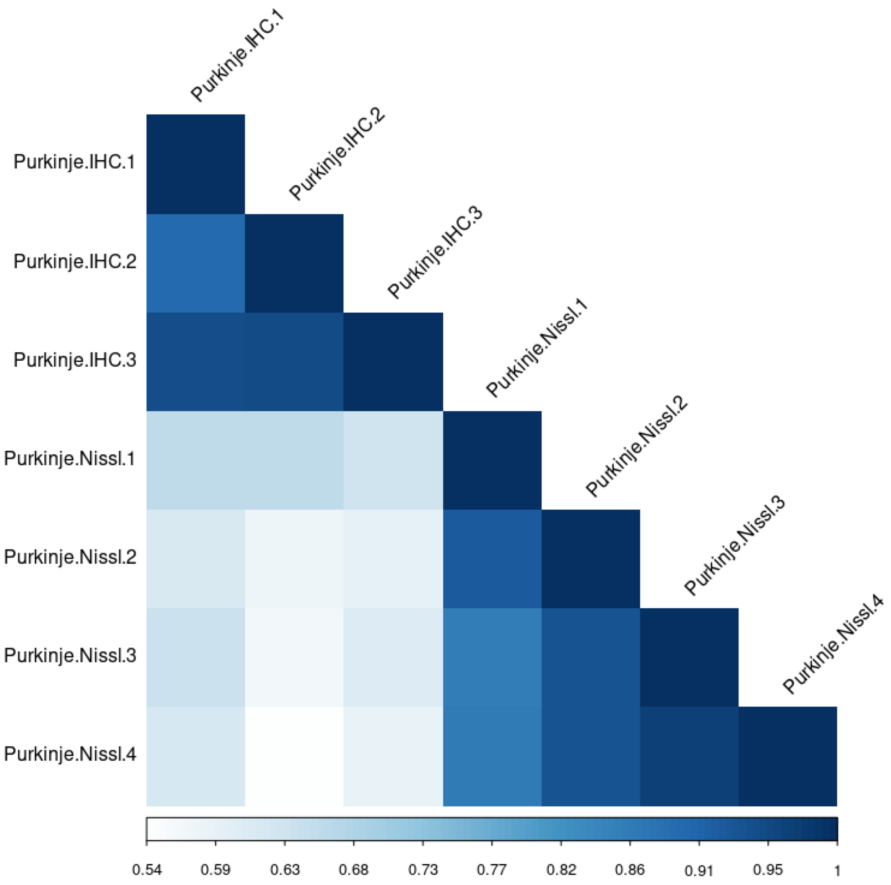


Figure C.14: Pearson Correlation Coefficient Heatmap - IHC vs. Nissl.
While protein abundance correlated well within replicates, there was only a slight correlation in protein abundance between IHC stained and Nissl stained Purkinje cells.

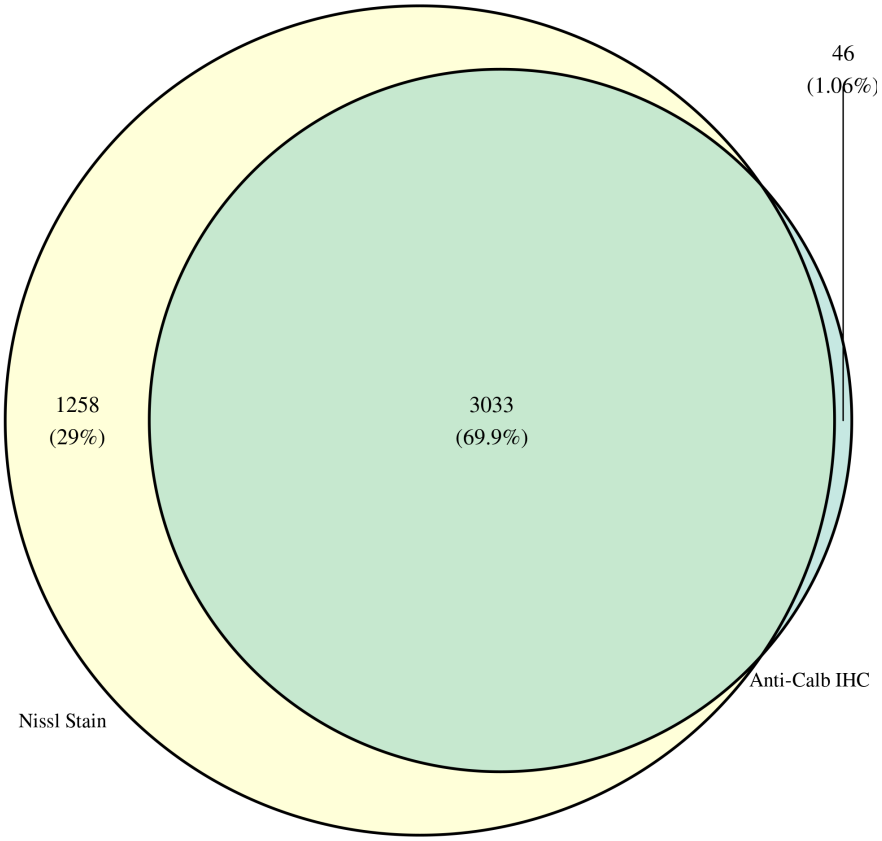


Figure C.15: Venn Diagram - Protein Overlap (IHC vs. Nissl).

D

Chapter 5 - Supplementary Data

Contents

D.1 Functional Classification 292
D.2 Absolute Proteomics - Betz Cells 294
D.3 Absolute Quantification - Betz Cells, Purkinje Cells, DNNs, and L5 Motor Neurons 299
D.4 Relative Proteomics - Betz Cells, L5 Neurons, DNNs, and Purkinje Cells. 302
D.5 Betz Cell Comparisons - 1D Annotation Enrichment . 310
D.6 Additional IHC Data 316
D.7 Excessive Contamination of Neuropil Proteins in L5 Neuronal Samples 319

D.1 Functional Classification

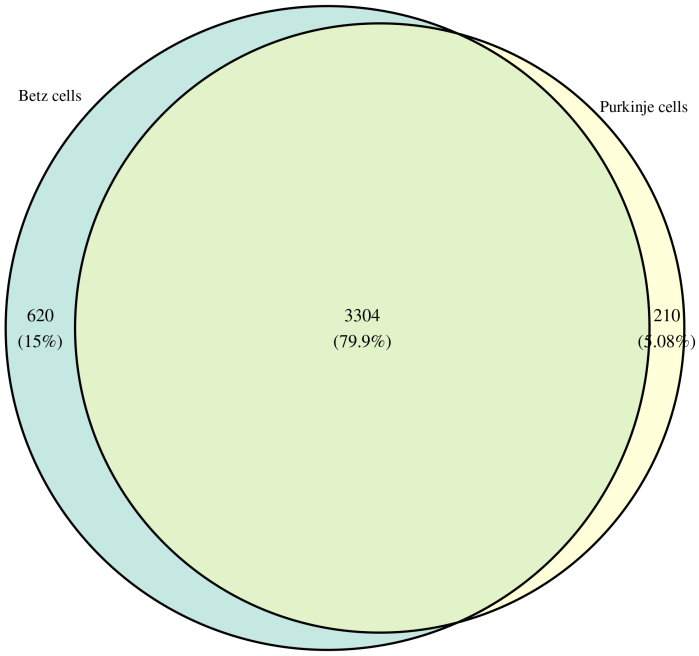


Figure D.1: Venn Diagram - Betz Cells vs. Purkinje Cells.
Number of proteins identified in each sample. As expected, there was a large overlap in identifiable proteins between these two large neurons.

Proteomaps [482] was generated using the proteins found within each cell fraction in an attempt to visualise the global proteome. Within this, each protein was designated a polygon coloured coded to its biological function and sized according to its relative abundance. Unfortunately, not all proteins are suitable for treemapping. Within the Betz cell samples, only 44.95% (1,692) of proteins were mapped. Nevertheless, visualising this subset of the Betz data provides a visual overview of the proteins identified within Betz cells - Figure D.2. Gene Ontology Functional Classification via PANTHER [269] was used to identify the protein pathways found with the Betz cell samples to gain further insights into the Betz cell proteome. Of the 3,924 proteins within the Betz cell samples, 1,688 had annotations for particular protein pathways - in total 145 pathways were identified within the Betz cell (data not shown). Both Betz and Purkinje cells contained roughly the same proportion of proteins relating to specific MF, BP, and CC Gene Ontology terms;

this is unsurprising considering the hierarchical organisation of Gene Ontology and the fact that there was a large overlap of proteins present in each sample (79.9%).

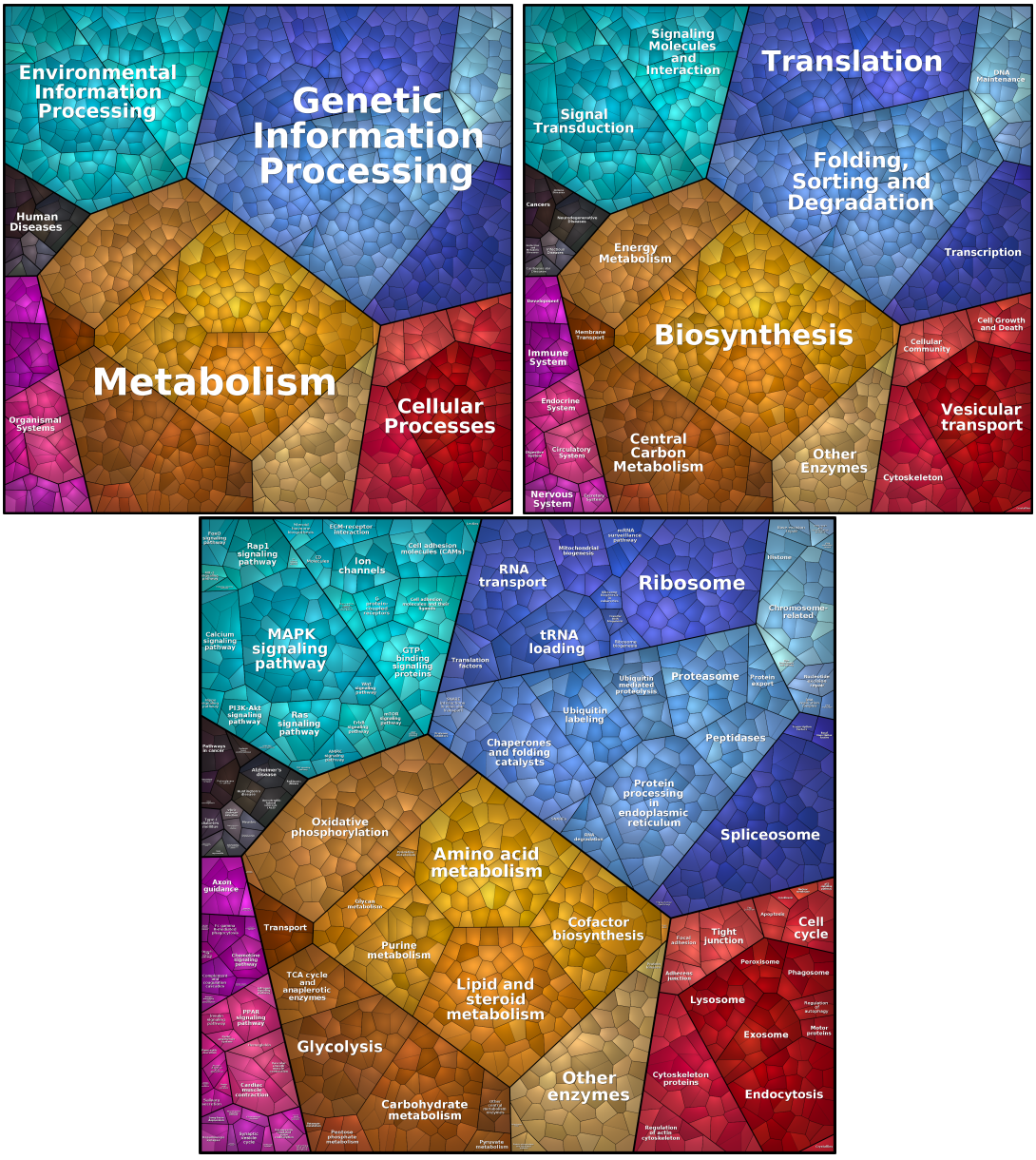


Figure D.2: Proteomap of Proteins found in Betz Cells.
An expression-weighted treemap summarising the proteins and their associated pathways that were found within the microdissected Betz cells. Each polygon represents one protein and the size reflects its relative expression. Classes: Teal - Environmental Information Processing, Blue - Genetic Information Processing, Red - Cellular Processes, Yellow - Metabolism, Pink - Organismal Systems, Black - Human Diseases.

D.2 Absolute Proteomics - Betz Cells

1D enrichment was performed on the iBAQ (Intensity Based Absolute Quantification) values within each sample to investigate which particular proteins/protein groups had a tenancy to be altered significantly compared to the total distribution of proteins. Within the Betz cell samples, 189 Biological Process, 91 Cellular Component, 88 Molecular Function, and 37 KEGG terms were significantly differentially distributed across the proteomap (Figure D.3). The most abundant proteins tended to be related to metabolism: Oxidative Phosphorylation (1D Score: 0.7522; Benj Hoch FDR: 5.84E-34), NADH dehydrogenase activity (1D Score: 0.7358, Benj Hoch FDR: 5.42E-13), Gluconeogenesis (1D Score: 0.6157, Benj Hoch FDR: 3.74E-06). With proteins found in the proton-transporting ATP synthase complex and mitochondrial membrane being disproportionately present among the most abundant proteins, along with proteins isolated to in the nuclear nucleosome (1D Score: 0.916; Benj Hoch FDR: 0.018), cytosolic large ribosomal subunit (1D Score: 0.7923; Benj Hoch FDR: 1.42E-19) and cytosolic small ribosomal subunit (1D Score: 0.7923; Benj Hoch FDR: 1.64E-15) - See Figure D.5 and D.6.

Next, the distribution for Betz cell proteins to be highly or lowly abundant as compared to the iBAQ values of the Purkinje cell using 2D enrichment - see Figure D.4. Many of the significantly enriched terms in the Betz samples followed the same abundance within the Purkinje cell sample, with proteins relating to Oxidative phosphorylation (Benj Hoch FDR: <0.1E-13) and ATP synthase complex (Benj Hoch FDR: <0.1E-13) being the most abundant; most likely attributed to the fact that their large size most likely requires a considerable amount of energy production. While there was an overall correlation between the iBAQ values between the Betz and Purkinje samples (PCC: 0.758, p value: 2.2e-16) there does exist several terms with differing concordance. Compared to the Purkinje cell samples, the Betz cell samples contained a higher abundance of proteins relating to neurotransmitter transport (No. of proteins: 52, Benj Hoch FDR: 1.40E-02), synaptic vesicle (No. of proteins: 48, Benj Hoch FDR: 9.62E-03), synaptic transmission (No. of proteins:

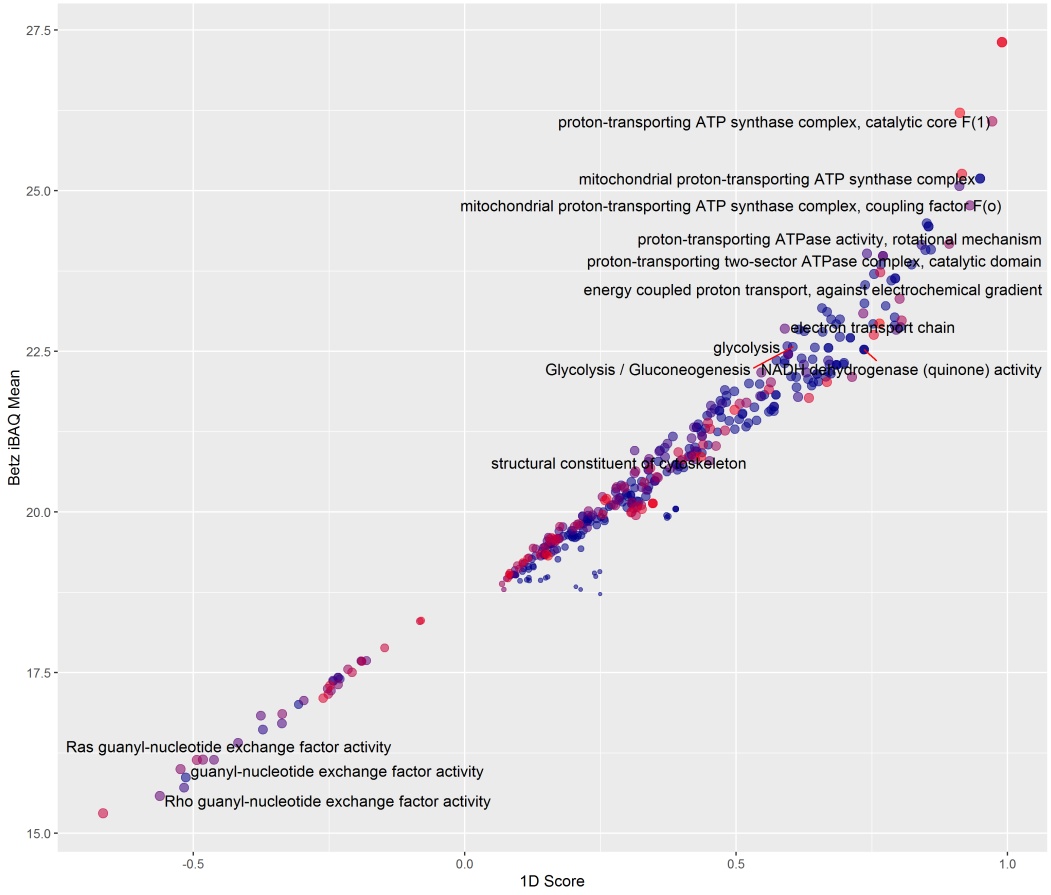


Figure D.3: 1D Annotation Enrichment of Betz Cell iBAQ Values. Comparing the distribution of proteins within Betz cells. Proteins relating to energy generation and the ATP synthase complex were among the most abundant proteins in the Betz cells. Each point represents a significantly enriched GO term. Size of point: number of proteins involved in that term. Colour: Darker colours (Blue) have smaller p values.

130, Benj Hoch FDR: 3.63E-04). Conversely, within the Betz cells samples, proteins found in the nuclear speck (No. of proteins: 50, Benj Hoch FDR: 4.99E-05) or Cajal body (No. of proteins: 12, Benj Hoch FDR: 3.04E-03) tended to be less abundant when compared to Purkinje cells. The Purkinje cell samples had several terms that were slightly enriched and had a negative correlation with the Betz cell samples: mRNA export from the nucleus (No. of proteins: 28, Benj Hoch FDR: 1.03E-02), mRNA transport (No. of proteins: 46, Benj Hoch FDR: 5.45E-03), mRNA splice site selection (No. of proteins: 18, Benj Hoch FDR: 1.96E-02), RNA splicing (No. of proteins: 122, Benj Hoch FDR: 2.15E-11) and spliceosomal

complex (No. of proteins: 66, Benj Hoch FDR: 2.68E-07). Finally, there were several pathways relating to cell junction adhesion that was lowered in the Purkinje cell when compared to the Betz cell.

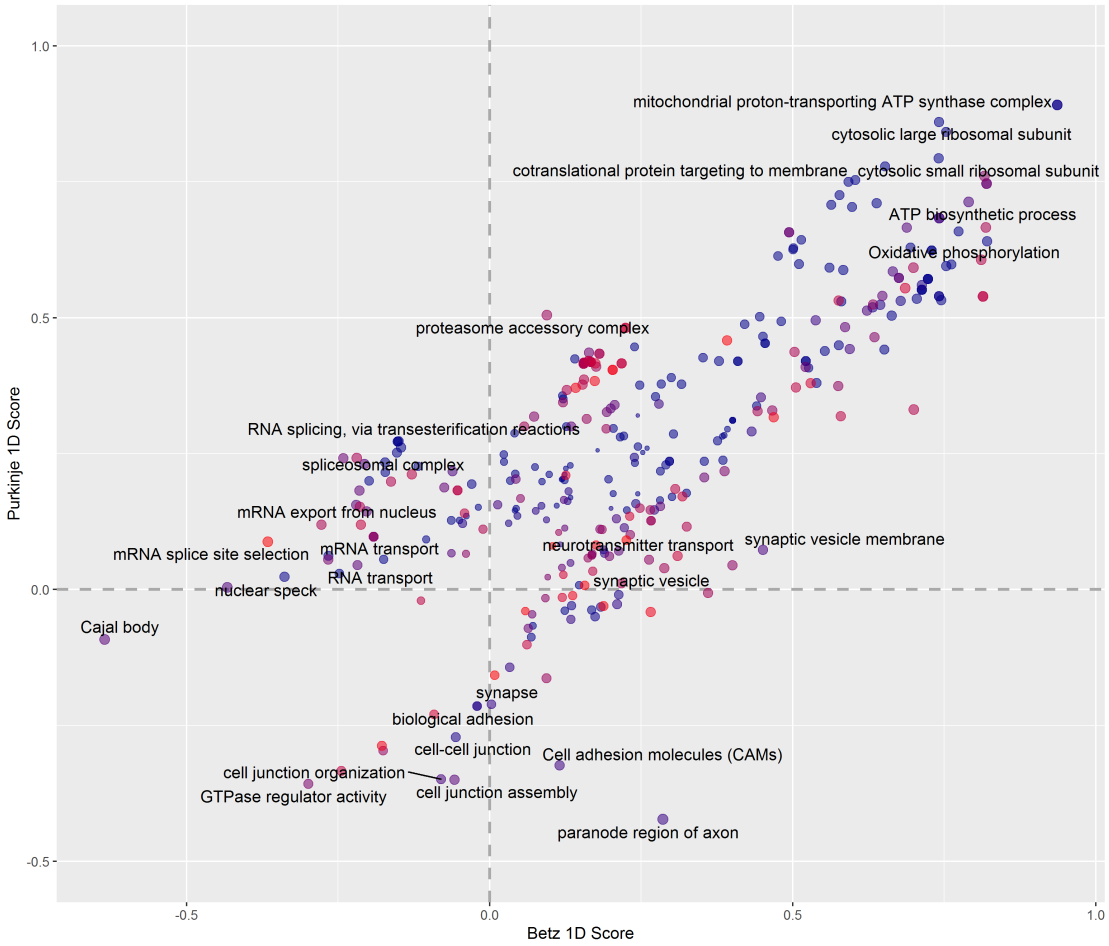


Figure D.4: 2D Annotation Enrichment of Betz/Purkinje cell iBAQ Values. Comparing the distribution of proteins within each cell-type. Overall, there was a strong correlation between the distribution of different protein pathways between Betz and Purkinje cells. However, proteins relating to mRNA transport and splicing were high in Purkinje cells but were among the least abundant proteins in Betz cells. Similarly, Purkinje cells had a low abundance of cell adhesion proteins, while these were found in moderate amounts in Betz cells. Each point represents a significantly enriched GO term. Size of point: number of proteins involved in that term. Colour: Darker colours (Blue) have smaller p values.

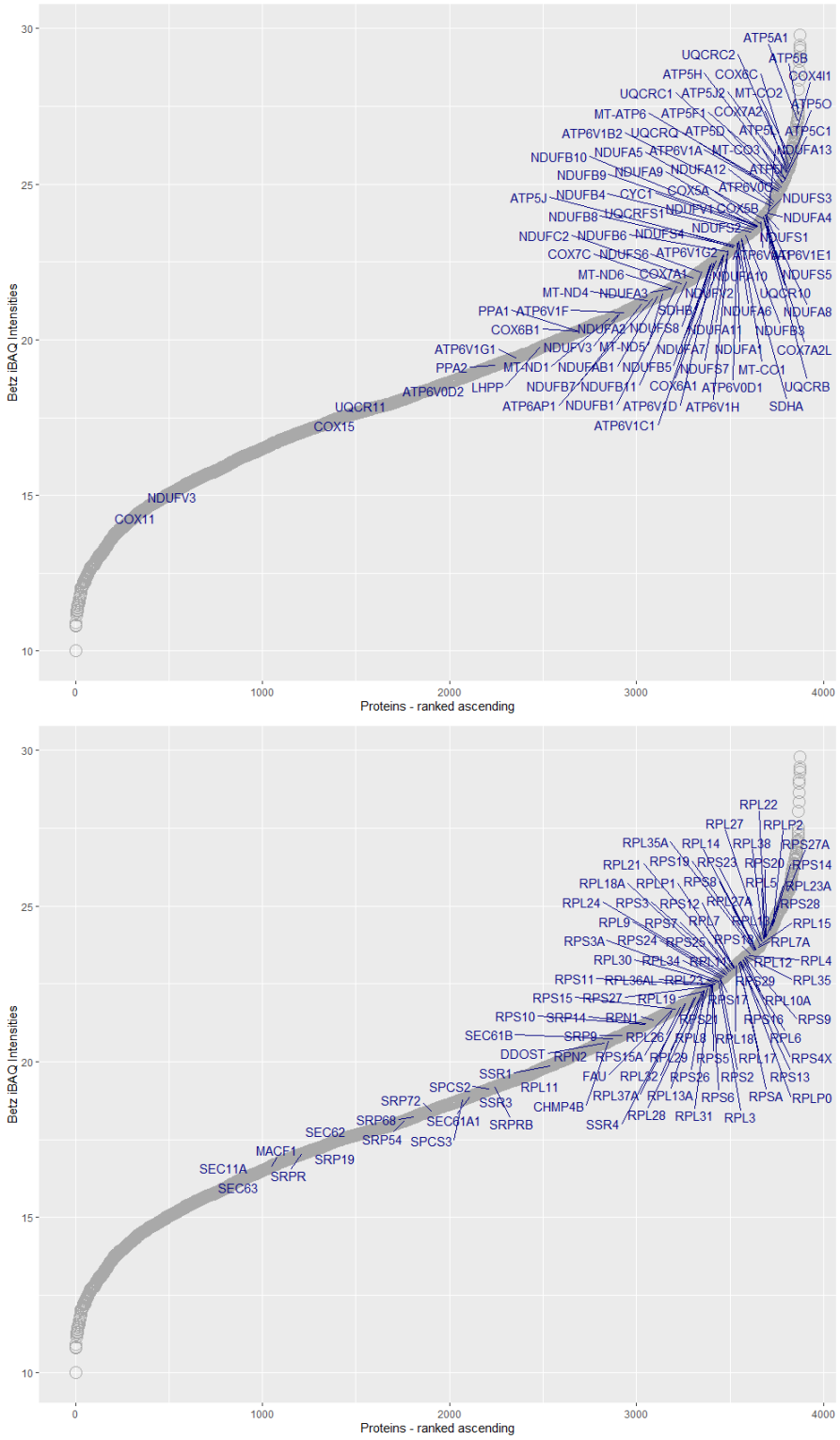


Figure D.5: Oxidative Phosphorylation- and Endoplasmic Reticulum-Related Proteins are among the most abundant Proteins in Betz Cells.

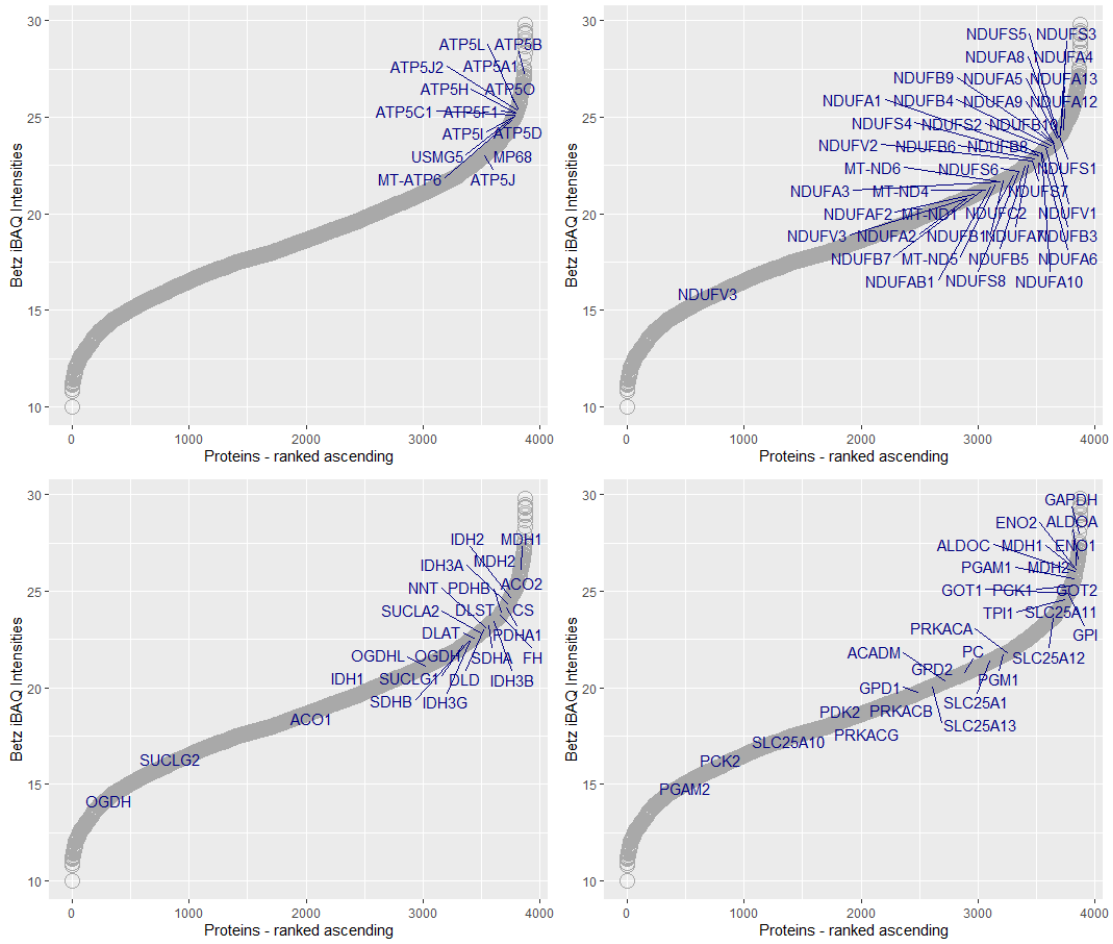


Figure D.6: ATP-, NADH-, TCA-, and Metabolism-Related Proteins are among the most abundant Proteins in Betz Cells.

D.3 Absolute Quantification - Betz Cells, Purkinje Cells, DNNs, and L5 Motor Neurons

When comparing the distribution of protein abundances (iBAQ) within cell-types - and the associated GO terms - it was found that Betz cell protein abundances were highly correlated with that of the DNN (PCC: 0.937), moderately correlated with the Purkinje cell (PCC: 0.819), and had a lowered correlation for the L5 neurons (PCC: 0.665). Although Betz cell and Dentate Nucleus had similar proteomes patterns, proteins relating to apical/basal polarity tended to be expressed at lower levels in Betz cells while they were moderately expressed in the Dentate Nucleus. When compared to Purkinje cells, the results echoed the previous comparison, in the proteins relating to mitochondrion organisation and function were moderately expressed in Betz cells but tended to be expressed at a lower level in Purkinje cells. Unsurprisingly, when compared to Betz cells, Purkinje cells had a significant increase in proteins relating to GABA vesicle transport and an increased expression in inositol 1,4,5-triphosphate-sensitive calcium release - a feature that was also seen in the previous Betz - Purkinje comparison. Although the sample pattern of mitochondrial protein abundance was seen in the Betz cell when compared to the L5 samples it was clear that there were several differences in global protein abundances.

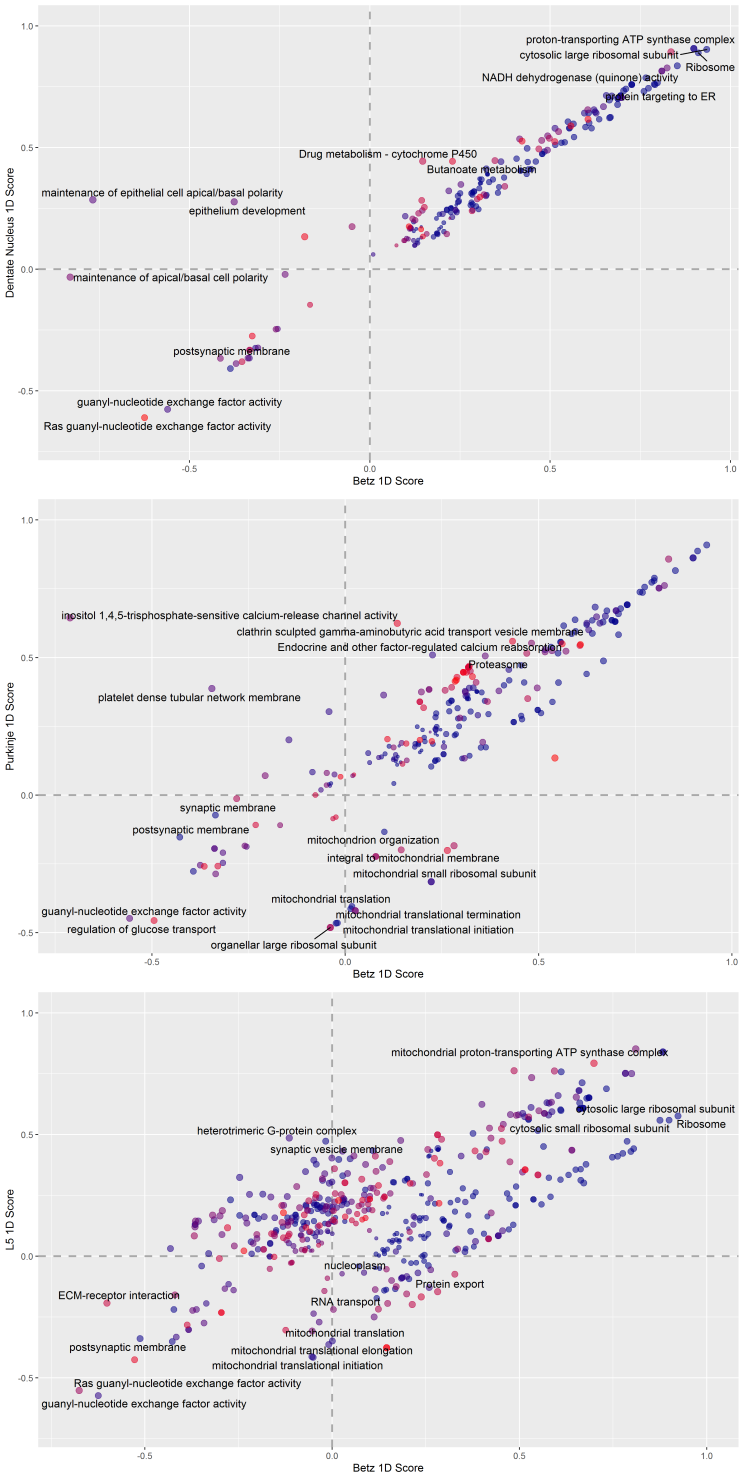


Figure D.7: 2D Annotation: Betz Cells vs. L5 Neurons, DNNs, and Purkinje Cells. Comparing the distribution of proteins within samples. Betz cells contain a similar distribution in proteins to DNNs (top panel), and Purkinje cells (middle panel), but large differences were seen when comparing the distribution of proteins between Betz cells and L5 neuron samples (bottom panel). Size: number of proteins involved in that term. Colour: Darker colours (Blue) have smaller p values.

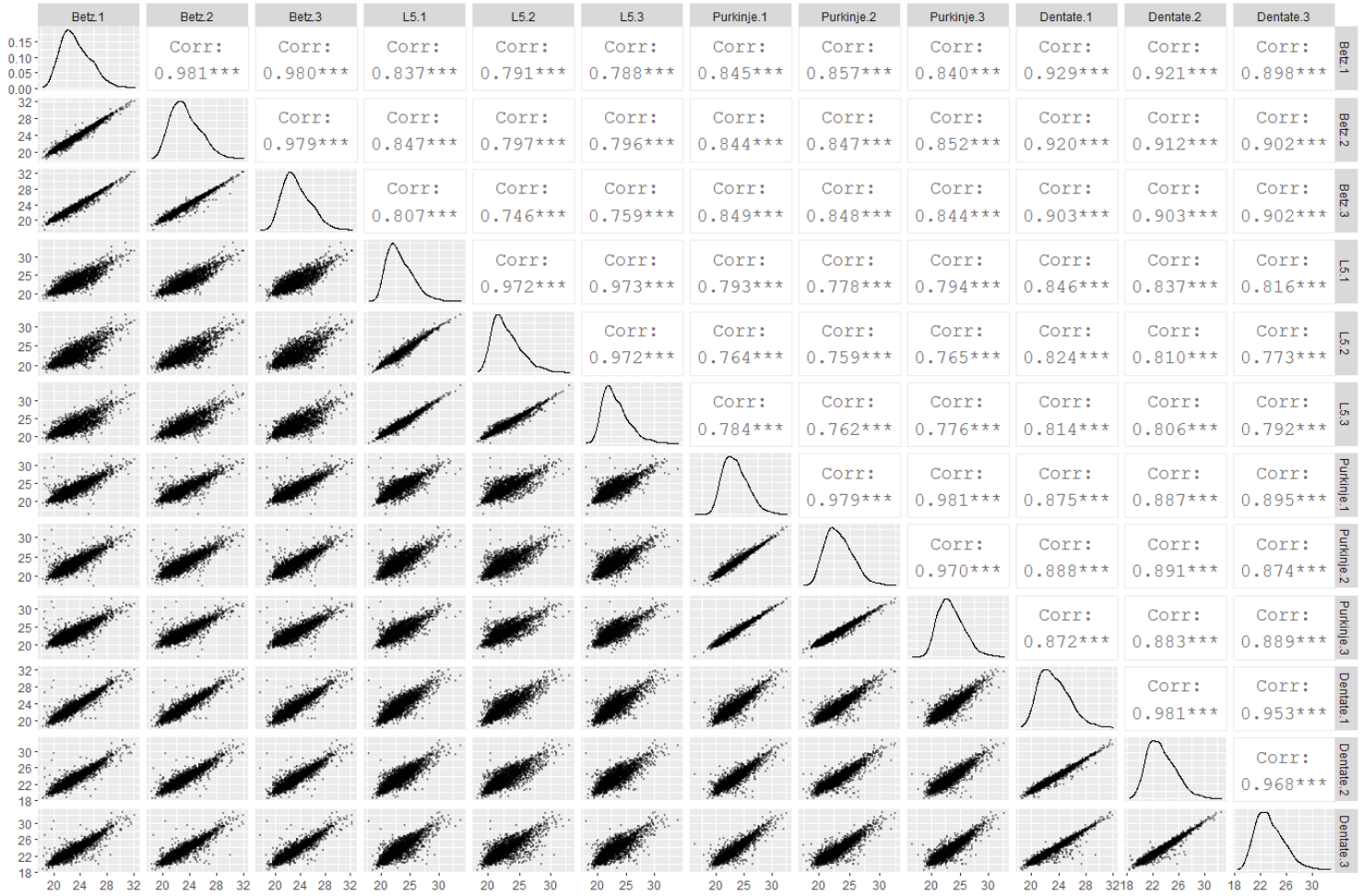


Figure D.8: Multi-scatter Graph - LCM Proteomics of 4 Cell-Types.
Overall, a strong correlation in protein abundance was seen within replicates and across samples. L5 samples was slightly less correlated to the remaining samples.

D.4 Relative Proteomics - Betz Cells, L5 Neurons, DNNs, and Purkinje Cells.

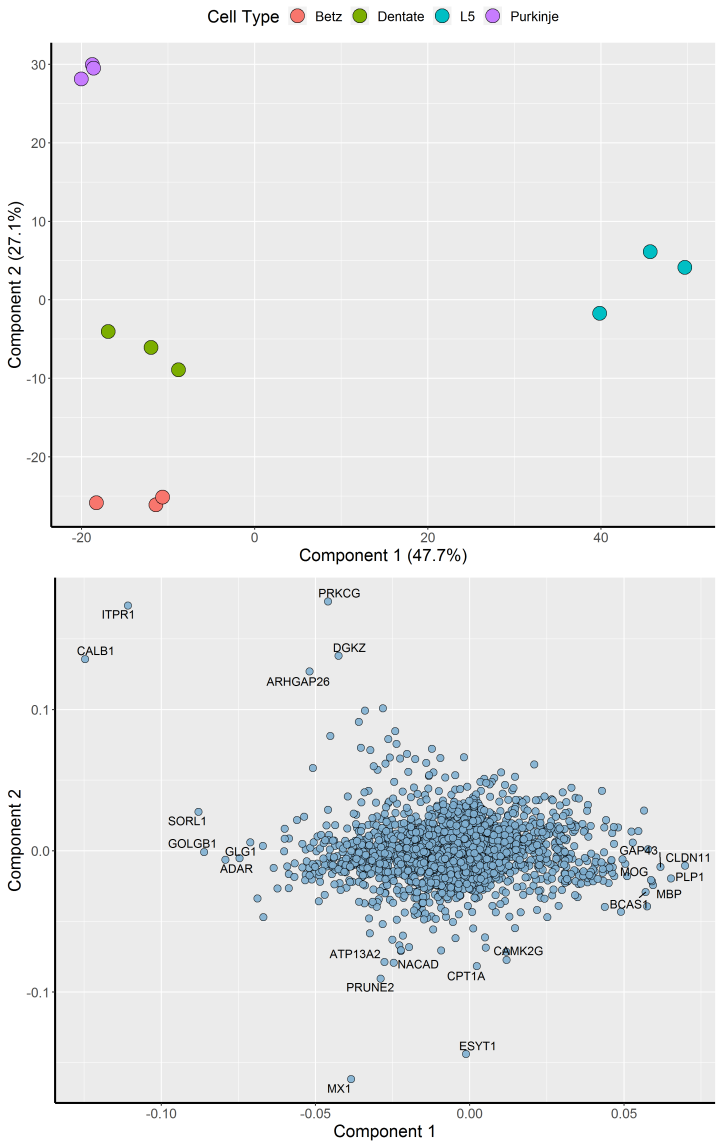


Figure D.9: PCA - Betz Cells, L5 Neurons, DNNs, and Purkinje Cells. Top: PCA plot showing separation of 4 cell-types. Unbiased PCA shows that all 4 cell types clustered among their individual groups with L5 samples begin considerably different from the remaining samples. Bottom: The proteins responsible for driving this variation. The proteins most responsible for the difference between Purkinje cells and the remaining samples were CALB1, ITPR1, PRKCG, and ARHGAP26; all known Purkinje cell markers. A large number of neuropil-rich proteins were found to be driving the separation between L5 neurons and the rest of the samples.

