

Identification and functional validation of pathogenic small nucleotide variants in patients with craniosynostosis

Rebecca S. Tooze

Exeter College, University of Oxford

DPhil Medical Sciences

Supervisors: Professor Andrew O. M. Wilkie, Associate Professor Stephen R. F. Twigg, Associate Professor Jenny Taylor

Word Count: 49,920

Submitted: 28th April 2023









EXETER
COLLEGE
OXFORD

Acknowledgements

This work could not have been possible without the guidance and mentorship of my supervisors, Professor Andrew Wilkie and Associate Professor Stephen Twigg who have provided unparalleled technical advice and challenged my thinking throughout. I am immensely grateful for the opportunities they have allowed me whilst undertaking my DPhil, including undertaking an 8-week bioinformatics course, a 7-week internship with Elsevier, opportunities to speak at international conferences, and an ability to learn and experiment with a variety of techniques. Despite Covid-19 interrupting a large portion of the first and second year of my DPhil, both Andrew and Steve were extremely supportive and facilitated my progress in what was a very unsettling time. I am also hugely grateful to my third supervisor, Associate Professor Jenny Taylor, for her continued technical advice. It has been invaluable to gain her expertise and insight from members of her research group who were often able to provide a solution to my bioinformatic struggles.

I am also incredibly thankful to Dr Eduardo Calpena and Dr Hana Mlcochova who helped get me started with many of the techniques used throughout my DPhil – I am incredibly grateful for both your time and patience. I also wish to thank Mr Yang Pei, who was extremely tolerant to my continued bioinformatic questions in the early days of my DPhil. I have been fortunate to be a part of an incredibly collaborative and welcoming lab, who have all provided advice, expertise, support, and laughter throughout the process – I therefore also wish to thank Dr Dagmara Korona, Mr Isaac Walton, Professor Anne Goriely, Dr Katherine Wood, Dr Dale Maxwell, and Spencer Tong. Additionally, I am enormously thankful to Ms Jill Brown, Ms Jackie Sloane-Stanley, Mr Philip Hublitz, Mr Kevin Clark, Mr Oliver Burns, Dr Jana Koth, and Dr Fena Ochs, for their skills and guidance throughout many of my projects.

I also wish to thank my thesis committee supervisors, Professor Rajesh Thakker and Dr Nicola Whiffen, who were able to offer clinical, wet-lab, and computational expertise throughout the process. Thank you both for the time you spent listening to me talk and for continuing to offer a fresh perspective on my data.

My sincere thanks to all of our collaborators for their assistance in clinical information and getting blood and DNA samples from patients. I also wish to express my gratitude to all patients and their families who have been involved in my studies – their enthusiasm towards our research and willingness to partake has really made all the difference.

This work could not have been possible without my scholarship from the Radcliffe Department of Medicine, in partnership with the Medical Research Council (MRC) and the Usher Cunningham Scholarship (Exeter College, University of Oxford). Additionally, I am grateful for funding I received from the Oxford Biomedical Research Centre (BRC) which paid for a biomedical data science training programme for 8-weeks; I was also thankful to receive an additional award from Exeter College, Oxford, which facilitated my attendance and presentation at the European Society of Human Genetics conference in Vienna (2022). I have also been fortunate to benefit from funding from my supervisors via the MRC, Wellcome Trust, VTCT Foundation, and Oxford BRC.

Last, I would like to thank my family and friends, but particularly my fiancé (Mr Michael Trim), mother (Mrs Vivien Tooze) and father (Mr Derek Tooze) for their unconditional love and support. I am immensely grateful to you for your encouragement, motivation, and patience, for always making me smile, keeping me grounded, and for listening to me chat endlessly about science – I couldn't have done this without you!

Statement of originality

I declare that this thesis constitutes original work. However, owing to the nature of experimental projects, some workflows and experiments were conducted by my colleagues. In chapter 3, the 100kGP data were mined jointly with Dr Eduardo Calpena and Mr Yang Pei, but variants identified through my independent analyses are discussed. The workflow used to analyse compound heterozygous variants was written by Dr Edoardo Giacopuzzi (section 3.2.3.2). In chapter 4, an initial subset of 297 samples underwent library preparation by Dr Eduardo Calpena for use in my targeted capture analysis (section 4.2.3, table 4.2). Dr Simon McGowan facilitated the upload of all variants analysed using amplimap onto CranioVar. In chapter 5, the variant identified in *SPRY1* (also in chapter 3) was originally flagged by Dr Eduardo Calpena in the 100kGP data. In chapter 6, targeted resequencing was initially conducted by Dr Kerry Miller (section 6.2.1) with additional samples screened by me in the resequencing analysis (chapter 4). Dr Sigrid Swagemakers identified the first *de novo PRRX1* variant in a patient with craniosynostosis (section 6.1). Collaborators (Dr Andrew Timberlake, Dr Tiong Tan, Dr Charlotte Ockeleon, and Dr Corinne Collet) identified additional patients with variants in *PRRX1* and contributed dideoxy sequencing traces (figure 6.2). Four substitutions in a vector containing *Prrx1* were originally made by Dr Kerry Miller (R96M, D54A, F113L, A147T) with an additional mutagenesis construct made by Associate Professor Stephen Twigg (R150H). Dr Fena Ochs facilitated image analysis of immunofluorescence slides using a ScanR microscope (figure 6.10b). In chapter 7, the WIMM Genome Engineering Facility aided the design of the CRISPR guide RNA and donor, while the WIMM Transgenics Facility provided mouse embryonic stem cells and conducted blastocyst injections for generation of chimeras (section 7.2.9.2 and 7.2.10). The WIMM flow cytometry facility (namely Mr Kevin Clark but also Mr Paul Sopp and Mr Craig Waugh) facilitated single cell sorting of GFP+ mouse embryonic stem cells after transfection (section 7.2.9.2). Additionally, general animal welfare and husbandry was managed by the John Radcliffe Biomedical Services staff (section 7.2.10).

Publications

Content from five of my thesis chapters (chapters 3–6 and 8) have in parallel been accepted for publication. Three of these are joint first-author publications; however, information included in this thesis constitutes original material that was written or produced by me. Statements about the contributions of each joint first-author are detailed below and in the statement of originality.

First/ Joint-first (*) author

Chapter 3

Hyder, Z.,* Calpena, E.,* Pei, Y.,* **Tooze, R. S.,*** Brittain, H., Twigg, S. R. F., Cilliers, D., Morton, J. E. V., Mccann, E., Weber, A., Wilson, L. C., Douglas, A. G. L., Mcgowan, R., Need, A., Bond, A., Tavares, A. L. T., Thomas, E. R. A., Genomics England Research, C., Hill, S. L., Deans, Z. C., Boardman-Pretty, F., Caulfield, M., Scott, R. H. & Wilkie, A. O. M. 2021. Evaluating the performance of a clinical genome sequencing program for diagnosis of rare genetic disease, seen through the lens of craniosynostosis. *Genet Med,* 23**,** 2360-2368.

> This publication was largely written by Dr Zerin Hyder and has therefore not been included in my thesis. Dr Eduardo Calpena, Mr Yang Pei and myself were involved in the formal analysis of data from the 100kGP, writing of the methods, and review and editing of the manuscript. In chapter 3 I have acknowledged, where appropriate, contributions to the total uplift in genetic diagnoses achieved by my colleagues. However, this chapter has been re-written to include only data and methods completed independently.

Chapter 4 and 8

Tooze, R. S., Calpena, E., Weber, A., Wilson, L. C., Twigg, S. R. F., Wilkie, A. O. M. 2023. Review of recurrently mutated genes in craniosynostosis supports expansion of diagnostic gene panels. *Genes, 14*, 615.

Results from my resequencing analysis (chapter 4) were included in a publication aiming to address the need to continually update PanelApp inline with recent publications. The analysis of the literature and identification of additional genes that should be promoted to Green status on PanelApp (chapter 8), alongside the figures, tables, and text produced were my own and have therefore been included, in full, in my thesis.

Chapter 5

Tooze, R. S.,* Calpena, E.,* Twigg, S. R. F., D'arco, F., Genomics England Research Consortium, Wakeling, E. L. & Wilkie, A. O. M. 2022. Craniosynostosis, inner ear, and renal anomalies in a child with complete loss of SPRY1 (sprouty homolog 1) function. *J Med Genet*. Epub ahead of print: 21 Dec 2022

Chapter 5 refers to the identification of a homozygous nonsense variant in Spry-1, which was originally found by Dr Eduardo Calpena from the 100kGP data. Dr Eduardo Calpena also taught me how to conduct western blot and RT-PCR analyses of protein and RNA and was therefore included as a joint first-author on this publication. However, all experimental work performed and used to generate figures and tables in the publication were completed independently. Additionally, the text included in this publication were written by me and have therefore been included in their entirety in chapter 5.