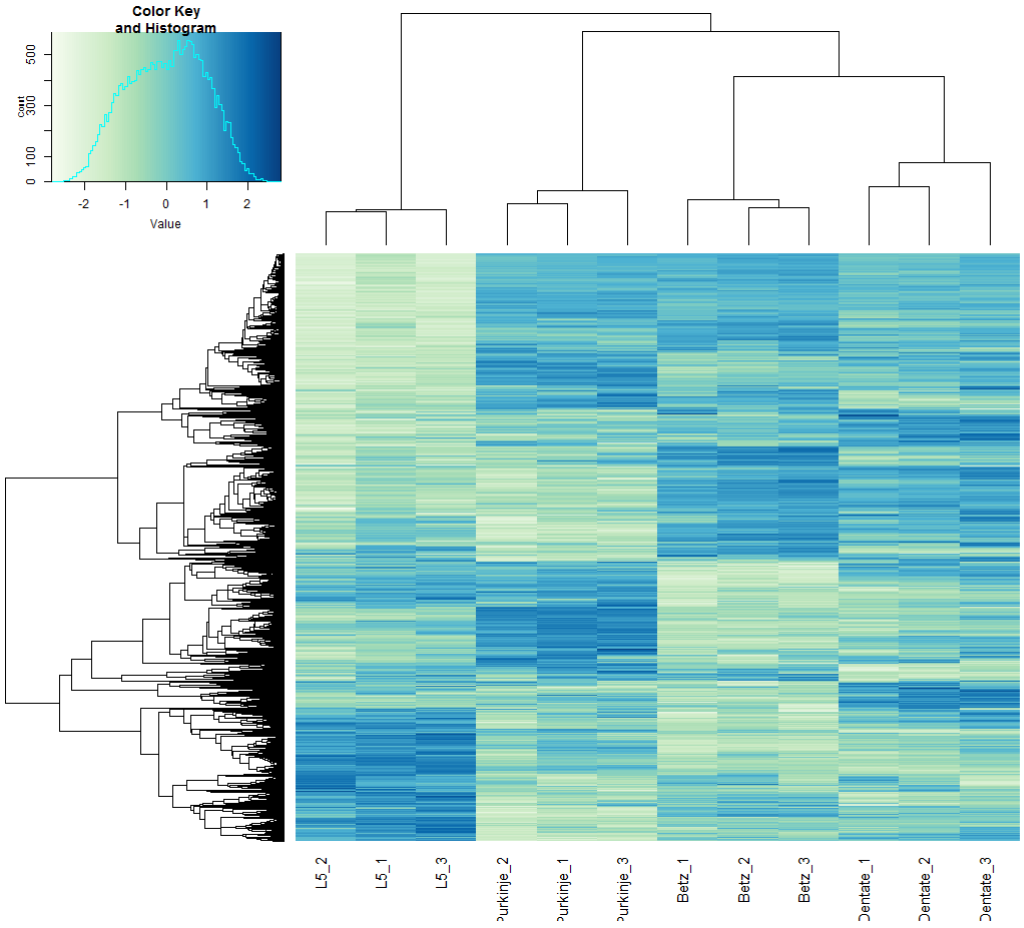


Figure D.10: Hierarchical Clustering - Betz, L5, Purkinje and DNNs. Clustering confirmed that DNNs and Betz cells contained very similar proteome, followed by Purkinje cells, and that L5 neurons were the most dissimilar from the other microdissected neurons. Using this information it was possible to identify ten unique clusters which represented cell-type specific protein enrichment.

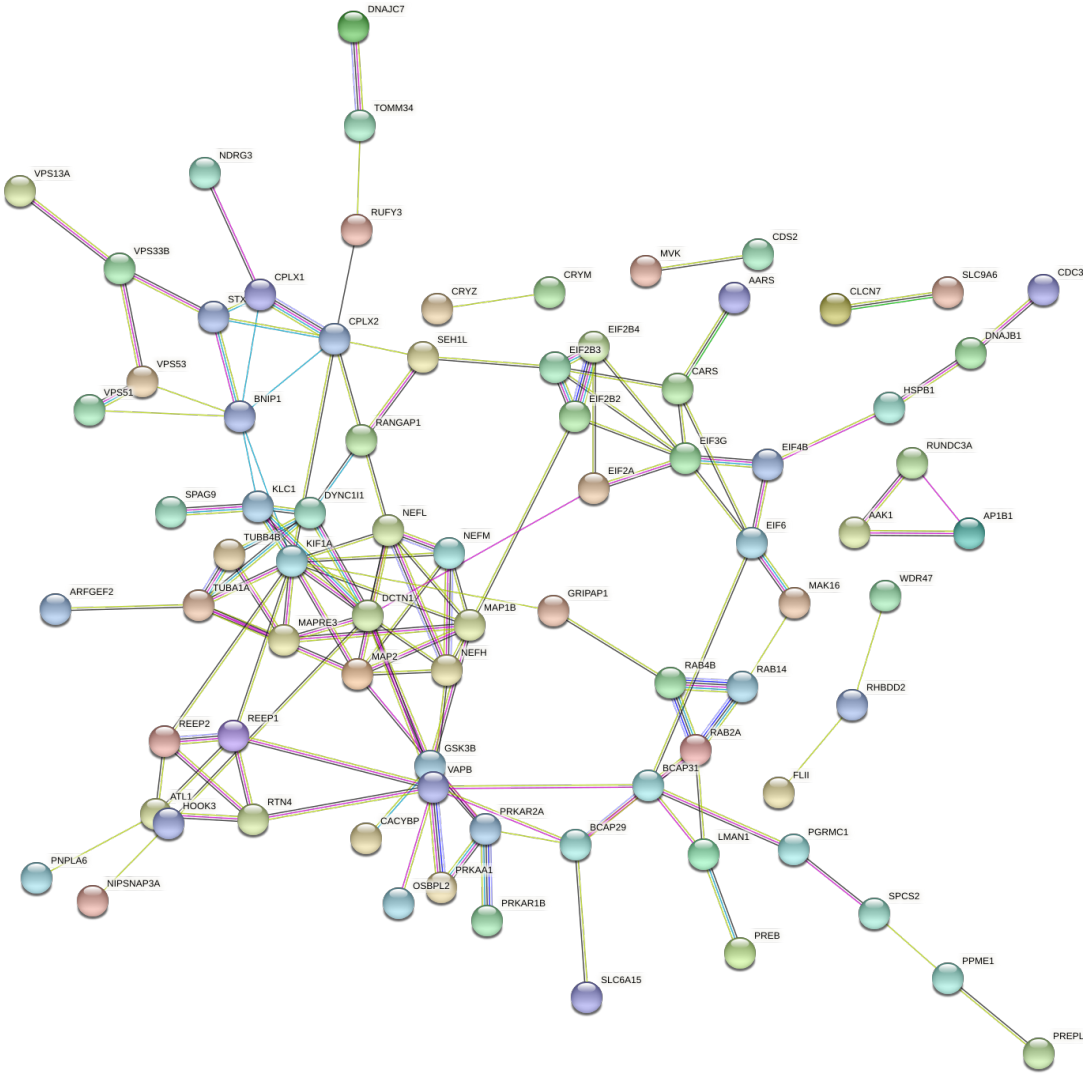


Figure D.11: Protein-Protein Interaction Network of 116 Betz DEPs from Cluster 4. This interaction network highlights that the 116 Betz DEPs from Cluster 4 are unlikely to be a result of chance/bioinformatic error, and that their exists a functional relationship between the proteins identified in this cluster.

Rank No.	Protein Name	Gene Name	z-score
1	Abhydrolase domain-containing protein 16A	ABHD16A	1.86752
2	Vesicle transport protein SEC20	BNIP1	1.74194
3	Ras-related protein Rab-2A	RAB2A	1.72557
4	H(+)/Cl(-) exchange transporter 7	CLCN7	1.70455
5	Brefeldin A-inhibited guanine nucleotide-exchange protein 2	ARFGEF2	1.69185
6	Cytochrome c oxidase subunit 7A1, mitochondrial	COX7A1	1.66976
7	Sodium/hydrogen exchanger 6	SLC9A6	1.64765
8	Receptor expression-enhancing protein 1	REEP1	1.64017
9	GRIP1-associated protein 1	GRIPAP1	1.61856
10	Sodium-dependent neutral amino acid transporter B(0)AT2	SLC6A15	1.61461
11	ADP-ribose pyrophosphatase, mitochondrial	NUDT9	1.60937
12	Protein prune homolog 2	PRUNE2	1.59101
13	Ran GTPase-activating protein 1	RANGAP1	1.58249
14	Translation initiation factor eIF-2B subunit gamma	EIF2B3	1.57588
15	Protein phosphatase methylesterase 1	PPME1	1.5752
16	Hsp90 co-chaperone Cdc37	CDC37	1.56747
17	Protein disulfide-isomerase TMX3	TMX3	1.5567
18	B-cell receptor-associated protein 31	BCAP31	1.55337
19	cAMP-dependent protein kinase type I-beta regulatory subunit	PRKAR1B	1.5506
20	Eukaryotic translation initiation factor 2A	EIF2A	1.54468
21	Cytoplasmic dynein 1 intermediate chain 1	DYNC1I1	1.54439
22	Protein MAK16 homolog	MAK16	1.5389
23	Translation initiation factor eIF-2B subunit beta	EIF2B2	1.5379
24	RUN domain-containing protein 3A	RUNDC3A	1.53607
25	Calcyclin-binding protein	CACYBP	1.52894
26	Kinesin-like protein KIF1A	KIF1A	1.52321
27	Cysteine–tRNA ligase, cytoplasmic	CARS	1.49435

Table D.1 continued from previous page

Rank No.	Protein Name	Gene Name	z-score
28	Complexin-2	CPLX2	1.49351
29	Vesicle-associated membrane protein-associated protein B/C	VAPB	1.48367
30	FSD1-like protein	FSD1L	1.4803
31	Phosphatidylserine synthase 1	PTDSS1	1.47441
32	Protein O-linked-mannose beta-1,4-N-acetylglucosaminyltransferase 2	POMGNT2	1.46862
33	Translation initiation factor eIF-2B subunit delta	EIF2B4	1.46593
34	Microtubule-associated protein 1B	MAP1B	1.43999
35	5-AMP-activated protein kinase catalytic subunit alpha-1	PRKAA1	1.43719
36	Vacuolar protein sorting-associated protein 13A	VPS13A	1.43153
37	Ras-related protein Rab-4B	RAB4B	1.42456
38	Eukaryotic translation initiation factor 3 subunit G	EIF3G	1.423
39	Tubulin alpha-1A chain	TUBA1A	1.41422
40	Protein LSM12 homolog	LSM12	1.41271
41	Tubulin beta-4B chain	TUBB4B	1.41087
42	28S ribosomal protein S35, mitochondrial	MRPS35	1.39564
43	Neuropathy target esterase	PNPLA6	1.38654
44	Annexin A6	ANXA6	1.36595
45	Transmembrane prolyl 4-hydroxylase	P4HTM	1.36358
46	Vacuolar protein sorting-associated protein 51 homolog	VPS51	1.36232
47	DnaJ homolog subfamily C member 7	DNAJC7	1.35716
48	Alanine–tRNA ligase, cytoplasmic	AARS	1.34211
49	Solute carrier family 41 member 3	SLC41A3	1.34044
50	Protein Hook homolog 3	HOOK3	1.33376
51	tRNA (guanine(26)-N(2))-dimethyltransferase	TRMT1	1.31839
52	Mitochondrial import receptor subunit TOM34	TOMM34	1.31777
53	Protein angel homolog 1	ANGEL1	1.31345

Table D.1 continued from previous page

Rank No.	Protein Name	Gene Name	z-score
54	Glycogen synthase kinase-3 beta	GSK3B	1.30895
55	Prolactin regulatory element-binding protein	PREB	1.29601
56	PITH domain-containing protein 1	PITHD1	1.29216
57	GPI transamidase component PIG-S	PIGS	1.29067
58	Microtubule-associated protein RP/EB family member 3	MAPRE3	1.28926
59	Eukaryotic translation initiation factor 6	EIF6	1.28684
60	Vacuolar protein sorting-associated protein 53 homolog	VPS53	1.28373
61	DBIRD complex subunit ZNF326	ZNF326	1.26808
62	Atlastin-1	ATL1	1.26658
63	B-cell receptor-associated protein 29	BCAP29	1.24434
64	Kinesin light chain 1	KLC1	1.2371
65	ATP-dependent RNA helicase DHX36	DHX36	1.2322
66	Phosphatidate cytidylyltransferase 2	CDS2	1.22927
67	Reticulon-4	RTN4	1.2249
68	Ras-related protein Rab-14	RAB14	1.18552
69	Microtubule-associated protein 2	MAP2	1.17686
70	Protein RUFY3	RUFY3	1.16479
71	Tumor protein p63-regulated gene 1-like protein	TPRG1L	1.15574
72	Retinol dehydrogenase 14	RDH14	1.14424
73	Heat shock protein beta-1	HSPB1	1.1345
74	Centrosomal protein of 170 kDa	CEP170	1.12417
75	Protein NipSnap homolog 3A	NIPSNAP3A	1.12285
76	Nuclear receptor coactivator 7	NCOA7	1.11299
77	Neurofilament heavy polypeptide	NEFH	1.10409
78	Protein flightless-1 homolog	FLII	1.08993
79	Mitochondrial chaperone BCS1	BCS1L	1.08926

Table D.1 continued from previous page

Rank No.	Protein Name	Gene Name	z-score
80	Receptor expression-enhancing protein 2	REEP2	1.0851
81	Rhomboid domain-containing protein 2	RHBDD2	1.08061
82	cAMP-dependent protein kinase type II-alpha regulatory subunit	PRKAR2A	1.07266
83	Neurofilament light polypeptide	NEFL	1.0706
84	Twinfilin-2	TWF2	1.06079
85	Neurofilament medium polypeptide	NEFM	1.05678
86	Syntaxin-7	STX7	1.03397
87	MAP7 domain-containing protein 1	MAP7D1	1.03251
88	Protein NDRG3	NDRG3	1.01575
89	Saccharopine dehydrogenase-like oxidoreductase	SCCPDH	1.01497
90	Oxysterol-binding protein-related protein 2	OSBPL2	1.00638
91	Rab-like protein 3	RABL3	0.997639
92	Dynactin subunit 1	DCTN1	0.995767
93	DnaJ homolog subfamily B member 1	DNAJB1	0.98015
94	Thioredoxin reductase 2, mitochondrial	TXNRD2	0.970669
95	Protein ERGIC-53	LMAN1	0.966072
96	Bcl-2-like protein 13	BCL2L13	0.958375
97	Prolyl endopeptidase-like	PREPL	0.94047
98	Ankyrin repeat domain-containing protein 17	ANKRD17	0.93892
99	Signal peptidase complex subunit 2	SPCS2	0.930606
100	Complexin-1	CPLX1	0.911065
101	C-Jun-amino-terminal kinase-interacting protein 4	SPAG9	0.890995
102	Quinone oxidoreductase	CRYZ	0.877003
103	Glycerophosphodiester phosphodiesterase 1	GDE1	0.848841
104	AP-1 complex subunit beta-1	AP1B1	0.80668
105	AP2-associated protein kinase 1	AAK1	0.797031

Table D.1 continued from previous page

Rank No.	Protein Name	Gene Name	z-score
106	Ketimine reductase mu-crystallin	CRYM	0.740455
107	WD repeat-containing protein 7	WDR7	0.727945
108	Membrane-associated progesterone receptor component 1	PGRMC1	0.719216
109	WD repeat-containing protein 47	WDR47	0.679432
110	Nucleoporin SEH1	SEH1L	0.644333
111	Mevalonate kinase	MVK	0.6349
112	Eukaryotic translation initiation factor 4B	EIF4B	0.627582
113	Centrosomal protein of 170 kDa protein B	CEP170B	0.596301
114	Vacuolar protein sorting-associated protein 33B	VPS33B	0.591846
115	Neuronal membrane glycoprotein M6-b	GPM6B	0.583325
116	MAP7 domain-containing protein 2	MAP7D2	0.523414

Table D.1: Ranked List of 116 DEPs found in the Betz Cell Cluster.

D.5 Betz Cell Comparisons - 1D Annotation Enrichment

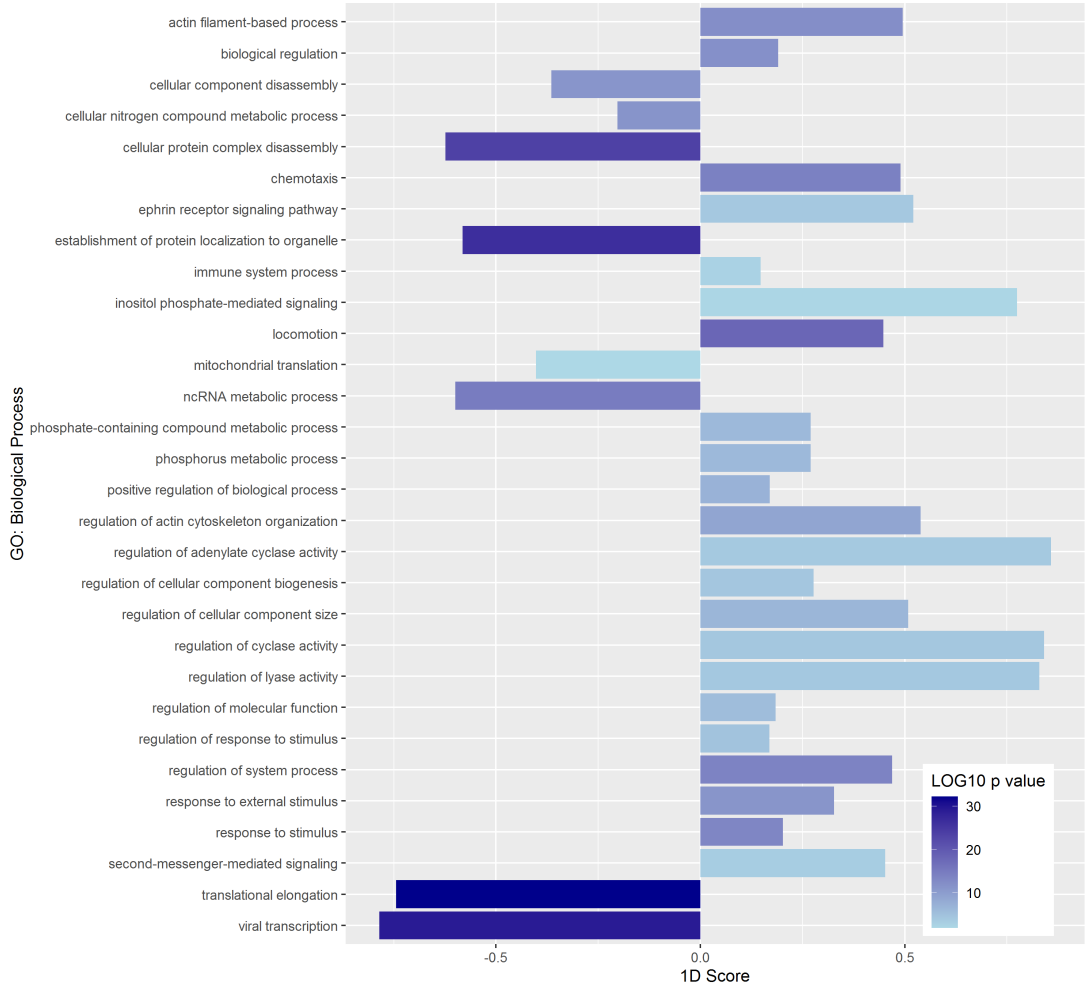


Figure D.12: 1D Annotation Enrichment - Betz Cells vs. Purkinje Cells (BP). Betz cells contained a higher enrichment of proteins involved in protein disassembly, protein localisation, protein synthesis, and ncRNA processing. Negative 1D Score: GO term enriched in Betz cells, Positive 1D Score: GO term enriched in Purkinje cells.

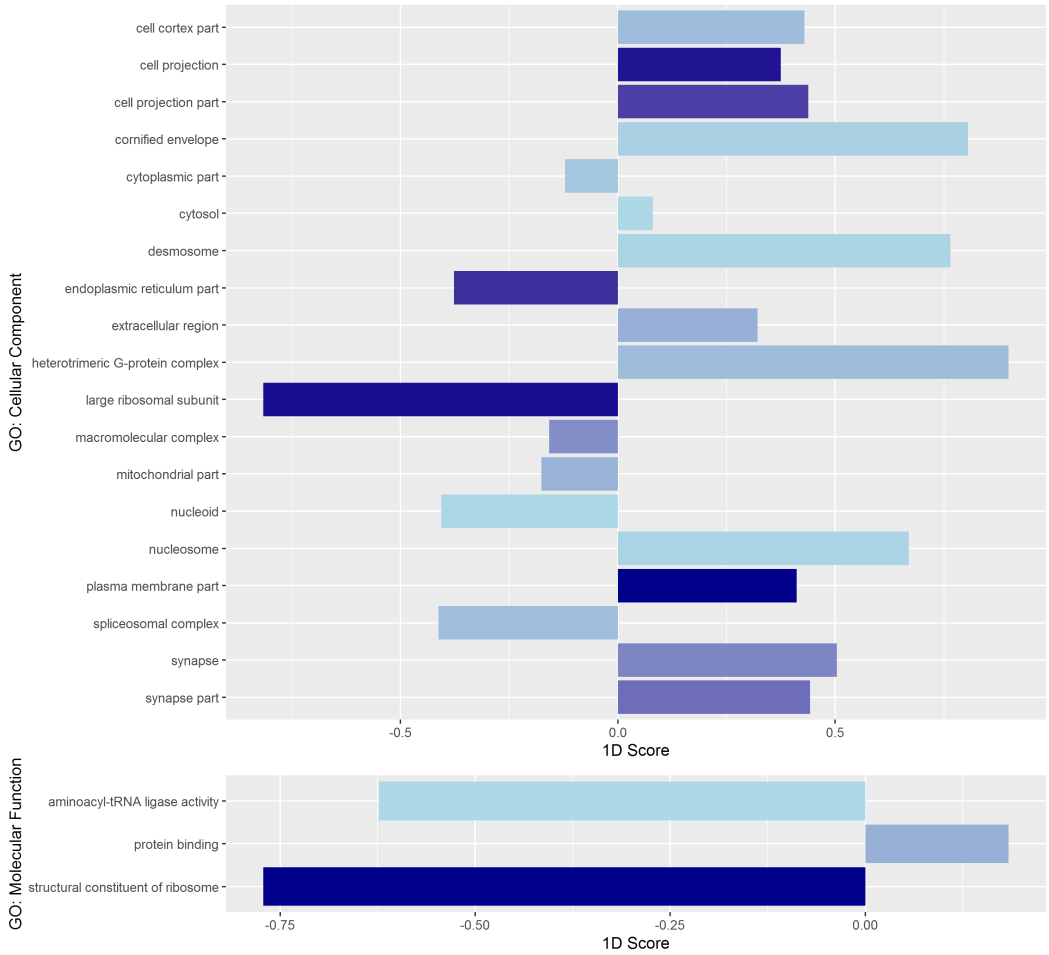


Figure D.13: 1D Annotation Enrichment - Betz Cells vs. Purkinje Cells (MF & CC). Betz cells contained a higher enrichment of proteins involved in protein binding, tRNA ligase activity, and an enrichment of proteins found in the endoplasmic reticulum, mitochondria, and spliceosome. Negative 1D Score: GO term enriched in Betz cells, Positive 1D Score: GO term enriched in Purkinje cells.

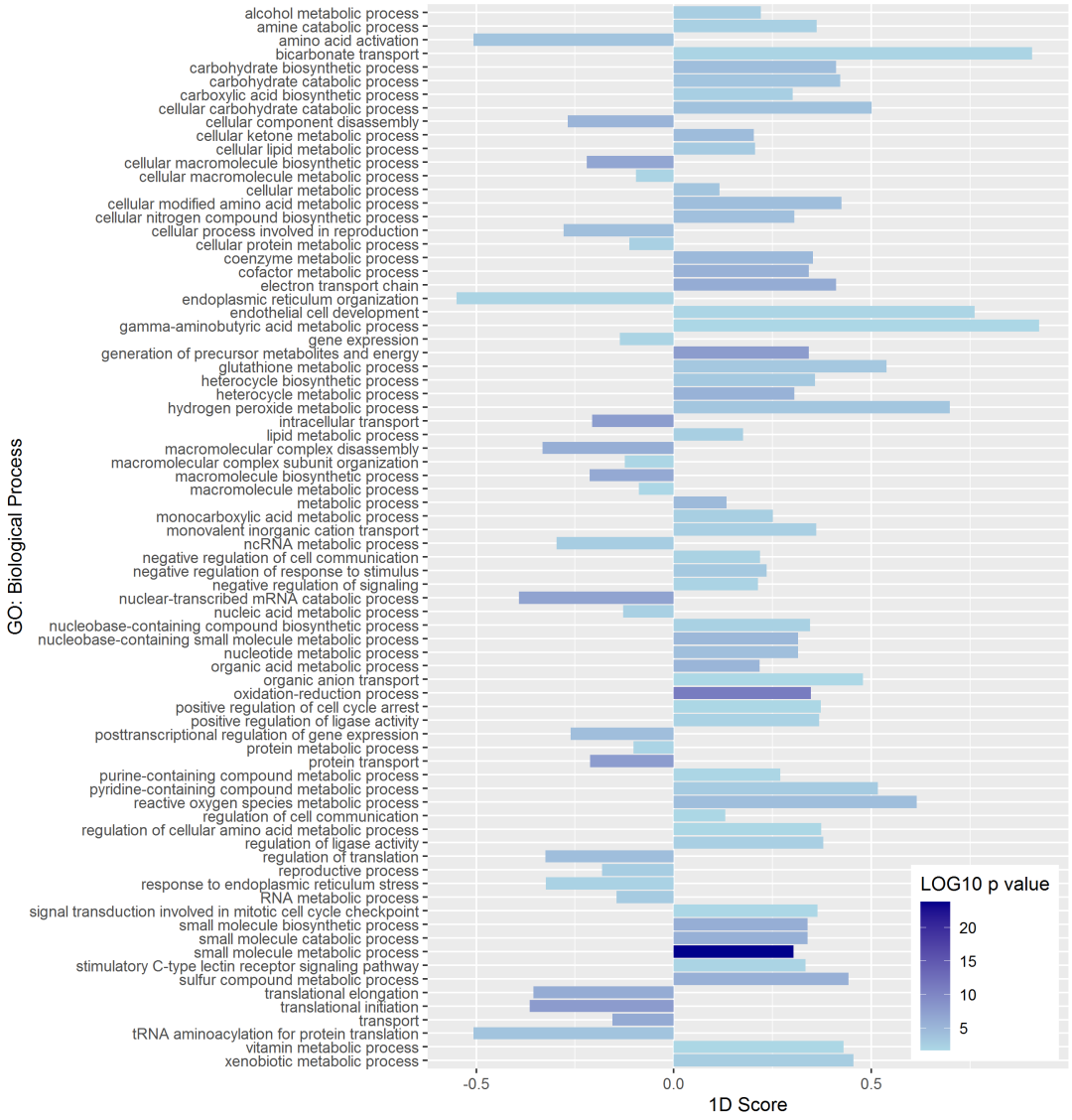


Figure D.14: 1D Annotation Enrichment - Betz Cells vs. DNNs (BP). A large number of BP terms were found to be enriched in either Betz cell or DNNs. Betz cells contained an enrichment in proteins involved in tRNA aminoacylation, intracellular transport, endoplasmic reticulum organisation, protein synthesis, and macromolecule disassembly. Negative 1D Score: GO term enriched in Betz cells, Positive 1D Score: GO term enriched in DNNs.

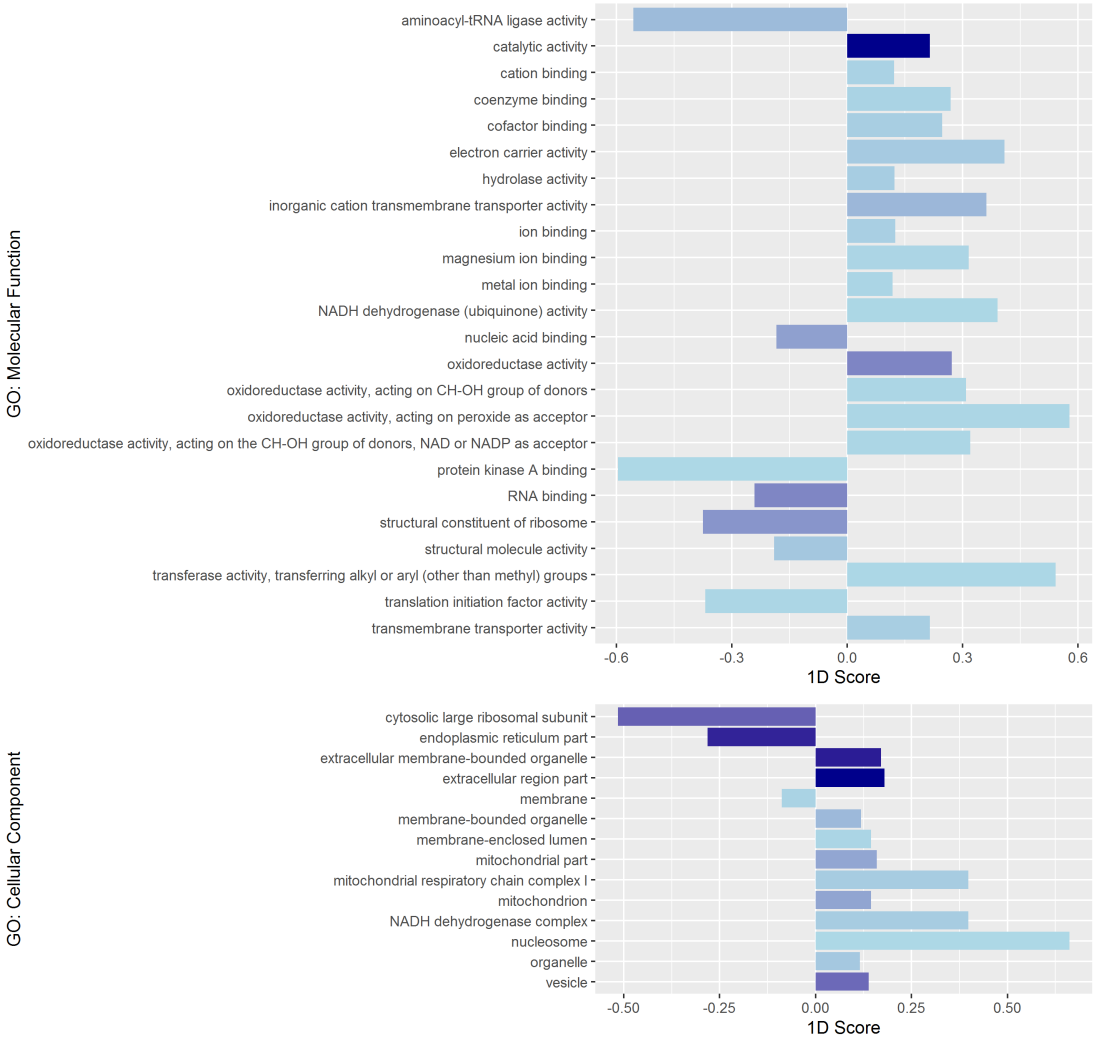


Figure D.15: 1D Annotation Enrichment - Betz Cells vs. DNNs (MF & CC). Betz cells contained an enrichment in proteins involved in tRNA ligase activity, protein kinase A binding, RNA binding, and protein synthesis. However, interestingly, DNNs had an enrichment in mitochondria and electron transport chain proteins over Betz cells. Negative 1D Score: GO term enriched in Betz cells, Positive 1D Score: GO term enriched in DNNs.

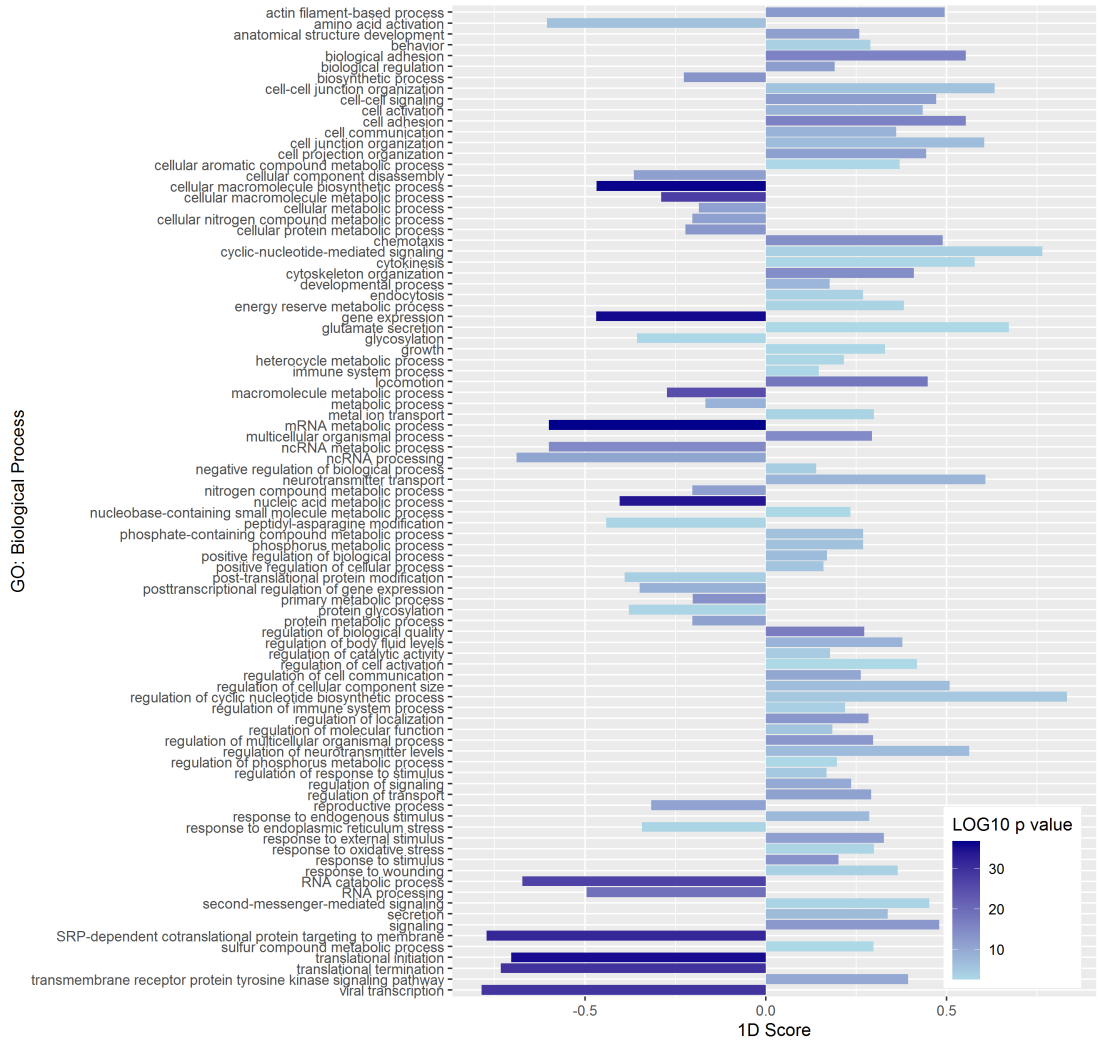


Figure D.16: 1D Annotation Enrichment - Betz Cells vs. L5 Neurons (BP). A large number of BP terms were found to be enriched in either Betz cell or L5 neurons. Betz cells contained an enrichment in proteins involved in response to endoplasmic reticulum stress, protein synthesis, macromolecule biosynthesis, macromolecule disassembly, and mRNA processing. Negative 1D Score: GO term enriched in Betz cells, Positive 1D Score: GO term enriched in L5 neurons.

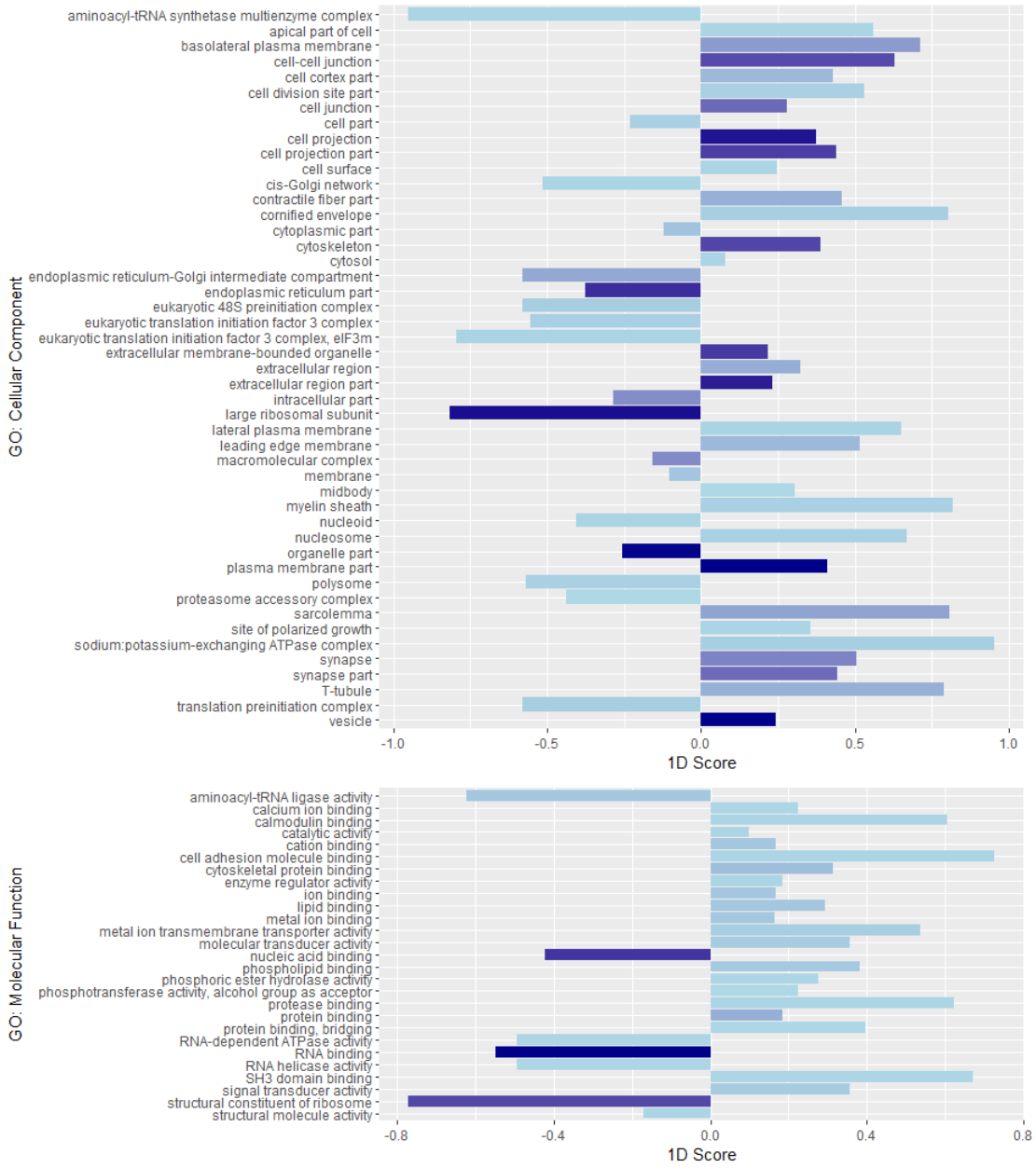


Figure D.17: 1D Annotation Enrichment - Betz Cells vs. L5 Neurons (MF & CC). A large number of MF/CC terms were found to be enriched in either Betz cell or L5 neurons. Betz cells contained an enrichment in proteins found in the proteasome, endoplasmic reticulum-Golgi Apparatus compartment, aminoacyl-tRNA synthase multienzyme complex, and large ribosomal subunit. These proteins were also found to be involved in RNA helicase activity, RNA binding, ATPase activity, and aminoacyl-tRNA activity. Negative 1D Score: GO term enriched in Betz cells, Positive 1D Score: GO term enriched in L5 neurons.

D.6 Additional IHC Data

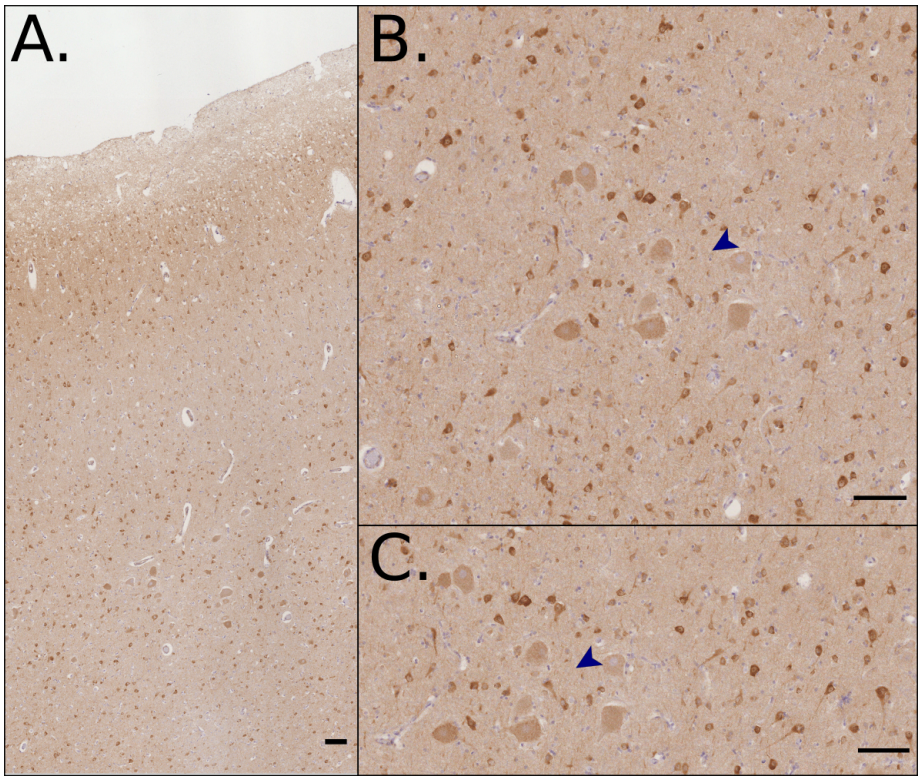


Figure D.18: ARHGAP26 Staining of the Primary Motor Cortex.
A: Anti-ARHGAP26 shows diffuse staining across the primary motor cortex. Betz cells (Blue arrows) were immuno-positive, however, they were difficult to identify due to neuropil staining and other neurons in L5 containing higher expression of ARHGAP26.

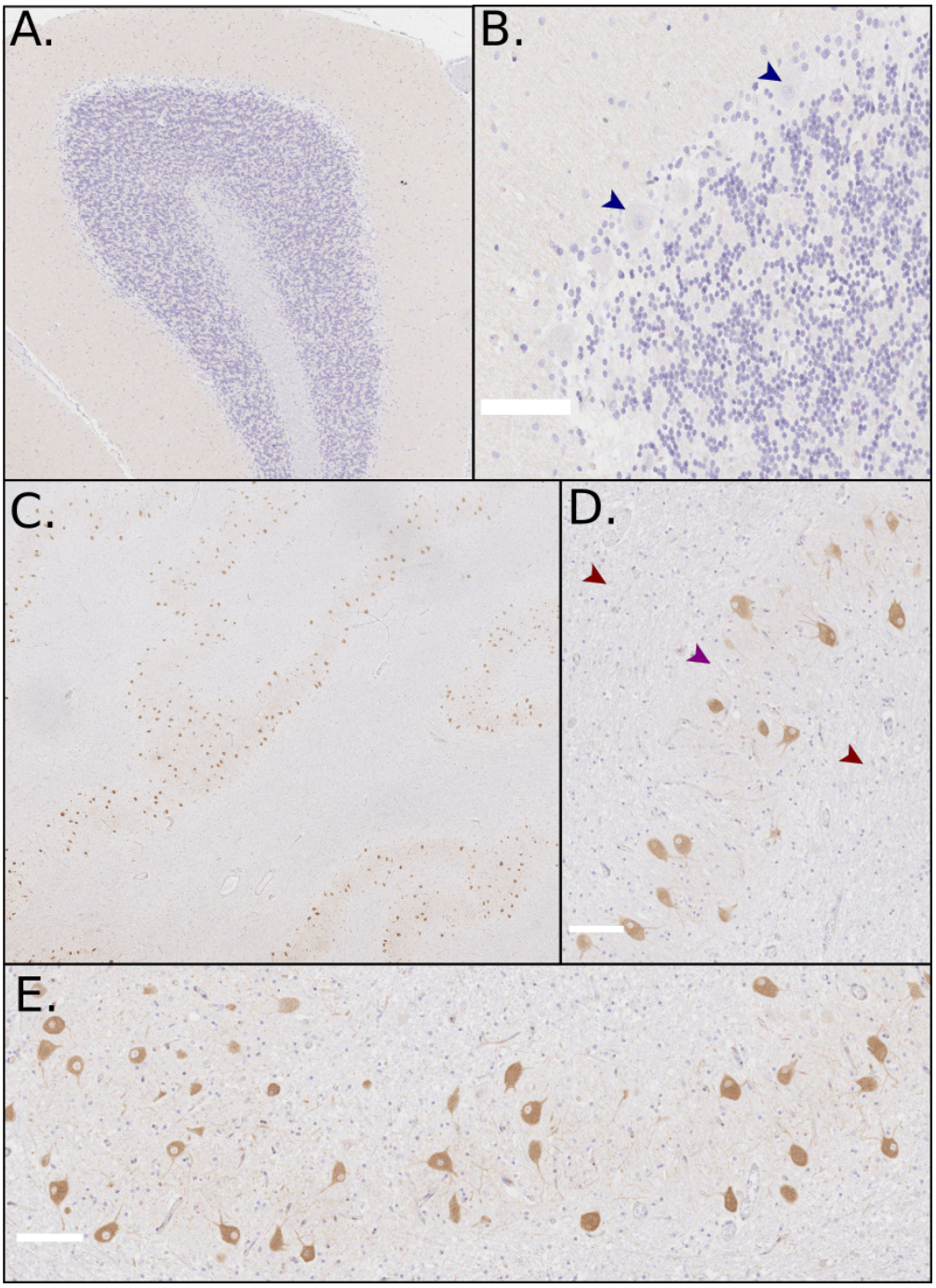


Figure D.19: ESYT1 Staining of the Cerebellum.
A: Limited staining can be seen in the cerebellum. White matter, Granule cell layer, and Purkinje cell layer (B) were negative for ESYT1. Conversely, ESYT1 expression was high among Dentate Nucleus neurons (C:E). Red arrow: white matter around Dentate Nucleus. Purple arrow: Specific staining of Dentate Nucleus neurons. Scalebar: 100µm.

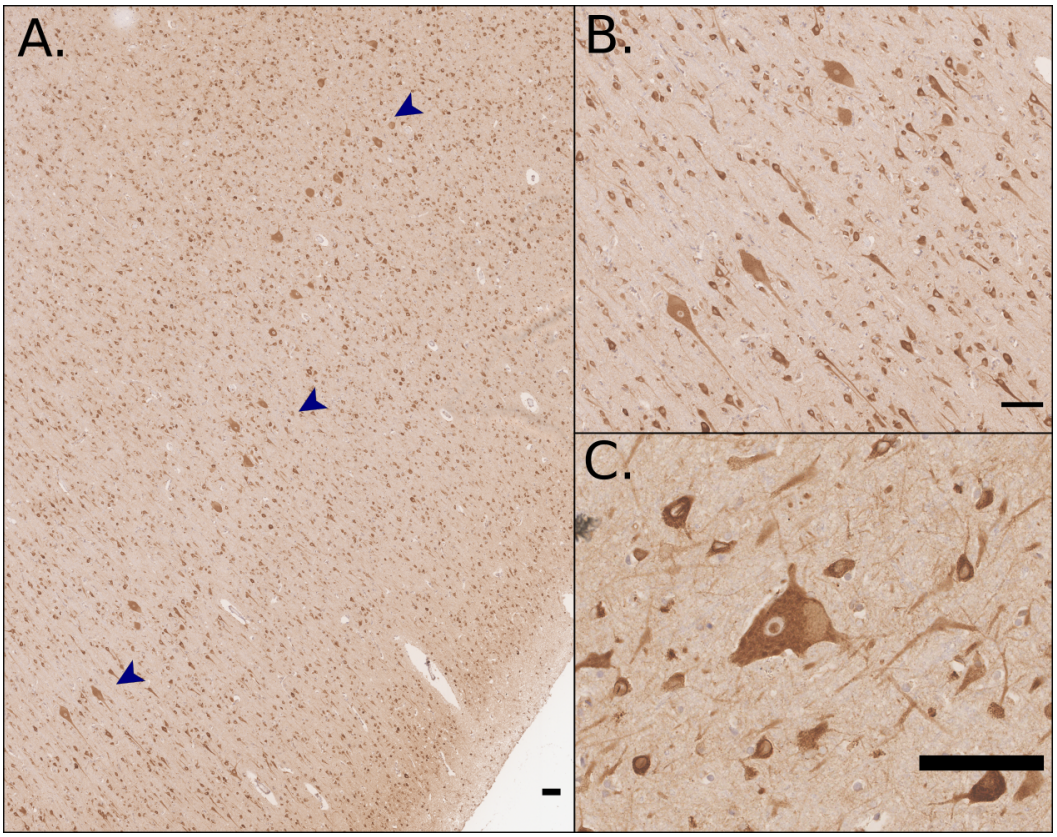
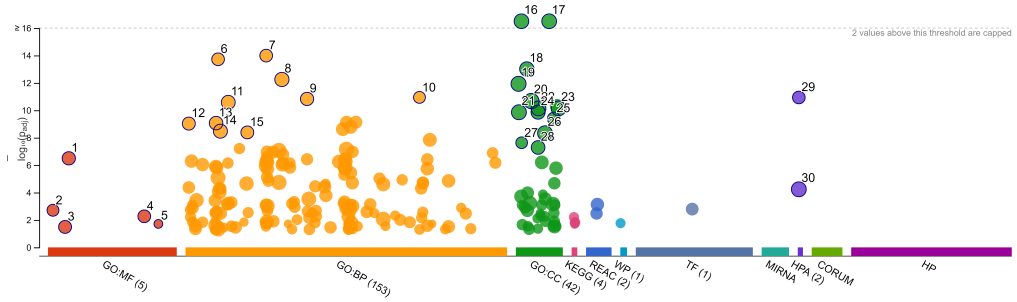


Figure D.20: ER81 Staining of the Primary Motor Cortex.
The large soma of the Betz cell could be visulised at both low magnification (A) and high magnification (B-C). Other neurons across the primary motor cortex were stained, including superfical layers. Scalebar: 100µm.

D.7 Excessive Contamination of Neuropil Proteins in L5 Neuronal Samples



ID	Sour...	Term ...	Term Na...	p_{adj} (query...
1	GO:MF	GO:0008092	cytoskeletal protein binding	3.175×10^{-7}
2	GO:MF	GO:0003779	actin binding	1.956×10^{-3}
3	GO:MF	GO:0005102	signaling receptor binding	3.221×10^{-2}
4	GO:MF	GO:0050839	cell adhesion molecule binding	5.407×10^{-3}
5	GO:MF	GO:0098632	cell-cell adhesion mediator activity	1.938×10^{-2}
6	GO:BP	GO:0007155	cell adhesion	1.893×10^{-14}
7	GO:BP	GO:0022610	biological adhesion	1.021×10^{-14}
8	GO:BP	GO:0032879	regulation of localization	5.610×10^{-13}
9	GO:BP	GO:0040011	locomotion	1.472×10^{-11}
10	GO:BP	GO:0098609	cell-cell adhesion	1.137×10^{-11}
11	GO:BP	GO:0009653	anatomical structure morphogenesis	2.605×10^{-11}
12	GO:BP	GO:0000902	cell morphogenesis	9.312×10^{-10}
13	GO:BP	GO:0006928	movement of cell or subcellular component	8.567×10^{-10}
14	GO:BP	GO:0007399	nervous system development	3.323×10^{-9}
15	GO:BP	GO:0016477	cell migration	4.095×10^{-9}
16	GO:CC	GO:0005886	plasma membrane	1.143×10^{-27}
17	GO:CC	GO:0071944	cell periphery	2.968×10^{-27}
18	GO:CC	GO:0030054	cell junction	9.696×10^{-14}
19	GO:CC	GO:0005576	extracellular region	1.168×10^{-12}
20	GO:CC	GO:0031982	vesicle	2.111×10^{-11}
21	GO:CC	GO:0005615	extracellular space	1.388×10^{-10}
22	GO:CC	GO:0043230	extracellular organelle	7.203×10^{-11}
23	GO:CC	GO:1903561	extracellular vesicle	7.203×10^{-11}
24	GO:CC	GO:0042995	cell projection	1.362×10^{-10}
25	GO:CC	GO:0098590	plasma membrane region	4.078×10^{-10}
26	GO:CC	GO:0070062	extracellular exosome	4.696×10^{-9}
27	GO:CC	GO:0005911	cell-cell junction	2.317×10^{-8}
28	GO:CC	GO:0043005	neuron projection	5.136×10^{-8}
29	HPA	HPA:0100213	cerebral cortex; neuropil[High]	1.156×10^{-11}
30	HPA	HPA:0100212	cerebral cortex; neuropil[>Medium]	5.800×10^{-5}

Figure D.21: Manhattan Plot: Summary of Proteins Up-Regulated in L5 Neurons. Proteins involved in extracellular functions were among the most enriched in L5 neurons when compared to all other cell-types. Importantly, an increase in extracellular/neuropil-related proteins was identified suggesting significant contamination - likely a result of the microdissection approach used for isolating L5 neurons. All points represent significantly enriched pathways and are colour-coded by their representative pathway type.

E

Chapter 6 - Supplementary Data

Contents

E.1 Developing a High-Throughput LCM Proteomic Workflow 321
E.2 LCM-seq Optimisations 327
E.3 LCM-seq Quality Control 333
E.4 LCM-seq Analysis 338
E.4.1 Healthy Betz Cells vs. Healthy Purkinje Cells 338
E.4.2 Healthy Betz Cells vs. Healthy L5 Neurons 344
E.4.3 Healthy Betz Cells vs. ALS Betz Cells 351
E.5 Microdissected Betz Cells / Allen Institute Clusters - Comparison 360

E.1 Developing a High-Throughput LCM Proteomic Workflow

The 60-minute gradient used in our RWS method generates rich data sets from minimal laser microdissected material. While 60 minutes is commonly used for the majority of LC-MS/MS protocols, this can quickly become problematic when attempting to analyse large numbers of samples, as this gradient is applied individually to each sample which often requires samples to be run in batches. These batches and the extended time taken to run all samples in a medium to large-sized

experiment can result in significant bias' as unwanted/uncontrolled data variation accumulates over time due to instrumental data drift [483]. To address this, our LCM proteomic approach was adapted to be used on a mass spectrometer with can be used in conjunction with substantially shorter gradient times (>4 minutes)[1]. While a reduced gradient would typically result in fewer identified peptides, the rationale behind this equipment change is to utilise the 4-dimensional data generated by trapped ion mobility spectrometry (TIMS). The additional measurement of ion mobility should help in resolving peptides with similar features, increasing the number of identifiable proteins. Coupled with the RWS workflow, we hope to generate data comparable to that of the Orbitrap Fusion Lumos which was used in the LCM proteomic experiments seen in Chapters 4 and 5.

To compare performance between the two mass spectrometry methods Betz cells ($\bar{X}$: 146.67) and Purkinje cells ($\bar{X}$: 483.83) were microdissected in triplicate, processed and analysed with both the Orbitrap Fusion Lumos and TimsTOF machines. The overall performance of the TimsTOF was considerably lower, albeit, still competitive, compared to the depth of the LCM proteomic method previously mentioned. However, with a gradient time of only 14 minutes, this dramatically reduces the time taken to analyse one sample. Impressively, this results in identifying 2.38 proteins per 1000µm^3 per one gradient hour. Overall, 49.5% of proteins (1,996) were shared between the two mass spectrometry methods. While the TimsTOF approach yielded shallower proteomic depth, it was able to identify 74 proteins (1.8%) that were not found in the Lumos samples. Reproducibility was high among all samples; Betz and Purkinje replicate on the Lumos had a mean PPC of 0.955 and 0.976, respectively; with the Betz and Purkinje replicates on TimsTOF having a mean PPC of 0.954 and 0.849, respectively. No GO enrichments were found among proteins identified by only one approach.

[1]Simon Davis devised the scientific rationale behind attempting to improve the sensitivity, and therefore performance of the RWS approach. Simon designed and optimised the workflow for high-throughput analysis. During the development of this approach, Simon's optimised parameters were tested on laser microdissected cerebellar granule cells to determine how the high-throughput workflow would perform on the Betz/Purkinje cell cohort.

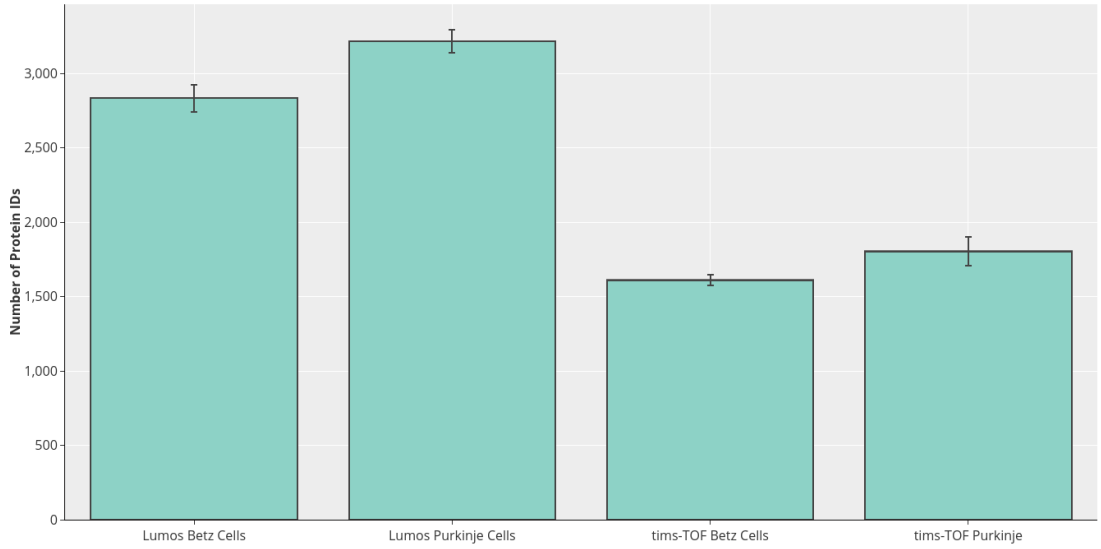


Figure E.1: Lumos / TimsTOF comparison.
Turquoise bars denote the number of proteins identified per sample. Vertical bars: standard deviation. Lumos: samples analysed using the Orbitrap Fusion Lumos. timsTOF: sampels analysed using the timsTOF Pro. Amount microdissected: 200,000µm^2.

Promisingly, performing differential expression analysis between Betz and Purkinje cells on both Lumos and TimsTOF samples resulting in similar observations as seen in chapter 5 - (data not shown). With ESYT1, MX1, DST, OXR1, and NEFL being among the most differential expressed in Betz cells and known Purkinje cell markers (ITPR1, CALB1, CA8, ARHGAP26, and PRKCG) being the most differentially expressed in Purkinje cells. As no obvious biological bias was seen between the two mass spectrometry methods, several attempts were made to adjust settings/parameters on the TimsTOF in an attempt to optimise performance and achieve similar proteomic depths to the Lumos. First, several gradient times were tested to see how that may affect proteomic depth. Furthermore, several sensitivity thresholds were also tested to determine the optimum settings - see Table E.1. As this was primarily an optimisation experiment, individual cells were not microdissected; instead, whole areas of the cerebellar Granule layer was microdissected (200,000µm^2).

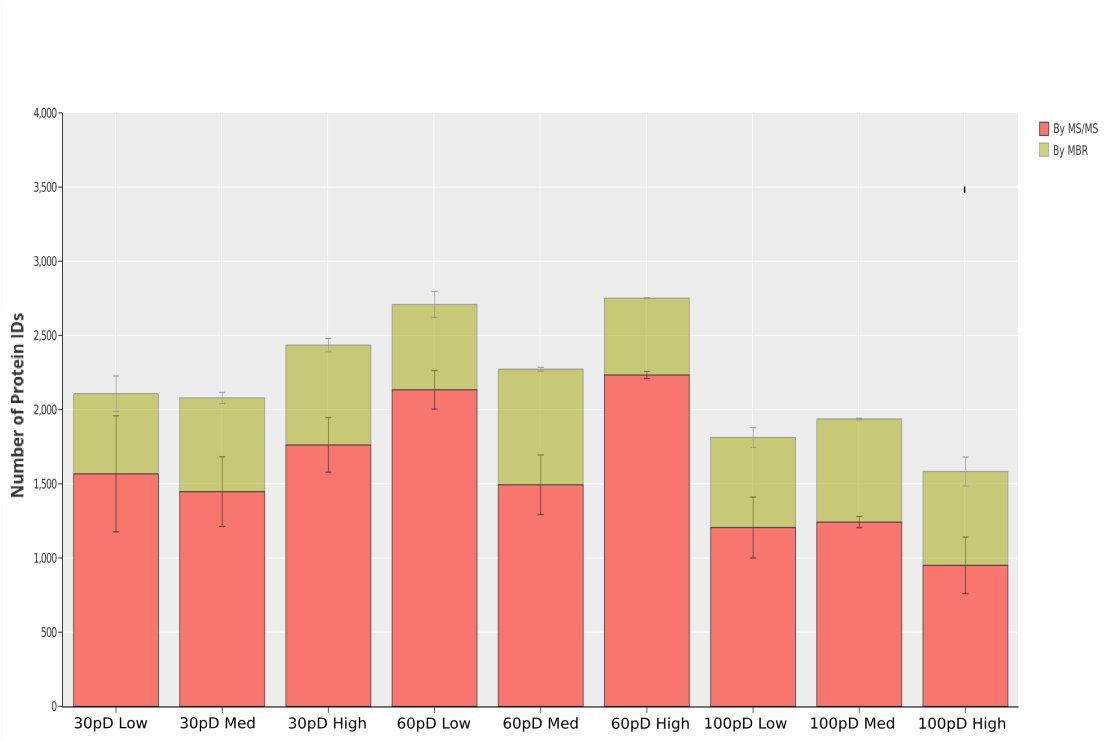


Figure E.2: TimsTOF Optimisation.
30pd = 48 minute gradient, 60pD = 24 minutes, 100pD = 14.4 minutes pD = maximum number of samples analysed per day i.e. 30pD = 30 samples per day. Red bars denote the number of proteins identified solely by MS/MS, with yellow-green bars showing the number of proteins identified by MaxQuant's Matched-Between-Run (MBRs) algorithm. Low/High/Med reflect intensity targets and threshold - see Table E.1. Vertical bars: standard deviation.

Name	Samples per day	Target Intensity	Intensity Threshold
30pD Default	30	20000	2500
30pD Low	30	16000	1600
30pD High	30	24000	2400
60pD medium	60	15000	1500
60pD high	60	17500	1500
60pD low	60	12500	1500
100pD medium	100	15000	1500
100pD high	100	17500	1500
100pD low	100	12500	1500

Table E.1: Parameters used for optimising the high-throughput approach. Target Intensity: ions are repeatedly targeted for MS/MS until their sum intensity is >= this value. Intensity threshold: The threshold that an ion needs to meet to be targeted for MS/MS in any one MS1 scan.

With the microdissected Granule cell areas, it was possible to achieve a maximum proteomic depth of 2,781 proteins with the 60pD high setting, followed by $\bar{X}$: 2,751 with the 60pD low settings, and finally $\bar{X}$: 2,708 proteins with the 60pD low settings. There was a pronounced decrease in proteomic depth in the 100pD samples ($\bar{X}$: 1777.33), likely indicating that a 14.4-minute gradient is too short to resolve a high number of peptides. Next, the above settings were tested on individual pooled cells rather than whole areas to better replicate the material to be used for future planned experiments. For the succeeding experiment, only 100 Purkinje cells (≈60,000µm^2) were microdissected. The rationale behind this is two-fold. First, the additional ion mobility data acquired by the TimsTOF equipment may result in diminishing returns when injecting more input material into the mass spectrometer. Second, the ability to resolve increasing amounts of proteomic data from smaller amounts of tissue is fundamental for clinical/medical studies. Developing a workflow for high-throughput micro-proteomics on very small amounts of material is critical to expanding LCM proteomics to wider audiences. In addition to using microdissected Purkinje cells, other configurations of the TimsTOF instrument was performed to see if the performance could be boosted. For example, the liquid chromatography component of the TimsTOF machine (named EvoSep herein) was replaced with the one used for the Lumos (named Dionex Nanoflow herein). The Dionex Nanoflow uses smaller particles, has a larger column, and is operated at a lower flow rate (400 nL/min vs. 1 µL/min) - all of which aims to increase chromatographic resolution and the concentration of the analytes, increasing MS sensitivity., and potentially identifying more proteins.

Combining the Dionex Nanoflow with the TimsTOF and optimising the sensitivity settings from the machine resulted in ≈ 2,000 proteins being identified from 100 microdissected Purkinje cells. Although comparable to earlier investigations using 100 microdissected Purkinje cells (See 4.3.3) it still falls below the number of proteins identified from 100 Purkinje cells once the RWS approach had been optimised (≈ 3700 proteins, see 4.3.6). Nevertheless, 2,000 proteins is still an impressive amount of protein considering the small amount of the input material, additionally, this

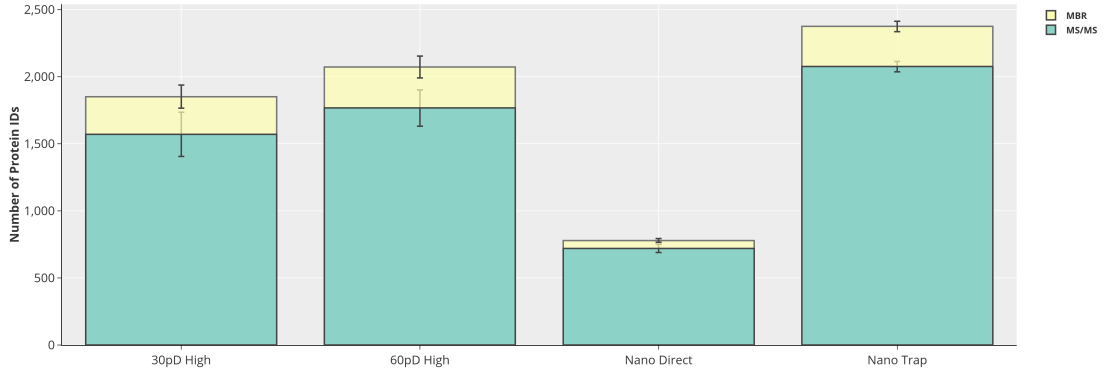


Figure E.3: TimsTOF Optimisation 2.
Turquoise bars denote the number of proteins identified solely by MS/MS, with cream bars showing the number of proteins identified by MaxQuant's Matched-Between-Run (MBRs) algorithm. 30pd = 48 minute gradient, 60pD = 24 minutes. High: Target intensity: 17,500; Threshold intensity: 1500 - see Table E.1

reduction in performance may be a necessary trade-off for decreasing the time taken to perform mass spectrometry on a sample in a high-throughput setting (17-minute gradient vs. 1-hour gradient).

E.2 LCM-seq Optimisations

As described in Section 2.2.5, the rich data generated by transcriptomics has produced novel and invaluable molecular insights into the nature of the CNS. Coupled with LCM, it has the ability to gain the same observations at a cell-type-specific level. In particular, the aforementioned, LCM-seq has been shown to produce impressive results with microdissected neurons, both from human and rodent brains. A significant challenge with investigating post-mortem nervous material with approaches such as LCM-seq is that the autolytic process that occurs after death causes biochemical degradation and may skew results in downstream applications. For RNA analysis, this is particularly challenging as RNA is already prone to such degradation; as such, strict quality control measures need to be placed to avoid wasting resources on material that is heavily degraded. RNA integrity numbers (RINs), based on the ratio of 28S and 18S ribosomal RNA are commonly used to filter degraded samples for analysis [484]. It is now commonplace in many transcriptomic workflows to use a Bioanalyser or equivalent to automatically calculate RINs. While a useful measurement, others have argued about its robustness. As most applications isolate, enrich, and sequence mRNA transcripts some have debated the relevancy of using RIN for RNA-seq; as the degradation of mRNA and rRNA may not be a linear or correlated process. Moreover, the process of first isolating RNA for RIN quality control can itself degrade RNA molecules or lead to significant loss of sample. Finally, while the ratio of 28S:18S may be reflective of rRNA degradation, the associated RIN produced by Agilent's Bioanalyzers are the result and calculation of inputting several additional measurements using proprietary algorithms that are difficult to replicate.

Another important variable for successful RNA-sequencing is ensuring the RNA to be analysed is representative of your starting material. Although many library preparation kits only require small amounts of material, typically in the nanogram range, this small amount of RNA material must be as representative of your global mRNA population as possible - a potential bottleneck for single-cell approaches. To maximise the quality and quantity of isolated mRNA from microdissected neurons

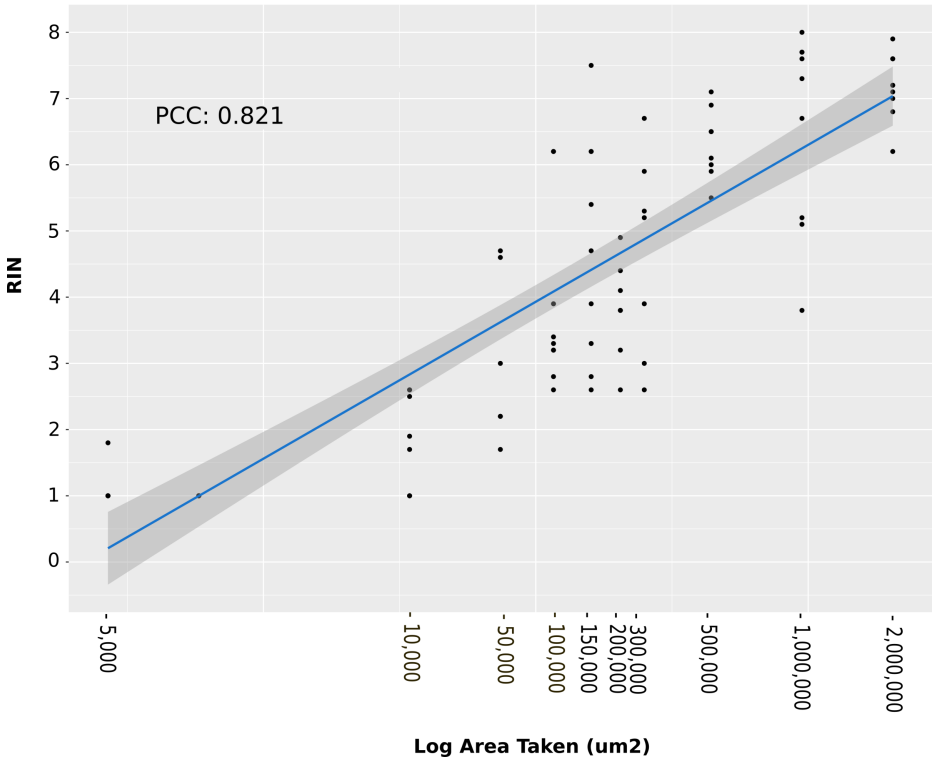


Figure E.4: RIN changes depending on the size of the Material Microdissected. Increasing amounts of cerebellar granule cells was microdissected from the same histology section, in quick succession.

various optimisation procedures were performed to improve LCM-seq performance on our tissue. First, to determine how much total RNA can be extracted by LCM a titration experiment was designed using microdissected Granule cell areas.

To investigate the deleterious effect of LCM on RNA, RINs from the above experiment were calculated using Aligent's Bioanalyzer; interestingly, a positive correlation was found between RIN and the amount of area microdissected, see Figure E.4, even though all samples originated from the same case. This may be a result of the minuscule amounts of RNA extracted from microdissection being challenging to purify, quantify, and analyse; potentially leading to a severe loss of material due to the number of pipetting steps upstream of RIN calculation. This is likely a reason why single-cell approaches often omit this. Although not analysing single cells *per se*, our pooled microdissected material often involves input of similar sizes, as slices of cells are dissected rather than whole cells. As such, it was decided

to skip RIN calculation and focus on improving quality and quantity using the SMART-Seq2/LCM-seq sample protocol that omits RNA extraction steps. Quality was assessed by the average fragment size of generated cDNA. Moreover, to ensure generated cDNA was of sufficient quality for RNA-seq the cDNA was used as a template to amplify several transcripts by RT-qPCR.

As a starting point, to determine how much mRNA can be obtained on our material with LCM-seq and to establish the minimum amount needed for representative cDNA; a titration experiment was performed similarly to the Purkinje cell titration described in Section 4.3.1, although Purkinje cells were extracted in smaller numbers (1 - 100) to minimise time spent microdissecting and because minute amounts of RNA is amplified by the pre-amplification steps in the SMART-Seq2 protocol. Learning from the optimisation steps for the RWS protocol; catapulting directly into a dry adhesive cap was compared to catapulting directly into a Lobind DNA/RNA cap filled with 0.2% (v/v) triton to act as a lysis buffer. A clear performance improvement was seen in retrieving mRNA between catapulting into lysis and dry caps, mirroring the results seen in RWS optimisation. While it was possible to generate enough cDNA to continue with the Nextera XT library construction (1 nanogram), additional experiments were performed in an attempt to maximise cDNA yield to ensure that the cDNA generated was representative of the total amount of mRNA transcripts found within microdissected samples.

Next, the composition of the lysis buffer was altered to encourage the lysis of cells once catapulted. For this, Granule cell areas were microdissected (100,000µm^2) and catapulted into caps containing either 0.2%, 0.4%, or 1% triton with ribonuclease inhibitors. Buffer RLT containing guanidine isothiocyanate was also tested as a lysis buffer. In addition to testing a variety of lysis buffers, the staining procedure (Nissl staining) was adapted with ribonuclease inactivation in mind; see 3.3.1. A striking increase in cDNA yield was seen across Granule cells that were catapulted into caps containing higher concentrations of Triton, ranging from 43 - 115 nanograms. Buffer RLT performed worst than expected, however, it was noted that, unlike the

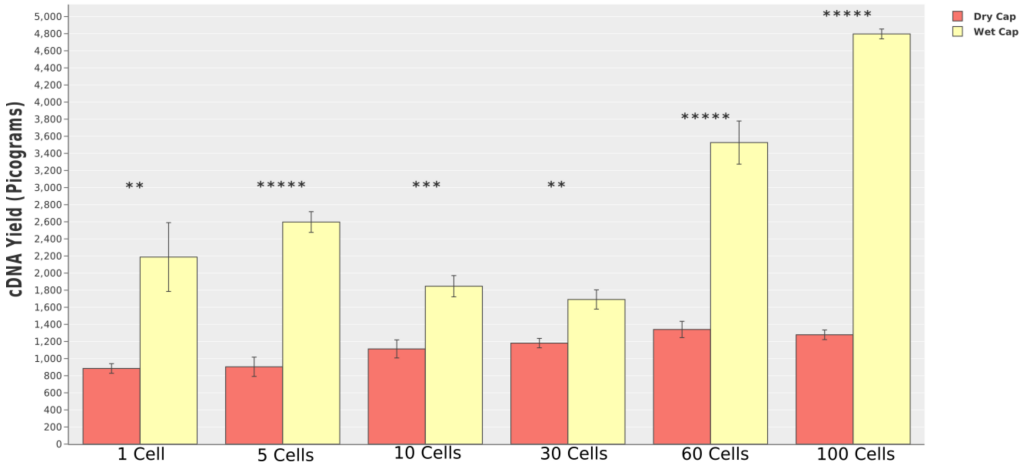


Figure E.5: Purkinje Cell Titration - cDNA Yield from LCM-seq. The amount of cDNA (picograms) obtained by microdissecting Purkinje cells. Overall, injecting into a wet cap (filled with lysis buffer) resulted in higher cDNA yeilds than injecting into dry caps - echoing the results from the optimisation steps used in the development of the RWS workflow. Red bars: cDNA yield after catapulting directly into the cap with lysis buffer, and adding lysis later. Cream bars: cDNA yield after catapulting directly into lysis buffer (0.1% triton). Vertical bars: standard deviation. $*****$ = p value: < 0.0001, $***$ = p value: < 0.001, $**$ = p value: < 0.01.

triton-based buffers, RLT was noticeably more difficult to smear across the cap of the tube - likely only lysing and capturing cells catapulted at the exact location where the RLT was positioned. Follow-up experiments using a larger volume or better smearing technique may result in a considerable increase in RNA yield. cDNA increased linearly with Triton concentration. In an attempt to reduce RNA degradation, tissue was snap-frozen after sectioning using a cryobar chilled to -48°C (i.e. "snap freezing"), however, this resulted in poor tissue adhesion and sub-optimise yields. Pre-chilled LCM slides were also tested in an attempt to preserve RNA quality but also resulted in poor tissue adhesion, making LCM difficult.

Similar to the development of the RWS approach, several parameters were tested to optimise the LCM-seq protocol to maximise the amount of cDNA generated from microdissected post-mortem neurons. Increasing cDNA yield allows fewer neurons to be microdissected, speeding up the LCM process which is known to rapidly degrade RNA. Moreover, generating an abundance of cDNA also ensures other quality control steps (quantification, fragment analysis, qPCR) can be performed

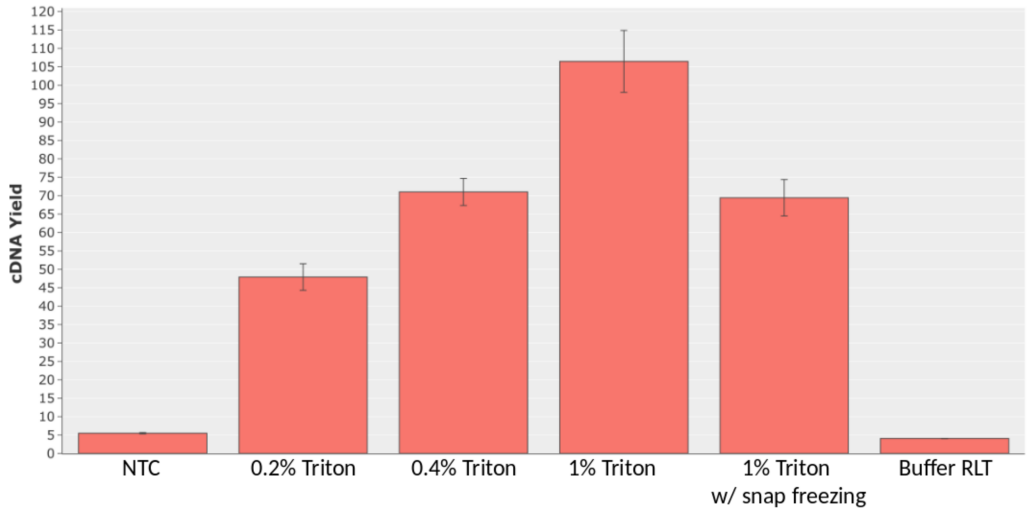


Figure E.6: Optimising Lysis buffer - cDNA Yield from LCM-seq.
Several lysis buffers were tested to maximise the amount of cDNA that could be synthesised by the SMART-Seq2 protocol. NTC: negative control. Snap freezing: prior to LCM, tissue sections were immediately frozen after sectioning by placing on a -48°C cryobar. Vertical bars: standard deviation. cDNA yield: in nanograms.

without depleting precious material. Finally, LCM-seq only requires 1 nanogram of material; generating a high yield may result in cDNA that is more representative of the neurons microdissected, rather than only containing mRNA from the most abundant transcripts. Significant steps were taken to reduce RNAse activity, overall, a 2163% increase in cDNA yield was seen achieved when microdissecting 100,000µm^2 of neuronal material - see Figure E.7. As $\approx$ 100 nanograms of cDNA was consistently generated on in-house samples when microdissecting this amount, this was set as the threshold for the subsequent LCM-seq experiment.