Chapter 6

Tooze, R. S.,* Miller, K. A.,* Swagemakers, S. M. A.,* Calpena, E., Mcgowan, S. J., Boute, O., Collet, C., Johnson, D., Laffargue, F., Leeuw, N. D., Morton, J. V., Peter Noons, Ockeloen, C. W., Phipps, J. M., Tan, T. Y., Timberlake, A. T., Vanlerberghe, C., Wall, S. A., Weber, A., Wilson, L. C., Zackai, E. H., Mathijssen, I. M. J., Twigg, S. R. F. & Wilkie, A. O. M. 2023b. Pathogenic variants in the paired-related homeobox 1 gene (PRRX1) cause craniosynostosis with incomplete penetrance. *Genet Med*. Accepted: 10 Apr 2023.

Content from chapter 6 was included in a publication that has been accepted in Genetics in Medicine. In this publication I am joint first-author with Dr Kerry Miller and Dr Sigrid Swagemakers. Dr Sigrid Swagemakers identified the first *PRRX1* variant in a patient with craniosynostosis, which led to the screening of unsolved patients to identify additional cases by Dr Kerry Miller. In this thesis chapter, the text has been re-written to reflect what was already known prior to my involvement in the project. The experimental work completed by Dr Kerry Miller, including preparation of plasmids and resequencing analysis by PCR have been acknowledged. However, all experimental work included in the publication, and this thesis, to generate figures and tables were produced by me; additionally, any text included in my thesis for chapter 6 is original work.

Co-author

- Lippincott, M. F., Xu, W., Smith, A. A., Miao, X., Lafont, A., Shennib, O., Farley, G. J., Sabbagh, R., Delaney, A., Stamou, M., Plummer, L., Salnikov, K., Georgopoulos, N. A., Mericq, V., Quinton, R., Mau-Them, F. T., Nambot, S., Hamad, A., Brittain, H., **Tooze, R. S.**, Calpena, E., Wilkie, A. O. M., Willems, M., Crowley, W. F., Balasubramanian, R., Lamarche-Vane, N., Davis, E. E. & Seminara, S. B. 2022. The p190 RhoGAPs, ARHGAP35, and ARHGAP5 are implicated in GnRH neuronal development: Evidence from patients with idiopathic hypogonadotropic hypogonadism, zebrafish, and in vitro GAP activity assay. *Genet Med,* 24**,** 2501-2515.

Conference oral presentations

- ESHG 2021 – Evaluating the performance of a clinical genome sequencing programme for diagnosis of rare genetic disease, seen through the lens of craniosynostosis.
- ESHG 2022 – Pathogenic variants in the paired-related homeobox 1 (PRRX1) gene are associated with craniosynostosis.

Table of Contents

Identification and functional validation of pathogenic small nucleotide variants in patients with craniosynostosis 1

Acknowledgements 2

Statement of originality 4

Publications 5

First/ Joint-first (*) author 5

Chapter 3 5

Chapter 4 and 8 6

Chapter 5 6

Chapter 6 7

Co-author 8

Conference oral presentations 8

Table of Contents 9

Figures and Tables 16

Abbreviations 18

Abstract 22

Chapter 1 23

An introduction to craniosynostosis, its aetiology, and the identification of pathogenic variants 23

1.1 General introduction 24

1.2 Anatomy and craniosynostosis 25

1.3 Non-syndromic craniosynostosis 28

1.3.1 Sagittal craniosynostosis 29

1.3.2 Metopic craniosynostosis 30

1.3.3 Coronal craniosynostosis 31

1.3.4 Lambdoid craniosynostosis 32

1.3.5 Multi-suture synostosis 33

1.4 Syndromic craniosynostosis 33

1.4.1 The first described craniosynostosis syndrome 34

1.4.2 Other common syndromes with a monoallelic cause 34

1.4.3 Recessive syndromes 35

1.5 Environmental causes 36

1.6 The importance of the sutural mesenchyme 37

1.6.1 *Gli1*+ expressing stem cells 37

1.6.2 *Axin2* as a marker of stemness **38**
1.6.3 The requirement for PRRX1 within the developing suture **38**
1.6.4 CTSK marks undifferentiated stem cells **39**
1.7 Suture development during embryogenesis ***39***
1.7.1 Stem cell specification and migration **39**
1.7.2 Lineage commitment **40**
1.7.3 Boundary formation and integrity **40**
1.7.4 Osteogenic proliferation/differentiation **41**
1.7.4.1 The hedgehog signalling pathway 41
1.7.4.2 The FGF/FGFR signalling axis 43
1.7.5 Resorption and homeostasis **46**
1.8 Genes and pathways implicated in craniosynostosis ***47***
1.8.1 Developmental genes and their role in craniosynostosis **47**
1.8.2 The long tail of rarer genetic diagnoses **48**
1.9 Next generation sequencing ***50***
1.10 Strategies for analysing WGS data ***52***
1.10.1 Protein coding variants **52**
1.10.2 Segregation patterns **53**
1.10.3 Identifying *de novo* variation **54**
1.11 Research aims ***55***
Chapter 2 ..... ***56***
General methodology ***56***
2.1 DNA Extraction **57**
2.2 Dideoxy-sequencing **57**
2.3 MiSeq Next Generation Sequencing **58**
2.4 Whole Genome Sequencing **58**
2.5 Centrifugation **59**
2.6 Cell Culture **59**
2.6.1 Lymphoblastoid Cell Culture 59
2.6.2 Fibroblast Cells 59
2.6.3 Adherent Cell Culture 60
2.7 Western Blot **60**
2.8 Reverse transcriptase PCR (RT-PCR) **61**
2.9 Quantitative real-time PCR (qPCR) **61**
Chapter 3 ..... ***62***
Increasing the diagnostic yield in craniosynostosis through systematic mining of the 100,000 Genomes Project Data ***62***

3.1 Introduction *63*
3.1.1 The UK 100,000 Genomes Project (100kGP) **63**
3.1.2 Aims of this chapter **64**
3.2 Methods *65*
3.2.1 Craniosynostosis disease cohort **65**
3.2.2 The GE Interpretation Pipeline **66**
3.2.2.1 Tiering Data 66
3.2.2.2 Exomiser 66
3.2.3 Independent Bioinformatic Analyses **67**
3.2.3.1 Recessive/ X-linked Analyses 67
3.2.3.2 Compound heterozygous variants 68
3.2.3.3 Families with more than one affected individual 68
3.2.3.4 DNMs 68
3.2.3.5 Exomiser 69
3.2.3.6 Indels 69
3.2.3.7 Cryptic splice site variants 70
3.2.3.8 UTR variants 70
3.2.3.9 Burden Analysis 70
3.2.4 Validation of a variant in *BCL11B* **71**
3.2.4.1 Deep Sequencing 71
3.3 Results *74*
3.3.1 Homozygosity/hemizygosity: A homozygous variant in *TMCO1* in a family with reported consanguinity **74**
3.3.1.1 *TMCO1* 74
3.3.2 Compound heterozygous variants in *SPTBN1* **75**
3.3.3 Analysis of segregating variants in multiply affected families **75**
3.3.4 DNM analysis: four missing diagnoses and one candidate disease gene **76**
3.3.4.1 *PTCH1* 77
3.3.4.2 *TCF12* 77
3.3.5 Analysis of the top Exomiser hits identifies a variant in *SMAD2* **77**
3.3.5.1 *SMAD2* 78
3.3.6 Indels are frequently missed from the GE/GMC triaged bioinformatics pipelines **78**
3.3.6.1 *ARID1B* 79
3.3.6.2 *BCL11B* 79
3.3.6.3 *MEGF8* 80
3.3.6.4 *MMP21* 80
3.3.7 No likely pathogenic UTR or splicing variants were identified **81**
3.3.8 Two candidate disease genes were identified through a burden analysis of the craniosynostosis cohort and the wider rare disease cohort **81**
3.4 Discussion *83*
3.4.1 Recessive analysis **83**
3.4.1.1 *TMCO1* 84
3.4.1.2 *SPTBN1* 85
3.4.2 *De novo* variant analysis **86**

3.4.2.1 A pathogenic heterozygous frameshift variant in *PTCH1* 87
3.4.2.2 A *de novo TCF12* variant in one member of a multiply affected family 88
3.4.2.3 *ANKRD28* was highlighted as a candidate disease gene for further scrutiny 88

3.4.3 Analysis of pathogenic indels 89
3.4.3.1 *MEGF8* 89
3.4.3.2 *ARID1B* 91
3.4.3.3 *BCL11B* 91
3.4.3.4 *MMP21* 91
3.4.3.5 Difficulties of identifying indel variants within the 100kGP cohort 92

3.4.4 Exomiser 93

3.4.5 Burden Analysis 94
3.4.5.1 *HSP90AA1* 94
3.4.5.2 *TFCP2* 95

3.4.6 Diagnostic yield 96

Chapter 4 97

Panel-based sequencing for the identification of SNVs and CNVs in a cohort of 617 patients with craniosynostosis 97