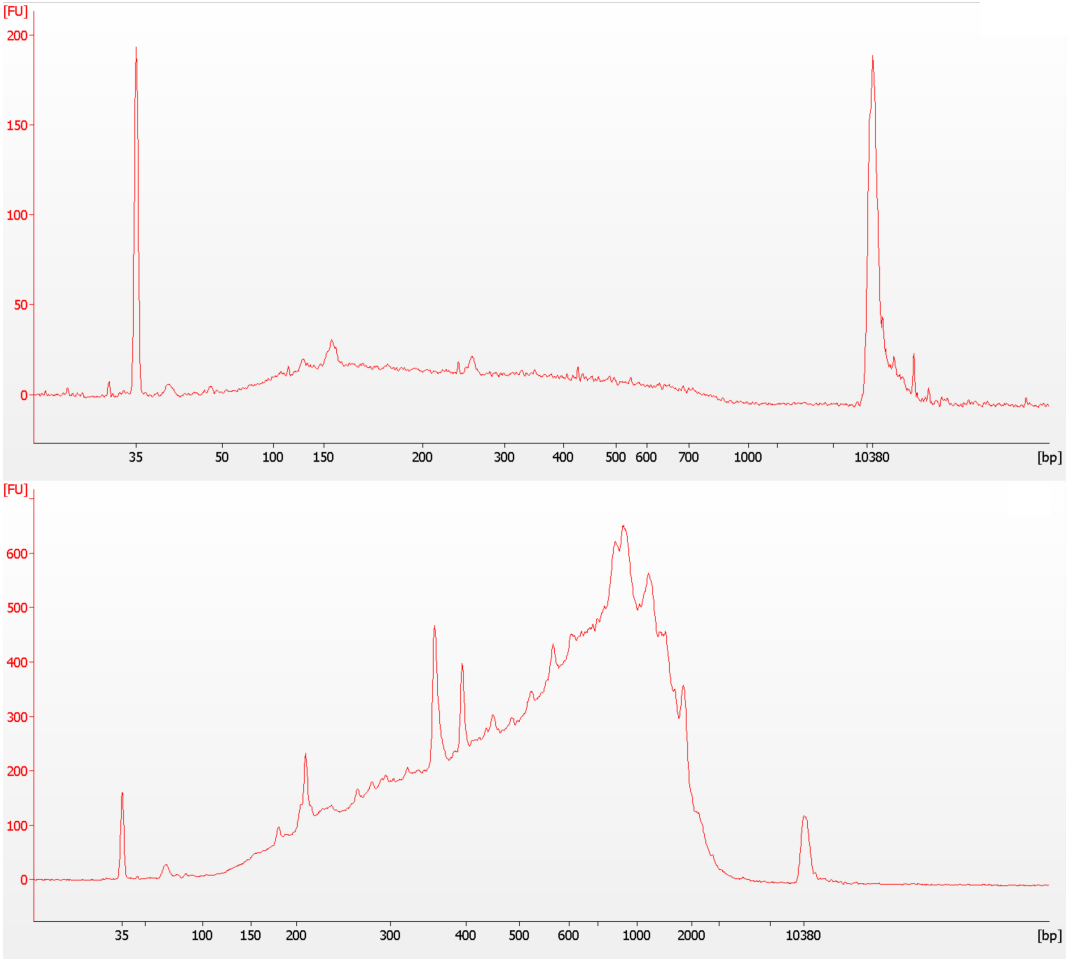


Figure E.7: cDNA Yield Before and After Optimisation.
Top panel: cDNA generated from microdissecting Purkinje cells (100,000µm^2) before optimisation. The broad peak around 150-400 bp indicates marked RNA degradation. Bottom panel: cDNA generated from microdissecting Purkinje cells (100,000µm^2) after optimisation. cDNA showed little degradation. y-axis: FU (Fluorescent units); x-axis: bp (base pairs).

E.3 LCM-seq Quality Control

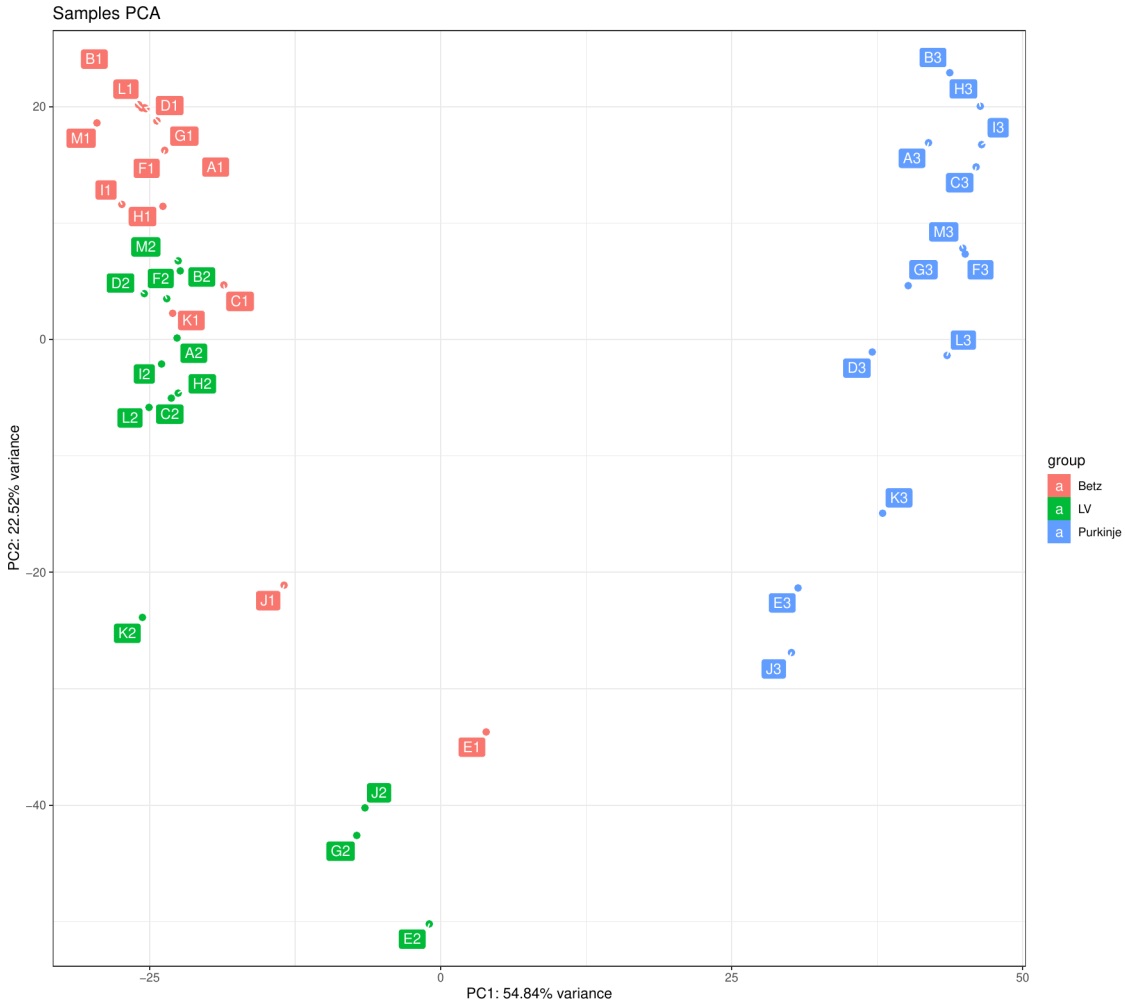


Figure E.8: PCA - LCM-seq Results (with Outliers).
Principal component analysis reveals a subset of samples that differ significantly from their respective groups.

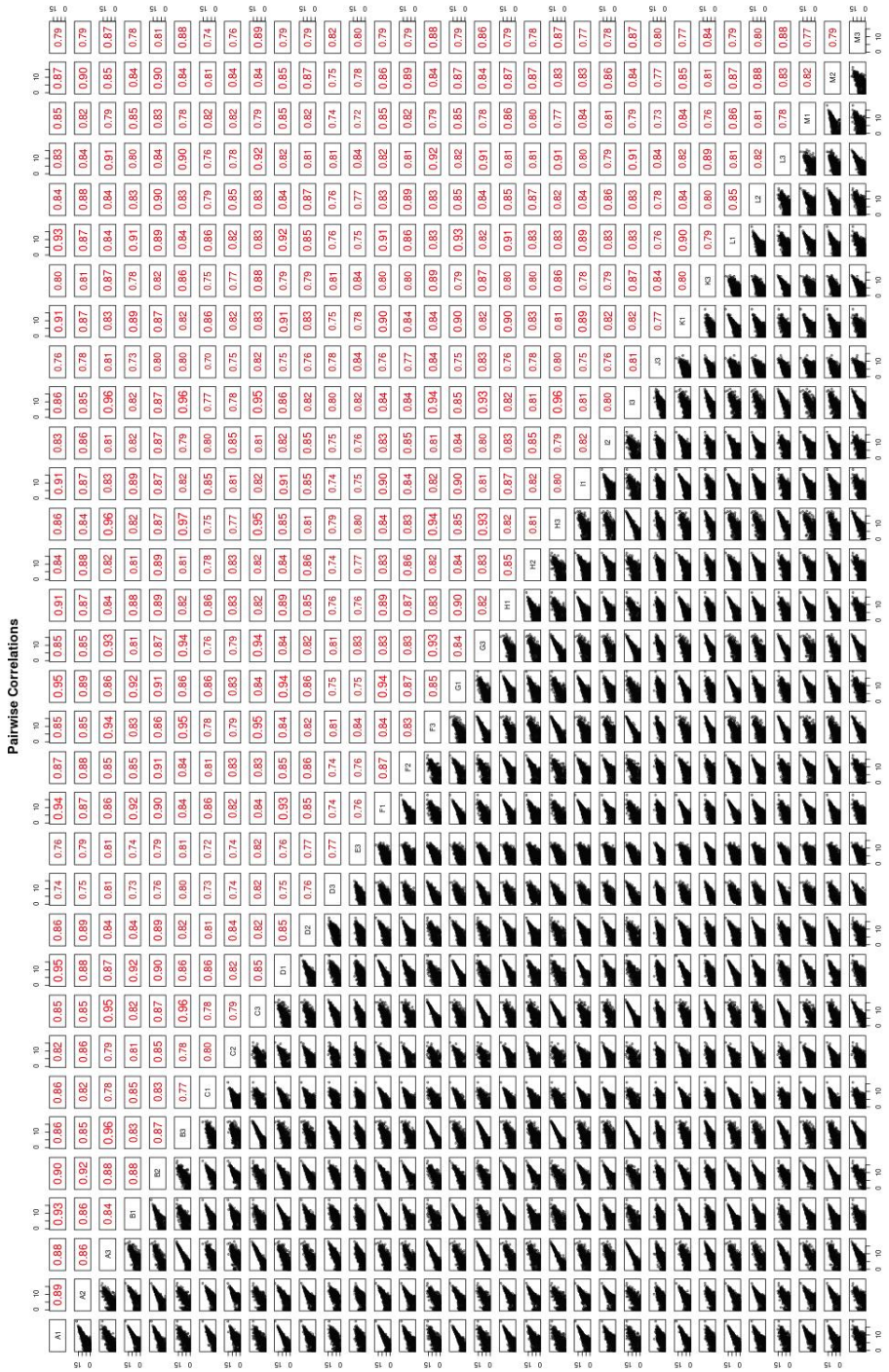


Figure E.9: Correlation Matrix of LCM Samples.

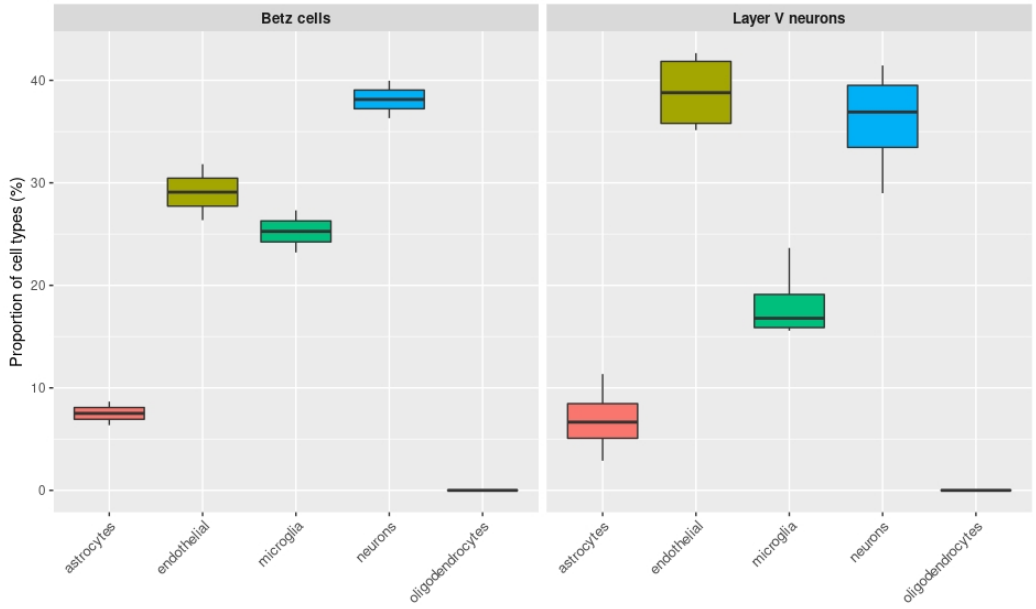


Figure E.10: MuSiC Results of LCM-seq Cohort - Outliers Only.
Comparing the expression of the outlier samples removed from analysis with MuSiC shows a large increase in non-neuronal transcripts. The lack of enrichment of neuronal transcripts indicate these samples have significant contamination from neighbouring cell-types.

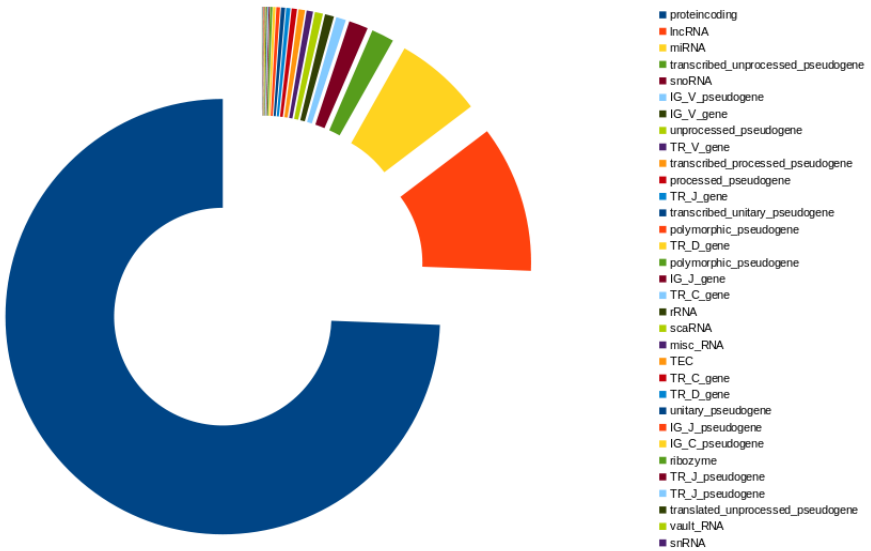


Figure E.11: Proportion of All RNA Species Identified.
The large number of protein-coding RNA confirms that mRNA has been selectively enriched in our samples following the SMART-Seq2 protocol.

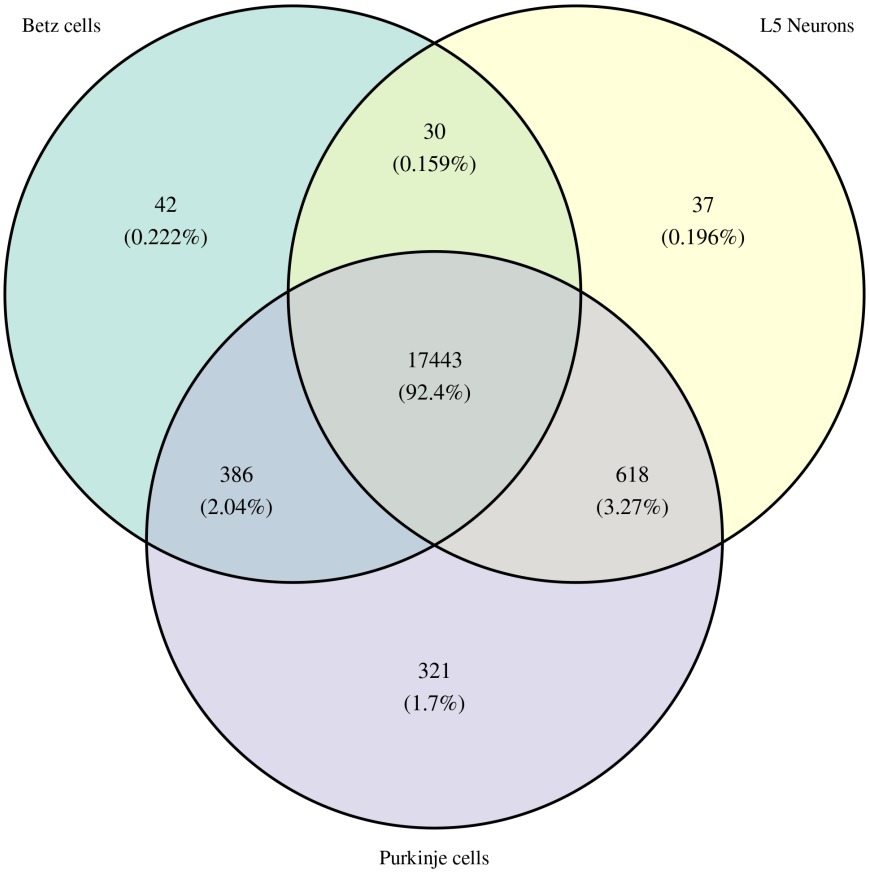


Figure E.12: Venn Diagram - All Protein-Coding Genes in Data-Set.
Overlap of all the protein-coding genes (mRNA transcripts) identified by LCM-seq and used for subsequent analysis. As expected, the vast majority of mRNA transcripts were shared across the 3 cell-types investigated.

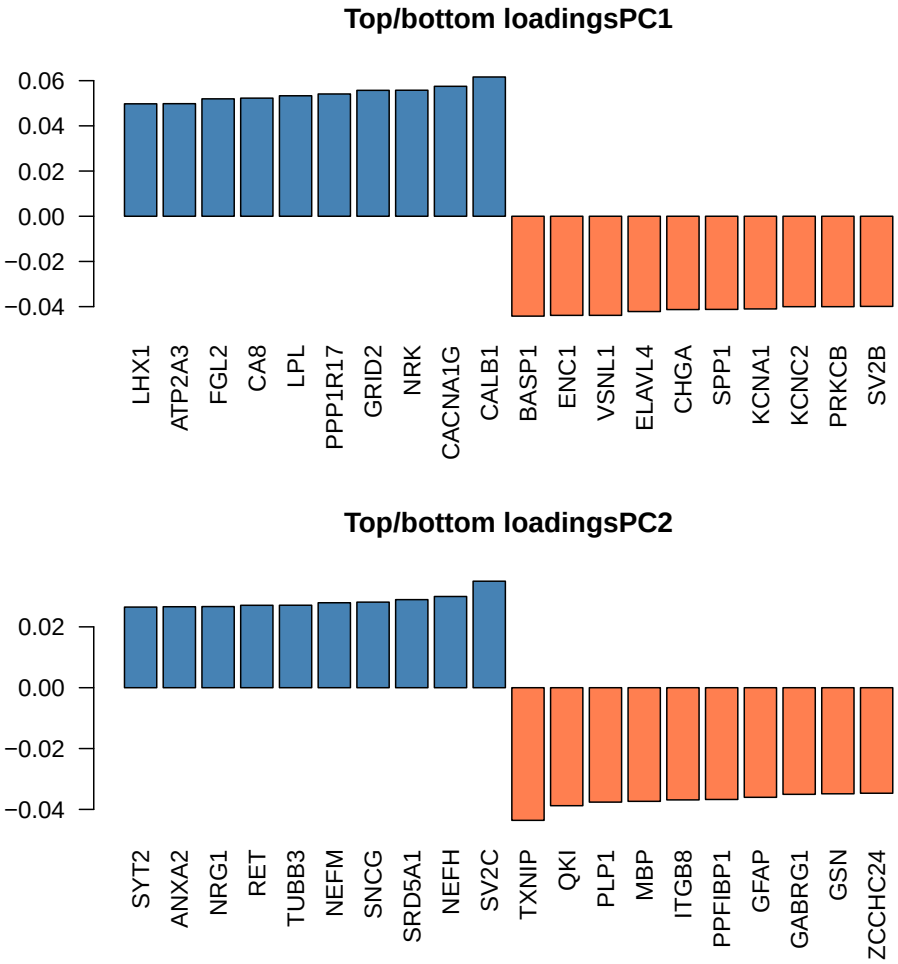


Figure E.13: Top-Weighted PCA Transcripts.
PC1: The top 20 transcripts responsible for the variation/separation between Purkinje cell and L5/Betz samples. PC2: The top 20 transcripts responsible for the variation/separation between Betz cell and L5 samples.

E.4 LCM-seq Analysis

E.4.1 Healthy Betz Cells vs. Healthy Purkinje Cells

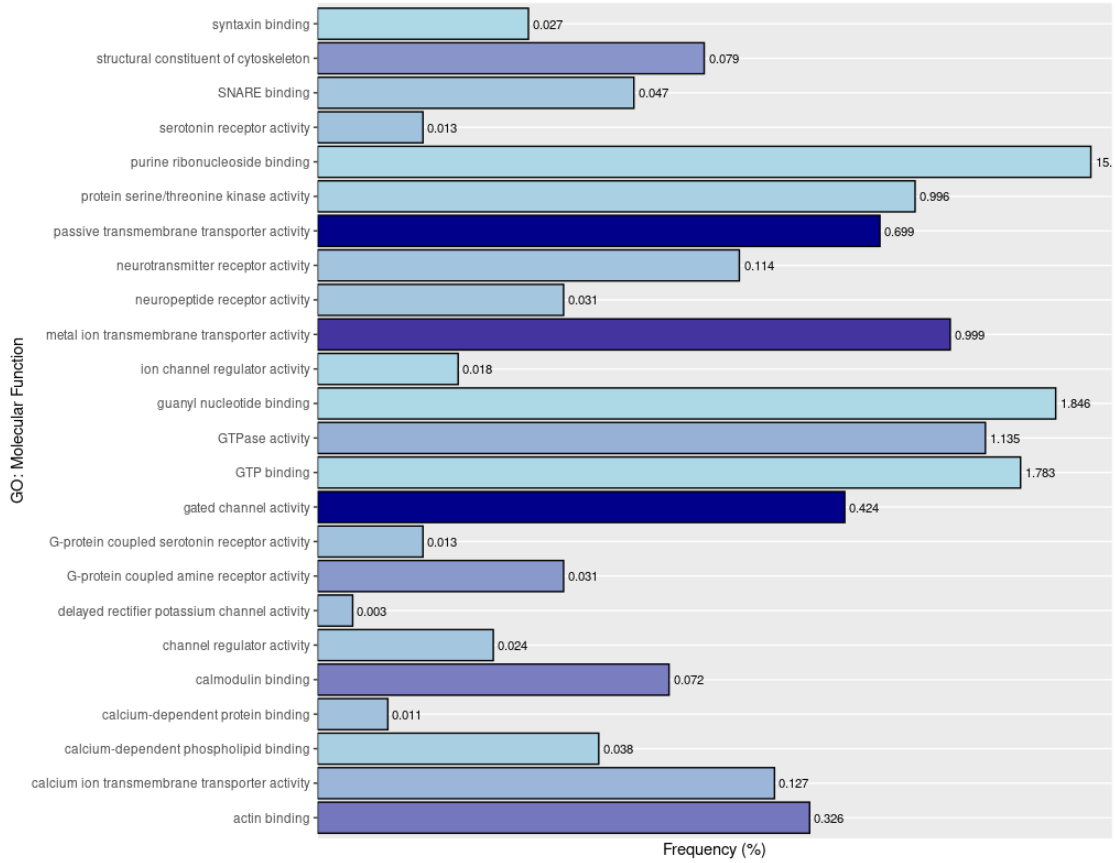


Figure E.14: REViGO Output - Betz Cells vs Purkinje Cells - MF.
A reduced list of GOMF pathways that were significantly enriched in Betz cell samples. Darker colour denotes smaller p values. Frequency: The proportion of this GO term in wider human proteome annotation database. Smaller percentages = rarer/more specific GO terms.

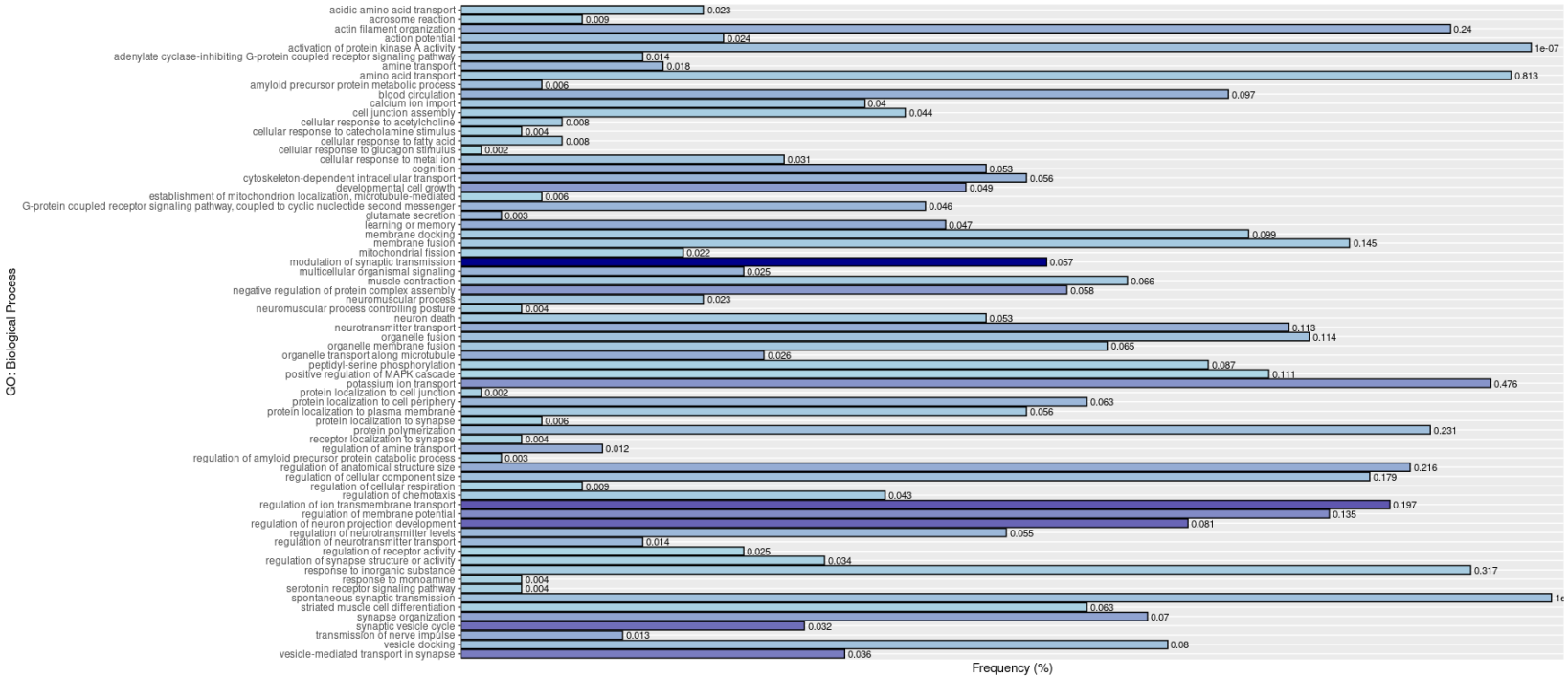


Figure E.15: REViGO Output - Betz Cells vs Purkinje Cells - BP.
A reduced list of GOBP pathways that were significantly enriched in Betz cell samples. Darker colour denotes smaller p values. Frequency: The proportion of this GO term in wider human proteome annotation database. Smaller percentages = rarer/more specific GO terms.

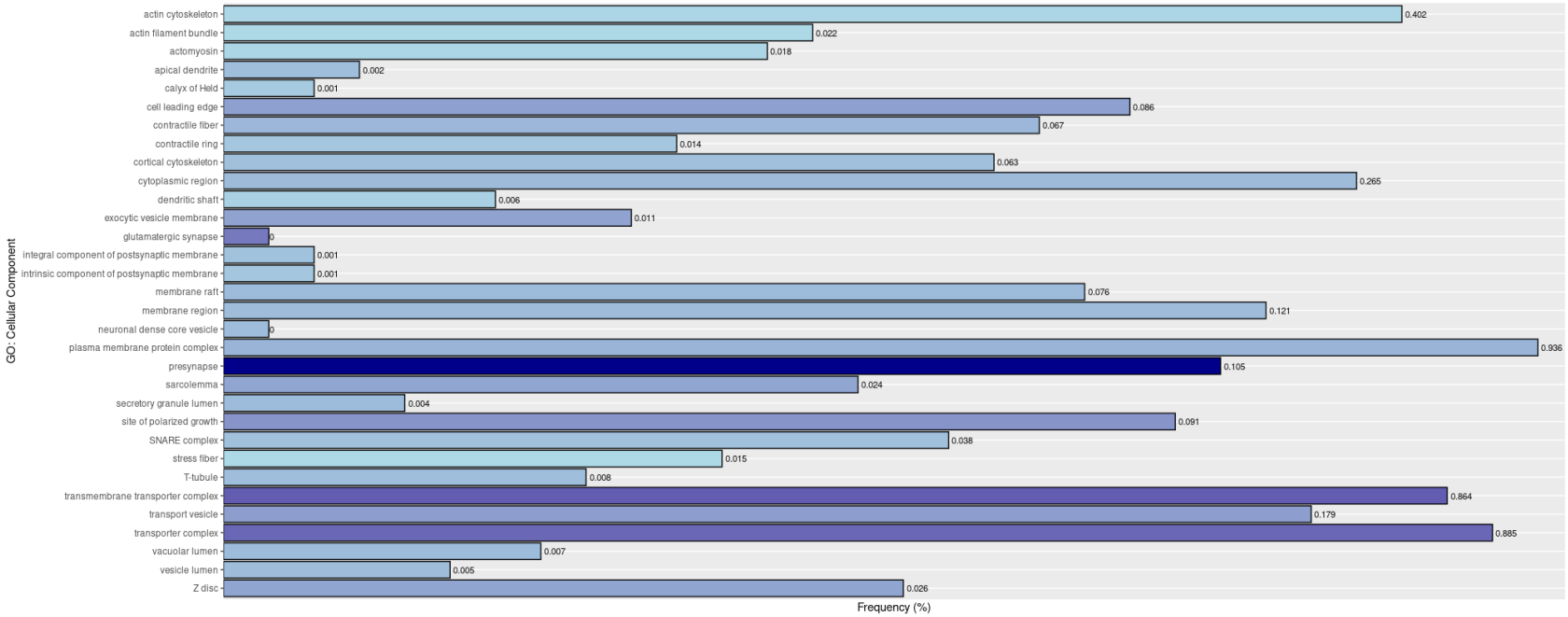


Figure E.16: REViGO Output - Betz Cells vs Purkinje Cells - CC.
A reduced list of GOCC pathways that were significantly enriched in Betz cell samples. Darker colour denotes smaller p values. Frequency: The proportion of this GO term in wider human proteome annotation database. Smaller percentages = rarer/more specific GO terms.

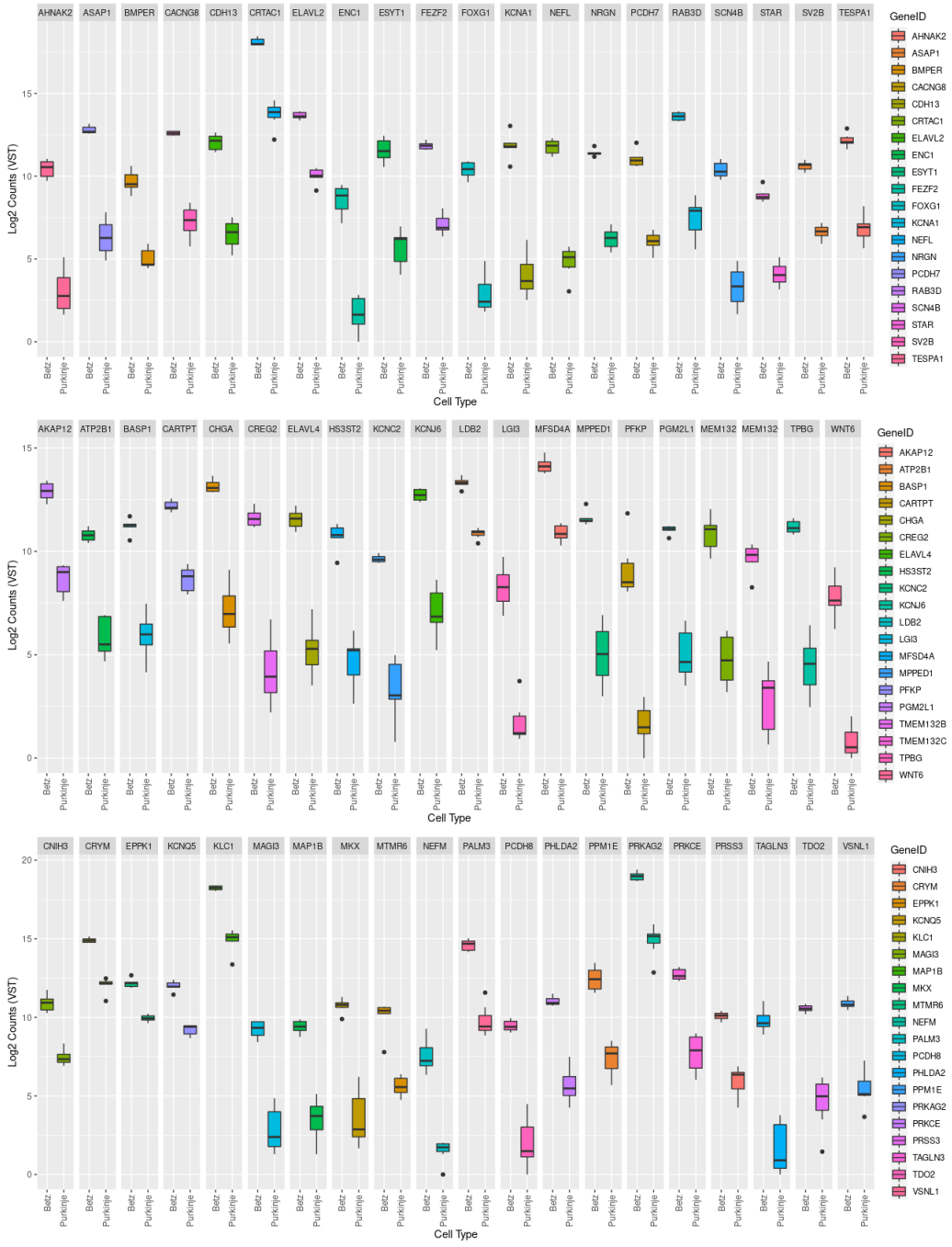


Figure E.17: Betz Cells vs. Purkinje Cells LCM-seq Comparison: Top 60 Betz DEGs. A DESeq2 stat ranked list of the top performing 60 DEGs found in Betz cell samples compared to Purkinje cells. This list may help identify individual transcripts/proteins that could serve as future Betz cell markers. Top panel: Top 20 DEGs, Middle panel: Top 21-40 DEGs, Bottom panel: Top 41-60 DEGs.

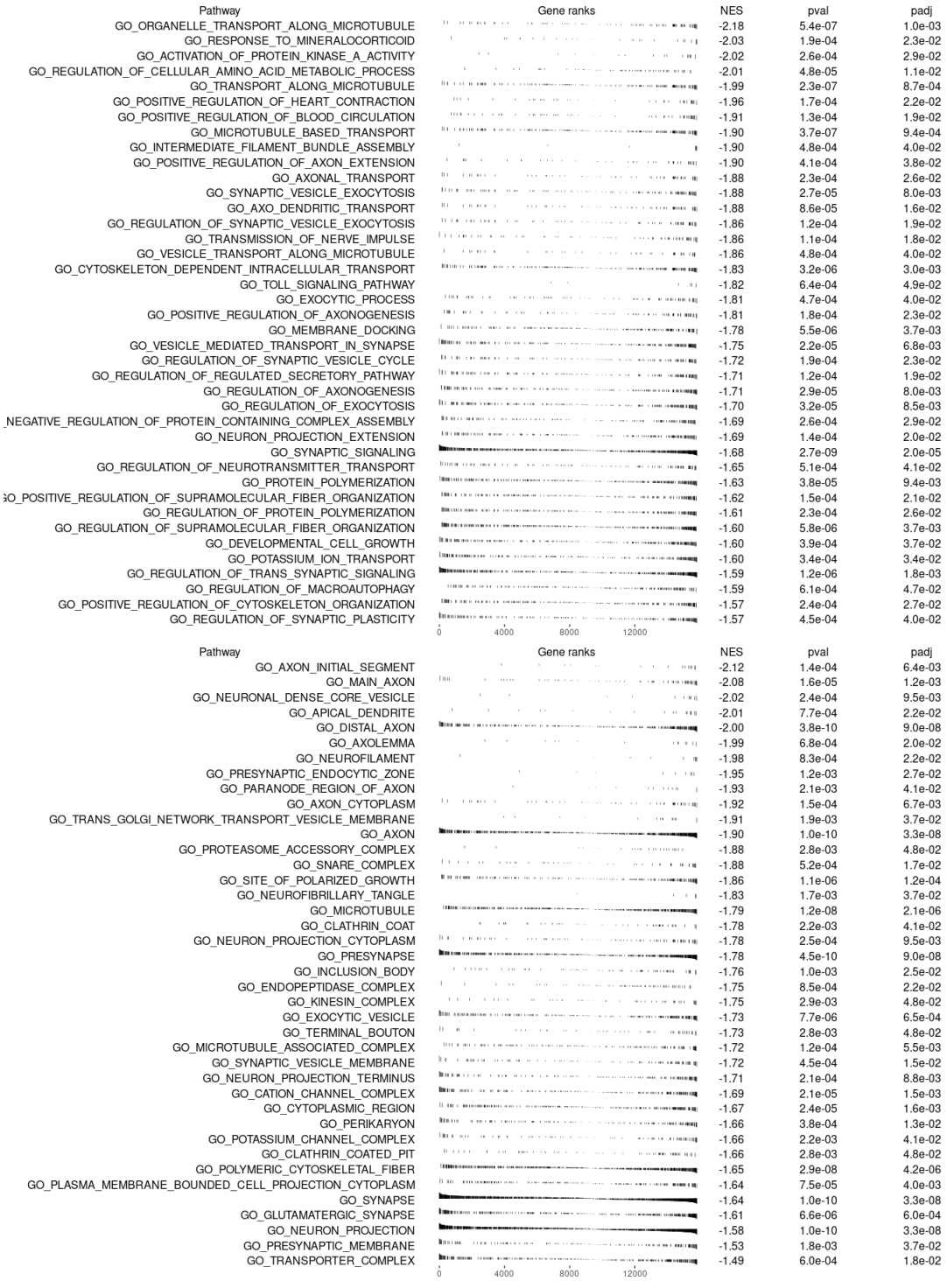


Figure E.18: The Top GO Enriched Pathways from GSEA - Betz Cells vs. Purkinje Cells.

Top figure: Enriched GOBP terms. Bottom figure: Enriched GOCC terms. NES: Normalised enrichment score, negative values are more enriched in Betz cell samples, positive NESs represent a higher enrichment in Purkinje cell samples.

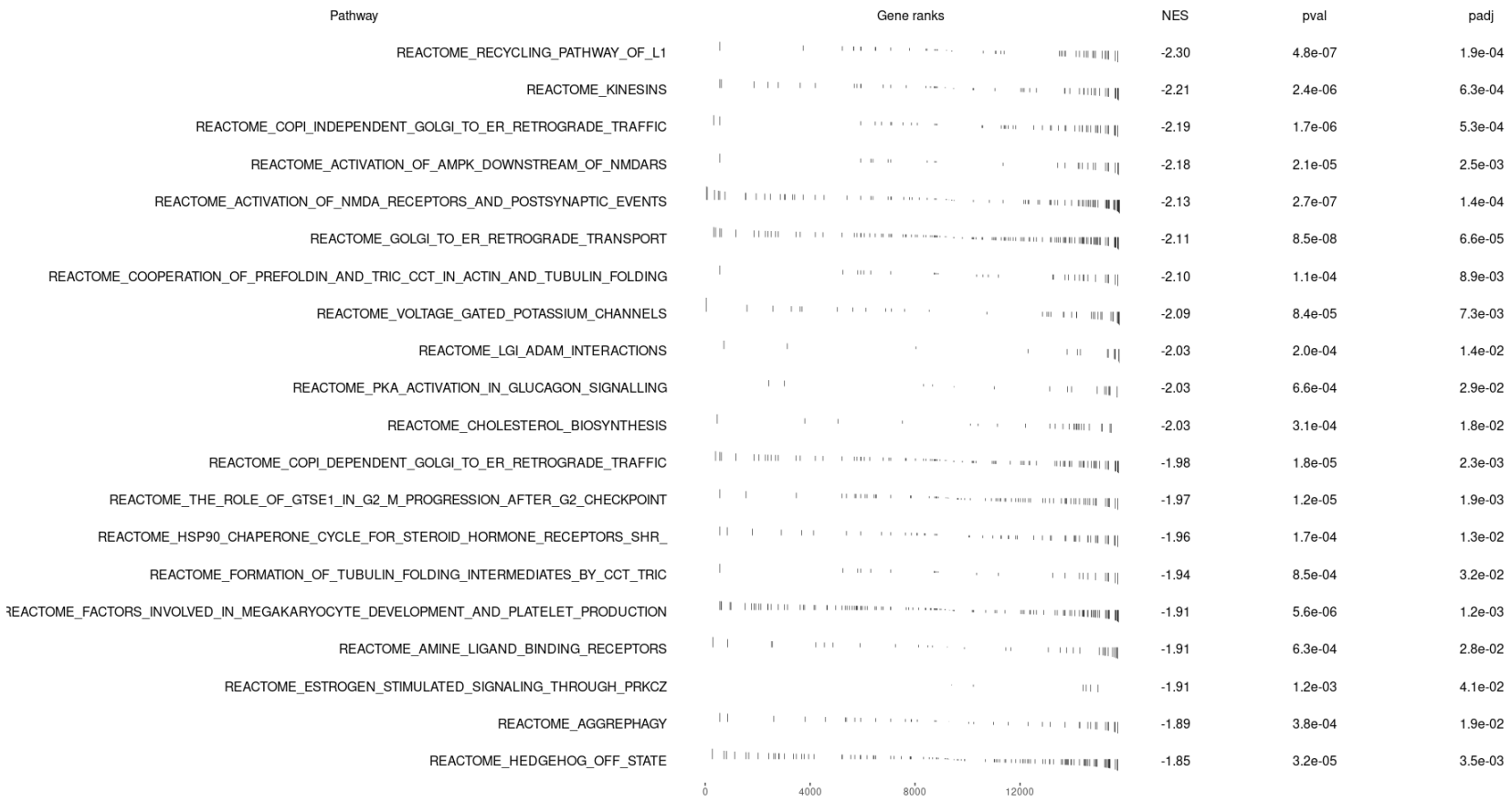


Figure E.19: The Top GO Enriched Pathways from GSEA - Betz Cells vs. Purkinje Cells.
NES: Normalised enrichment score, negative values are more enriched in Betz cells samples and positive NESs are more enriched in Purkinje cell samples.

E.4.2 Healthy Betz Cells vs. Healthy L5 Neurons

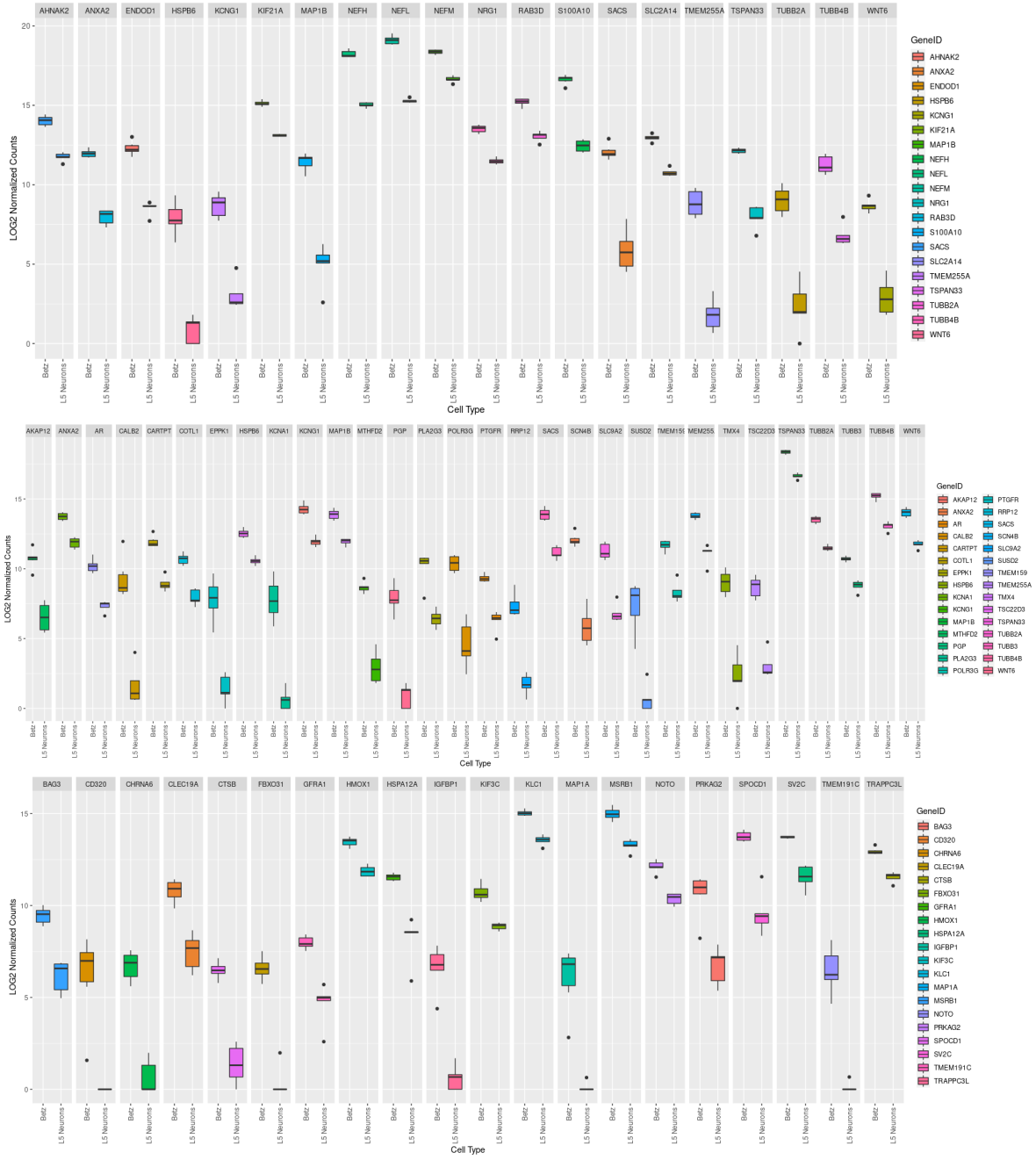


Figure E.20: Betz Cells vs. L5 Neurons Comparison: Top 60 Betz DEGs. A DESeq2 stat ranked list of the top performing 60 DEGs found in Betz cell samples compared to L5 neurons. This list may help identify individual transcripts/proteins that could serve as future Betz cell markers, especially since the two neuronal cell-types are closely related. Top panel: Top 20 DEGs, Middle panel: Top 21-40 DEGs, Bottom panel: Top 41-60 DEGs.

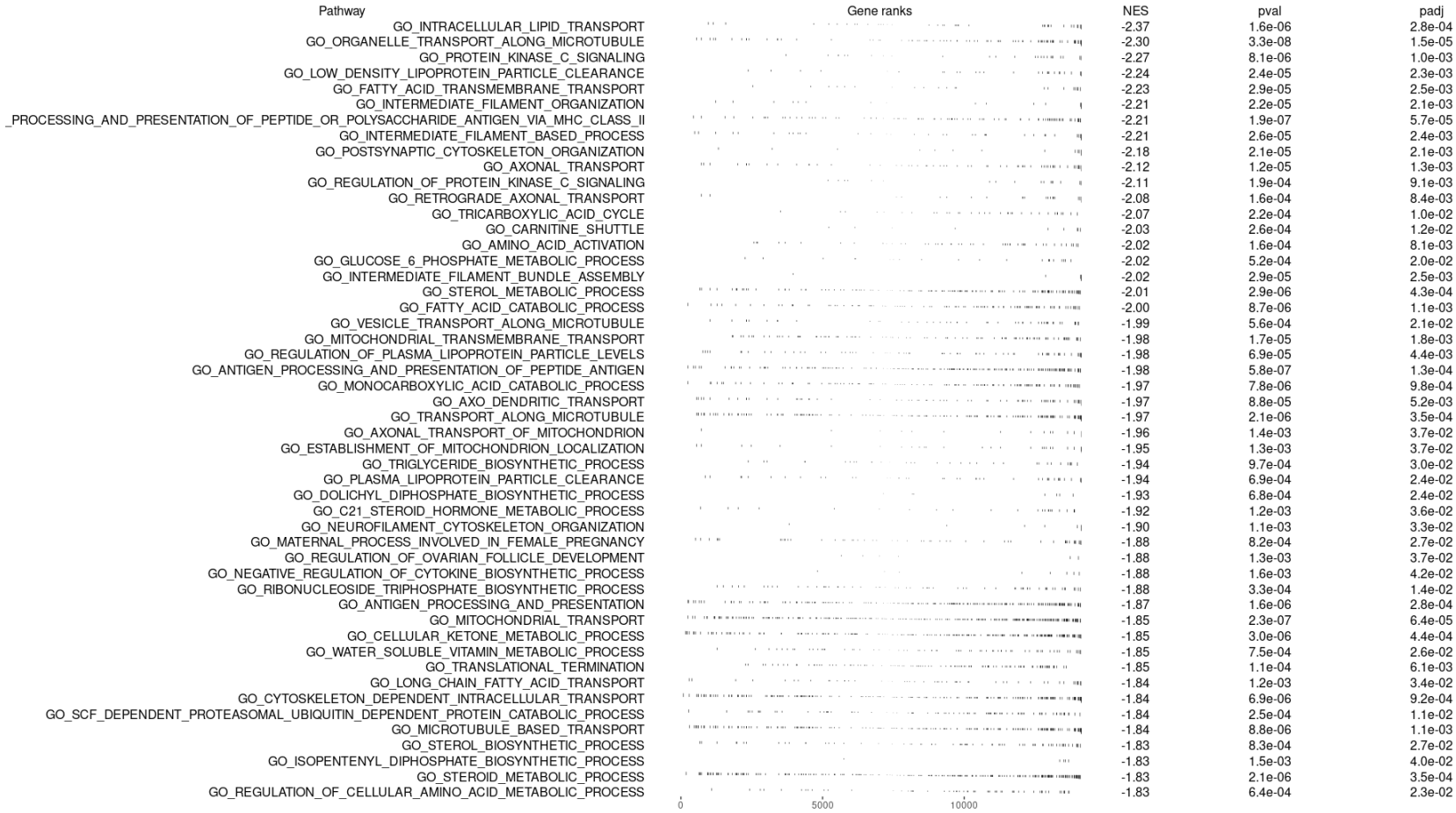


Figure E.21: Top Enriched GOBP Terms - Betz Cells vs. L5 Neurons.
NES: Normalised enrichment score, negative values are more enriched in Betz cells samples and positive NESs are more enriched in L5 neurons.

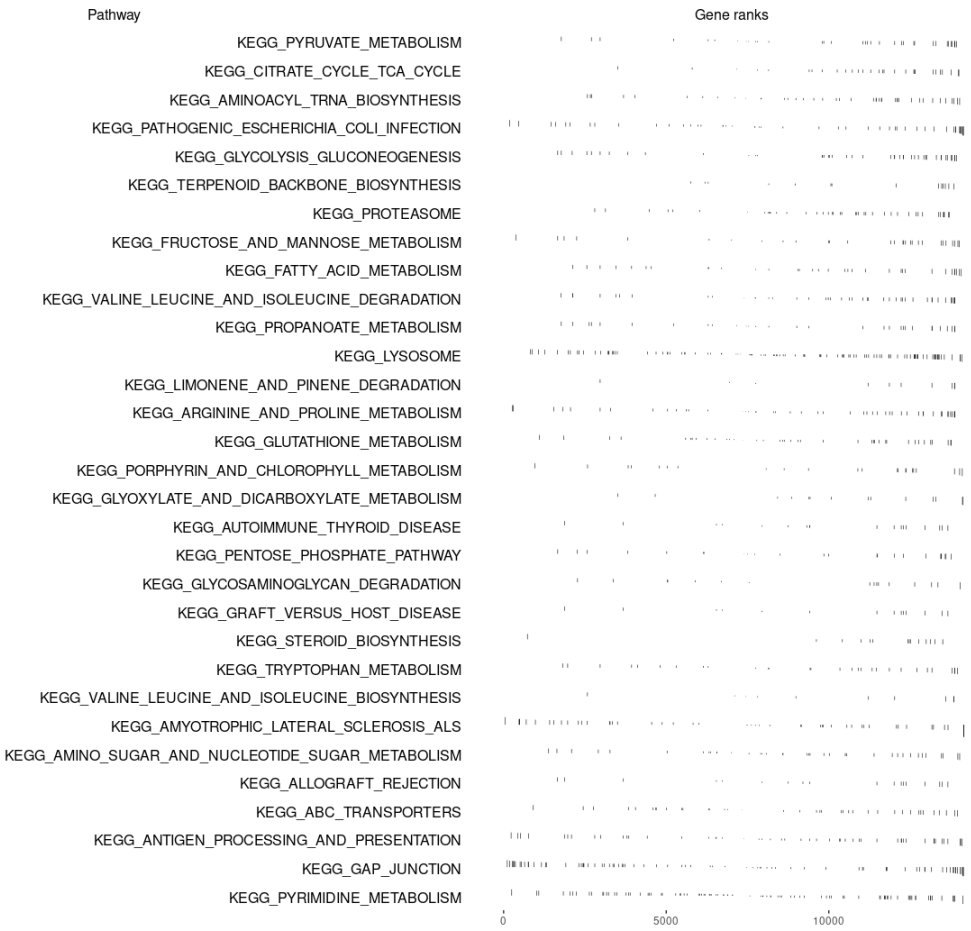


Figure E.22: Top Enriched KEGG Terms - Betz Cells vs. L5 Neurons.
NES: Normalised enrichment score, negative values are more enriched in Betz cells samples and positive NESs are more enriched in L5 neurons.

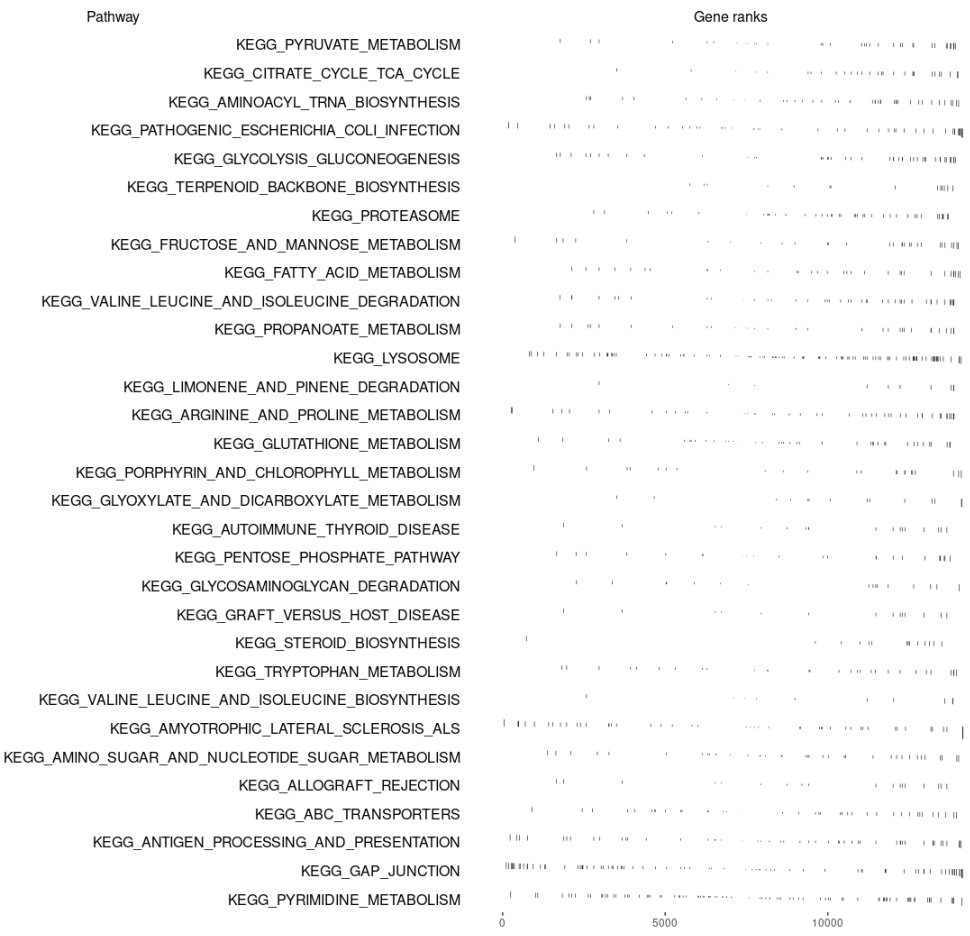


Figure E.23: Top Enriched Reactome Terms - Betz Cells vs. L5 Neurons.
NES: Normalised enrichment score, negative values are more enriched in Betz cells samples and positive NESs are more enriched in L5 neurons.

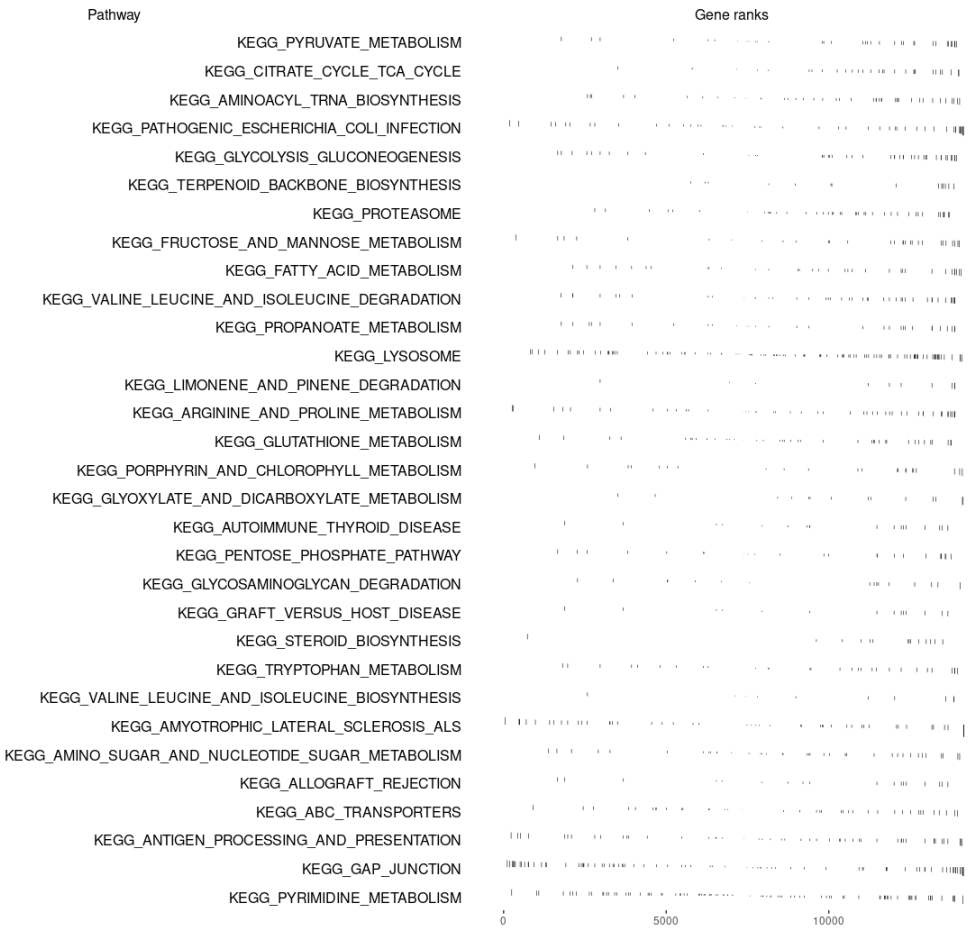


Figure E.24: Top Enriched WIKI Terms - Betz Cells vs. L5 Neurons.
NES: Normalised enrichment score, negative values are more enriched in Betz cells samples and positive NESs are more enriched in L5 neurons.

GO Type	GO Term	Fold Enrichment	Adj. p value
GOBP	retrograde axonal transport	8.13	1.46E-03
	negative regulation of anion transmembrane transport	7.27	4.67E-02
	NADPH regeneration	6.79	2.17E-02
	protein kinase C signaling	6.26	5.87E-03
	glucose 6-phosphate metabolic process	5.94	1.79E-02
	low-density lipoprotein particle clearance	4.92	4.75E-02
	vesicle transport along microtubule	4.86	4.52E-03
	organelle transport along microtubule	4.76	4.53E-05
	axonal transport	4.28	5.36E-03
	axo-dendritic transport	3.89	5.87E-03
	vesicle cytoskeletal trafficking	3.49	2.16E-02
	negative regulation of protein polymerization	3.44	2.21E-02
	transport along microtubule	3.13	5.45E-04
	antigen processing and presentation of exogenous peptide antigen	3.12	5.45E-04
	regulation of intracellular pH	3.09	4.91E-02
	cytoskeleton-dependent intracellular transport	2.95	5.45E-04
	monosaccharide biosynthetic process	2.94	4.91E-02
	microtubule-based transport	2.86	8.89E-04
	negative regulation of protein-containing complex assembly	2.75	2.16E-02
	alcohol biosynthetic process	2.68	1.40E-02
	negative regulation of supramolecular fiber organization	2.55	4.62E-02
	regulation of cellular ketone metabolic process	2.44	2.76E-02
	glucose metabolic process	2.25	4.91E-02
	microtubule-based movement	2.12	5.87E-03
	monosaccharide metabolic process	2.07	4.62E-02
	alcohol metabolic process	2.05	1.09E-02
	Golgi vesicle transport	2.02	1.09E-02

GO Type	GO Term	Fold Enrichment	Adj. p value
	organic hydroxy compound metabolic process	1.82	1.61E-02
GOMF	dynein heavy chain binding	7.95	1.52E-02
	structural constituent of cytoskeleton	4.62	1.39E-05
	clathrin binding	3.76	4.53E-03
	microtubule binding	2.43	1.70E-03
	tubulin binding	2.29	4.60E-04
	voltage-gated ion channel activity	2.28	3.65E-02
	GTPase activity	2.16	3.54E-03
	nucleoside binding	2.13	2.18E-03
	ribonucleoside binding	2.11	2.18E-03
GOCC	clathrin coat of trans-Golgi network vesicle	8.70	4.43E-03
	trans-Golgi network transport vesicle membrane	7.46	8.09E-03
	endolysosome	6.96	4.22E-03
	proteasome regulatory particle, base subcomplex	6.96	2.83E-02
	endolysosome membrane	6.96	2.83E-02
	axon initial segment	6.53	1.44E-02
	pore complex	5.97	8.09E-03
	outer mitochondrial membrane protein complex	5.22	1.51E-02
	perinuclear endoplasmic reticulum	5.22	3.07E-02
	proteasome regulatory particle	4.97	3.77E-02
	axon cytoplasm	4.94	1.62E-04
	clathrin vesicle coat	4.47	2.83E-02
	main axon	3.59	5.77E-03
	vesicle coat	3.42	2.05E-02
	proteasome complex	3.37	1.44E-02

Table E.2: Gene Ontology Over-Representation Analysis of the DEGs Consistently Enriched in Betz Cells.

E.4.3 Healthy Betz Cells vs. ALS Betz Cells

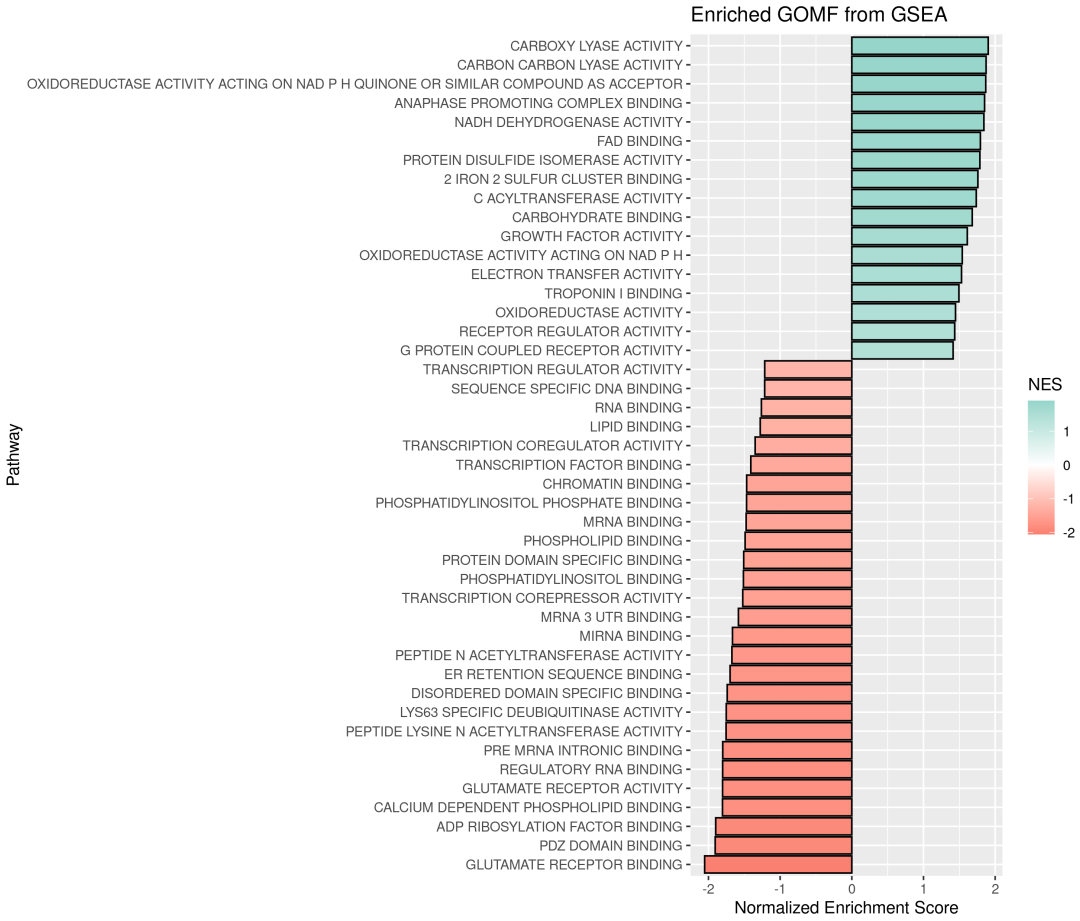


Figure E.25: Top Enriched GOMF Terms - Healthy Betz Cells vs. ALS Betz Cells. NES: Normalised enrichment score. Positive NESs: downregulated in ALS Betz cells. Negative NESs: upregulated in ALS Betz cells.

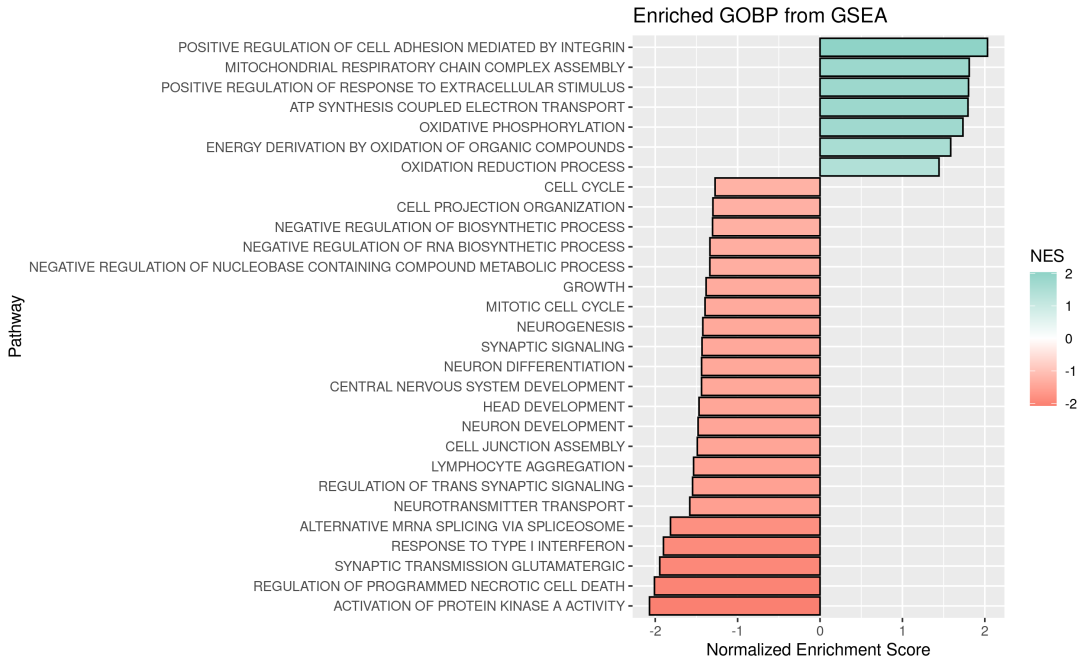


Figure E.26: Top Enriched GOBP Terms - Healthy Betz Cells vs. ALS Betz Cells. NES: Normalised enrichment score. Positive NESs: downregulated in ALS Betz cells. Negative NESs: upregulated in ALS Betz cells.

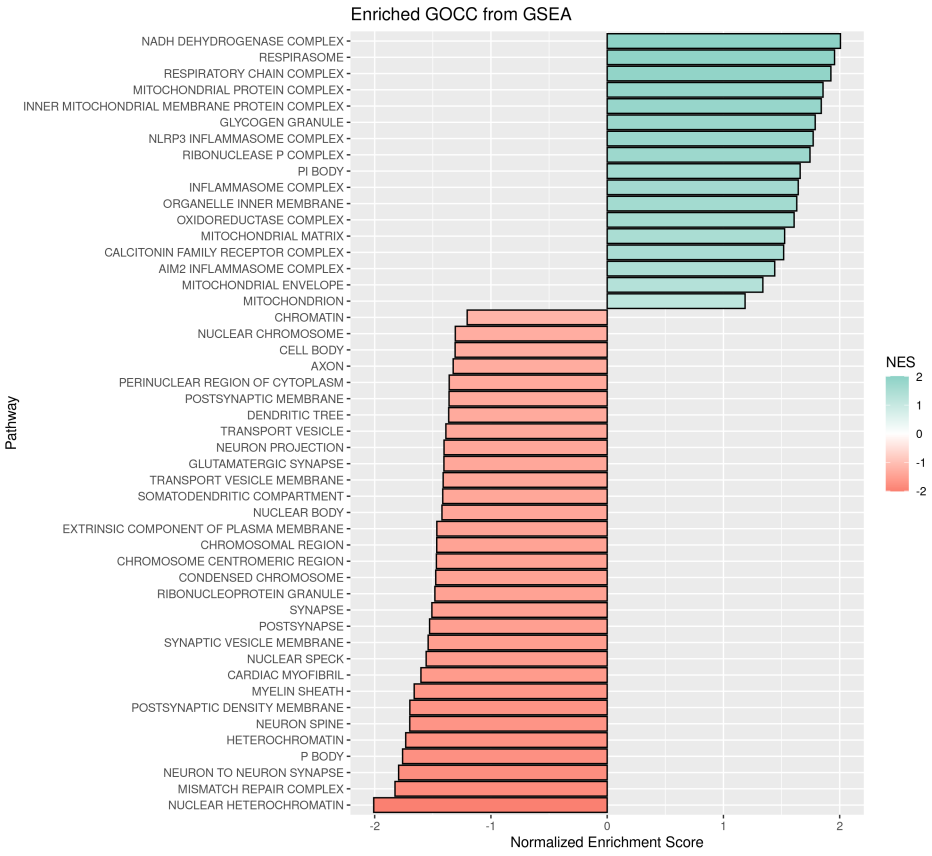


Figure E.27: Top Enriched CC Terms - Healthy Betz Cells vs. ALS Betz Cells. NES: Normalised enrichment score. Positive NESs: downregulated in ALS Betz cells. Negative NESs: upregulated in ALS Betz cells.

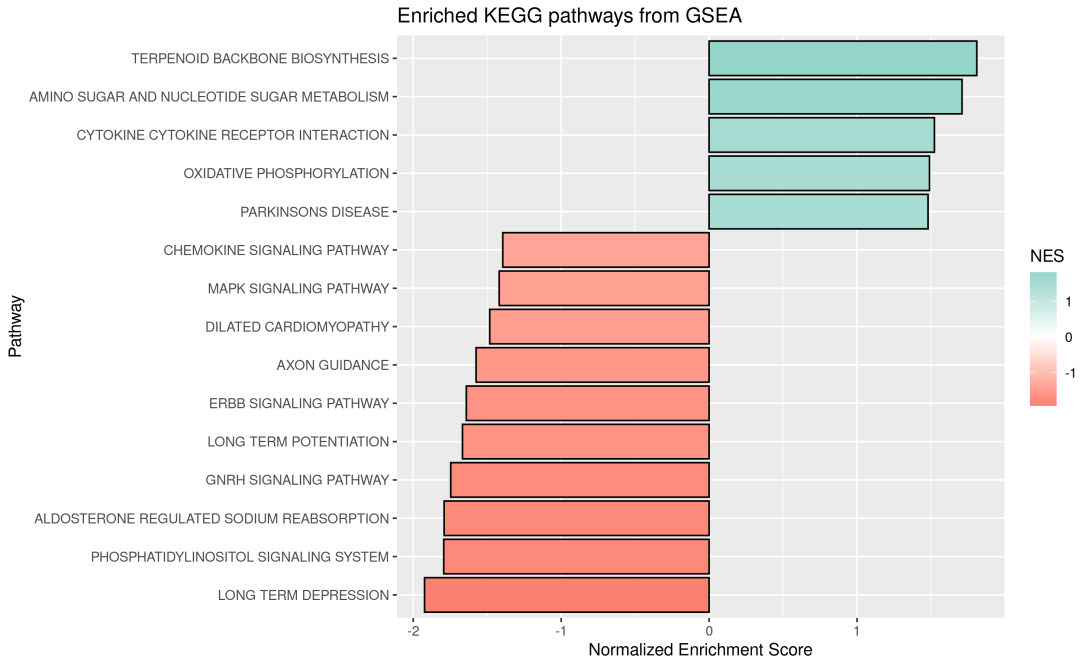


Figure E.28: Top Enriched KEGG Terms - Healthy Betz Cells vs. ALS Betz Cells. NES: Normalised enrichment score. Positive NESs: downregulated in ALS Betz cells. Negative NESs: upregulated in ALS Betz cells.

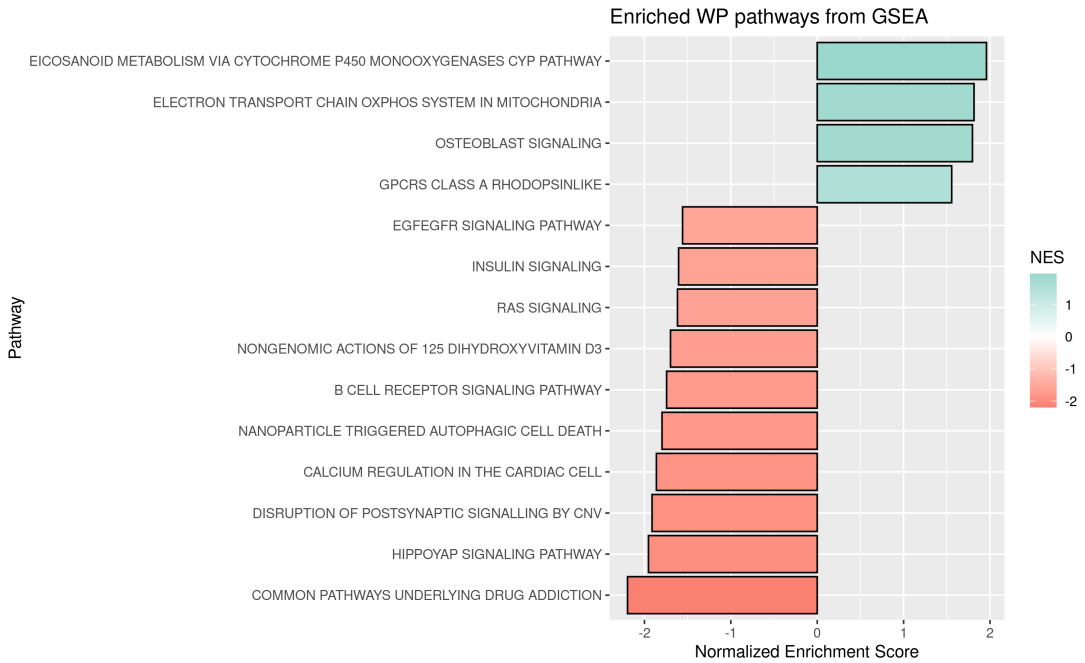


Figure E.29: Top Enriched WIKI Terms - Healthy Betz Cells vs. ALS Betz Cells. NES: Normalised enrichment score. Positive NESs: downregulated in ALS Betz cells. Negative NESs: upregulated in ALS Betz cells.

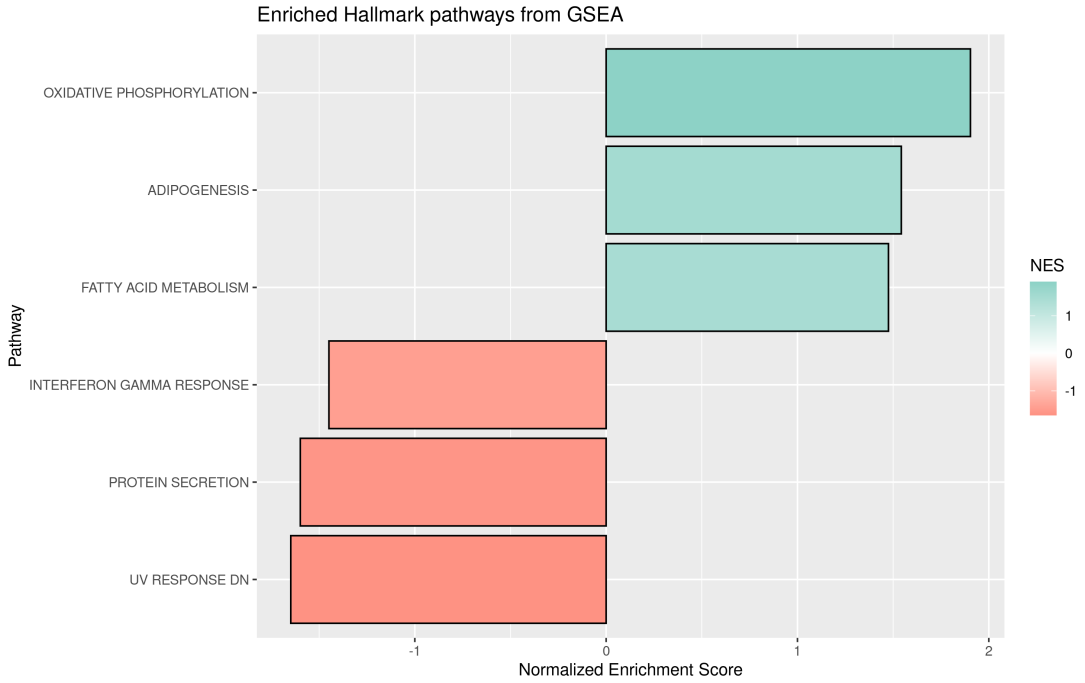


Figure E.30: Top Enriched Hallmark Terms - Healthy Betz Cells vs. ALS Betz Cells. NES: Normalised enrichment score. Positive NESs: downregulated in ALS Betz cells. Negative NESs: upregulated in ALS Betz cells.

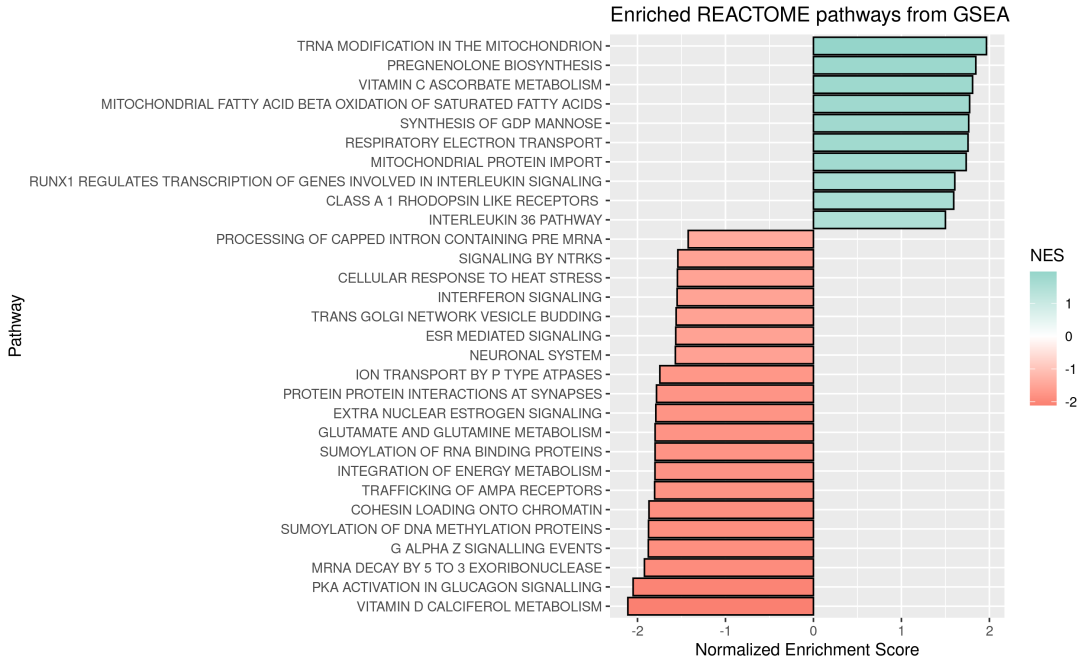


Figure E.31: Top Enriched Reactome Terms - Healthy Betz Cells vs. ALS Betz Cells. NES: Normalised enrichment score. Positive NESs: downregulated in ALS Betz cells. Negative NESs: upregulated in ALS Betz cells.