4.1 Introduction 98

4.2 Methods 101

4.2.1 Probe design 101

4.2.2 Unsolved craniosynostosis cases 101

4.2.3 Targeted capture next-generation sequencing 102

4.2.4 Identification of SNVs 102

4.2.5 Coverage analysis 103

4.2.6 RT-PCR 104

4.3 Results 104

4.3.1 Sequencing metrics 104

4.3.2 Six likely pathogenic SNVs: *ALX4*, *MSX2*, *PRRX1*, *SMAD3*, and two *SOX6* 104

4.3.3 Candidate variants 113
4.3.3.1 *ALPL* 113
4.3.3.2 *ALX4* 114
4.3.3.3 *LMX1B* 115
4.3.3.4 *MSX2* 116
4.3.3.5 *NFIA* 117
4.3.3.6 *PDGFRA* 118
4.3.3.7 *SMAD3* 119
4.3.3.8 *TFCP2* 119

4.3.4 Four CNVs were identified: Two *PRRX1* deletions, and a *PDGFRA* and *SPRY1* copy number gain 125

4.4 Discussion 131

4.4.1 *PRRX1* 131

4.4.2 *SOX6* 131

4.4.3 *ALX4* **133**
4.4.4 *MSX2* **136**
4.4.5 *SMAD3* **137**
4.4.6 Candidate variants **137**
4.4.7 Success of targeted resequencing for genetic discovery **138**
Chapter 5 ..... ***140***
Complete loss of Spry-1 (sprouty homolog 1) function in a child with craniosynostosis, inner ear, and renal anomalies ***140***
5.1 Introduction ***141***
5.1.1 Case report **145**
5.2 Methods ***148***
5.2.1 Genetic investigation and variant filtering of whole-genome sequencing data **148**
5.2.2 Functional analysis of cDNA and protein **148**
5.2.2.1 RT-PCR 148
5.2.2.2 Next Generation Sequencing (MiSeq) 149
5.2.2.3 Western blot 149
5.2.3 Targeted resequencing of *SPRY1* **150**
5.3 Results ***150***
5.3.1 Molecular analysis **150**
5.3.2 Functional validation of the escape from nonsense-mediate decay **153**
5.4 Discussion ***156***
Chapter 6 ..... ***159***
Pathogenic variants in the paired-related homeobox 1 (PRRX1) gene cause craniosynostosis ***159***
6.1 Introduction ***160***
6.2 Methods ***165***
6.2.1 Targeted sequencing of *PRRX1* **165**
6.2.2 Isolation of breakpoints in two patients with deletions in *PRRX1* **166**
6.2.3 Plasmid construction **168**
6.2.3.1 *PRRX1a* missense constructs 168
6.2.3.2 pGL2-Luciferase-TN-C construct 169
6.2.4 Dual luciferase reporter assay **170**
6.2.5 Western Blot **171**
6.2.6 Immunofluorescence **171**
6.3 Results ***172***
6.3.1 Identification of rare heterozygous single nucleotide or indel variants of *PRRX1* in craniosynostosis **172**
6.3.1.1 Resequencing analysis 172

6.3.1.2 Coverage analysis 173
6.3.1.3 Additional cases through collaboration 174
6.3.2 A dual luciferase assay was unable to establish a pathogenic mechanism for any missense variant 180
6.3.3 PRRX1 missense variants affect nuclear localisation 188
6.3.4 Clinical phenotype in individuals with heterozygous *PRRX1* loss-of-function variants 191
6.4 Discussion 195
Chapter 7 202
A novel Carpenter syndrome-like phenotype associated with a homozygous variant in PTCH1 202
7.1 Introduction 203
7.2 Methods 208
7.2.1 Patient-derived lymphoblastoid and fibroblast cells 208
7.2.2 Variant confirmation from gDNA 208
7.2.3 RT-PCR 208
7.2.4 WGS 208
7.2.5 Rare disease analysis – 100kGP 209
7.2.6 Western blot 209
7.2.7 qPCR 209
7.2.8 Plasmid construction 210
7.2.9 CRISPR targeting 212
7.2.9.1 Construction of the pX458-SpCas9-Ptch1 construct 212
7.2.9.2 Culturing mESCs 213
7.2.9.3 Colony screening 214
7.2.10 Mouse model 215
7.3 Results 215
7.3.1 Case Report (III-1) 215
7.3.2 Case Report (III-2) 216
7.3.3 RT-PCR did not identify any splicing defects in *RAB23* 218
7.3.4 WGS identified a single homozygous missense variant in PTCH1 218
7.3.5 Analysis of patient RNA and protein from lymphoblastoid and fibroblast cells 227
7.3.6 A homozygous PTCH1 missense variant in a patient with intellectual disability 227
7.3.7 Cohort screening of unsolved patients 227
7.3.8 Assessment of the effect of p.(Phe434Ser) in a PTCH1 expression vector 228
7.3.9 Establishing a mouse model of p.(Phe434Ser) 229

7.4 Discussion *231*
7.4.1 Proposed pathophysiological mechanism 231
7.4.2 Phenotype considerations 233
7.4.3 A second homozygous missense variant in PTCH1 235
7.4.4 Conclusion 236
Chapter 8 ..... *238*
Conclusions and Final Remarks *238*
8.1 Identification of diagnostic and candidate variants *239*
8.1.1 Analysis of the 100kGP craniosynostosis cohort 239
8.1.2 Resequencing analysis 240
8.1.3 Avenues for further research 244
8.2 Functional validation of variants in three genes *244*
8.2.1 *SPRY1* 244
8.2.2 *PRRX1* 245
8.2.3 *PTCH1* 247
8.3 Final remarks *248*
References *249*

Figures and Tables

Figure 1.1 Schematic of a new-born skull……**26**
Figure 1.2 Schematic of the human skull and craniofacial skeleton……**27**
Figure 1.3 An illustration of intramembranous ossification……**28**
Figure 1.4 Schematic representation of skull deformities arising from the premature fusion of individual sutures……**30**
Figure 1.5 Schematic of the FGF/FGFR signalling axis……**44**
Figure 1.6 A schematic of the main pathways involved in suture development……**48**
Figure 3.1 Filtering strategies and workflows used in bioinformatic analyses……**73**
Table 3.1 Diagnostic variants identified within the 100kGP……**82**
Table 3.2 Candidate genes and the associated variants identified through independent analyses of the 100kGP craniosynostosis cohort/CranioDB……**82**
Table 4.1 Category of genes included in the resequencing panel……**101**
Table 4.2 Samples included in the resequencing analysis by diagnosis……**108**
Figure 4.1 Sequencing confirmation from IGV and dideoxy sequencing……**109**
Figure 4.2 RT-PCR analysis of aberrant splicing defects in *SMAD3*……**110**
Table 4.3 Diagnostic SNVs from 617 samples in the resequencing analysis……**111**
Table 4.4 Candidate SNVs identified in the resequencing analysis which were confirmed by dideoxy-sequencing……**121**
Figure 4.3 Boxplot of normalised gene coverage for included samples (*n* = 479)……**126**
Table 4.5 Identified copy number variants from 479 samples……**127**
Figure 4.4 Boxplots of the normalised exon and probe coverage for *PRRX1*……**128**
Figure 4.5 Boxplot of the normalised exon and probe coverage for *PDGFRA*……**129**
Figure 4.6 Boxplot of normalised probe coverage for *SPRY1*……**130**
Figure 4.7 *ALX4* variants identified in patients with craniosynostosis and parietal foramina……**135**
Figure 5.1 Inhibitory mechanism of Sprouty proteins on the RAS-MAPK pathway….**142**
Figure 5.2 Clinical features of proband II-1 homozygous for *SPRY1* nonsense variant……**147**
Figure 5.3 Family pedigree and homozygosity analysis of WGS data from the proband II-1……**151**
Table 5.1 Variants prioritised after filtering by allele frequency and CADD score……**152**
Figure 5.4 *SPRY1* exon structure and predicted consequence of p.(Leu27)* nonsense variant……**154**