Gene Name	stat	adj p value	LogFC
FOS	6.042413358	1.09E-05	5.907715637
RECQL	5.902769267	1.72E-05	5.845309434
GHSR	5.069861802	0.001148213	5.103069906
MCUB	5.031152371	0.001171804	7.062902807
ULBP1	4.937067157	0.00163382	6.529846322
CSNK2B	4.821927456	0.00245631	1.50635624
RUNX1	4.806890593	0.00245631	5.284556304
NANS	4.612271797	0.004519909	1.443561097
TMEM120A	4.610670804	0.004519909	1.33806497
ZNF567	4.607543914	0.004519909	2.452463428
ACSM6	4.576504959	0.004726021	6.686508157
GSAP	4.510102734	0.005564377	3.639488689
ZNF214	4.490918977	0.005564377	4.648569188
FAM81B	4.489679283	0.005564377	6.400093324
RFC4	4.483839099	0.005564377	3.070013465
LXN	4.390568417	0.007088532	5.837537297
TMEM45A	4.34584376	0.008336618	5.39047842
SERTAD1	4.26750286	0.010557947	4.280850656
TRAPPC3L	4.226288993	0.011080348	3.953676444
LRR1	4.225710625	0.011080348	5.552844685
COA7	4.126957551	0.01621255	3.058929186
ISCA1	4.115925568	0.01635633	0.823947844
PXDN	4.101962651	0.016879179	5.556384233
SLC19A1	4.034934417	0.020727361	3.595950283
HMBS	4.027062776	0.020883811	1.749048141
DNASE1L2	3.984186672	0.023816328	4.88392171
PLAC8L1	3.96202474	0.025517217	5.419295647
PKN1	3.887659734	0.03317334	4.131028944
PCP4	3.850034828	0.036229888	1.604639619
BTNL2	3.846652111	0.036229888	7.280687351
TRPC5	3.844922729	0.036229888	3.062732979
CCDC13	3.834307566	0.036509351	4.580981693
KAZALD1	3.833011609	0.036509351	3.983695608
POP1	3.815898698	0.038367825	2.783177243
TSPAN19	3.808240982	0.038814356	5.19775279
GLIPR1L1	3.793612012	0.039308642	6.188219356
ACTL6A	3.791204343	0.039308642	2.545121599
HAUS4	3.783919689	0.039645698	6.45429894
MEST	3.780200106	0.039645698	0.87105489
OTOS	3.724709793	0.044508045	4.705125358
LY75	3.71508938	0.044508045	7.822715436
CLEC2D	3.710130381	0.044508045	2.827934371
PTX3	3.702145563	0.044508045	2.607826606

Table E.3 continued from previous page

Gene Name	stat	adj p value	LogFC
MEIOB	3.700394374	0.044508045	6.657184626
ATF3	3.699852224	0.044508045	5.984276487
MAGOHB	3.699477856	0.044508045	2.674522667
F12	3.688981062	0.045731627	5.366241712
STARD9	3.682246236	0.04630502	3.050096395
OSCAR	3.675607085	0.046418987	3.976913929
CAPN2	3.66906524	0.046418987	1.122723011
SLC7A6OS	3.667814723	0.046418987	4.882182412
B3GNT7	3.651298809	0.047753895	7.466996076
FANCL	3.647394446	0.047753895	4.651270662
POMC	3.612909168	0.052895327	6.485064438
PPIP5K2	3.575562662	0.05727029	1.117095062
PHF5A	3.539158897	0.064318326	3.776945112
ZFPM2	3.523399105	0.066783224	1.060524279
TGIF1	3.51065342	0.068577497	5.971504628
MOV10L1	3.498843345	0.07028254	5.216781588
TMEM41B	3.485851706	0.072436698	0.964283078
RGS1	3.480087366	0.07301546	6.82463269
LIPI	3.476501066	0.073258999	5.964732102
PAQR5	3.464484283	0.073720036	7.027361092
IL1RAP	3.463603203	0.073720036	3.042441861
PLK3	3.461555697	0.073720036	3.148913744
EFHB	3.455587437	0.073720036	4.90669971
ERGIC1	3.453307756	0.073720036	2.512580399
MTX3	3.452852531	0.073720036	0.908234719
ZNF180	3.450087587	0.073720036	2.978828847
CIDEA	3.442427126	0.073720036	2.145075825
MELTF	3.438110081	0.073720036	6.09121715
SLC22A31	3.434456542	0.073720036	5.127216149
ZNF121	3.434001992	0.073720036	7.193636666
CHMP7	3.432018498	0.073720036	1.440916978
KIF26B	3.419654967	0.075350456	2.434093195
SHCBP1	3.398847909	0.078069913	5.441541803
TACR1	3.393143952	0.078458549	2.273527835
P2RY12	3.388121848	0.079285074	7.357358336
DDIT3	3.375153661	0.081681778	0.916874187
CDC20B	3.364752739	0.08244968	6.536918367
PARM1	3.356842596	0.084211198	1.121577941
ZNF804A	3.34663112	0.085480983	1.060000918
GPN1	3.341229786	0.086534467	1.612016027
DHX8	3.32112786	0.089785201	1.349145272
PIGB	3.298276911	0.095188725	3.101292895

Table E.3 continued from previous page

Gene Name	stat	adj p value	LogFC
MOB1B	3.291130932	0.096271406	1.784971087
TIGD3	3.268670626	0.099251389	4.545932308
KLHDC4	-3.269211095	0.099251389	-2.217427121
NOL4	-3.269331515	0.099251389	-1.360572392
ZGRF1	-3.276765091	0.098328456	-2.34068606
NUMA1	-3.277713948	0.098328456	-1.211182209
CYB5RL	-3.27865336	0.098328456	-5.134437928
ZPBP	-3.28288274	0.098130986	-4.864342137
DDR1	-3.290142429	0.096271406	-2.823717197
ZNF217	-3.297095176	0.095188725	-6.22115883
C1D	-3.302177065	0.094761268	-1.548841358
KCNT2	-3.303207016	0.094761268	-4.451865134
RC3H2	-3.321217766	0.089785201	-1.550379944
ZNF782	-3.321796618	0.089785201	-2.384435578
NEXN	-3.326404832	0.089785201	-1.567429539
PEA15	-3.327263111	0.089785201	-1.340480543
AFF3	-3.348608777	0.085480983	-1.038668712
CDK19	-3.354818815	0.084211198	-2.477155154
SUN3	-3.367650665	0.082201424	-6.741982057
EFHD1	-3.371478296	0.081681778	-3.114455791
DESI2	-3.372162674	0.081681778	-1.853925174
HOMER3	-3.385658195	0.07938008	-3.813102375
FANCD2	-3.393395948	0.078458549	-3.301297196
ANKRA2	-3.40338227	0.077404822	-2.747483936
SLC47A1	-3.404708374	0.077404822	-0.955697482
MT1M	-3.413225394	0.075886165	-3.669932719
HSPA2	-3.414167598	0.075886165	-2.699734772
ZNF704	-3.421259929	0.075350456	-2.179775139
TADA2A	-3.430164669	0.073720036	-1.712065829
EMP2	-3.431805914	0.073720036	-5.004852496
FSIP1	-3.432665304	0.073720036	-2.719922888
CDK13	-3.438053961	0.073720036	-1.048051458
SEC11A	-3.463473458	0.073720036	-1.640524621
ALCAM	-3.470172183	0.073720036	-2.463587483
PLEKHG3	-3.48493636	0.072436698	-2.913421819
SYF2	-3.498508623	0.07028254	-1.001653484
ZNF451	-3.51832643	0.067341077	-0.857173342
C14orf132	-3.52678794	0.066658856	-1.596806673
ZMYM3	-3.546891121	0.063161655	-1.154749688
NCOA3	-3.59107458	0.054587194	-0.936356398
HTR2C	-3.59599921	0.054187541	-6.196305856
DET1	-3.596776508	0.054187541	-1.944747843

Table E.3 continued from previous page

Gene Name	stat	adj p value	LogFC
JAM2	-3.605321676	0.053523629	-3.30508183
CHDH	-3.611493416	0.052895327	-6.390504747
TNIP3	-3.628722238	0.050709843	-5.165024904
UROS	-3.650320215	0.047753895	-1.206213867
GRIA1	-3.658983505	0.047424046	-2.137001456
GPC4	-3.668881427	0.046418987	-3.482402578
NABP2	-3.709604277	0.044508045	-0.846545263
FHL5	-3.728130047	0.044508045	-5.835218897
UACA	-3.732209834	0.044508045	-3.40234674
ZSCAN29	-3.736794401	0.044508045	-2.159574877
ASXL2	-3.737574008	0.044508045	-1.307048349
HEG1	-3.76020103	0.042212969	-4.075209576
NR4A2	-3.791821252	0.039308642	-4.471342349
SLC39A12	-3.880750992	0.033371609	-4.102839113
GAB1	-3.937171032	0.027650706	-3.485748789
ISOC1	-3.984615459	0.023816328	-4.599503321
OVGP1	-4.048190274	0.020117365	-4.330057774
ISLR	-4.074261273	0.018491182	-3.352755099
PLPP3	-4.124837523	0.01621255	-3.022195774
ACBD7	-4.243819777	0.010927481	-2.702299393
ART4	-4.257764001	0.010634548	-7.661299921
MAOB	-4.313768188	0.00890193	-1.376492377
REPIN1	-4.332886581	0.008489178	-1.570810285
TUBA3D	-4.392734516	0.007088532	-6.549294026
NANOS1	-4.44498479	0.006036087	-2.405326651
CACNA2D2	-4.456828573	0.005997771	-1.42000158
FAM107A	-4.56835073	0.004726021	-2.23507245
DDX6	-4.647171037	0.004519909	-1.00876927
FAM184B	-5.108283027	0.001148213	-2.475170235

Table E.3: Full Ranked-List of DEGs from Betz Cell (ALS vs. Control) Comparison. Positive LogFC: DEG is downregulated in ALS Betz cells; negative LOG FC: DEG is upregulated in ALS Betz cells.

E.5 Microdissected Betz Cells / Allen Institute Clusters - Comparison

The aforementioned Allen Institute has recently produced a body of work that extensively studied the electrophysiology, morphology, epigenome and transcriptome of individual cells in the motor cortex of the mouse, non-human primate, and human brain [163]. Within this study, the authors found cells with Betz cell-like morphology in their L3-5 Exc FEZF2 ASGR2 and L5 Exc FEZF2 CSN1S1 transcriptomic clusters. To validate these findings, the cluster-specific marker genes (CSN1S1 and ASGR2) were immunohistochemically tested on our FFPE Betz cell material. Immuno-staining confirmed the presence of both ASGR2 and CSN1S1 in Betz cells, however, it is important to note that these proteins were also expressed in other cell-types in the motor cortex.

The Betz cell containing clusters (L3-5 Exc FEZF2 ASGR2 and L5 Exc FEZF2 CSN1S1), along with the other 128 clusters identified in their study were compared to the mRNA expression data of our microdissected Betz cell data. In total, ≈8,000 genes were shared between the LCM Betz cells and Betz containing snRNA-seq clusters. Promisingly, unbiased clustering revealed that the microdissected Betz samples were most similar to a group of transcriptomic clusters that contained ET neurons i.e. L3-5 Exc FEZF2 ASGR2 (PCC: 0.643) and Exc FEZF2 CSN1S1 (PCC: 0.684). To investigate how the underlying molecular pathways of these clusters compared to microdissected Betz cells the above clusters were analysed using 2D annotation enrichment - See Figure E.32. Overall, the combined abundance of individual transcripts and their associated molecular pathways were similar among all 3 data sets, with the majority of gene ontology terms having concordant expression. Among the most correlated pathways were those related to glutamate signalling, generation of an action potential, and organelle transport along microtubules. However, there was some anti-correlating behaviour between the microdissected Betz cells and above transcriptomic clusters, notably, terms relating to proteasomes, ribosomes, and mitochondria. However, the increase in these organelle related

pathways in microdissected Betz but decrease in the Betz-containing clustered may be a reflection on the data bias when comparing gene expression data from different source material i.e. single nuclei vs. cell slices, the latter, of course, containing intact cytoplasm.

The results from the Allen Institute highlighted the conserved physiology features of primate ET neurons. Principally, they highlighted an increase in axon guidance related genes across evolutionary distance to humans, particularly those from the ROBO, SLIT, and EPHRIN families. Within our data, 6 of the 14 axon guidance proteins identified by the Allen Institute to be conserved showed a significant enrichment within Betz cells compared to L5 samples e.g. GFRA1, IRS2, PRKCA, ROBO2, SEMA3D and SLIT3. Similarly, the results from this paper highlighted that potassium and calcium channels were enriched within primate ET neurons compared to IT neurons. When investigating the channel protein diversity among Betz cells several voltage-dependent chloride, sodium, and potassium channels were differentially expressed when compared to L5 and Purkinje cells, particularly KCNG1, SCN4A, KCNN4, KCNQ4, CLCN2, AND CLCN7. However, it is important to note that neurons within L5 were captured without bias, so the resulting data is likely to reflect a homogeneous mix of ET and IT neurons.

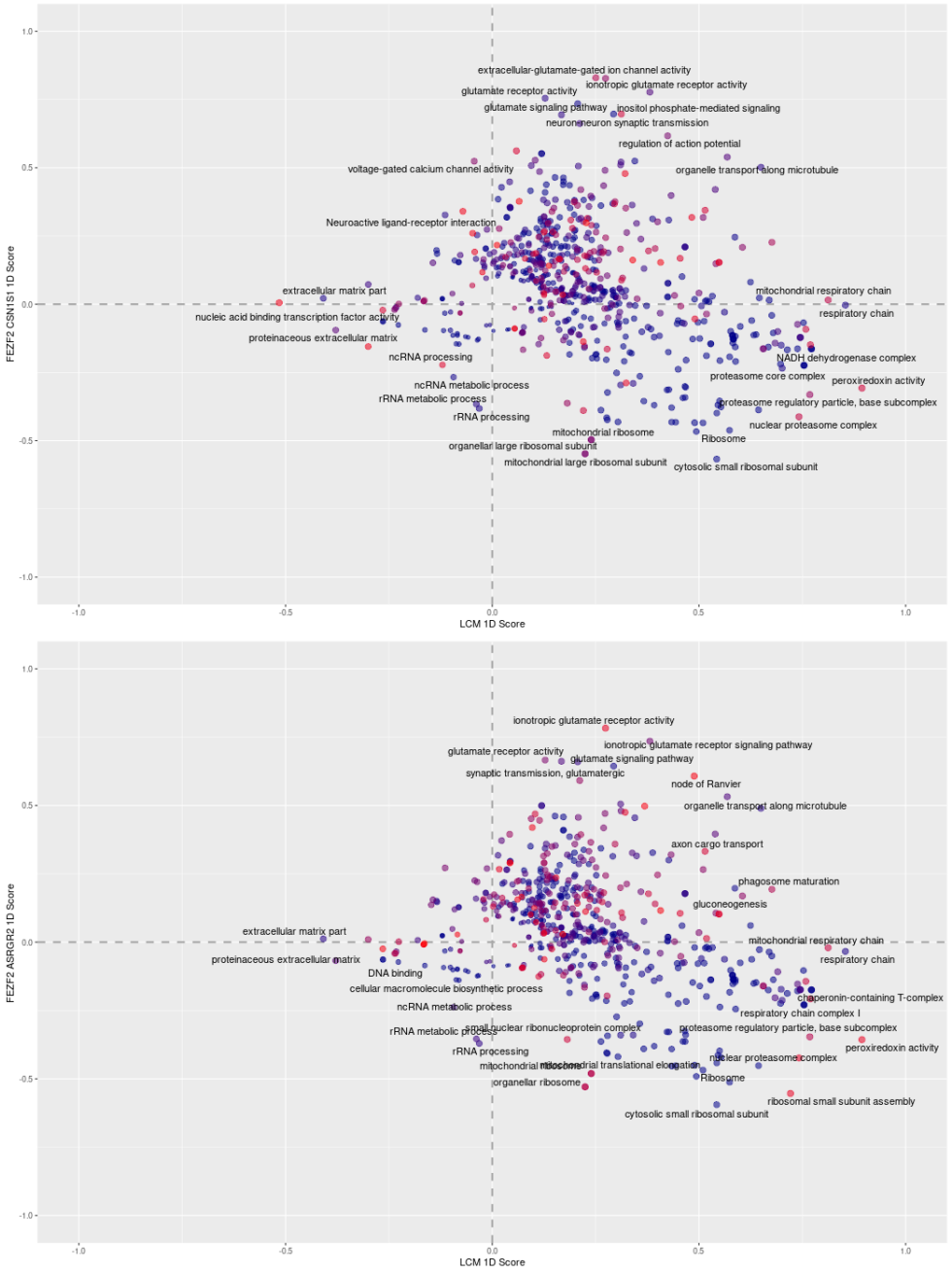


Figure E.32: 2D Annotation of Allen Institute and Microdissected Betz Cells. Comparing the distribution of mRNA of microdissected Betz cells to the Allen Institute's Betz-containing clusters. Overall, there was a strong correlation in protein distribution between microdissected Betz cells and the Allen Institute's CSN1S1 (top panel) and ASRGR2 (bottom panel) clusters. Microdissected Betz cells contained a higher amount of mitochondrial, ribosomal, and proteasomal transcripts. Although, it is not clear if this reflects the different transciptomic depths between the samples or the fact that some transcripts were depleted by the virtue of isolated nuclei being analysed rather than cell slices. Each point represents a significantly enriched GO term. Size of point: number of proteins involved in that term. Colour: Darker colours (Blue) have smaller p values.

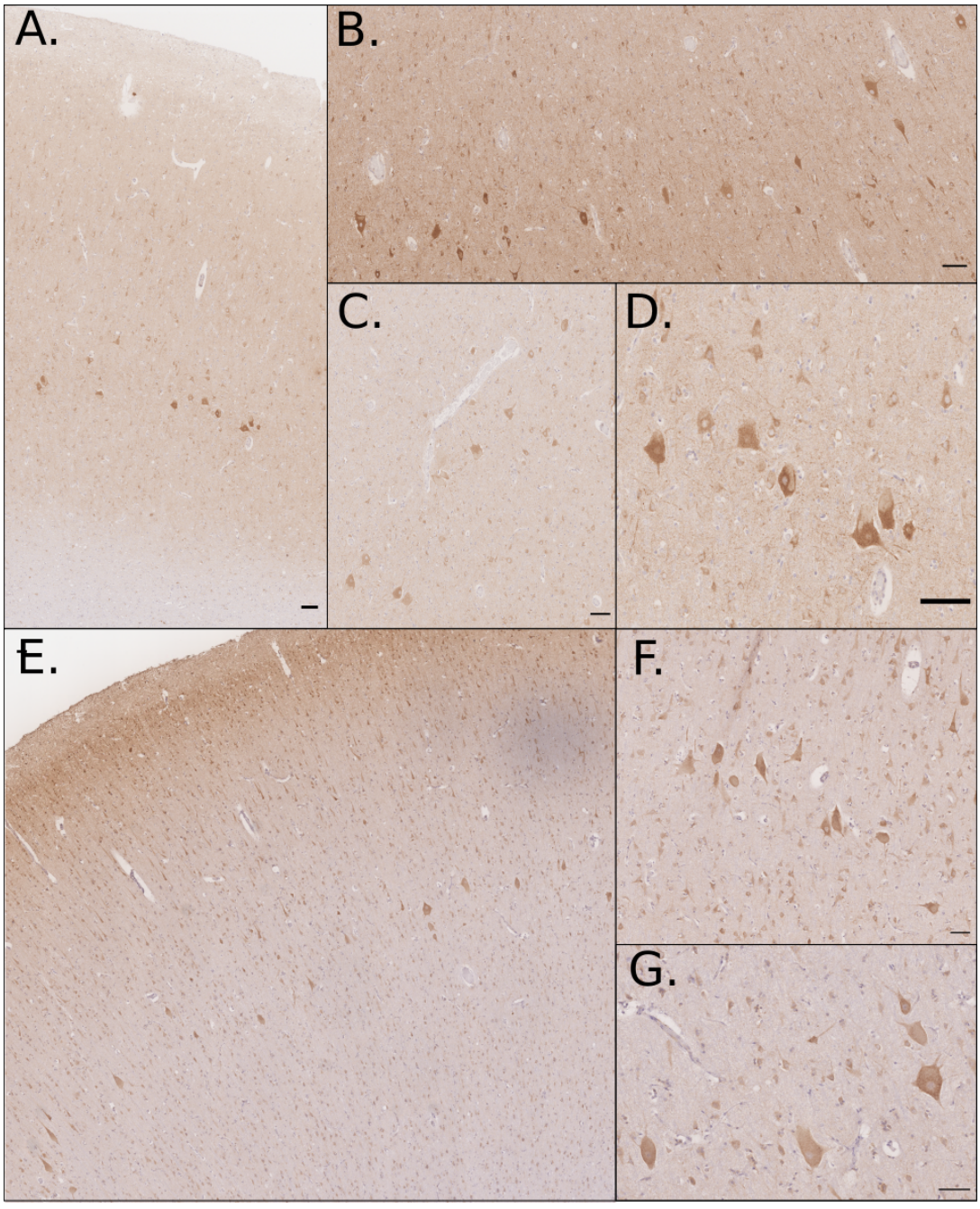


Figure E.33: ASGR2 & CSN1S1 IHC on Betz Cells.
Both AGSR2 (A-D) and CSN1S1 (E-G) showed strong immunoreactivity in Betz cells. Although surrounding non-Betz neurons also stained positively.

F

Betz-Purkinje Research Cohort

Contents

F.1 Full Post-Mortem Cohort used in this Thesis 365
F.2 Additional Information on Cohort 368

F.1 Full Post-Mortem Cohort used in this Thesis

See overleaf.

Disease State	Cell Type Microdissected	Chapter 4 - RWS Development	Chapter 5 - LCM Proteomics	Chapter 6 - LCM-seq	Chapter 6 - LCM Proteomics
Control_1	Betz cell	-	Yes	Yes	Yes
	L5 pyramidal neuron	-	Yes	Yes	Yes
	Purkinje cell	Yes	Yes	Yes	Yes
	Cerebellar Granule cells	Yes	-	-	-
	DNNs	-	Yes	-	-
Control_2	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
Control_3	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
Control_4	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
Control_5	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
Control _6	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
Control_7	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
ALS_1	Purkinje cell	Yes	-	-	-
	Molecular Layer	Yes	-	-	-
ALS_2	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes

Table F.1 continued from previous page

Disease State	Cell Type Microdissected	Chapter 4 - RWS Development	Chapter 5 - LCM Proteomics	Chapter 6 - LCM-seq	Chapter 6 - LCM Proteomics
	Purkinje cell	-	-	Yes	Yes
ALS_3	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
ALS_4	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
ALS_5	Betz cell	-	-	Yes	-
	L5 pyramidal neuron	-	-	Yes	-
	Purkinje cell	-	-	Yes	-
ALS_6	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
ALS_7	Betz cell	-	-	Yes	Yes
	L5 pyramidal neuron	-	-	Yes	Yes
	Purkinje cell	-	-	Yes	Yes
ALS_8	Betz cell	-	-	-	Yes
	L5 pyramidal neuron	-	-	-	Yes
	Purkinje cell	-	-	-	Yes

Table F.1: Full List of Post-Mortem Cases used in this Thesis.

F.2 Additional Information on Cohort

Case Name	Cell Type Microdissected	Age	Sex	PMI (Hours)	pH	cDNA Size (bp)
Control_1	Betz cell	55	M	18	6.05	469
	L5 pyramidal neuron					439
	Purkinje cell					511
Control_2	Betz cell	57	M	46	5.91	446
	L5 pyramidal neuron					504
	Purkinje cell					542
Control_3	Betz cell	56	M	49	5.97	370
	L5 pyramidal neuron					397
	Purkinje cell					477
Control_4	Betz cell	63	M	83	6.05	466
	L5 pyramidal neuron					478
	Purkinje cell					338
Control_5	Betz cell	70	F	34	5.63	423
	L5 pyramidal neuron					462
	Purkinje cell					362
Control _6	Betz cell	58	F	59	6	420
	L5 pyramidal neuron					496
	Purkinje cell					437
Control_7	Betz cell	38	F	41	6.06	459
	L5 pyramidal neuron					369
	Purkinje cell					418
ALS_1	-	54	M	72	-	-
ALS_2	Betz cell	59	M	22	6.17	387
	L5 pyramidal neuron					506
	Purkinje cell					561
ALS_3	Betz cell	86	M	36	5.78	454
	L5 pyramidal neuron					436
	Purkinje cell					520
ALS_4	Betz cell	76	F	23	5.7	377
	L5 pyramidal neuron					403
	Purkinje cell					309
ALS_5	Betz cell	69	F	108	6.15	422
	L5 pyramidal neuron					430
	Purkinje cell					359
ALS_6	Betz cell	75	F	29	5.72	383
	L5 pyramidal neuron					455
	Purkinje cell					367
ALS_7	Betz cell	60	M	45	6.14	360
	L5 pyramidal neuron					455
	Purkinje cell					375
ALS_8	-	77	F	31	5.46	-

Table F.2: Demographics of Post-Mortem Cohort used in this Thesis.}{PMI: post-mortem interval, cDNA size: average fragment size of generated cDNA from microdissected neurons; a proxy for RNA quality.

Bibliography

[1] Robert Hooke. *Micrographia, or, Some physiological descriptions of minute bodies made by magnifying glasses: with observations and inquiries thereupon.* 1665.

[2] Suzana Herculano-Houzel et al. "Cellular scaling rules for primate brains". In: *Proceedings of the National Academy of Sciences of the United States of America* 104.9 (2007), pp. 3562–3567.

[3] Suzana Herculano-Houzel. "The human brain in numbers: A linearly scaled-up primate brain". In: *Frontiers in Human Neuroscience* 3 (Nov. 2009).

[4] Suzana Herculano-Houzel, Paul R. Manger, and Jon H. Kaas. "Brain scaling in mammalian evolution as a consequence of concerted and mosaic changes in numbers of neurons and average neuronal cell size". In: *Frontiers in Neuroanatomy* 8 (Aug. 2014), p. 77.

[5] Frederico A.C. Azevedo et al. "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain". In: *Journal of Comparative Neurology* 513.5 (2009), pp. 532–541.

[6] Ghouse Ahmed Shariff. "Cell counts in the primate cerebral cortex". In: *Journal of Comparative Neurology* 98.3 (1953), pp. 381–400.

[7] D. P. Pelvig et al. "Neocortical glial cell numbers in human brains". In: *Neurobiology of Aging* 29.11 (2008), pp. 1754–1762.

[8] Birgitte Bo Andersen, Lise Korbo, and Bente Pakkenberg. "A quantitative study of the human cerebellum with unbiased stereological techniques". In: *Journal of Comparative Neurology* 326.4 (1992), pp. 549–560.

[9] Michel A. Hofman. "Evolution of the human brain: When bigger is better". In: *Frontiers in Neuroanatomy* 8 (2014), pp. 1–12.

[10] Dean Falk. *Evolution of the Primate Brain.* 2013, pp. 1–28.

[11] Katrin Amunts and Karl Zilles. "Architectonic Mapping of the Human Brain beyond Brodmann". In: *Neuron* 88.6 (Dec. 2015), pp. 1086–1107.

[12] Matthew F. Glasser et al. "The Human Connectome Project's neuroimaging approach". In: *Nature Neuroscience* 19.9 (2016), pp. 1175–1187.

[13] Bruce Fischl and Martin I. Sereno. "Microstructural parcellation of the human brain". In: *NeuroImage* 182.3 (2018), pp. 219–231.

[14] André M.M. Sousa et al. "Molecular and cellular reorganization of neural circuits in the human lineage". In: *Science* 358.6366 (Nov. 2017), pp. 1027–1032.

[15] Hongjun Fu, John Hardy, and Karen E. Duff. "Selective vulnerability in neurodegenerative diseases". In: *Nature Neuroscience* 21.10 (2018), pp. 1350–1358.

[16] Rodney J. Douglas and Kevan A.C. Martin. “Neuronal circuits of the neocortex”. In: *Annual Review of Neuroscience* 27 (2004), pp. 419–451.

[17] Luciano Custo Greig et al. “Molecular logic of neocortical projection neuron specification, development and diversity”. In: *Nature Reviews Neuroscience* 14.11 (2013), pp. 755–769.

[18] Bradley J. Molyneaux et al. “Neuronal subtype specification in the cerebral cortex”. In: *Nature Reviews Neuroscience* 8.6 (2007), pp. 427–437.

[19] Kenneth D. Harris and Gordon M.G. Shepherd. “The neocortical circuit: Themes and variations”. In: *Nature Neuroscience* 18.2 (2015), pp. 170–181.

[20] Gord Fishell and Bernardo Rudy. “Mechanisms of inhibition within the telencephalon: "Where the wild things are"”. In: *Annual Review of Neuroscience* 34 (2011), pp. 535–567.

[21] Adam Kepecs and Gordon Fishell. “Interneuron cell types are fit to function”. In: *Nature* 505.7483 (2014), pp. 318–326.

[22] Bernardo Rudy et al. “Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons”. In: *Developmental Neurobiology* 71.1 (2011), pp. 45–61.

[23] Robin Tremblay, Soohyun Lee, and Bernardo Rudy. “GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits”. In: *Neuron* 91.2 (2016), pp. 260–292.

[24] Michael N. Economo et al. “A platform for brain-wide imaging and reconstruction of individual neurons”. In: *eLife* 5 (2016).

[25] Euiseok J. Kim et al. “Three Types of Cortical Layer 5 Neurons That Differ in Brain-wide Connectivity and Function”. In: *Neuron* 88.6 (2015), pp. 1253–1267.

[26] Takako Kita and Hitoshi Kita. “The subthalamic nucleus is one of multiple innervation sites for long-range corticofugal axons: A single-axon tracing study in the rat”. In: *Journal of Neuroscience* 32.17 (2012), pp. 5990–5999.

[27] Mateo Vélez-Fort et al. “The stimulus selectivity and connectivity of layer six principal cells reveals cortical microcircuits underlying visual processing”. In: *Neuron* 83.6 (2014), pp. 1431–1443.

[28] Henry Markram et al. “Reconstruction and Simulation of Neocortical Microcircuitry”. In: *Cell* 163.2 (2015), pp. 456–492.

[29] Bosiljka Tasic et al. “Adult mouse cortical cell taxonomy revealed by single cell transcriptomics”. In: *Nature Neuroscience* 19.2 (2016), pp. 335–346.

[30] Rebecca D. Hodge et al. “Conserved cell types with divergent features in human versus mouse cortex”. In: *Nature* 573.7772 (Sept. 2019), pp. 61–68.

[31] Dimitry Ofengeim et al. “Single-Cell RNA Sequencing: Unraveling the Brain One Cell at a Time”. In: *Trends in Molecular Medicine* 23.6 (2017), pp. 563–576.

[32] Marcus Gry et al. “Correlations between RNA and protein expression profiles in 23 human cell lines”. In: *BMC Genomics* 10 (2009).

[33] Bruno Domon and Ruedi Aebersold. “Mass spectrometry and protein analysis”. In: *Science* 312.5771 (2006), pp. 212–217.

[34] Jürgen Cox and Matthias Mann. "Is Proteomics the New Genomics?" In: *Cell* 130.3 (2007), pp. 395–398.

[35] Andreas Wanninger. "Morphology is dead - long live morphology! Integrating MorphoEvoDevo into molecular EvoDevo and phylogenomics". In: *Frontiers in Ecology and Evolution* 3 (2015).

[36] Oliver Hobert, Inés Carrera, and Nikolaos Stefanakis. "The molecular and gene regulatory signature of a neuron". In: *Trends in Neurosciences* 33.10 (2010), pp. 435–445.

[37] Ed S. Lein et al. "Genome-wide atlas of gene expression in the adult mouse brain". In: *Nature* 445.7124 (Jan. 2007), pp. 168–176.

[38] Graciana Diez-Roux et al. "A high-resolution anatomical atlas of the transcriptome in the mouse embryo". In: *PLoS Biology* 9.1 (2011).

[39] Mike Hawrylycz et al. "Multi-scale correlation structure of gene expression in the brain". In: *Neural Networks* 24.9 (Nov. 2011), pp. 933–942.

[40] Hongkui Zeng et al. "Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures". In: *Cell* 149.2 (Apr. 2012), pp. 483–496.

[41] David Lähnemann et al. "Eleven grand challenges in single-cell data science". In: *Genome Biology* 21.1 (2020).

[42] Guo Cheng Yuan et al. "Challenges and emerging directions in single-cell analysis". In: *Genome Biology* 18.1 (2017).

[43] Jason M. Keil, Adel Qalieh, and Kenneth Y. Kwan. "Brain transcriptome databases: A user's guide". In: *Journal of Neuroscience* 38.10 (Mar. 2018), pp. 2399–2412.

[44] Hyo Jung Kang et al. "Spatio-temporal transcriptome of the human brain". In: *Nature* 478.7370 (Oct. 2011), pp. 483–489.

[45] Fuchou Tang et al. "mRNA-Seq whole-transcriptome analysis of a single cell". In: *Nature Methods* 6.5 (2009), pp. 377–382.

[46] Amit Zeisel et al. "Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq". In: *Science* 347.6226 (2015), pp. 1138–1142.

[47] Dmitry Usoskin et al. "Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing". In: *Nature Neuroscience* 18.1 (2015), pp. 145–153.

[48] Peng Hu et al. "Dissecting Cell-Type Composition and Activity-Dependent Transcriptional State in Mammalian Brains by Massively Parallel Single-Nucleus RNA-Seq". In: *Molecular Cell* 68.5 (2017), pp. 1006–1015.

[49] Amit Zeisel et al. "Molecular Architecture of the Mouse Nervous System". In: *Cell* 174.4 (2018), pp. 999–1014.

[50] Arpiar Saunders et al. "Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain". In: *Cell* 174.4 (2018), pp. 1015–1030.

[51] Spyros Darmanis et al. "A survey of human brain transcriptome diversity at the single cell level". In: *Proceedings of the National Academy of Sciences of the United States of America* 112.23 (2015), pp. 7285–7290.

[52] Gioele La Manno et al. “Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells”. In: *Cell* 167.2 (2016), pp. 566–580.

[53] Alex A. Pollen et al. “Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex”. In: *Nature Biotechnology* 32.10 (2014), pp. 1053–1058.

[54] Alex A. Pollen et al. “Molecular Identity of Human Outer Radial Glia during Cortical Development”. In: *Cell* 163.1 (2015), pp. 55–67.

[55] Blue B. Lake et al. “Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain”. In: *Science* 352.6293 (June 2016), pp. 1586–1590.

[56] Naomi Habib et al. “Massively parallel single-nucleus RNA-seq with DroNc-seq”. In: *Nature Methods* 14.10 (2017), pp. 955–958.

[57] Tomasz J. Nowakowski et al. “Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex”. In: *Science* 358.6368 (2017), pp. 1318–1323.

[58] Suijuan Zhong et al. “A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex”. In: *Nature* 555.7697 (2018), pp. 524–528.

[59] Xiaoying Fan et al. “Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis”. In: *Cell Research* 28.7 (2018), pp. 730–745.

[60] Damon Polioudakis et al. “A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation”. In: *Neuron* 103.5 (2019), pp. 785–801.

[61] Hansruedi Mathys et al. “Single-cell transcriptomic analysis of Alzheimer’s disease”. In: *Nature* 570.7761 (2019), pp. 332–337.

[62] Lucas Schirmer et al. “Neuronal vulnerability and multilineage diversity in multiple sclerosis”. In: *Nature* 573.7772 (2019), pp. 75–82.

[63] Volker Hovestadt et al. “Resolving medulloblastoma cellular architecture by single-cell genomics”. In: *Nature* 572.7767 (2019), pp. 74–79.

[64] Itay Tirosh et al. “Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma”. In: *Nature* 539.7628 (2016), pp. 309–313.

[65] Charmaine Lang et al. “Single-Cell Sequencing of iPSC-Dopamine Neurons Reconstructs Disease Progression and Identifies HDAC4 as a Regulator of Parkinson Cell Phenotypes”. In: *Cell Stem Cell* 24.1 (2019), pp. 93–106.

[66] Fabian Hosp and Matthias Mann. “A Primer on Concepts and Applications of Proteomics in Neuroscience”. In: *Neuron* 96.3 (2017), pp. 558–571.

[67] Dorte B. Bekker-Jensen et al. “An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes”. In: *Cell Systems* 4.6 (June 2017), pp. 587–599.

[68] Simon Davis et al. “Expanding Proteome Coverage with CHarge Ordered Parallel Ion aNalysis (CHOPIN) Combined with Broad Specificity Proteolysis”. In: *Journal of Proteome Research* 16.3 (Mar. 2017), pp. 1288–1299.

[69] Robert R. Kitchen et al. “Decoding neuroproteomics: Integrating the genome, translatome and functional anatomy”. In: *Nature Neuroscience* 17.11 (2014), pp. 1491–1499.

[70] Kirti Sharma et al. “Cell type- and brain region-resolved mouse brain proteome”. In: *Nature Neuroscience* 18.12 (2015), pp. 1819–1831.

[71] Steven L. Salzberg. “Open questions: How many genes do we have?” In: *BMC Biology* 16.1 (2018), pp. 10–12.

[72] Ole Nørregaard Jensen. “Modification-specific proteomics: Characterization of post-translational modifications by mass spectrometry”. In: *Current Opinion in Chemical Biology* 8.1 (2004), pp. 33–41.

[73] Fabio Sambataro and Maria Pennuto. “Post-translational modifications and protein quality control in motor neuron and polyglutamine diseases”. In: *Frontiers in Molecular Neuroscience* 10 (2017), p. 82.

[74] Yansheng Liu, Andreas Beyer, and Ruedi Aebersold. “On the Dependency of Cellular Protein Levels on mRNA Abundance”. In: *Cell* 165.3 (2016), pp. 535–550.

[75] Yanfang Guo et al. “How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes”. In: *Acta Biochimica et Biophysica Sinica* 40.5 (2008), pp. 426–436.

[76] John C. Rogers and Ryan D. Bomgarden. “Sample preparation for mass spectrometry-based proteomics; from proteomes to peptides”. In: *Advances in Experimental Medicine and Biology* 919 (2016), pp. 43–62.

[77] Evelina Sjöstedt et al. “Defining the human brain proteome using transcriptomics and antibody-based profiling with a focus on the cerebral cortex”. In: *PLoS ONE* 10.6 (2015).

[78] Mathias Uhlén et al. “Tissue-based map of the human proteome”. In: *Science* 347.6220 (2015).

[79] Linn Fagerberg et al. “Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics”. In: *Molecular and Cellular Proteomics* 13.2 (2014), pp. 397–406.

[80] Erik C.B. Johnson et al. “Deep proteomic network analysis of Alzheimer’s disease brain reveals alterations in RNA binding proteins and RNA splicing associated with disease”. In: *Molecular Neurodegeneration* 13.1 (2018).

[81] Lingyan Ping et al. “Global quantitative analysis of the human brain proteome in Alzheimer’s and Parkinson’s Disease”. In: *Scientific Data* 5 (2018).

[82] Nicholas T. Seyfried et al. “A Multi-network Approach Identifies Protein-Specific Co-expression in Asymptomatic and Symptomatic Alzheimer’s Disease”. In: *Cell Systems* 4.1 (Jan. 2017), pp. 60–72.

[83] Measho H. Abreha et al. “Quantitative Analysis of the Brain Ubiquitylome in Alzheimer’s Disease”. In: *Proteomics* 18.20 (2018).

[84] Fuyuki Kametani et al. “Mass spectrometric analysis of accumulated TDP-43 in amyotrophic lateral sclerosis brains”. In: *Scientific Reports* 6 (2016), pp. 1–15.

[85] Justin McKetney et al. “Proteomic Atlas of the Human Brain in Alzheimer’s Disease”. In: *Journal of Proteome Research* 18.3 (2019), pp. 1380–1391.

[86] Ying Zhu et al. “Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells”. In: *Nature Communications* 9.1 (2018).

[87] Harrison Specht and Nikolai Slavov. “Transformative Opportunities for Single-Cell Proteomics”. In: *Journal of Proteome Research* 17.8 (2018), pp. 2565–2571.

[88] Xuran Wang et al. “Bulk tissue cell type deconvolution with multi-subject single-cell expression reference”. In: *Nature Communications* 10.1 (2019).

[89] Yong Wang and Nicholas E. Navin. “Advances and Applications of Single-Cell Sequencing Technologies”. In: *Molecular Cell* 58.4 (2015), pp. 598–609.

[90] Lingxue Zhu et al. “A unified statistical framework for single cell and bulk rna sequencing data”. In: *Annals of Applied Statistics* 12.1 (2018), pp. 609–632.

[91] Priyojit Das and K. A.Abdul Nazeer. “A novel clustering method to identify cell types from single cell transcriptional profiles”. In: *Procedia Computer Science* 132 (Jan. 2018), pp. 983–992.