Figure 5.5 Identification of *SPRY1* nonsense variant and effect on expression of RNA and protein……...**155**
Table 5.2 Deep sequencing of cDNA from both parents and the affected child.....…...**155**
Figure 6.1 Structure of *PRRX1* and binding of homeodomain proteins to DNA……...**162**
Table 6.1 Phenotypic composition of cohorts of unsolved CRS probands……..……**167**
Figure 6.2 Pedigree figures and dideoxy sequencing confirmation of each variant….**175**
Figure 6.3 Structure, conservation, and variants in *PRRX1*……**176**
Figure 6.4 Analysis of coverage of *PRRX1* in a screen of 479 unsolved craniosynostosis patients……..……..….**177**
Figure 6.5 Analysis of the breakpoints of partial deletions including *PRRX1*……...**178**
Table 6.2 Rare heterozygous *PRRX1* variants identified in this and previous studies……..…. **179**
Figure 6.6 Dual luciferase assay using C2C12 cells transfected with wildtype and mutant *Prrx1* constructs and a TN-C luciferase reporter plasmid……..…….....**183**
Figure 6.7 Optimisation of dual luciferase assays using new TN-C constructs……...**184**
Figure 6.8 Relative luciferase activity of Prrx1 constructs……...**185**
Figure 6.9 Protein levels and relative luciferase activity of wildtype and mutant Prrx1 constructs……**187**
Figure 6.10 Immunofluorescence analysis of PRRX1 homeodomain missense substitutions……**190**
Table 6.3 Subjects with craniosynostosis analysed for rare, deleterious *PRRX1* variants in three screens…….....……**193**
Figure 6.11 Pre-operative clinical presentation of craniosynostosis in patients with *PRRX1* missense or loss-of-function variants……**194**
Figure 7.1 Schematic illustration of the interaction of PTCH1 with SHH and SMO....**205**
Figure 7.2 Phenotype of the index patient (III-1).....**217**
Figure 7.3 RT-PCR analysis of *RAB23*.....**220**
Figure 7.4 Pedigree and dideoxy sequence confirmation of NM_000264.5: c.1301T>C…….....**221**
Figure 7.5 Conservation of variant residues.....**222**
Figure 7.6 Structure and position of PTCH1 Phe434.....**223**
Table 7.1 Variants remaining after filtering WGS data.....**224**
Figure 7.7 Protein analysis of lymphoblastoid patient cells.....**228**
Figure 7.8 PCR analysis of two targeted mE14 clones: 2D1 and 2B6.....**230**
Figure 7.9 Interaction of Phe434 with neighbouring residues.....**232**
Table 8.1 Genes with sufficient evidence to be updated to Green PanelApp status..**242**

Abbreviations

Abbreviation	Meaning
100kGP	UK 100,000 Genomes Project
aa	Amino acid
aCGH	Array comparative genomic hybridisation
AF	Allele frequency
ALPL	Alkaline phosphatase
ANKRD28	Ankyrin repeat domain 28
ARID1B	AT-rich interactive domain-containing protein 1B
AXIN2	Axin-2
BAF	Brahma-associated factor
BBS9	Bardet-Biedl syndrome 9
BC	Bi-coronal
BCL11B	BAF chromatin remodelling complex subunit
bHLH	Basic helix-loop-helix
BL	Bi-lambdoid
BMP	Bone morphogenetic protein
Bp	Base pairs
CADD	Combined annotation-dependent depletion
CALD1	Caldesmon 1
CFNS	Craniofrontonasal syndrome
CGG	Clinical genetics group
CH	Calponin-homology domain
CHH	Congenital hypogonadotropic hypogonadism
CNV	Copy number variant
CRD	Cysteine-rich domain
CRISPR	Clustered regularly interspaced short palindromic repeat
CT	Computerised tomography
CTSK	Cathepsin K
CYP26B1	Cytochrome P450 26B1
DDD	Deciphering developmental disorders
DDG2P	Developmental disorder genotype-to-phenotype
DGV	Database for genomic variation
DHH	Desert hedgehog
DMEM	Dulbecco's modified eagle medium
DNA	Deoxyribonucleic acid
DNM	*De novo* mutation
E	Embryonic day
EBV	Epstein-Barr virus
ECD	Extracellular domain
EFNB1	Ephrin B1
Eks	Elbow-knee-synostosis
EMSA	Electrophoretic mobility shift assay
EN1	Engrailed 1
ENU	N-ethyl-N-nitrosourea
EPHB3	EPH receptor B3
ERF	ETS2 repressor factor
ESHG	European society of human genetics
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor

gDNA	Genomic DNA
GE	Genomics England
GeCIP	Genomics England clinical interpretation partnership
GFP	Green fluorescent protein
GIST	Gastrointestinal stromal tumours
GLI1	GLI family zinc finger 1
GMEM	Glasgow's minimum essential medium
GMC	Genomic medicine centre
GP130	Glycoprotein 130
GRB2	Growth factor receptor-bound protein 2
Hh	Hedgehog
HNRNPK	Heterogenous nuclear ribonucleoprotein K
HPO	Human phenotype ontology
HS	High sensitivity
HSP90AA1	Heat shock protein 90 alpha family class A member 1
ICP	Intracranial pressure
ID	Intellectual disability
IHH	Indian hedgehog
IL6ST	IL6-signal transducer
IL11	Interleukin 11
IL11RA	Interleukin receptor subunit alpha
IQR	Interquartile range
KLD	Kinase, ligase, and DpnI
KS	Kallmann syndrome
LDS	Loeys-Dietz syndrome
LIF	Leukaemia inhibitory factor
LMX1B	LIM homeobox transcription factor 1 beta
LOEUF	Loss-of-function observed/expected upper bound fraction
mAb	Monoclonal antibody
MAF	Minor allele frequency
MEGF8	Multiple EGF-like domains 8
mESCs	Mouse embryonic stem cells
MGRN1	RING superfamily E3 ligase
MMP21	Matrix metallopeptidase 21
MOI	Mode of inheritance
MRI	Magnetic resonance imaging
MSC	Mesenchymal stem cells
MSX2	Msh homeobox 2
NBCCS	Nevoid basal cell carcinoma syndrome
NFIA	Nuclear factor IA
NGS	Next generation sequencing
NHS	National health service
NLS	Nuclear localisation sequence
NMD	Nonsense-mediated decay
NPC1	Niemann-Pick C1
NS	Non-syndromic
NSC	Non-syndromic craniosynostosis
OAR	C-terminal otp, aristaless, and rax domain
P	Postnatal day
pAb	Polyclonal antibody
PAM	Protospacer adjacent motif
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDGFRA	Platelet-derived growth factor receptor alpha

pLI	Probability of loss-of-function intolerance
POR	Cytochrome P450 oxidoreductase
PolyPhen2	Polymorphism Phenotyping 2
PPP1C	Protein phosphatase 1 catalytic subunit
PRRX1	Paired related homeobox 1
PTB	Cbl-tyrosine kinase-binding domain
PTC	Premature termination codon
PTCH1	Patched 1
PTHLH	Parathyroid hormone-like hormone
qPCR	Quantitative polymerase chain reaction
RAB23	Member RAS oncogene family 23
RIPD	Researcher-identified potential diagnosis
RL-TK	Renilla-luciferase (TK promoter)
RNA	Ribonucleic acid
RND	Resistance nodulation division
ROH	Region of homozygosity
RT	Room temperature
RTK	Receptor tyrosine kinase
RT-PCR	Reverse transcriptase polymerase chain reaction
RUNX2	RUNX family transcription factor 2
SARA	SMAD anchor for receptor activation
SD	Standard deviation
SDM	Site directed mutagenesis
SHH	Sonic hedgehog
SIFT	Sorting intolerant from tolerant (SIFT)
SMAD	Mothers against decapentaplegic homolog
SMO	Smoothened
SNP	Single nucleotide polymorphism
SNV	Small nucleotide variant
SOS	Son of sevenless
SOX	SRY-box transcription factor
SP7	Osterix
SPRY	Sprouty
SPTBN1	βII-spectrin gene
SR	Spectrin repeat
SRC	Supraorbital regulatory centre
SRD	Serine-rich domain
SSD	Sterol sensing domain
TAR	Thrombocytopenia-absent radius syndrome
TCF12	Transcription factor 12
TFCP2	Transcription factor CP2
TGFβ	Transforming growth factor beta
TGFβR2	Transforming growth factor beta receptor
TGS	Tris-glycine-SDS (sodium dodecyl sulphate)
T-TBS	Tween-tris buffered saline
TLK2	Tousled-like kinase 2
TM	Transmembrane
TMCO1	Transmembrane coiled-coil domain 1
TN-C	Tenascin-c
TNSALP	Tissue non-specific alkaline phosphatase
TWIST1	Twist family bHLH transcription Factor 1
UC	Uni-coronal
UL	Uni-lambdoid
UTR	Untranslated region

UV	Ultraviolet
VCF	Variant call format
VEP	Variant effect predictor
WES	Whole exome sequencing
WGS	Whole genome sequencing
WIMM	Weatherall Institute of Molecular Medicine
WNT	Wingless-related integration site

Abstract

Craniosynostosis, the premature fusion of the cranial sutures, affects ~1 in 2,000 live births. It is a heterogenous disorder with documented monogenic, polygenic, chromosomal, and environmental causes. Although many patients with a genetically determined cause harbour a variant in one of just seven genes, or have a chromosomal abnormality, there is a long tail of rarer genetic diagnoses with over 60 genes known to be recurrently mutated. Despite the recent uplift in knowledge of the genetic causes of craniosynostosis and the biogenesis of the cranial suture, there remains a suspected 10% of undiagnosed patients with a putatively pathogenic monogenic cause. In my analyses I aimed to target this subset of patients to delineate novel genetic causes of craniosynostosis. Using data generated from genome and exome sequencing I developed systematic pipelines to screen for small nucleotide variants. This approach identified 8 variants that explained the phenotype of previously undiagnosed patients (*n* = 114) enrolled in the UK 100,000 Genomes Project (100kGP) and highlighted additional candidate genes for further scrutiny. The candidate genes were included alongside 35 additional genes with reasonable evidence for involvement in craniosynostosis in a panel-based resequencing analysis of 617 individuals without a genetic diagnosis. Through this approach, I identified 8 diagnostic variants and 15 additional candidate variants. Drawing on the identification of pathogenic variants from the UK 100kGP, my resequencing analysis, and genes identified through other sources, I undertook experiments to explore and functionally validate the likely pathophysiological mechanisms associated with variants in three genes (*PRRX1*, *PTCH1*, and *SPRY1*) by using immunofluorescence, mouse models, and RNA and protein assays. Overall, this work has achieved improved diagnostic pipelines for craniosynostosis, and rare disease more widely, and has also served to delineate three new genetic causes of craniosynostosis by providing a key evidence base to support variant pathogenicity.

Chapter 1

An introduction to craniosynostosis, its aetiology, and the identification of pathogenic variants

1.1 General introduction

Craniosynostosis, the premature fusion of one or more cranial sutures of the skull, is a clinically and genetically heterogenous congenital anomaly, affecting approximately one in 2,000 live births (Tønne et al., 2020, Cornelissen et al., 2016). Despite being one of the most prevalent craniofacial abnormalities (second to cleft lip with or without cleft palate), the variability in its causes and presentation can make identifying a genetic diagnosis extremely challenging. In cohort studies, monogenic, polygenic (Justice et al., 2012, Justice et al., 2022), chromosomal, and environmental (Sanchez-Lara et al., 2010) factors have all been identified as likely causes of craniosynostosis, with a burden of approximately 25% of the cases originating from monogenic causes alone (Wilkie et al., 2017). The majority of diagnoses are attributable to variants in just 7 genes (*EFNB1, ERF, FGFR2, FGFR3, SMAD6, TCF12* and *TWIST1*) (Timberlake et al., 2017b, Wilkie et al., 2017), yet variants in over 60 genes are known to be recurrently mutated in craniosynostosis more rarely (Twigg and Wilkie, 2015, Goos and Mathijssen, 2019). Despite advances in the throughput and capacity of sequencing technologies, and the huge uplift in knowledge surrounding the genes and molecular basis for pathogenicity of craniosynostosis, there remains a suspected 10% of patients with craniosynostosis and putative undiagnosed monogenic disorders (Wilkie et al., 2017).

Over the last decade, the Clinical Genetics group (CGG), based at the MRC Weatherall Institute of Molecular Medicine (WIMM), Oxford, have been recruiting craniosynostosis patients, and family members, into a research project aiming to identify new genetic causes of craniosynostosis. This work has been completed in collaboration with the four major specialist craniofacial centres in England (Oxford, Birmingham, Liverpool, and Great Ormond Street, London) and provides a large cohort in which to screen for novel genetic causes of craniosynostosis. In parallel, since its launch in 2012, the UK 100,000 Genomes Project (100kGP) has recruited over 120 families with craniosynostosis who

have received whole genome sequencing (WGS) (Smedley et al., 2021), providing additional cases to add to our cohort.

The aim of my project has been to identify and validate pathogenic small nucleotide variants (SNVs), up to 50 bp, in our cohort of genetically undiagnosed patients with craniosynostosis, through a combination of bioinformatic and functional approaches. In this thesis, I have investigated the molecular basis of pathogenicity of variants in *SPRY1, PRRX1,* and *PTCH1* and drawn attention to candidate genes identified from the 100kGP and other sources, requiring further functional validation. In this introductory chapter, I will outline the key stages in human skull and suture development, review current knowledge of the aetiology and genetics of craniosynostosis, and outline methods for the identification of pathogenic small nucleotide variants.

1.2 Anatomy and craniosynostosis

The mammalian skull vault primarily consists of five bones: the paired frontals, the paired parietals, and the unpaired occipital (Wu et al., 2017) (figure 1.1). Separating these bones are fibrous gaps, known as sutures, which allow for deformation of the skull during birth and permit growth of the underlying brain during development. Along the midline of the skull, the frontal bones are separated by the metopic suture and the parietal bones are delimited by the sagittal suture. Perpendicular to this, the left and right coronal sutures provide the boundary between the frontal and parietal bones, whereas the lambdoid sutures separate the parietal bones from the occipital bone.

The human skull is a composite structure made up of the neurocranium, which surrounds and protects the brain, and the viscerocranium, which forms the facial skeleton (Morriss-Kay and Wilkie, 2005, Yilmaz et al., 2019) (figure 1.2). Unlike most bones of the human skeleton which are formed through endochondral ossification, a majority of the

neurocranium and most of the facial bones arise through intramembranous ossification (figure 1.3). The cranial base, more specifically the chondrocranium, undergoes endochondral ossification (Opperman, 2000), whereby bone develops from a cartilaginous intermediate. In contrast, for intramembranous ossification the sutures remain in an unossified state, while the bone fronts develop compact and spongy bone directly from sheets of undifferentiated mesenchymal connective tissue (Yilmaz et al., 2019).

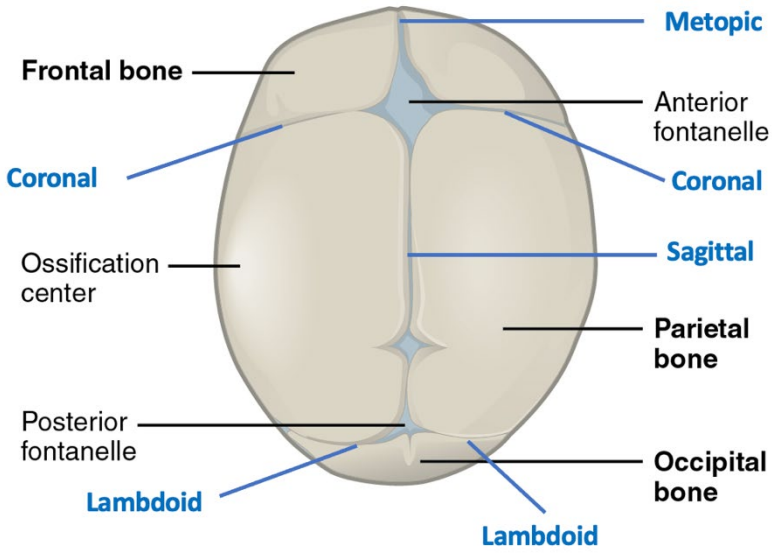


Figure 1.1 Schematic of a new-born skull. The six main sutures are highlighted in blue with the main bone plates annotated in black with bold text. Image adapted from Anatomy & Physiology, Connexions (image licensed under the Creative Commons Attribution Unported license).

The sagittal, metopic, and lambdoid sutures overlie areas in which the brain tissue does not lie close to the bony surface, namely the midline of the skull between the cerebral hemispheres and olfactory lobes (sagittal and metopic sutures) and between the cerebral hemispheres and the cerebellum (lambdoid) (Morriss-Kay and Wilkie, 2005). These three sutures arise from the narrowing of membranous gaps between the bones that they separate, however the coronal suture sits at a unique boundary in which the parietal bone overlaps the frontal bone from the outset. Despite the overlap in bone fronts at the coronal suture, this suture is extremely flexible and can be widened by the parietal bone sliding away from the frontal.

Unlike mice, in which the sagittal, coronal, and lambdoid sutures remain patent throughout their life (Grova et al., 2012), human calvarial sutures naturally fuse over time. This permits flexibility of the skull for continued growth of the head and the underlying brain during childhood. These sutures remain open until early adulthood (~22–26 years of age) (Idriz et al., 2015), with the exception of the metopic suture which fuses during the first year of life (Vu et al., 2001, Grova et al., 2012). The premature fusion of any one or more of these sutures results in craniosynostosis, which typically presents during the last third of pregnancy or up to the first year of life. Craniosynostosis prevents growth of the bones in a direction perpendicular to the suture itself; continued enlargement of the brain promotes compensatory overgrowth at the unfused sutures, resulting in the gradual dysmorphism of the skull. This can give rise to further complications, including raised intracranial pressure and hearing or sight abnormalities.