[92] Yuchen Yang et al. “SAFE-clustering: Single-cell Aggregated (from Ensemble) clustering for single-cell RNA-seq data”. In: *Bioinformatics* 35.8 (Sept. 2019). Ed. by Inanc Birol, pp. 1269–1277.

[93] Ruiqing Zheng et al. “SinNLRR: A robust subspace clustering method for cell type detection by non-negative and low-rank representation”. In: *Bioinformatics* 35.19 (Mar. 2019), pp. 3642–3650.

[94] Suhas Srinivasan, Nathan Johnson, and Dmitry Korkin. “A Hybrid Deep Clustering Approach for Robust Cell Type Profiling Using Single-cell RNA-seq Data”. In: *bioRxiv* (Jan. 2019), pp. 511–626.

[95] Ping Hu et al. “Single cell isolation and analysis”. In: *Frontiers in Cell and Developmental Biology* 4.OCT (2016).

[96] Sherrif F. Ibrahim and Ger Van Den Engh. “Flow cytometry and cell sorting”. In: *Advances in Biochemical Engineering/Biotechnology* 106 (2007), pp. 19–39.

[97] Sreemanti Basu et al. “Purification of specific cell population by fluorescence activated cell sorting (FACS)”. In: *Journal of Visualized Experiments* 41 (2010).

[98] Laura E. Mickelsen et al. “Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons”. In: *Nature Neuroscience* 22.4 (2019), pp. 642–656.

[99] Claire Bénédicte Rivara et al. “Stereologic characterization and spatial distribution patterns of Betz cells in the human primary motor cortex”. In: *Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology* 270.2 (Feb. 2003), pp. 137–151.

[100] C. Robert Bagnell. “Laser capture microdissection”. In: *Molecular Diagnostics: For the Clinical Laboratorian* (2005), pp. 219–224.

[101] Alfred Vogel et al. “Mechanisms of laser-induced dissection and transport of histologic specimens”. In: *Biophysical Journal* 93.12 (2007), pp. 4481–4500.

[102] Carlo Colantuoni et al. “Temporal dynamics and genetic control of transcription in the human prefrontal cortex”. In: *Nature* 478.7370 (Oct. 2011), pp. 519–523.

[103] Sofia Fertuzinhos et al. “Laminar and temporal expression dynamics of coding and noncoding RNAs in the mouse neocortex”. In: *Cell Reports* 6.5 (Mar. 2014), pp. 938–950.

[104] Zhisong He et al. “Comprehensive transcriptome analysis of neocortical layers in humans, chimpanzees and macaques”. In: *Nature Neuroscience* 20.6 (Apr. 2017), pp. 886–895.

[105] Mark S. Cembrowski et al. “Hipposeq: A comprehensive RNA-seq database of gene expression in hippocampal principal neurons”. In: *eLife* 5 (Apr. 2016).

[106] Eva Hedlund et al. “Global gene expression profiling of somatic motor neuron populations with different vulnerability identify molecules and pathways of degeneration and protection”. In: *Brain* 133.8 (2010), pp. 2313–2330.

[107] Chee Yeun Chung et al. “Cell type-specific gene expression of midbrain dopaminergic neurons reveals molecules involved in their vulnerability and protection”. In: 14.13 (2005), pp. 1709–1725.

[108] Filip Simunovic et al. “Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson’s disease pathology”. In: *Brain* 132.7 (2009), pp. 1795–1809.

[109] Smita Saxena, Erik Cabuy, and Pico Caroni. “A role for motoneuron subtype-selective ER stress in disease manifestations of FALS mice”. In: *Nature Neuroscience* 12.5 (2009), pp. 627–636.

[110] Lyndsay M. Murray et al. “Transcriptional profiling of differentially vulnerable motor neurons at pre-symptomatic stage in the Smn (2b/-) mouse model of spinal muscular atrophy”. In: *Acta neuropathologica communications* 3 (2015), p. 55.

[111] Christian S. Lobsiger, Séverine Boillée, and Don W. Cleveland. “Toxicity from different SOD1 mutants dysregulates the complement system and the neuronal regenerative response in ALS motor neurons”. In: *Proceedings of the National Academy of Sciences of the United States of America* 104.18 (2007), pp. 7319–7326.

[112] Banafsheh Kadkhodaei et al. “Transcription factor Nurr1 maintains fiber integrity and nuclear-encoded mitochondrial gene expression in dopamine neurons”. In: *Proceedings of the National Academy of Sciences of the United States of America* 110.6 (2013), pp. 2360–2365.

[113] Urmi Bandyopadhyay et al. “RNA-Seq Profiling of Spinal Cord Motor Neurons from a Presymptomatic SOD1 ALS Mouse”. In: *PLoS ONE* 8.1 (2013).

[114] William G Pembroke et al. “Temporal transcriptomics suggest that twin-peaking genes reset the clock”. In: *eLife* 4 (2015), pp. 1–15.

[115] Sabrina Zechel et al. “Topographical transcriptome mapping of the mouse medial ganglionic eminence by spatially resolved RNA-seq”. In: *Genome biology* 15.10 (2014), p. 486.

[116] Susanne Nichterwitz et al. “Laser capture microscopy coupled with Smart-seq2 for precise spatial transcriptomic profiling”. In: *Nature Communications* 7 (July 2016), p. 12139.

[117] Susanne Nichterwitz et al. “LCM-seq: A method for spatial transcriptomic profiling using laser capture microdissection coupled with PolyA-based RNA sequencing”. In: *Methods in Molecular Biology* 1649 (2018), pp. 95–110.

[118] Simone Picelli et al. “Full-length RNA-seq from single cells using Smart-seq2”. In: *Nature Protocols* 9.1 (2014), pp. 171–181.

[119] Julio Aguila et al. “LCM-seq identifies robust markers of vulnerable and resistant human midbrain dopamine neurons”. In: *bioRxiv* (2020), pp. 1–29.

[120] S Nichterwitz et al. “LCM-seq reveals unique transcriptional adaption mechanisms of resistant neurons in spinal muscular atrophy”. In: *bioRxiv* (2018), p. 356113.

[121] M. Nizzardo et al. “Synaptotagmin 13 is neuroprotective across motor neuron diseases”. In: *Acta Neuropathologica* 139.5 (2020), pp. 837–853.

[122] Ilary Allodi et al. “Modeling Motor Neuron Resilience in ALS Using Stem Cells”. In: *Stem Cell Reports* 12.6 (2019), pp. 1329–1341.

[123] Prospero Civita et al. “Laser capture microdissection and RNA-seq analysis: High sensitivity approaches to explain histopathological heterogeneity in human glioblastoma FFPE archived tissues”. In: *Frontiers in Oncology* 9.JUN (2019), pp. 1–12.

[124] Andrew E. Jaffe et al. “Profiling gene expression in the human dentate gyrus granule cell layer reveals insights into schizophrenia and its genetic risk”. In: *Nature Neuroscience* 23.April (2020).

[125] Filipa Bouçanova et al. “Disrupted function of lactate transporter MCT1, but not MCT4, in Schwann cells affects the maintenance of motor end-plate innervation”. In: *Glia* 69.1 (2021), pp. 124–136.

[126] Daniel Almeida et al. “59. LCM-Seq: Pyramidal Neuron Transcriptomic Profiling in the Prefrontal Cortex of Abused Suicides”. In: *Biological Psychiatry* 85.10 (2019), S25.

[127] Sabine Amon et al. “Ultra-sensitive proteome profiling of FACS-isolated cell populations by data-independent acquisition-MS: Application to human hematopoietic stem and progenitor cells”. In: *bioRxiv* (July 2018).

[128] Samuel A. Myers et al. “Streamlined protocol for deep proteomic profiling of FAC-sorted cells and its application to freshly isolated murine immune cells”. In: *Molecular and Cellular Proteomics* 18.5 (May 2019), pp. 995–1009.

[129] Sruti Rayaprolu et al. “Flow-cytometric microglial sorting coupled with quantitative proteomics identifies moesin as a highly-abundant microglial protein with relevance to Alzheimer’s disease”. In: *Molecular Neurodegeneration* 15.1 (May 2020), p. 28.

[130] Peter Feist and Amanda B. Hummon. “Proteomic challenges: Sample preparation techniques for Microgram-Quantity protein analysis from biological samples”. In: *International Journal of Molecular Sciences* 16.2 (2015), pp. 3537–3563.

[131] Rashaun S. Wilson and Angus C. Nairn. “Cell-type-specific proteomics: A neuroscience perspective”. In: *Proteomes* 6.4 (Dec. 2018), p. 51.

[132] Ying Zhu et al. “Proteomic Analysis of Single Mammalian Cells Enabled by Microfluidic Nanodroplet Sample Preparation and Ultrasensitive NanoLC-MS”. In: *Angewandte Chemie - International Edition* 57.38 (Sept. 2018), pp. 12370–12374.

[133] Ying Zhu et al. "Spatially resolved proteome mapping of laser capture microdissected tissue with automated sample transfer to nanodroplets". In: *Molecular and Cellular Proteomics* 17.9 (2018), pp. 1864–1874.

[134] Yiran Liang et al. "Spatially Resolved Proteome Profiling of <200 Cells from Tomato Fruit Pericarp by Integrating Laser-Capture Microdissection with Nanodroplet Sample Preparation". In: *Analytical Chemistry* 90.18 (Sept. 2018), pp. 11106–11114.

[135] Ying Zhu et al. "Proteome Profiling of 1 to 5 Spiked Circulating Tumor Cells Isolated from Whole Blood Using Immunodensity Enrichment, Laser Capture Microdissection, Nanodroplet Sample Processing, and Ultrasensitive nanoLC-MS". In: *Analytical Chemistry* 90.20 (2018), pp. 11756–11759.

[136] Maowei Dou et al. "Nanowell-mediated multidimensional separations combining nanoLC with SLIM IM-MS for rapid, high-peak-capacity proteomic analyses". In: *Analytical and Bioanalytical Chemistry* 411.21 (2019), pp. 5363–5372.

[137] Christopher S. Hughes et al. "Single-pot, solid-phase-enhanced sample preparation for proteomics experiments". In: *Nature Protocols* 14.1 (Jan. 2019), pp. 68–85.

[138] Christopher S Hughes et al. "Ultrasensitive proteome analysis using paramagnetic bead technology". In: *Molecular Systems Biology* 10.10 (Oct. 2014), p. 757.

[139] Jusal Quanico et al. "Development of liquid microjunction extraction strategy for improving protein identification from tissue sections". In: *Journal of Proteomics* 79 (Feb. 2013), pp. 200–218.

[140] Maxence Wisztorski et al. "Spatially-resolved protein surface microsampling from tissue sections using liquid extraction surface analysis". In: *Proteomics* 16.11-12 (2016), pp. 1622–1632.

[141] David G. Rizzo et al. "Enhanced Spatially Resolved Proteomics Using On-Tissue Hydrogel-Mediated Protein Digestion". In: *Analytical Chemistry* 89.5 (2017), pp. 2948–2955.

[142] Jusal Quanico et al. "Integrated mass spectrometry imaging and omics workflows on the same tissue section using grid-aided, parafilm-assisted microdissection". In: *Biochimica et Biophysica Acta - General Subjects* 1861.7 (July 2017), pp. 1702–1714.

[143] Fabrizio Donnarumma and Kermit K. Murray. "Laser ablation sample transfer for localized LC-MS/MS proteomic analysis of tissue". In: *Journal of Mass Spectrometry* 51.4 (2016), pp. 261–268.

[144] Marialaura Dilillo et al. "Mass Spectrometry Imaging, Laser Capture Microdissection, and LC-MS/MS of the Same Tissue Section". In: *Journal of Proteome Research* 16.8 (2017), pp. 2993–3001.

[145] Mariana Molina et al. "Enrichment of single neurons and defined brain regions from human brain tissue samples for subsequent proteome analysis". In: *Journal of Neural Transmission* 122.7 (2015), pp. 993–1005.

[146] Eleanor S. Drummond et al. "Proteomic analysis of neurons microdissected from formalin-fixed, paraffin-embedded Alzheimer's disease brain tissue". In: *Scientific Reports* 5.October (2015), p. 15456.

[147] Ugljesa Djuric et al. “Spatiotemporal proteomic profiling of human cerebral development”. In: *Molecular and Cellular Proteomics* 16.9 (2017), pp. 1558–1562.

[148] Teresa García-Berrocoso et al. “Single cell immuno-laser microdissection coupled to label-free proteomics to reveal the proteotypes of human brain cells after ischemia”. In: *Molecular and Cellular Proteomics* 17.1 (Jan. 2018), pp. 175–189.

[149] David I. Anderson et al. “The role of locomotion in psychological development”. In: *Frontiers in Psychology* 4 (July 2013), p. 440.

[150] Christian Laut Ebbesen and Michael Brecht. “Motor cortex - To act or not to act?” In: *Nature Reviews Neuroscience* 18.11 (2017), pp. 694–705.

[151] Russell H Swerdlow. “The Brain Atlas: A Visual Guide to the Human Central Nervous System, 3rd Edition”. In: *Journal of Neuro-Ophthalmology* 28.3 (2008), pp. 249–250.

[152] Adriana L. Natali and Bruno Bordoni. “Neuroanatomy, Corticospinal Cord Tract”. eng. In: *StatPearls* (Jan. 2018).

[153] Thomas A. Woolsey, Joseph Hanaway, and Mokhtar H. Gado. “The Brain Atlas: A Visual Guide to the Human Central Nervous System (4th edition)”. In: *Neuropathology and Applied Neurobiology* 43.6 (2017), pp. 548–548.

[154] Peter McColgan et al. “The human motor cortex microcircuit: insights for neurodegenerative disease”. In: *Nature Reviews Neuroscience* 21.8 (2020), pp. 401–415.

[155] Charles D. Gilbert and Torsten N. Wiesel. “Functional Organization of the Visual Cortex”. In: *Progress in Brain Research* 58.C (1983), pp. 209–218.

[156] W. Martin Usrey and David Fitzpatrick. “Specificity in the axonal connections of layer VI neurons in tree shrew striate cortex: Evidence for distinct granular and supragranular systems”. In: *Journal of Neuroscience* 16.3 (1996), pp. 1203–1218.

[157] Hillel Adesnik and Alexander Naka. “Cracking the Function of Layers in the Sensory Cortex”. In: *Neuron* 100.5 (2018), pp. 1028–1043.

[158] R. J. Douglas and K. A. Martin. “A functional microcircuit for cat visual cortex.” In: *The Journal of Physiology* 440.1 (1991), pp. 735–769.

[159] Helen Barbas and Miguel García-Cabezas. “Motor cortex layer 4: Less is more”. In: *Trends in Neurosciences* 38.5 (2015), pp. 259–261.

[160] Miguel Ángel García-Cabezas and Helen Barbas. “Area 4 has layer IV in adult primates”. In: *European Journal of Neuroscience* 39.11 (June 2014), pp. 1824–1834.

[161] Stewart Shipp. “The importance of being agranular: A comparative account of visual and motor cortex”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 360.1456 (2005), pp. 797–814.

[162] Taro Kiritani et al. “Hierarchical connectivity and connection-specific dynamics in the corticospinal-corticostriatal microcircuit in mouse motor cortex”. In: *Journal of Neuroscience* 32.14 (Apr. 2012), pp. 4992–5001.

[163] Trygve Bakken et al. “Evolution of cellular diversity in primary motor cortex of human, marmoset monkey, and mouse”. In: *bioRxiv* (Jan. 2020).

[164] Domagoj Džaja et al. “Neocortical calretinin neurons in primates: Increase in proportion and microcircuitry structure”. In: *Frontiers in Neuroanatomy* 8.SEP (2014).

[165] Bryan M. Hooks et al. “Organization of cortical and thalamic input to pyramidal neurons in mouse motor cortex”. In: *Journal of Neuroscience* 33.2 (2013), pp. 748–760.

[166] Roger N. Lemon. “Descending pathways in motor control”. In: *Annual Review of Neuroscience* 31.1 (July 2008), pp. 195–218.

[167] Michael N. Economo et al. “Distinct descending motor cortex pathways and their roles in movement”. In: *Nature* 563.7729 (2018), pp. 79–84.

[168] Zengcai V. Guo et al. “Maintenance of persistent activity in a frontal thalamocortical loop”. In: *Nature* 545.7653 (2017), pp. 181–186.

[169] Bosiljka Tasic et al. “Shared and distinct transcriptomic cell types across neocortical areas”. In: *bioRxiv* 20.7729 (2017), pp. 229–542.

[170] C. E. Catsman-Berrevoets and H. G.J.M. Kuypers. “A search for corticospinal collaterals to thalamus and mesencephalon by means of multiple retrograde fluorescent tracers in cat and rat”. In: *Brain Research* 218.1-2 (1981), pp. 15–33.

[171] M. Steriade and G. Yossif. “Afferent and recurrent collateral influences on cortical somatosensory neurons”. In: *Experimental Neurology* 56.2 (1977), pp. 334–360.

[172] Zoltán Molnár and Amanda F.P. Cheung. “Towards the classification of subpopulations of layer V pyramidal projection neurons”. In: *Neuroscience Research* 55.2 (2006), pp. 105–115.

[173] Ekkehard M. Kasper et al. “Pyramidal neurons in layer 5 of the rat visual cortex. III. Differential maturation of axon targeting, dendritic morphology, and electrophysiological properties”. In: *Journal of Comparative Neurology* 339.4 (1994), pp. 495–518.

[174] Bernardo Rudy et al. “Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons”. In: *Developmental Neurobiology* 71.1 (2011), pp. 45–61.

[175] Sergiy V. Kushchayev et al. “The discovery of the pyramidal neurons: Vladimir Betz and a new era of neuroscience”. In: *Brain* 135.1 (Jan. 2012), pp. 285–300.

[176] Walshe F. M. R. “The giant cells of betz, the motor cortex and the pyramidal tract: a critical review”. In: *Brain* 65.4 (1942), pp. 409–461.

[177] Ita R. Kaiserman-Abramof and Alan Peters. “Some aspects of the morphology of Betz cells in the cerebral cortex of the cat”. In: *Brain Research* 43.2 (1972), pp. 527–546.

[178] J. M. Crawford and D. R. Curtis. “Pharmacological studies on feline Betz cells”. In: *The Journal of Physiology* 186.1 (1966), pp. 121–138.

[179] J. W. Phillis and J. J. Limacher. “Substance P excitation of cerebral cortical Betz cells”. In: *Brain Research* 69.1 (1974), pp. 158–163.

[180] Jon H. Kaas. “Evolution of somatosensory and motor cortex in primates”. In: *Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology* 281.1 (2004), pp. 1148–1156.

[181] N. L. Badlangana et al. “Observations on the giraffe central nervous system related to the corticospinal tract, motor cortex and spinal cord: What difference does a long neck make?” In: *Neuroscience* 148.2 (2007), pp. 522–534.

[182] Bob Jacobs et al. “The neocortex of cetartiodactyls. II. Neuronal morphology of the visual and motor cortices in the giraffe (Giraffa camelopardalis)”. In: *Brain Structure and Function* 220.5 (Sept. 2015), pp. 2851–2872.

[183] Khang Cheng Ho et al. “Brief communication: Correlation of neuronal cell body size in motor cortex and hippocampus with body height, body weight, and axonal length”. In: *International Journal of Neuroscience* 65.1-4 (1992), pp. 147–153.

[184] Heiko Braak et al. “Pathological TDP-43 changes in Betz cells differ from those in bulbar and spinal α-motoneurons in sporadic amyotrophic lateral sclerosis”. In: *Acta Neuropathologica* 133.1 (Jan. 2017), pp. 79–90.

[185] A. M. Lassek and G. L. Rasmussen. “The human pyramidal tract: A fiber and numerical analysis”. In: *Archives of Neurology And Psychiatry* 42.5 (Nov. 1939), pp. 872–876.

[186] Arthur Marvel Lassek. “The pyramidal tract of the monkeys. A betz cell and pyramidal tract enumeration”. In: *Journal of Comparative Neurology* 74.2 (1941), pp. 193–202.

[187] Shin ichi Terao et al. “Age-related changes of the myelinated fibers in the human corticospinal tract: a quantitative analysis”. In: *Acta Neuropathologica* 88.2 (1994), pp. 137–142.

[188] Rudolf Nieuwenhuys, Jan Voogd, and Christiaan van Huijzen. “Telencephalon: Neocortex”. In: *The Human Central Nervous System* (2008). Ed. by Rudolf Nieuwenhuys, Jan Voogd, and Christiaan van Huijzen, pp. 491–679.

[189] Cécile Vogt Oskar Vogt. “Morphologische gestaltungen unter normalen und pathogenen bedingungen”. In: *J Psychol Neurol* 50 (1942), pp. 162–524.

[190] Bob Jacobs et al. “Comparative morphology of gigantopyramidal neurons in primary motor cortex across mammals”. In: *Journal of Comparative Neurology* 526.3 (2018), pp. 496–536.

[191] Madge E. Scheibel et al. “Basilar dendrite bundles of giant pyramidal cells”. In: *Experimental Neurology* 42.2 (1974), pp. 307–319.

[192] Madge E. Scheibel, Uwami Tomiyasu, and Arnold B. Scheibel. “The aging human betz cell”. In: *Experimental Neurology* 56.3 (1977), pp. 598–609.

[193] John A. Kiernan and Arthur J. Hudson. “Changes in sizes of cortical and lower motor neurons in amyotrophic lateral sclerosis”. In: *Brain* 114.2 (1991), pp. 843–853.

[194] Bob Jacobs et al. “Comparative neuronal morphology of the cerebellar cortex in afrotherians, carnivores, cetartiodactyls, and primates”. In: *Frontiers in Neuroanatomy* 8 (Apr. 2014), p. 24.

[195] Johannes Tigges, James G. Herndon, and Alan Peters. “Neuronal population of area 4 during the life span of the rhesus monkey”. In: *Neurobiology of Aging* 11.3 (1990), pp. 201–208.

[196] Fitzpatrick D Purves D Augustine GJ. “Upper Motor Neurons That Initiate Complex Voluntary Movements”. In: *Neuroscience. 2nd edition* (2001).

[197] Javier H. Jara et al. “Retrograde labeling, transduction, and genetic targeting allow cellular analysis of corticospinal motor neurons: Implications in health and disease”. In: *Frontiers in Neuroanatomy* 8 (2014).

[198] Shoichi Sasaki and Makoto Iwata. “Ultrastructural study of Betz cells in the primary motor cortex of the human brain”. In: *Journal of Anatomy* 199.6 (2001), pp. 699–708.

[199] Heiko Braak. “A primitive gigantopyramidal field buried in the depth of the cingulate sulcus of the human brain”. In: *Brain Research* 109.2 (June 1976), pp. 219–233.

[200] Chet C. Sherwood et al. “Evolution of specialized pyramidal neurons in primate visual and motor cortex”. In: *Brain, Behavior and Evolution* 61.1 (2003), pp. 28–44.

[201] Gundela Meyer. “Forms and spatial arrangement of neurons in the primary motor cortex of man”. In: *Journal of Comparative Neurology* 262.3 (1987), pp. 402–428.

[202] Ronald P. Hammer, Uwamie Tomiyasu, and Arnold B. Scheibel. “Degeneration of the human Betz cell due to amyotrophic lateral sclerosis”. In: *Experimental Neurology* 63.2 (1979), pp. 336–346.

[203] Vivian T. Nguyen et al. “Comparative neocortical neuromorphology in felids: African lion, African leopard, and cheetah”. In: *Journal of Comparative Neurology* 528.8 (2020), pp. 1392–1422.

[204] Martin Deschênes, André Labelle, and Pierre Landry. “Morphological characterization of slow and fast pyramidal tract cells in the cat”. In: *Brain Research* 178.2-3 (1979), pp. 251–274.

[205] J. J. Sloper, R. W. Hiorns, and T. P. Powell. “A qualitative and quantitative electron microscopic study of the neurons in the primate motor and somatic sensory cortices.” In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 285.1006 (1979), pp. 141–171.

[206] Johannes Tigges. “Novel inclusion bodies in Betz cells of cortical area 4 of aged rhesus monkeys”. In: *The Anatomical Record* 233.1 (1992), pp. 162–168.

[207] H. Braak and E. Braak. “The pyramidal cells of Betz within the cingulate and precentral gigantopyramidal field in the human brain - A golgi and pigmentarchitectonic study”. In: *Cell and Tissue Research* 172.1 (Sept. 1976), pp. 103–119.

[208] George D. Pappas. “Texture of the Nervous System of Man and the Vertebrates—1899, Volume 1”. In: *Surgical Neurology* 59.6 (2003), p. 524.

[209] J. P. de Ruiter. “The influence of post-mortem fixation delay on the reliability of the Golgi silver impregnation”. In: *Brain Research* 266.1 (Apr. 1983), pp. 143–147.

[210] Asao Hirano et al. “The fine structure of some intraganglionic alterations: Neurofibrillary Tangle, Granulovacuolar Bodies and “Rod-Like” Structures as seen in Guam Amyotrophic Lateral Sclerosis and parkinsonism-Dementia Complex”. In: *Journal of Neuropathology and Experimental Neurology* 27.2 (1968), pp. 167–182.

[211] Asao Hirano et al. “Fine structural observations of neurofilamentous changes in amyotrophic lateral sclerosis”. In: *Journal of Neuropathology and Experimental Neurology* 43.5 (1984), pp. 461–470.

[212] I. H. Coriat and Alfred W. Campbell. “Histological Studies on the Localization of Cerebral Function”. In: *The American Journal of Psychology* 17.4 (1906), p. 600.

[213] A. M. Lassek. “The human pyramidal tract: XVI. Reaction of individual axons in selected cases with acute cerebral lesions”. In: *Journal of Neuropathology and Experimental Neurology* 7.1 (1948), pp. 89–93.

[214] P. Landry, C. J. Wilson, and S. T. Kitai. “Morphological and electrophysiological characteristics of pyramidal tract neurons in the rat”. In: *Experimental Brain Research* 57.1 (Jan. 1984), pp. 177–190.

[215] Guo-Fang -F Tseng and David A. Prince. “Heterogeneity of rat corticospinal neurons”. In: *Journal of Comparative Neurology* 335.1 (Sept. 1993), pp. 92–108.

[216] Soraya Ghafari, Elnaz Golalipour, and Mohammad Jafar Golalipour. “Alterations of the Giant Pyramidal Neurons (Betz Cells) in Brain Cortex of Rat Offspring Born from Gestational Diabetic Dams: A Morphometric Study”. In: *International Journal of Morphology* 33.3 (2015), pp. 1120–1125.

[217] Simon Davis et al. “Development of a Sensitive, Scalable Method for Spatial, Cell-Type-Resolved Proteomics of the Human Brain”. In: *Journal of Proteome Research* 18.4 (Feb. 2019), pp. 1787–1795.

[218] Lucas Graybuck et al. “Prospective, brain-wide labeling of neuronal subclasses with enhancer-driven AAVs”. In: *bioRxiv* (2019), p. 525014.

[219] Amanda L. Brown and Doug W. Smith. “Improved RNA preservation for immunolabeling and laser microdissection”. In: *Rna* 15.12 (2009), pp. 2364–2374.

[220] E. Henneman, G. Somjen, and D. O. Carpenter. “Functional Significance of Cell Size in Spinal Motoneurons”. In: *Journal of neurophysiology* 28 (1965), pp. 560–580.

[221] K. A. Jellinger. “The Motoneurone and its Muscle Fibres”. English. In: *European Journal of Neurology* 14.10 (2007), e13–e13.

[222] Nicole A. Young, Christine E. Collins, and Jon H. Kaas. “Cell and neuron densities in the primary motor cortex of primates”. In: *Frontiers in Neural Circuits* (2013).

[223] Jon H. Kaas. “Why is brain size so important: Design problems and solutions as neocortex gets bigger or smaller”. In: *Brain and Mind* 1.1 (2000), pp. 7–23.

[224] K. Brodmann and Laurence J. Garey. *Brodmann’s localisation in the cerebral cortex: The principles of comparative localisation in the cerebral cortex based on cytoarchitectonics*. 2006, pp. 1–298.

[225] Rickye Heffner and Bruce Masterton. “Variation in form of the pyramidal tract and its relationship to digital dexterity”. In: *Brain, Behavior and Evolution* 12.3 (1975), pp. 161–200.

[226] Bob Jacobs et al. “Neocortical neuronal morphology in the newborn giraffe (Giraffa camelopardalis tippelskirchi) and African elephant (Loxodonta africana)”. In: *Journal of Comparative Neurology* 524.2 (2016), pp. 257–287.

[227] A. J. Buller, J. C. Eccles, and Rosamond M. Eccles. “Interactions between motoneurones and muscles in respect of the characteristic speeds of their responses”. In: *The Journal of Physiology* 150.2 (Feb. 1960), pp. 417–439.

[228] S. Salmons and Gerta Vrbová. "The influence of activity on some contractile characteristics of mammalian fast and slow muscles". In: *The Journal of Physiology* 201.3 (May 1969), pp. 535–549.

[229] A. Lundberg and P. Voorhoeve. "Effects from the Pyramidal Tract on Spinal Reflex Arcs". In: *Acta Physiologica Scandinavica* 56.3-4 (Nov. 1962), pp. 201–219.

[230] Adriano Chiò et al. "Prognostic factors in ALS: A critical review". In: *Amyotrophic Lateral Sclerosis* 10.5-6 (2009), pp. 310–323.

[231] Glenda Halliday et al. "The progression of pathology in longitudinally followed patients with Parkinson's disease". In: *Acta Neuropathologica* 115.4 (2008), pp. 409–415.

[232] Rodolfo Savica, Walter A. Rocca, and J. Eric Ahlskog. "When does Parkinson disease start?" In: *Archives of Neurology* 67.7 (2010), pp. 798–801.

[233] Nimeshan Geevasinga et al. "Pathophysiological and diagnostic implications of cortical dysfunction in ALS". In: *Nature Reviews Neurology* 12.11 (Nov. 2016), pp. 651–661.

[234] Yvonne D. Rollins, Björn Oskarsson, and Steven P. Ringel. "Primary Lateral Sclerosis". In: *International Neurology: A Clinical Approach* (2010), pp. 203–204.

[235] Erik P. Pioro et al. "Detection of cortical neuron loss in motor neuron disease by proton magnetic resonance spectroscopic imaging in vivo". In: *Neurology* 44.10 (1994), pp. 1933–1938.

[236] Ole Gredal et al. "Unchanged total number of neurons in motor cortex and neocortex in amyotrophic lateral sclerosis: A stereological study". In: *Journal of Neuroscience Methods* 95.2 (2000), pp. 171–176.

[237] Mette Helene Toft, Ole Gredal, and Bente Pakkenberg. "The size distribution of neurons in the motor cortex in amyotrophic lateral sclerosis". In: *Journal of Anatomy* 207.4 (2005), pp. 399–407.

[238] Grant Coan and Cassie S. Mitchell. "An assessment of possible neuropathology and clinical relationships in 46 sporadic amyotrophic lateral sclerosis patient autopsies". In: *Neurodegenerative Diseases* 15.5 (2015), pp. 301–312.

[239] S. Murayama, T. W. Bouldin, and K. Suzuki. "Immunocytochemical and ultrastructural studies of upper motor neurons in amyotrophic lateral sclerosis". In: *Acta Neuropathologica* 83.5 (1992), pp. 518–524.

[240] Kuninobu Nihei, Ann C. McKee, and Neil W. Kowall. "Patterns of neuronal degeneration in the motor cortex of amyotrophic lateral sclerosis patients". In: *Acta Neuropathologica* 86.1 (1993), pp. 55–64.

[241] Barlay Genç et al. "Apical dendrite degeneration, a novel cellular pathology for Betz cells in ALS". In: *Scientific Reports* 7 (2017).

[242] F. Udaka, M. Kameyama, and M. Tomonaga. "Degeneration of betz cells in motor neuron disease. A Golgi study". In: *Acta Neuropathologica* 70.3-4 (1986), pp. 289–295.

[243] Javier H. Jara et al. "AAV2 mediated retrograde transduction of corticospinal motor neurons reveals initial and selective apical dendrite degeneration in ALS". In: *Neurobiology of Disease* 47.2 (Aug. 2012), pp. 174–183.

[244] Christopher M. Henstridge, Eleanor Pickett, and Tara L. Spires-Jones. "Synaptic pathology: A shared mechanism in neurological disease". In: *Ageing Research Reviews* 28 (2016), pp. 72–84.

[245] D. S. Horoupian et al. "Dementia and motor neuron disease: Morphometric, biochemical, and Golgi studies". In: *Annals of Neurology* 16.3 (1984), pp. 305–313.

[246] Elena Abati et al. "Glial cells involvement in spinal muscular atrophy: Could SMA be a neuroinflammatory disease?" In: *Neurobiology of Disease* 140 (2020).

[247] Shoichi Sasaki and Makoto Iwata. "Ultrastructural change of synapses of Betz cells in patients with amyotrophic lateral sclerosis". In: *Neuroscience Letters* 268.1 (1999), pp. 29–32.

[248] Matthew J. Fogarty et al. "Size-Dependent Vulnerability of Lumbar Motor Neuron Dendritic Degeneration in SOD1G93A Mice". In: *Anatomical Record* 303.5 (Sept. 2020), pp. 1455–1471.

[249] Mukesh Gautam et al. "Mitochondria, ER, and nuclear membrane defects reveal early mechanisms for upper motor neuron vulnerability with respect to TDP-43 pathology". In: *Acta Neuropathologica* 137.1 (2019), pp. 47–69.

[250] Emily E. Handley et al. "Synapse Dysfunction of Layer v Pyramidal Neurons Precedes Neurodegeneration in a Mouse Model of TDP-43 Proteinopathies". In: *Cerebral Cortex* 27.7 (2017), pp. 3630–3647.

[251] P. Hande Özdinler et al. "Corticospinal motor neurons and related subcerebral projection neurons undergo early and specific neurodegeneration in hSOD1G93A transgenic ALS mice". In: *Journal of Neuroscience* 31.11 (Mar. 2011), pp. 4166–4177.

[252] David Gordon et al. "Single-copy expression of an amyotrophic lateral sclerosis-linked TDP-43 mutation (M337V) in BAC transgenic mice leads to altered stress granule dynamics and progressive motor dysfunction". In: *Neurobiology of Disease* 121.May 2018 (2019), pp. 148–162.

[253] Samuel M. Chou and Forbes H. Norris. "Issues & Opinions: Amyotrophic lateral sclerosis: Lower motor neuron disease spreading to upper motor neurons". In: *Muscle & Nerve* 16.8 (1993), pp. 864–869.

[254] Mehdi A.J. van den Bos et al. "Pathophysiology and diagnosis of ALS: Insights from advances in neurophysiological techniques". In: *International Journal of Molecular Sciences* 20.11 (2019).

[255] Mark R. Baker. "ALS - Dying forward, backward or outward?" In: *Nature Reviews Neurology* 10.11 (2014), p. 660.

[256] Stefka Tyanova, Tikira Temu, and Juergen Cox. "The MaxQuant computational platform for mass spectrometry-based shotgun proteomics". In: *Nature Protocols* 11.12 (2016), pp. 2301–2319.

[257] Joshua E. Elias and Steven P. Gygi. "Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry". In: *Nature Methods* 4.3 (Mar. 2007), pp. 207–214.

[258] Jürgen Cox and Matthias Mann. "Quantitative, high-resolution proteomics for data-driven systems biology". In: *Annual Review of Biochemistry* 80 (2011), pp. 273–299.

[259] Alexey I. Nesvizhskii and Ruedi Aebersold. “Interpretation of shotgun proteomic data: The protein inference problem”. In: *Molecular and Cellular Proteomics* 4.10 (Oct. 2005), pp. 1419–1440.

[260] Jürgen Cox et al. “Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ”. In: *Molecular and Cellular Proteomics* 13.9 (Sept. 2014), pp. 2513–2526.

[261] Andy T Kong et al. “MSFragger: Ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics”. In: *Nature Methods* 14.5 (Apr. 2017), pp. 513–520.

[262] Guo Ci Teo et al. “Fast Deisotoping Algorithm and Its Implementation in the MSFragger Search Engine”. In: *Journal of Proteome Research* 20.1 (Jan. 2021), pp. 498–505.

[263] Fengchao Yu et al. “Fast Quantitative Analysis of timsTOF PASEF Data with MSFragger and IonQuant”. In: *Molecular and Cellular Proteomics* 19.9 (Sept. 2020), pp. 1575–1585.

[264] Felipe da Veiga Leprevost et al. “Philosopher: a versatile toolkit for shotgun proteomics data analysis”. In: *Nature Methods* 17.9 (July 2020), pp. 869–870.

[265] Chris Bielow, Guido Mastrobuoni, and Stefan Kempa. “Proteomics Quality Control: Quality Control Software for MaxQuant Results”. In: *Journal of Proteome Research* 15.3 (Mar. 2016), pp. 777–787.

[266] Stefka Tyanova et al. “The Perseus computational platform for comprehensive analysis of (prote)omics data”. In: *Nature Methods* 13.9 (Aug. 2016), pp. 731–740.

[267] Stefka Tyanova and Juergen Cox. “Perseus: A bioinformatics platform for integrative analysis of proteomics data in cancer research”. In: *Methods in Molecular Biology* 1711 (2018), pp. 133–148.

[268] Shisheng Wang et al. “NAguideR: Performing and prioritizing missing value imputations for consistent bottom-up proteomic analyses”. In: *Nucleic Acids Research* 48.14 (2020).

[269] Huaiyu Mi et al. “PANTHER version 16: A revised family classification, tree-based classification tool, enhancer regions and extensive API”. In: *Nucleic Acids Research* 49.D1 (2021), pp. D394–D403.

[270] Fran Supek et al. “Revigo summarizes and visualizes long lists of gene ontology terms”. In: *PLoS ONE* 6.7 (2011).

[271] Uku Raudvere et al. “G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update)”. In: *Nucleic Acids Research* 47.W1 (2019), pp. 191–198.

[272] Peter Bankhead et al. “QuPath: Open source software for digital pathology image analysis”. In: *Scientific Reports* 7.1 (Dec. 2017).

[273] Ignacio Arganda-Carreras et al. “Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification”. In: *Bioinformatics* 33.15 (Aug. 2017). Ed. by Robert Murphy, pp. 2424–2426.

[274] Michael W. Pfaffl. “A new mathematical model for relative quantification in real-time RT–PCR”. In: *Nucleic Acids Research* 29.9 (2001), E45.

[275] Jo Vandesompele et al. “Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.” In: *Genome biology* 3.7 (June 2002), pp. 1–12.

[276] Kuanwei Sheng and Chenghang Zong. “Single-cell RNA-Seq by multiple annealing and tailing-based quantitative single-cell RNA-Seq (MATQ-Seq)”. In: *Methods in Molecular Biology* 1979 (2019), pp. 57–71.

[277] Simon Andrews et al. “FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics”. In: *Babraham Institute* 1.1 (2015), p. 1.

[278] a. M. Bolger, M. Lohse, and B. Usadel. “Trimmomatic: A flexible read trimming tool for Illumina NGS data”. In: *Bioinformatics* 30.15 (2014), pp. 2114–2120.

[279] Alexander Dobin et al. “STAR: Ultrafast universal RNA-seq aligner”. In: *Bioinformatics* 29.1 (2013), pp. 15–21.

[280] Liguo Wang, Shengqin Wang, and Wei Li. “RSeQC: Quality control of RNA-seq experiments”. In: *Bioinformatics* 28.16 (2012), pp. 2184–2185.

[281] Yang Liao, Gordon K. Smyth, and Wei Shi. “FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features”. In: *Bioinformatics* 30.7 (2014), pp. 923–930.

[282] Michael I. Love, Wolfgang Huber, and Simon Anders. “Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2”. In: *Genome Biology* 15.12 (2014).

[283] Gennady Korotkevich et al. “Fast gene set enrichment analysis”. In: *bioRxiv* (2016).

[284] Yungui He et al. “Targeted deletion of a single Sca8 ataxia locus allele in mice causes abnormal gait, progressive loss of motor coordination, and Purkinje cell dendritic deficits”. In: *Journal of Neuroscience* 26.39 (2006), pp. 9975–9982.

[285] Yoshihito Ishida et al. “Vulnerability of Purkinje Cells Generated from Spinocerebellar Ataxia Type 6 Patient-Derived iPSCs”. In: *Cell Reports* 17.6 (2016), pp. 1482–1490.

[286] Günter Siegemund et al. “Fluorine Compounds, Organic”. In: *Ullmann’s Encyclopedia of Industrial Chemistry* (Jan. 2016), pp. 1–56.

[287] Ryan Kelly et al. “Single Cell Proteome Mapping of Tissue Heterogeneity Using Microfluidic Nanodroplet Sample Processing and Ultrasensitive LC-MS.” In: *Journal of biomolecular techniques : JBT* 30.Suppl (2019), S61.

[288] Ying Zhu et al. “Proteomic Analysis of Single Mammalian Cells Enabled by Microfluidic Nanodroplet Sample Preparation and Ultrasensitive NanoLC-MS”. In: *Angewandte Chemie - International Edition* 57.38 (2018), pp. 12370–12374.

[289] Sarah M. Williams et al. “Automated Coupling of Nanodroplet Sample Preparation with Liquid Chromatography-Mass Spectrometry for High-Throughput Single-Cell Proteomics”. In: *Analytical Chemistry* 92.15 (2020), pp. 10588–10596.

[290] Kerui Xu et al. “Benchtop-compatible sample processing workflow for proteome profiling of $<$ 100 mammalian cells”. In: *Analytical and Bioanalytical Chemistry* 411.19 (2019), pp. 4587–4596.