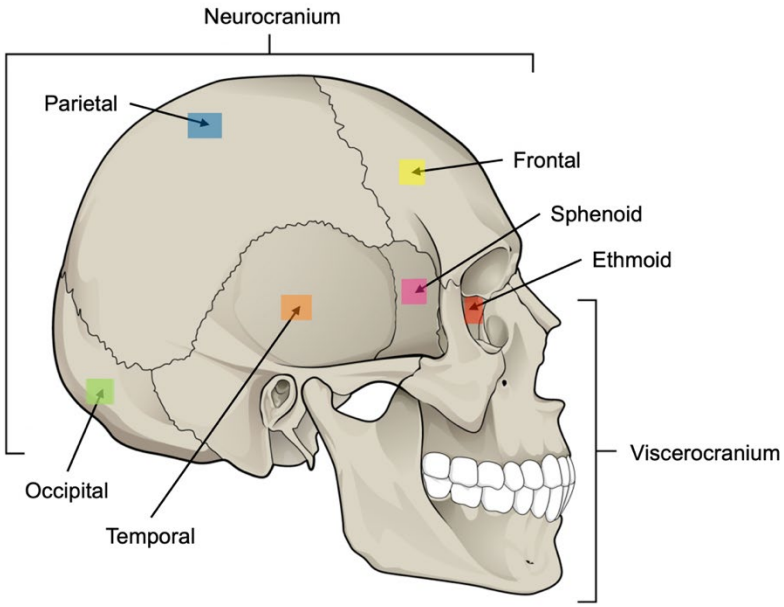


Figure 1.2 Schematic of the human skull and craniofacial skeleton. The neurocranium (which surrounds and protects the brain) and the majority of the viscerocranium (which forms the face skeleton) are formed by intramembranous ossification. Adapted from Anatomy & Physiology by Lindsay M. Biga, et al. Licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

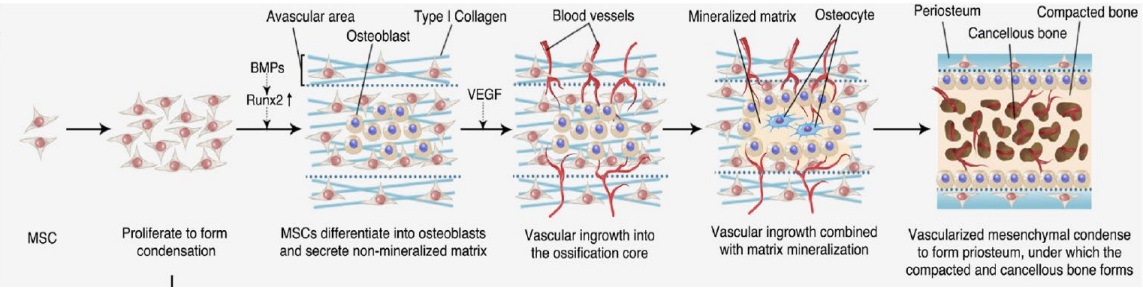


Figure 1.3 An illustration of intramembranous ossification. Mesenchymal stem cells (MSC) proliferate to form a cluster of cells which differentiate into osteoblasts. Secreted osteoid traps osteoblasts within ossification centres which subsequently form osteoclasts. Trabecular matrix and periosteum forms. Compact bone develops superficial to the trabecular bone and crowded blood vessels condense. Image obtained from (Zhu et al., 2021). Licensed under a Creative Commons Attribution-NonCommercial-NoDerivates 4.0 International License.

1.3 Non-syndromic craniosynostosis

Craniosynostosis is a heterogenous condition, defined by the presence or absence of additional clinical features (syndromic or non-syndromic, respectively) and the pattern of suture fusion (coronal, sagittal, metopic, and lambdoid). The majority of cases present with non-syndromic craniosynostosis (NSC), with estimates between 65–85% of a cross-sectional cohort (Wilkie et al., 2017, Timberlake and Persing, 2018) displaying NSC. NSC is sporadic in more than 95% of affected cases (Timberlake and Persing, 2018) and displays a highly variable clinical presentation, suggestive of a multifactorial disease where both genetic and environmental factors play a role. Owing to this mixed manifestation, it is much harder to identify a genetic diagnosis in patients with NSC. Despite this, the notable differences in phenotype, population incidence, proportion of familial cases, and male-female ratios between the different non-syndromic sutural synostoses are suggestive of distinct genetic aetiologies. In support, trio-based whole exome sequencing has identified genetic causes of both sagittal and metopic NSC, implicating *SMAD6* (a negative regulator of BMP [bone morphogenetic protein] signalling) variants as causative of 10% of familial non-syndromic midline synostoses (Timberlake et al., 2016). Variants in negative regulators of WNT (wingless-related integration site), BMP, TGFβ (transforming growth factor beta), and FGF (fibroblast

growth factor) signalling have also been identified as possible causes of NSC (Timberlake et al., 2017b).

1.3.1 Sagittal craniosynostosis

Sagittal synostosis is the most common type of NSC (~40% cases) (Wilkie et al., 2017, Tønne et al., 2020) and typically presents as a long, narrow skull, termed scaphocephaly (otherwise referred to as dolichocephaly) (figure 1.4). This is defined by a cephalic index ratio below 0.75 (maximum head width/ maximal head length) (Boyadjiev, 2007). Sagittal synostosis displays a largely male predominance with a male to female ratio of 3.5:1 (Lajeunie et al., 1996, Tønne et al., 2020). The majority of cases are sporadic, but an estimated 2–6% of cases are documented to be familial (Hunter and Rudd, 1976, Lajeunie et al., 1996, Boyadjiev, 2007). To date, there has been no evidence to suggest that paternal or maternal age correlates with the prevalence of scaphocephaly, suggesting that *de novo* mutations in the parental germline are unlikely to explain the majority of sporadic cases. The aetiology of non-syndromic sagittal synostosis (NSS) is not well understood, although both genetic and environmental factors are suspected. Environmental factors such as twinning (Lajeunie et al., 2005), teratogenic exposure (for example, maternal smoking) (Carmichael et al., 2008), increased parity, and intrauterine head constraint (Sanchez-Lara et al., 2010) have all been identified as potential risk factors for sagittal synostosis. Genome wide association studies have also implicated variants in regions near *BMP2* or within *BBS9*, as indicators of increased susceptibility to NSS (Justice et al., 2012). Following this, a common polymorphism downstream of *BMP2* (rs1884302C) was hypothesised to augment the penetrance of rare *SMAD6* variants (Timberlake et al., 2016), but analyses in an independent cohort did not support this conclusion (Calpena et al., 2020). As such, despite sagittal synostosis making up a large proportion of a cohort of patients with craniosynostosis, the complexity of its presentation and aetiology makes identifying a genetic cause challenging. Indeed, in a

cohort of 666 patients, only 8% of individuals with sagittal synostosis were considered solved with a monogenic cause; the remaining cases were likely attributable to environmental causes (Wilkie et al., 2017).

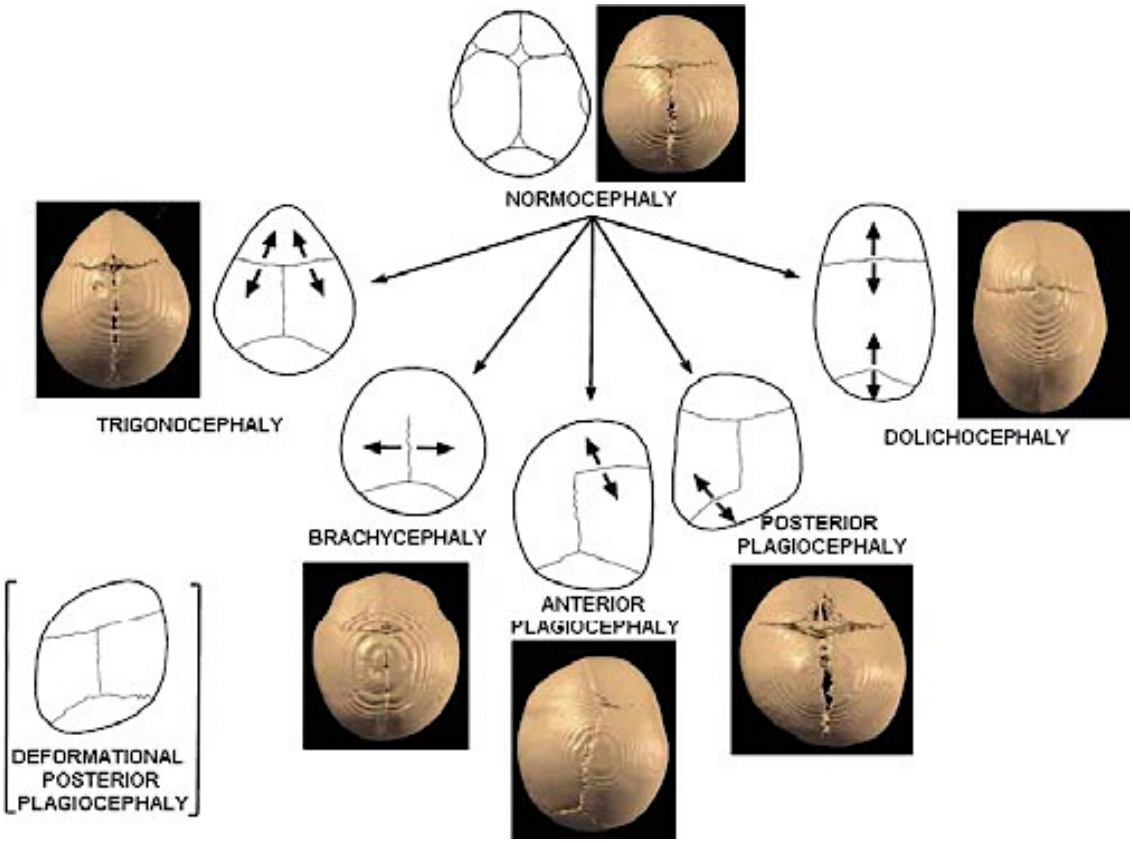


Figure 1.4 Schematic representation of skull deformities arising from the premature fusion of individual sutures. The arrows indicate the compensatory overgrowth of the patent sutures resulting in the gradual deformation of the skull. 3D-CT images of patients are also provided. Image reproduced from (Boyadjiev, 2007).

1.3.2 Metopic craniosynostosis

The second most common form of craniosynostosis is metopic synostosis, accounting for 20–34% of NSC diagnoses (Tønne et al., 2020, Cornelissen et al., 2016, Wilkie et al., 2017). Fusion of the metopic suture results in trigonocephaly (figure 1.4), characterised by a triangular appearance of the head and a forehead ridge. The prognosis for a patient with metopic synostosis is good, with an estimated 70% of patients displaying no other clinical manifestation or complication (Boyadjiev, 2007). Metopic synostosis displays a

male to female ratio of 3.3:1 (Lajeunie et al., 1998, Tønne et al., 2020) and is familial in ~10% of cases (Jehee et al., 2005). However, twin studies suggest a multifactorial mode of inheritance (MOI), given that concordance in monozygotic twins was found to be only 43% (Lajeunie et al., 2005). Environmental factors include sodium valproate ingestion (a medication for epilepsy), which was found in one study to result in metopic craniosynostosis in all 17 patients in whom their mothers were known to take valproate during pregnancy (Lajeunie et al., 2001). Multiple pregnancy and pre-term delivery have also been associated with metopic synostosis, increasing the risk by between 2.2-fold and 4.6-fold, respectively (Sanchez-Lara et al., 2010). Alongside sagittal synostosis, variants in *SMAD6* (Timberlake et al., 2016, Calpena et al., 2020) and variants in the regions near *BMP2* or within *BBS9* were also shown to be enriched in patients with metopic synostosis (Justice et al., 2020).

1.3.3 Coronal craniosynostosis

Fusion of a single coronal suture accounts for 6–11% of NSC cases (Tønne et al., 2020, Cornelissen et al., 2016) and results in anterior plagiocephaly (figure 1.4), whereby the skull appears flattened on one side; fusion of both coronal sutures (bi-coronal synostosis) is characterised by a broad and flat skull, termed brachycephaly (figure 1.4). Coronal synostosis has a largely female preponderance (male to female ratio of 1:2.1) (Tønne et al., 2020) and a higher proportion of familial cases (10–14%) (Hunter and Rudd, 1976, Lajeunie et al., 1995), indicating a greater genetic component associated with coronal synostosis compared to fusion of the sagittal or metopic sutures. In support, diagnostic success rate in a 13-year birth cohort considering fusion of any suture was highest in individuals with coronal synostosis (13% for uni-coronal and 60% for bi-coronal) with transcription factor 12 (*TCF12*) being most frequently implicated (Wilkie et al., 2017). *TCF12* encodes a basic helix-loop-helix (bHLH) transcription factor, capable of forming homo- or hetero-dimers with other members of the bHLH family (including *TWIST1* [twist

family bHLH transcription factor 1]) to regulate early developmental processes like mesoderm formation and neurogenesis. During cranial vault development, studies have shown elevated *TCF12* expression in ossifying cells of the frontal and parietal bones and inside the emerging cranial suture (Blumel et al., 2019, Teng et al., 2018). Coinciding with sites of expression, variants in *TCF12* cause coronal synostosis (Sharma et al., 2013, Di Rocco et al., 2014). Given the high proportion of cases attributable to variants in *TCF12*, targeted molecular testing of any child who displays coronal synostosis is recommended (Wilkie et al., 2017).

1.3.4 Lambdoid craniosynostosis

The rarest form of craniosynostosis is fusion of the lambdoid suture, accounting for approximately 1% of cases (Wilkie et al., 2017). Bilateral lambdoid synostosis results in a flattened and widened occipital bone; the majority of lambdoid synostoses are however uni-lambdoid, presenting as asymmetric posterior plagiocephaly (figure 1.4). Due to its rarity, there are few studies that have investigated associated gender, environmental, or genetic factors. Nevertheless, familial lambdoid synostosis has been documented in two families (Fryburg et al., 1995) and a further description of sporadic bi-lambdoid synostosis was reported in a patient with developmental delay, distinctive facial features (broad forehead, arched eyebrows, coarse hair, flat nasal bridge, telecanthus, epicanthus, up-swept nose, long philtrum, flattened upper lip and a pointed chin) and heart anomalies (Shimojima et al., 2016). In the latter case, the patient was found to have inherited a 16p13.11 microduplication from her father, alongside harbouring a *de novo* 19p13.3 microdeletion which explained many of her clinical features (although craniosynostosis is not a frequent phenotype of 19p13.3 microdeletions (Shimojima et al., 2016)). Additionally, a single case of bi-lambdoid synostosis was documented in a patient with prominent scalp veins, an unusual hair whorl at the apex of the skull, hypertelorism, telecanthus, short nose, small and low set ears, and microretrognathia.

Genetic testing of FGF receptors (*FGFR1–3*) and *TWIST1*, alongside molecular karyotyping did not identify a likely genetic diagnosis (Rieubland et al., 2011). Moreover, lambdoid synostosis (unilateral or bilateral) has also been documented in a subset of patients with variants in *ERF* (Twigg et al., 2013, Glass et al., 2019, Korberg et al., 2020). Indeed, in a cohort of patients with *ERF* variants 13/16 probands displayed lambdoid synostosis, which typically presented as a feature of pansynostosis (fusion of all sutures) but also in combination with fusion of the sagittal suture (Glass et al., 2019); only one of the probands displayed isolated fusion of the lambdoid suture (Glass et al., 2019).

1.3.5 Multi-suture synostosis

Fusion of multiple sutures account for 5% of NSC and are associated with a higher re-operation rate (5 times greater than fusion of a single suture) (Boyadjiev, 2007). Complex craniosynostosis (more than two sutures fused) is a frequent cause of raised increased intracranial pressure (ICP) and developmental delay; indeed, in two-thirds of patients with coronal and sagittal suture fusion the patients were documented with raised ICP (Renier et al., 1997). Furthermore, 75% of these patients also displayed a Chiari I malformation (Renier et al., 1997), demonstrating a greater need for clinical monitoring of patients with multi-suture synostosis. Multi-suture synostoses are reported to display a greater genetic component than fusion of a single suture, which is reflected in the greater diagnostic success rate of 15% of non-syndromic multi-suture cases (Wilkie et al., 2017).

1.4 Syndromic craniosynostosis

A smaller proportion of patients with craniosynostosis display syndromic craniosynostosis (~30%) (Tønne et al., 2020, Wilkie et al., 2017). Syndromes display a greater genetic component and in a cohort analysis of patients with craniosynostosis (*n*

= 666), a positive genetic diagnosis was identified in 69% of syndromic patients, compared to 5% of NSC (Wilkie et al., 2017). There are more than 180 syndromes which manifest craniosynostosis and which predominately display distinct features of the face and limbs (Twigg and Wilkie, 2015, Goos and Mathijssen, 2019). The genetic basis for many of these syndromes has been delineated, including a recurrent heterozygous variant, encoding p.(Pro250Arg), in *FGFR3* which was identified in 24/666 patients with craniosynostosis (Wilkie et al., 2017) and is associated with Muenke syndrome. This is the most commonly identified single genetic variant associated with a craniosynostosis syndrome and is characterised by coronal synostosis, hearing loss, developmental delay, and carpal/tarsal fusions (Doherty et al., 2007).

1.4.1 The first described craniosynostosis syndrome

Additionally, the first documented variant in craniosynostosis was a missense variant of the homeodomain within Msh homeobox 2 (*MSX2*), encoding p.(Pro148His) (Jabs et al., 1993), which causes Boston-type craniosynostosis (cranial abnormalities, headaches, poor vision, and seizures). Interestingly, heterozygous loss-of-function variants in *MSX2* result in skull vault defects, characteristically in the form of an enlarged parietal foramina (Mavrogiannis et al., 2006, Wuyts et al., 2000). However, a recurrent gain-of-function variant in *MSX2* at residue 148, including p.(Pro148His) (Jabs et al., 1993), p.(Pro148Leu) (Florisson et al., 2013), and p.(Pro148Thr) (see chapter 4), is known to cause Boston-type craniosynostosis (Wuyts et al., 2000).