[291] Eva Griesser et al. “Quantitative Profiling of the Human Substantia Nigra Proteome from Laser-capture Microdissected FFPE Tissue”. In: *Molecular and Cellular Proteomics* 19.5 (Mar. 2020), pp. 839–851.

[292] Melanie Christine Föll et al. “Reproducible proteomics sample preparation for single FFPE tissue slices using acid-labile surfactant and direct trypsinization”. In: *Clinical Proteomics* 15.1 (Mar. 2018), p. 11.

[293] Lionel Moulédous et al. “Proteomic analysis of immunostained, laser-capture microdissected brain samples”. In: *Electrophoresis* 24.1-2 (Jan. 2003), pp. 296–302.

[294] Michael A. Tangrea et al. “Effect of immunohistochemistry on molecular analysis of tissue samples: Implications for microdissection technologies”. In: *Journal of Histochemistry and Cytochemistry* 59.6 (June 2011), pp. 591–600.

[295] Jintang He et al. “Immunohistochemical staining, laser capture microdissection, and filter-aided sample preparation-assisted proteomic analysis of target cell populations within tissue samples”. In: *Electrophoresis* 34.11 (June 2013), pp. 1627–1636.

[296] Juergen Cox and Matthias Mann. “1D and 2D annotation enrichment: a statistical method integrating quantitative proteomics with complementary high-throughput data.” In: *BMC bioinformatics* 13 Suppl 1.Suppl 16 (2012), S12.

[297] Frency Varghese et al. “IHC profiler: An open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples”. In: *PLoS ONE* 9.5 (2014).

[298] Egidio D’Angelo et al. “The cerebellar Golgi cell and spatiotemporal organization of granular layer activity”. In: *Frontiers in Neural Circuits* APR 2013 (2013).

[299] Jia Jia Liu et al. “BPAG1n4 is essential for retrograde axonal transport in sensory neurons”. In: *Journal of Cell Biology* 163.2 (2003), pp. 223–229.

[300] Chi Lun Chang et al. “Feedback regulation of receptor-induced ca2+ signaling mediated by e-syt1 and nir2 at endoplasmic reticulum-plasma membrane junctions”. In: *Cell Reports* 5.3 (2013), pp. 813–825.

[301] Vaisaly R Nath et al. “ Extended synaptotagmin regulates membrane contact site structure and lipid transfer function in vivo ”. In: *EMBO reports* 21.9 (2020).

[302] Haijia Yu et al. “Extended synaptotagmins are Ca2+-dependent lipid transfer proteins at membrane contact sites”. In: *Proceedings of the National Academy of Sciences of the United States of America* 113.16 (2016), pp. 4362–4367.

[303] Koto Kikuma et al. “Extended synaptotagmin localizes to presynaptic ER and promotes neurotransmission and synaptic growth in drosophila”. In: *Genetics* 207.3 (2017), pp. 993–1006.

[304] Otto Haller and Georg Kochs. “Human MxA protein: An interferon-induced dynamin-like GTPase with broad antiviral activity”. In: *Journal of Interferon and Cytokine Research* 31.1 (2011), pp. 79–87.

[305] Marc P. Lussier et al. “MxA, a member of the dynamin superfamily, interacts with the ankyrin-like repeat domain of TRPC”. In: *Journal of Biological Chemistry* 280.19 (2005), pp. 19393–19400.

[306] Damian Szklarczyk et al. “STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets”. In: *Nucleic Acids Research* 47.D1 (2019), pp. 607–613.

[307] Reiji Hommyo et al. “Expression of CRYM in different rat organs during development and its decreased expression in degenerating pyramidal tracts in amyotrophic lateral sclerosis”. In: *Neuropathology* 38.3 (2018), pp. 247–259.

[308] Carl M. Hurt et al. “REEP1 and REEP2 proteins are preferentially expressed in neuronal and neuronal-like exocytotic tissues”. In: *Brain Research* 1545 (2014), pp. 12–22.

[309] Julia Falk et al. “Functional Mutation Analysis Provides Evidence for a Role of REEP1 in Lipid Droplet Biology”. In: *Human Mutation* 35.4 (2014), pp. 497–504.

[310] Seong H. Park et al. “Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule interactions with the tubular ER network”. In: *Journal of Clinical Investigation* 120.4 (2010), pp. 1097–1110.

[311] Yi Guo et al. “A Rare KIF1A Missense Mutation Enhances Synaptic Function and Increases Seizure Activity”. In: *Frontiers in Genetics* 11 (2020).

[312] Yu Ding et al. “Caprin-2 enhances canonical Wnt signaling through regulating LRP5/6 phosphorylation”. In: *Journal of Cell Biology* 182.5 (Sept. 2008), pp. 865–872.

[313] Silvana B. Rosso and Nibaldo C. Inestrosa. “WNT signalling in neuronal maturation and synaptogenesis”. In: *Frontiers in Cellular Neuroscience* 7 (June 2013).

[314] Maria Rosita Pantò et al. “Role of the Otx1 gene in cell differentiation of mammalian cortex”. In: *European Journal of Neuroscience* 19.10 (2004), pp. 2893–2902.

[315] Z. Zhenjiawang et al. “BART: A transcription factor prediction tool with query gene sets or epigenomic profiles”. In: *Bioinformatics* 34.16 (2018), pp. 2867–2869.

[316] Péter Szocsics et al. “Perisomatic innervation and neurochemical features of giant pyramidal neurons in both hemispheres of the human primary motor cortex”. In: *Brain Structure and Function* 226.1 (2021), pp. 281–296.

[317] Siddhesh S. Kamat et al. “Immunomodulatory lysophosphatidylserines are regulated by ABHD16A and ABHD12 interplay”. In: *Nature Chemical Biology* 11.2 (2015), pp. 164–171.

[318] Fei Tang et al. “RNF185, a novel mitochondrial ubiquitin E3 ligase, regulates autophagy through interaction with BNIP1”. In: *PLoS ONE* 6.9 (2011).

[319] Celeste M. Karch et al. “Selective genetic overlap between amyotrophic lateral sclerosis and diseases of the frontotemporal dementia spectrum”. In: *JAMA Neurology* 75.7 (2018), pp. 860–875.

[320] Shimo Li et al. “The expression and localization of Prune2 mRNA in the central nervous system”. In: *Neuroscience Letters* 503.3 (2011), pp. 208–214.

[321] Steven G. Potkin et al. “Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer’s Disease”. In: *PLoS ONE* 4.8 (2009).

[322] Gratien Dalpé et al. “Dystonin is essential for maintaining neuronal cytoskeleton organization”. In: *Molecular and Cellular Neurosciences* 10.5-6 (Apr. 1998), pp. 243–257.

[323] Laetitia Francelle et al. “Loss of the thyroid hormone-binding protein Crym renders striatal neurons more vulnerable to mutant huntingtin in Huntington’s disease”. In: *Human Molecular Genetics* 24.6 (2015), pp. 1563–1573.

[324] Francois Guillemot et al. “Molecular mechanisms of cortical differentiation”. In: *European Journal of Neuroscience* 23.4 (2006), pp. 857–868.

[325] Robert F. Hevner et al. “Beyond laminar fate: Toward a molecular classification of cortical projection/pyramidal neurons”. In: *Developmental Neuroscience* 25.2-4 (2003), pp. 139–151.

[326] H. Yoneshima et al. “ER81 is expressed in a subpopulation of layer 5 neurons in rodent and primate neocortices”. In: *Neuroscience* 137.2 (2006), pp. 401–412.

[327] Rusty L. Montgomery et al. “Histone deacetylases 1 and 2 control the progression of neural precursors to neurons during brain development”. In: *Proceedings of the National Academy of Sciences of the United States of America* 106.19 (2009), pp. 7876–7881.

[328] Vincent T. Cunliffe. “Histone deacetylase 1 is required to repress Notch target gene expression during zebrafish neurogenesis and to maintain the production of motoneurones in response to hedgehog signalling”. In: *Development* 131.12 (2004), pp. 2983–2995.

[329] Yongcheol Cho and Valeria Cavalli. “HDAC signaling in neuronal development and axon regeneration”. In: *Current Opinion in Neurobiology* 27 (2014), pp. 118–126.

[330] Anne Laurence Boutillier, Laura Tzeplaeff, and Luc Dupuis. “The dark side of HDAC inhibition in ALS”. In: *EBioMedicine* 41 (2019), pp. 38–39.

[331] Simona Sanna et al. “HDAC1 inhibition ameliorates TDP-43-induced cell death in vitro and in vivo”. In: *Cell Death and Disease* 11.5 (2020).

[332] Wei Zhi Wang et al. “High quality RNA from multiple brain regions simultaneously acquired by laser capture microdissection”. In: *BMC Molecular Biology* 10.1 (July 2009), p. 69.

[333] J. Gerard G. Borst and John Soria Van Hoeve. “The calyx of held synapse: From model synapse to auditory relay”. In: *Annual Review of Physiology* 74 (2012), pp. 199–224.

[334] Vladislav A. Petyuk et al. “The human brainome: Network analysis identifies HSPA2 as a novel Alzheimer’s disease target”. In: *Brain* 141.9 (2018), pp. 2721–2739.

[335] Alessandra Torraco et al. “ISCA1 mutation in a patient with infantile-onset leukodystrophy causes defects in mitochondrial [4Fe-4S] proteins”. In: *Human Molecular Genetics* 27.15 (2018), pp. 2739–2754.

[336] Andrew F. Russo. “Overview of Neuropeptides: Awakening the Senses?” In: *Headache* 57 (2017), pp. 37–46.

[337] Lisa M. Landino et al. “Repair of peroxynitrite damage to tubulin by the thioredoxin reductase system”. In: *Free Radical Biology and Medicine* 36.4 (Feb. 2004), pp. 497–506.

[338] L. Van Den Bosch et al. “The role of excitotoxicity in the pathogenesis of amyotrophic lateral sclerosis”. In: *Biochimica et Biophysica Acta - Molecular Basis of Disease* 1762.11-12 (2006), pp. 1068–1082.

[339] Massimo Stefani. “Generic cell dysfunction in neurodegenerative disorders: Role of surfaces in early protein misfolding, aggregation, and aggregate cytotoxicity”. In: *Neuroscientist* 13.5 (2007), pp. 519–531.

[340] Michael T. Lin and M. Flint Beal. “Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases”. In: *Nature* 443.7113 (2006), pp. 787–795.

[341] Nichole A. Reyes et al. “Blocking the mitochondrial apoptotic pathway preserves motor neuron viability and function in a mouse model of amyotrophic lateral sclerosis”. In: *Journal of Clinical Investigation* 120.10 (2010), pp. 3673–3679.

[342] Alexandro Gianforcaro and Mazen J. Hamadeh. “Vitamin D as a potential therapy in amyotrophic lateral sclerosis”. In: *CNS Neuroscience and Therapeutics* 20.2 (2014), pp. 101–111.

[343] William Camu et al. “Vitamin D confers protection to motoneurons and is a prognostic factor of amyotrophic lateral sclerosis”. In: *Neurobiology of Aging* 35.5 (2014), pp. 1198–1205.

[344] Michael I. Sandstrom and George V. Rebec. “Extracellular ascorbate modulates glutamate dynamics: Role of behavioral activation”. In: *BMC Neuroscience* 8 (2007).

[345] Fiona E. Harrison and James M. May. “Vitamin C function in the brain: vital role of the ascorbate transporter SVCT2”. In: *Free Radical Biology and Medicine* 46.6 (Mar. 2009), pp. 719–730.

[346] Tijs Vandoorne, Katrien De Bock, and Ludo Van Den Bosch. “Energy metabolism in ALS: an underappreciated opportunity?” In: *Acta Neuropathologica* 135.4 (Apr. 2018), pp. 489–509.

[347] Tetsuya Akiyama et al. “Aberrant axon branching via Fos-B dysregulation in FUS-ALS motor neurons”. In: *EBioMedicine* 45 (2019), pp. 362–378.

[348] Jia Liu and Fei Wang. “Role of neuroinflammation in amyotrophic lateral sclerosis: Cellular mechanisms and therapeutic implications”. In: *Frontiers in Immunology* 8 (2017).

[349] Cindy Maurel et al. “A role for SUMOylation in the Formation and Cellular Localization of TDP-43 Aggregates in Amyotrophic Lateral Sclerosis”. In: *Molecular Neurobiology* 57.3 (2020), pp. 1361–1373.

[350] Audrey Dangoumau et al. “Protein SUMOylation, an emerging pathway in amyotrophic lateral sclerosis”. In: *International Journal of Neuroscience* 123.6 (2013), pp. 366–374.

[351] Ashok Verma. “Altered RNA metabolism and amyotrophic lateral sclerosis”. In: *Annals of Indian Academy of Neurology* 14.4 (2011), pp. 239–244.

[352] Orietta Pansarasa et al. “RNA Metabolism and Therapeutics in Amyotrophic Lateral Sclerosis”. In: *Amyotrophic Lateral Sclerosis - Recent Advances and Therapeutic Challenges* (2020).

[353] Byung Woo Kim et al. “DNA damage accumulates and responses are engaged in human ALS brain and spinal motor neurons and DNA repair is activatable in iPSC-derived motor neurons with SOD1 mutations”. In: *Acta Neuropathologica Communications* 8.1 (2020).

[354] Nicolas Paquet et al. “HSSB1 (NABP2/OBFC2B) is regulated by oxidative stress”. In: *Scientific Reports* 6 (2016).

[355] Anastassios Vourekas et al. “The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing”. In: *Genes and Development* 29.6 (2015), pp. 617–629.

[356] Kaiqiang Fu et al. “Biological and RNA regulatory function of MOV10 in mammalian germ cells”. In: *BMC Biology* 17.1 (2019).

[357] Xiaobing Huang and Garry Wong. “An old weapon with a new function: PIWI-interacting RNAs in neurodegenerative diseases”. In: *Translational Neurodegeneration* 10.1 (2021).

[358] Keiko Tsuji Wakisaka et al. “Novel roles of Drosophila FUS and Aub responsible for piRNA biogenesis in neuronal disorders”. In: *Brain Research* 1708 (2019), pp. 207–219.

[359] Tino Prell and Julian Grosskreutz. “The involvement of the cerebellum in amyotrophic lateral sclerosis”. In: *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 14.7-8 (2013), pp. 507–515.

[360] Safa Al-Sarraj et al. “P62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS”. In: *Acta Neuropathologica* 122.6 (2011), pp. 691–702.

[361] Kirstin Zimmermann et al. “NOSTRIN: A protein modulating nitric oxide release and subcellular distribution of endothelial nitric oxide synthase”. In: *Proceedings of the National Academy of Sciences of the United States of America* 99.26 (2002), pp. 17167–17172.

[362] Derek A. Drechsel et al. “Nitric oxide-mediated oxidative damage and the progressive demise of motor neurons in ALS”. In: *Neurotoxicity Research* 22.4 (2012), pp. 251–264.

[363] M. Dragunow and R. Faull. “The use of c-fos as a metabolic marker in neuronal pathway tracing”. In: *Journal of Neuroscience Methods* 29.3 (1989), pp. 261–265.

[364] Yasuhiro Manabe et al. “Enhanced Fos expression in rat lumbar spinal cord cultured with cerebrospinal fluid from patients with amyotrophic lateral sclerosis”. In: *Neurological Research* 21.3 (1999), pp. 309–312.

[365] Jianhua Zhang et al. “C-Fos Regulates Neuronal Excitability and Survival”. In: *Nature Genetics* 30.4 (2002), pp. 416–420.

[366] Varun Rawat et al. “c-Fos Protects Neurons Through a Noncanonical Mechanism Involving HDAC3 Interaction: Identification of a 21-Amino Acid Fragment with Neuroprotective Activity”. In: *Molecular Neurobiology* 53.2 (2016), pp. 1165–1180.

[367] Fred A. Mettler. “Physiologic effects of bilateral simultaneous frontal lesions in the primate”. In: *Journal of Comparative Neurology* 81.2 (1944), pp. 105–136.

[368] R. S. Snider. “Connections of the Cerebral Cortex”. In: *Neurology* 15.1 (1965), pp. 99–99.

[369] K. S. Lashley and George Clark. “The cytoarchitecture of the cerebral cortex of ateles: A critical examination of architectonic studies”. In: *Journal of Comparative Neurology* 85.2 (1946), pp. 223–305.

[370] Juan C. Nañagas. “Anatomical studies on the motor cortex of Macacus rhesus”. In: *Journal of Comparative Neurology* 35.1 (1922), pp. 67–96.

[371] M. Illert, A. Lundberg, and R. Tanaka. “Integration in descending motor pathways controlling the forelimb in the cat - 2. Convergence on neurones mediating disynaptic cortico-motoneuronal excitation”. In: *Experimental Brain Research* 26.5 (1976), pp. 521–540.

[372] H. W. Yang and R. N. Lemon. “An electron microscopic examination of the corticospinal projection to the cervical spinal cord in the rat: Lack of evidence for cortico-motoneuronal synapses”. In: *Experimental Brain Research* 149.4 (2003), pp. 458–469.

[373] B. Alstermark, J. Ogawa, and T. Isa. “Lack of Monosynaptic Corticomotoneuronal EPSPs in Rats: Disynaptic EPSPs Mediated Via Reticulospinal Neurons and Polysynaptic EPSPs Via Segmental Interneurons”. In: *Journal of Neurophysiology* 91.4 (2004), pp. 1832–1839.

[374] Hitoshi Maeda et al. “Corticospinal axons make direct synaptic connections with spinal motoneurons innervating forearm muscles early during postnatal development in the rat”. In: *Journal of Physiology* 594.1 (2016), pp. 189–205.

[375] Rickye S. Heffner and R. Bruce Masterton. “The role of the corticospinal tract in the evolution of human digital dexterity”. In: *Brain, Behavior and Evolution* 23.3-4 (1983), pp. 165–183.

[376] Roger N. Lemon et al. “Direct and indirect pathways for corticospinal control of upper limb motoneurons in the primate”. In: *Progress in Brain Research* 143 (Jan. 2004), pp. 263–279.

[377] E. Palmer and P. Ashby. “Corticospinal projections to upper limb motoneurones in humans.” In: *The Journal of Physiology* 448.1 (1992), pp. 397–412.

[378] Alain Maertens De Noordhout et al. “Corticomotoneuronal synaptic connections in normal man. An electrophysiological study”. In: *Brain* 122.7 (1999), pp. 1327–1340.

[379] H. Kuypers. “The Descending Pathways to the Spinal Cord, their Anatomy and Function”. In: *Progress in Brain Research* 11.C (1964), pp. 178–202.

[380] Johannes Tigges, James G. Herndon, and Alan Peters. “Axon terminals on betz cell somata of area 4 in rhesus monkey throughout adulthood”. In: *The Anatomical Record* 232.2 (1992), pp. 305–315.

[381] William Bevan-Lewis. “On the comparative structure of the cortex cerebri”. In: *Brain* 1.1 (1878), pp. 79–96.

[382] Peter Ebinger. "A cytoarchitectonic volumetric comparison of the area gigantopyramidalis in wild and domestic sheep". In: *Anatomy and Embryology* 147.2 (1975), pp. 167–175.

[383] Cameron B. Johnson et al. "Neocortical neuronal morphology in the Siberian Tiger (Panthera tigris altaica) and the clouded leopard (Neofelis nebulosa)". In: *Journal of Comparative Neurology* 524.17 (2016), pp. 3641–3665.

[384] Wladimir Betz. "Anatomischer Nachweis zweier Gehirncentra". In: *Centralblatt für die medizinischen Wissenschaften* 12 (1874), pp. 578–580, 595–599.

[385] T. A. Yousry et al. "Localization of the motor hand area to a knob on the precentral gyrus. A new landmark". In: *Brain* 120.1 (1997), pp. 141–157.

[386] B. Boroojerdi et al. "Localization of the motor hand area using transcranial magnetic stimulation and functional magnetic resonance imaging". In: *Clinical Neurophysiology* 110.4 (1999), pp. 699–704.

[387] I. Ferrer and M. Perera. "Structure and nerve cell organisation in the cerebral cortex of the dolphin Stenella coeruleoalba a Golgi study - With special attention to the primary auditory area". In: *Anatomy and Embryology* 178.2 (1988), pp. 161–173.

[388] Rui Furutani. "Laminar and cytoarchitectonic features of the cerebral cortex in the Risso's dolphin (Grampus griseus), striped dolphin (Stenella coeruleoalba), and bottlenose dolphin (Tursiops truncatus)". In: *Journal of Anatomy* 213.3 (2008), pp. 241–248.

[389] Camilla Butti et al. "The neocortex of cetartiodactyls: I. A comparative Golgi analysis of neuronal morphology in the bottlenose dolphin (Tursiops truncatus), the minke whale (Balaenoptera acutorostrata), and the humpback whale (Megaptera novaeangliae)". In: *Brain Structure and Function* 220.6 (2015), pp. 3339–3368.

[390] Camilla Butti et al. "The Cerebral Cortex of the Pygmy Hippopotamus, Hexaprotodon liberiensis (Cetartiodactyla, Hippopotamidae): MRI, Cytoarchitecture, and Neuronal Morphology". In: *Anatomical Record* 297.4 (2014), pp. 670–700.

[391] Kaeley Anderson et al. "The morphology of supragranular pyramidal neurons in the human insular cortex: A quantitative golgi study". In: *Cerebral Cortex* 19.9 (2009), pp. 2131–2144.

[392] D. Kent Morest and R. R. Morest. "Perfusion-fixation of the brain with chrome-osmium solutions for the rapid Golgi method". In: *American Journal of Anatomy* 118.3 (1966), pp. 811–831.

[393] Tokuzo Kojima. "On the brain of the sperm whale (Physeter catodon L.)" In: *Scientific Reports of the Whales Research Institute* 6.3 (1951), pp. 49–72.

[394] V. S. Kesarev, L. I. Malofeyeva, and O. V. Trykova. "Ecological specificity of cetacean neocortex". eng. In: *Journal fur Hirnforschung* 18.6 (1977), pp. 447–460.

[395] Patrick R. Hof and Chet C. Sherwood. "Morphomolecular neuronal phenotypes in the neocortex reflect phylogenetic relationships among certain mammalian orders". In: *Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology* 287.1 (Nov. 2005), pp. 1153–1163.

[396] Patrick R. Hof and Estel Van Der Gucht. “Structure of the cerebral cortex of the humpback whale, Megaptera novaeangliae (Cetacea, Mysticeti, Balaenopteridae)”. In: *Anatomical Record* 290.1 (Jan. 2007), pp. 1–31.

[397] Patrick E. Hof, Rebecca Chanis, and Lori Marino. “Cortical complexity in cetacean brains”. In: *Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology* 287.1 (Nov. 2005), pp. 1142–1152.

[398] K. W.S. Ashwell, L. L. Zhang, and L. R. Marotte. “Cyto- and chemoarchitecture of the cortex of the tammar wallaby (Macropus eugenii): Areal organization”. In: *Brain, Behavior and Evolution* 66.2 (July 2005), pp. 114–136.

[399] Chet C Sherwood et al. “Histological Asymmetries of Primary Motor Cortex Predict Handedness in Chimpanzees (Pan troglodytes) CHET”. In: 503.4 (2007), pp. 525–537.

[400] Quentin Welniarz, Isabelle Dusart, and Emmanuel Roze. “The corticospinal tract: Evolution, development, and human disorders”. In: *Developmental Neurobiology* 77.7 (2017), pp. 810–829.

[401] Sandra A. Heldstab et al. “Manipulation complexity in primates coevolved with brain size and terrestriality”. In: *Scientific Reports* 6 (Apr. 2016).

[402] Leah A. Krubitzer and Adele M.H. Seelke. “Cortical evolution in mammals: The bane and beauty of phenotypic variability”. In: *Proceedings of the National Academy of Sciences of the United States of America* 109.Suppl.1 (June 2012), pp. 10647–10654.

[403] William D. Hopkins et al. “Evolution of the central sulcus morphology in primates”. In: *Brain, Behavior and Evolution* 84.1 (2014), pp. 19–30.

[404] W. J. Spain, P. C. Schwindt, and W. E. Crill. “Anomalous rectification in neurons from cat sensorimotor cortex in vitro”. In: *Journal of Neurophysiology* 57.5 (1987), pp. 1555–1576.

[405] P. C. Schwindt et al. “Slow conductances in neurons from cat sensorimotor cortex in vitro and their role in slow excitability changes”. In: *Journal of Neurophysiology* 59.2 (1988), pp. 450–467.

[406] P. C. Schwindt, W. J. Spain, and W. E. Crill. “Influence of anomalous rectifier activation on afterhyperpolarizations of neurons from cat sensorimotor cortex in vitro”. In: *Journal of Neurophysiology* 59.2 (1988), pp. 468–481.

[407] P. C. Schwindt, W. J. Spain, and W. E. Crill. “Long-lasting reduction of excitability by a sodium-dependent potassium current in cat neocortical neurons”. In: *Journal of Neurophysiology* 61.2 (1989), pp. 233–244.

[408] W. J. Spain, P. C. Schwindt, and W. E. Crill. “Two transient potassium currents in layer V pyramidal neurones from cat sensorimotor cortex.” In: *The Journal of Physiology* 434.1 (1991), pp. 591–607.

[409] W. J. Spain, P. C. Schwindt, and W. E. Crill. “Post-inhibitory excitation and inhibition in layer V pyramidal neurones from cat sensorimotor cortex.” In: *The Journal of Physiology* 434.1 (1991), pp. 609–626.

[410] P. C. Schwindt, W. J. Spain, and W. E. Crill. “Effects of intracellular calcium chelation on voltage-dependent and calcium-dependent currents in cat neocortical neurons”. In: *Neuroscience* 47.3 (1992), pp. 571–578.

[411] P. C. Schwindt, W. J. Spain, and W. E. Crill. "Calcium-dependent potassium currents in neurons from cat sensorimotor cortex". In: *Journal of Neurophysiology* 67.1 (1992), pp. 216–226.

[412] W. J. Spain. "Serotonin has different effects on two classes of Betz cells from the cat". In: *Journal of Neurophysiology* 72.4 (1994), pp. 1925–1937.

[413] E. D. Adrian and G. Moruzzi. "Impulses in the pyramidal tract". In: *The Journal of Physiology* 97.2 (1939), pp. 153–199.

[414] J. W. Lance and R. L. Manning. "Origin of the pyramidal tract in the cat". In: *The Journal of Physiology* 124.2 (May 1954), pp. 385–399.

[415] E. V. EVARTS. "Relation of Discharge Frequency To Conduction Velocity in Pyramidal Tract Neurons". In: *Journal of neurophysiology* 28 (1965), pp. 216–228.

[416] P. G. Guyenet and G. K. Aghajanian. "Excitation of neurons in the nucleus locus coeruleus by substance P and related peptides". In: *Brain Research* 136.1 (1977), pp. 178–184.

[417] K. Krnjević and J. W. Phillis. "Acetylcholine-sensitive cells in the cerebral cortex". In: *The Journal of Physiology* 166.2 (1963), pp. 296–327.

[418] J. W. Phillis and J. J. Limacher. "Excitation of cerebral cortical neurons by various polypeptides". In: *Experimental Neurology* 43.2 (1974), pp. 414–423.

[419] C. G. Phillips. "Intracellular Records From Betz Cells in the Cat". In: *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences* 41.1 (1956), pp. 58–69.

[420] Pierre Landry, André Labelle, and Martin Deschênes. "Intracortical distribution of axonal collaterals of pyramidal tract cells in the cat motor cortex". In: *Brain Research* 191.2 (1980), pp. 327–336.

[421] Kunitaro Takahashi, Kisou Kubota, and Masatake Uno. "Recurrent Facilitation in Cat Pyramidal Tract Cells". In: *Journal of Neurophysiology* 30.1 (1967), pp. 22–34.

[422] Hisao Suzuki and Yasuo Tukahara. "Recurrent inhibition of the betz cell". In: *Japanese Journal of Physiology* 13.4 (1963), pp. 386–398.

[423] K. C. Gatter, J. J. Sloper, and T. P.S. Powell. "An electron microscopic study of the termination of intracortical axons upon betz cells in area 4 of the monkey". In: *Brain* 101.3 (1978), pp. 543–553.

[424] C. STEFANIS and H. JASPER. "Recurrent Collateral Inhibition in Pyramidal Tract Neurons". In: *Journal of neurophysiology* 27 (1964), pp. 855–877.

[425] N. Ichinohe et al. "A voltage-gated potassium channel, Kv3.1b, is expressed by a subpopulation of large pyramidal neurons in layer 5 of the macaque monkey cortex". In: *Neuroscience* 129.1 (2004), pp. 179–185.

[426] David Soares et al. "Expression of Kv3.1b potassium channel is widespread in macaque motor cortex pyramidal cells: A histological comparison between rat and macaque". In: *Journal of Comparative Neurology* 525.9 (2017), pp. 2164–2174.

[427] Massimo Matelli, Giuseppe Luppino, and Giacomo Rizzolatti. "Patterns of cytochrome oxidase activity in the frontal agranular cortex of the macaque monkey". In: *Behavioural Brain Research* 18.2 (Nov. 1985), pp. 125–136.

[428] Todd M. Preuss and Jon H. Kaas. “Parvalbumin-like immunoreactivity of layer V pyramidal cells in the motor and somatosensory cortex of adult primates”. In: *Brain Research* 712.2 (1996), pp. 353–357.

[429] Dirk Hausen et al. “Pyramidal cells ensheathed by perineuronal nets in human motor and somatosensory cortex”. In: *NeuroReport* 7.11 (1996), pp. 1725–1729.

[430] Andrew T.N. Tebbenkamp et al. “The 7q11.23 Protein DNAJC30 Interacts with ATP Synthase and Links Mitochondria to Brain Development”. In: *Cell* 175.4 (2018), pp. 1088–1104.

[431] M-Marsel -M Mesulam and Changiz Geula. “Acetylcholinesterase-rich neurons of the human cerebral cortex: Cytoarchitectonic and ontogenetic patterns of distribution”. In: *Journal of Comparative Neurology* 306.2 (1991), pp. 193–220.

[432] Shogo Tominaga et al. “Immunocytochemical localization of α-, βI-, βII- and γ-subspecies of protein kinase C in the motor and premotor cortices of the rhesus monkey”. In: *Neuroscience Research* 16.4 (1993), pp. 275–286.

[433] Mark R. Cookson et al. “Poly(ADP-ribose) polymerase is found in both the nucleus and cytoplasm of human CNS neurons”. In: *Brain Research* 834.1-2 (1999), pp. 182–185.

[434] Haruhiko Akiyama et al. “Distribution of phosphate-activated glutaminase in the human cerebral cortex”. In: *Journal of Comparative Neurology* 297.2 (1990), pp. 239–252.

[435] J. P. Donoghue, R. J. Wenthold, and R. A. Altschuler. “Localization of glutaminase-like and aspartate aminotransferase-like immunoreactivity in neurons of cerebral neocortex”. In: *Journal of Neuroscience* 5.10 (1985), pp. 2597–2608.

[436] M. Otsuka et al. “Measurement of Γ-Aminobutyric Acid in Isolated Nerve Cells of Cat Central Nervous System”. In: *Journal of Neurochemistry* 18.2 (1971), pp. 287–295.

[437] A. Muñoz, T. M. Woods, and E. G. Jones. “Laminar and cellular distribution of AMPA, kainate, and NMDA receptor subunits in monkey sensory-motor cortex”. In: *Journal of Comparative Neurology* 407.4 (1999), pp. 472–490.

[438] T. L. Williams et al. “An immunocytochemical study of the distribution of AMPA selective glutamate receptor subunits in the normal human motor system”. In: *Neuroscience* 74.1 (1996), pp. 185–198.

[439] Pamela J. Shaw et al. “Low expression of GluR2 AMPA receptor subunit protein by human motor neurons”. In: *NeuroReport* 10.2 (1999), pp. 261–265.

[440] Sean G. Carriedo, Hong Z. Yin, and John H. Weiss. “Motor neurons are selectively vulnerable to AMPA/kainate receptor-mediated injury in vitro”. In: *Journal of Neuroscience* 16.13 (1996), pp. 4069–4079.

[441] Ingrid Moeller et al. “Distribution of AT4 receptors in the Macaca fascicularis brain”. In: *Brain Research* 712.2 (1996), pp. 307–324.

[442] J. Michael Crawford. “The sensitivity of cortical neurones to acidic amino acids and acetylcholine”. In: *Brain Research* 17.2 (1970), pp. 287–296.

[443] Albert Gjedde, William R. Bauer, and Dean F. Wong. “Neurokinetics: The dynamics of neurobiology in vivo”. In: *Neurokinetics: The Dynamics of Neurobiology in Vivo* (2011), pp. 1–343.

[444] Miguel Marín-Padilla. "Cajal-Retzius cells and the development of the neocortex". In: *Trends in Neurosciences* 21.2 (1998), pp. 64–71.

[445] S. Sasaki and S. Maruyama. "Immunocytochemical and ultrastructural studies of the motor cortex in amyotrophic lateral sclerosis". In: *Acta Neuropathologica* 87.6 (1994), pp. 578–585.

[446] Koichi Wakabayashi et al. "Lewy bodies in Betz cells of the motor cortex in a patient with Parkinson's disease". In: *Acta Neuropathologica* 105.2 (2003), pp. 189–192.

[447] K. Tsuchiya et al. "Constant involvement of the Betz cells and pyramidal tract in multiple system atrophy: A clinicopathological study of seven autopsy cases". In: *Acta Neuropathologica* 99.6 (2000), pp. 628–636.

[448] Kuniaki Tsuchiya et al. "Constant and severe involvement of Betz cells in corticobasal degeneration is not consistent with pyramidal signs: A clinicopathological study of ten autopsy cases". In: *Acta Neuropathologica* 109.4 (2005), pp. 353–366.

[449] Arnulf H. Koeppen and Joseph E. Mazurkiewicz. "Friedreich ataxia: Neuropathology revised". In: *Journal of Neuropathology and Experimental Neurology* 72.2 (Feb. 2013), pp. 78–90.

[450] K. Oyanagi et al. "Galactosialidosis: neuropathological findings in a case of the late-infantile type". In: *Acta Neuropathologica* 82.5 (Sept. 1991), pp. 331–339.

[451] Kuniaki Tsuchiya et al. "Pathological heterogeneity of the precentral gyrus in Pick's disease: A study of 16 autopsy cases". In: *Acta Neuropathologica* 112.1 (2006), pp. 29–42.

[452] L. CROME. "Absence of Betz cells in diffuse gliosis of the brain without demyelination". In: *The Journal of pathology and bacteriology* 63.4 (1951), pp. 746–748.

[453] Santiago Giménez-Roldán et al. "Historical setting and neuropathology of lathyrism: Insights from the neglected 1944 report by Oliveras de la Riva". In: *Journal of the History of the Neurosciences* 28.4 (2019), pp. 361–386.

[454] Michael C. Levin et al. "Autoimmunity due to molecular mimicry as a cause of neurological disease". In: *Nature Medicine* 8.5 (2002), pp. 509–513.

[455] Miguel Marin-Padilla. "Pyramidal cell abnormalities in the motor cortex of a child with Down's syndrome. A Golgi study". In: *Journal of Comparative Neurology* 167.1 (1976), pp. 63–81.

[456] S. Araki et al. "Neuropathological analysis in spinal muscular atrophy type II". In: *Acta Neuropathologica* 106.5 (2003), pp. 441–448.

[457] Kewei Huang and Yi Luo. "Adult spinal muscular athrophy. A report of four cases". In: *Journal of the Neurological Sciences* 61.2 (1983), pp. 249–259.

[458] K. Gierga et al. "Spinocerebellar ataxia type 6 (SCA6): Neurodegeneration goes beyond the known brain predilection sites". In: *Neuropathology and Applied Neurobiology* 35.5 (Oct. 2009), pp. 515–527.

[459] Franziska Hoche et al. "Spinocerebellar ataxia type 2 (SCA2): Identification of early brain degeneration in one monozygous twin in the initial disease stage". In: *Cerebellum* 10.2 (June 2011), pp. 245–253.

[460] Siew Ping Han et al. “Differential subcellular distributions and trafficking functions of hnRNP A2/B1 spliceoforms”. In: *Traffic* 11.7 (2010), pp. 886–898.

[461] Emanuele Buratti et al. “Nuclear factor TDP-43 can affect selected microRNA levels”. In: *FEBS Journal* 277.10 (2010), pp. 2268–2281.

[462] Clotilde Lagier-Tourenne, Magdalini Polymenidou, and Don W. Cleveland. “TDP-43 and FUS/TLS: Emerging roles in RNA processing and neurodegeneration”. In: *Human Molecular Genetics* 19.R1 (2010).

[463] Sami J. Barmada et al. “Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models”. In: *Nature Chemical Biology* 10.8 (2014), pp. 677–685.

[464] Magdalini Polymenidou et al. “Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43”. In: *Nature Neuroscience* 14.4 (2011), pp. 459–468.

[465] Eun Joo Lee et al. “Identification of piRNAs in the central nervous system”. In: *Rna* 17.6 (2011), pp. 1090–1099.

[466] Felix Geser et al. “Evidence of multisystem disorder in whole-brain map of pathological TDP-43 in amyotrophic lateral sclerosis”. In: *Archives of Neurology* 65.5 (2008), pp. 636–641.

[467] Yue Shan Piao et al. “Neuropathology with clinical correlations of sporadic amyotrophic lateral sclerosis: 102 Autopsy cases examined between 1962 and 2000”. In: *Brain Pathology* 13.1 (2003), pp. 10–22.

[468] Johannes Brettschneider et al. “Stages of pTDP-43 pathology in amyotrophic lateral sclerosis”. In: *Annals of Neurology* 74.1 (2013), pp. 20–38.

[469] Mathieu Clément-Ziza et al. “Stabilization of RNA during laser capture microdissection by performing experiments under argon atmosphere or using ethanol as a solvent in staining solutions”. In: *Rna* 14.12 (2008), pp. 2698–2704.

[470] H. Auer et al. “The effects of frozen tissue storage conditions on the integrity of RNA and protein”. In: *Biotechnic and Histochemistry* 89.7 (2014), pp. 518–528.

[471] Ove J.R. Gustafsson, Georgia Arentz, and Peter Hoffmann. “Proteomic developments in the analysis of formalin-fixed tissue”. In: *Biochimica et Biophysica Acta - Proteins and Proteomics* 1854.6 (June 2015), pp. 559–580.

[472] Sameh Magdeldin and Tadashi Yamamoto. “Toward deciphering proteomes of formalin-fixed paraffin-embedded (FFPE) tissues”. In: *Proteomics* 12.7 (Apr. 2012), pp. 1045–1058.

[473] Judy Toews et al. “Mass spectrometric identification of formaldehyde-induced peptide modifications under in vivo protein cross-linking conditions”. In: *Analytica Chimica Acta* 618.2 (June 2008), pp. 168–183.

[474] Tue Bjerg Bennike et al. “Comparing the proteome of snap frozen, RNAlater preserved, and formalin-fixed paraffin-embedded human tissue samples”. In: *EuPA Open Proteomics* 10 (2016), pp. 9–18.

[475] Xiaogang Jiang et al. “Development of efficient protein extraction methods for shotgun proteome analysis of formalin-fixed tissues”. In: *Journal of Proteome Research* 6.3 (Mar. 2007), pp. 1038–1047.

[476] Tong Guo et al. “Proteome analysis of microdissected formalin-fixed and paraffin-embedded tissue specimens”. In: *Journal of Histochemistry and Cytochemistry* 55.7 (July 2007), pp. 763–772.

[477] Darryl Erik Palmer-Toy et al. “Efficient method for the proteomic analysis of fixed and embedded tissues”. In: *Journal of Proteome Research* 4.6 (Nov. 2005), pp. 2404–2411.

[478] S. I. Hwang et al. “Direct cancer tissue proteomics: A method to identify candidate cancer biomarkers from formalin-fixed paraffin-embedded archival tissues”. In: *Oncogene* 26.1 (Jan. 2007), pp. 65–76.

[479] Robert W. Sprung et al. “Equivalence of protein inventories obtained from formalin-fixed paraffin-embedded and frozen tissue in multidimensional liquid chromatography-tandem mass spectrometry shotgun proteomic analysis”. In: *Molecular and Cellular Proteomics* 8.8 (2009), pp. 1988–1998.

[480] Marshall S. Scicchitano et al. “Protein extraction of formalin-fixed, paraffin-embedded tissue enables robust proteomic profiles by mass spectrometry”. In: *Journal of Histochemistry and Cytochemistry* 57.9 (Sept. 2009), pp. 849–860.

[481] Ruilin Wu et al. “NanoTPOT: Enhanced Sample Preparation for Quantitative Nanoproteomic Analysis”. In: *Analytical Chemistry* 92.9 (Apr. 2020), pp. 6235–6240.

[482] Wolfram Liebermeister et al. “Visual account of protein investment in cellular functions”. In: *Proceedings of the National Academy of Sciences of the United States of America* 111.23 (2014), pp. 8488–8493.

[483] Qin Liu et al. “Addressing the batch effect issue for LC/MS metabolomics data in data preprocessing”. In: *Scientific Reports* 10.1 (2020).

[484] Andreas Schroeder et al. “The RIN: An RNA integrity number for assigning integrity values to RNA measurements”. In: *BMC Molecular Biology* 7 (2006).