1.4.2 Other common syndromes with a monoallelic cause

Heterozygous variants in *FGFR2* (including p.(Ser252Trp) and p.(Pro253Arg)) are associated with Apert syndrome (Wilkie et al., 1995), a disorder that presents with bi-coronal or multi-suture synostosis, severe syndactyly, and skin abnormalities. Variants in the alternatively spliced exon 9 of *FGFR2*, corresponding to the third immunoglobulin-

like domain, were also identified in Crouzon syndrome (Reardon et al., 1994, Kan et al., 2002). Crouzon displays similar clinical features to Apert syndrome but with no documented hand or feet anomalies. Saethre-Chotzen syndrome (craniosynostosis alongside brachydactyly, soft tissue syndactyly and facial dysmorphism) is defined by a spectrum of variants in *TWIST1*, a partner protein to *TCF12*, resulting in haploinsufficiency of *TWIST1* (El Ghouzzi et al., 1997, Zhou et al., 2018, Howard et al., 1997). Additional syndromes which are associated with recurrent genetic variants are described in reviews by Twigg and Wilkie (2015) and Goos and Mathijssen (2019).

1.4.3 Recessive syndromes

A smaller proportion of craniosynostosis syndromes are associated with recessive inheritance. Variants in *RAB23* (member RAS oncogene family 23) and *MEGF8* (multiple EGF like domains 8) are associated with Carpenter syndrome, characterised by multi-suture craniosynostosis, polysyndactyly, and obesity (Jenkins et al., 2007, Twigg et al., 2012) due to elevated hedgehog signalling (see section 1.7.4.1 and chapter 7 for further information). Other genes in which craniosynostosis is associated with a biallelic MOI are *POR* (cytochrome P450 oxidoreductase) (Flück et al., 2004) and *IL11RA* (interleukin 11 receptor subunit alpha) (Kespohl et al., 2021), which cause Antley-Bixler syndrome (craniosynostosis with abnormal steroidogenesis) and a Crouzon-like syndrome (multi-suture synostosis with midface hypoplasia, hearing loss, and dental anomalies), respectively (Nieminen et al., 2011, Keupp et al., 2013). POR is a flavoprotein that donates electrons to all microsomal P450 enzymes, which are required for retinoic acid degradation. Retinoic acid has been shown to control osteoblast differentiation in association with increased expression of the master regulator of osteogenesis, *RUNX2* (runt-related transcription factor 2) (Green et al., 2017, Escobar et al., 2021). Loss of *POR* in mouse and human studies results in elevated blood retinoic acid, suggesting that retinoic acid metabolism is a key contributor to its pathology (Ribes et al., 2007, Fukami

et al., 2010). In support, homozygous and compound heterozygous variants in *CYP26B1* (cytochrome P450 26B1), which encodes a P450 enzyme, cause mineralisation defects of the skull owing to elevated osteoblast differentiation to terminal osteocytes (Laue et al., 2011, Morton et al., 2016, Grand et al., 2021). In contrast, the proposed pathophysiological mechanism for variants in *IL11RA* is likely due to a bone resorption defect owing to an abnormality in osteoclast function (Nieminen et al., 2011). Indeed, mice deficient in *Il11ra* displayed increased trabecular bone volume and osteoclast precursors with a cell-autonomous defect in differentiation (Sims et al., 2005).

1.5 Environmental causes

Suspected environmental causes of craniosynostosis are well-documented in non-syndromic sagittal and metopic synostoses. Studies suggest that clearly defined environmental factors contribute to 2.5% of total cases (Wilkie et al., 2017). These factors include known environmental exposure to teratogens, for example, smoking, alcohol use, maternal thyroid disorders, and ingestion of sodium valproate during pregnancy (Singh et al., 2010, Lajeunie et al., 2001, Carmichael et al., 2008, Yilmaz et al., 2019). In addition, broader epidemiological factors are suspected to cause craniosynostosis, including external pressure on the foetal skull as a result of malposition (Wilkie et al., 2010) or twinning (Lajeunie et al., 2005), and increased parity (Sanchez-Lara et al., 2010); of note, these factors are not included in the 2.5% of total cases, as it would be difficult to ascertain the causative nature of epidemiological associations. Understanding the relative contribution of these environmental factors in craniosynostosis, compared to a likely genetic cause is complicated. A recent twin study suggested that craniosynostosis likely occurs much earlier on in the pregnancy and therefore arises from a disturbance in suture development (suggesting a possible genetic cause), as opposed to environmental constraint (Cinca et al., 2022). However, although a higher risk of developing craniosynostosis is afforded to monozygotic twins, there is not 100%

concordance (Sanchez-Lara et al., 2010, Cinca et al., 2022), consolidating that genetics does not explain the entire phenotype.

1.6 The importance of the sutural mesenchyme

Bone tissue undergoes constant turnover, a process that is sustained by resident stem cells. For the regeneration and repair of long bones, formed by endochondral ossification, a major stem cell population is known to reside within the bone marrow (Sacchetti et al., 2007). However, the stem-cell niche essential for craniofacial development has, until more recently, been less certain. Given the limited space for bone marrow within the calvaria and the distinct mechanism by which these bones are formed (intramembranous ossification), it has long been proposed that the sutural mesenchyme may act as the regulating centre required for growth and repair. This suggestion has been supported by the identification of specific markers expressed during early craniofacial development by resident cell populations within the suture – namely, *Gli1, Axin2*, *Prrx1*, and *Ctsk*-expressing cells (Zhao et al., 2015, Maruyama et al., 2016, Wilk et al., 2017, Debnath et al., 2018, Farmer et al., 2021).

1.6.1 *Gli1*+ expressing stem cells

Gli1 (a known marker of active hedgehog signalling) was identified as a marker of mesenchymal stem cells in the adult mouse incisor (Zhao et al., 2014), and later found to be expressed in stem cells of the sutural mesenchyme (Zhao et al., 2015). Localisation studies identified *Gli1*+ cells in the periosteum, dura matter, and sutural mesenchyme at birth; after 21 days, these cells were gradually restricted to the cranial sutures and remained until adulthood. Consistent with sites of expression, genetic ablation of *Gli1*-expressing cells in 1-month old mice, using a tamoxifen induced *Gli1-CreER^{T2};DTA$^{flox/flox}$* mouse, caused coronal synostosis (Zhao et al., 2015), further supporting a requirement for *Gli1* in the maintenance of the undifferentiated suture mesenchyme.

1.6.2 *Axin2* as a marker of stemness

Axin family genes are negative regulators of canonical WNT signalling, through their ability to degrade β-catenin. Studies have shown that AXIN2 is involved in cranial morphogenesis, with *Axin2* knockout mice developing craniosynostosis of the coronal and metopic sutures at early postnatal stages (Yu et al., 2005). *Axin2*-expressing cells have been shown to localise to the midline of sutures, to be slow-cycling cells (an indicator of stem cell quiescence) and to be highly responsive to calvarial injury (Maruyama et al., 2016). Gene-profiling analyses show a high level of *Gli1* expression in the *Axin2*+ cells, suggesting an overlap of these two populations. Of note, *Gli1* is a more ubiquitous marker of the entire suture mesenchyme, whereas *Axin2* is only found in 15% of these cells, localising specifically to the midline suture.

1.6.3 The requirement for PRRX1 within the developing suture

Post-natal skeletal stem cells expressing the homeobox transcription factor *Prrx1* (previously referred to as *Prx1*) have been shown to reside exclusively in the calvarial suture niche and decrease in number with age (Wilk et al., 2017), suggesting a requirement for *Prrx1* for suture patency during early development. Despite this, global ablation of post-natal *Prrx1*+ cells did not disrupt craniofacial growth or induce craniosynostosis, unlike global ablation of *Gli1*. This is hypothesised to be due to the much sparser distribution of *Prrx1*+ cells in the post-natal suture mesenchyme, compared to *Gli1*+, which may result in a less significant phenotype. As such, it is likely that PRRX1 acts much earlier in development and indeed has been shown to be expressed in the mouse coronal suture at embryonic day (E)15.5 (Farmer et al., 2021). In support, null mutants in mice (*Prrx1*$^{-/-}$) display cleft palate defects of multiple bones of the face and lateral skull and absent supraoccipital bones, alongside additional limb and vertebral anomalies (short, thickened endochondral bones, and incomplete vertebral arches) (Martin et al., 1995, ten Berge et al., 1998).

1.6.4 CTSK marks undifferentiated stem cells

Single-cell RNA sequencing, together with lineage tracing, have also identified *Ctsk*-expressing cells as a marker of stem cells at E14.5 (Debnath et al., 2018) and likely a key factor in the maintenance of the sutural mesenchyme. Like the more widespread *Gli1* expression found in the suture and surrounding tissue, *Ctsk*+ cells label the cranial sutures, alongside the overlying periosteum and underlying dura matter at postnatal day (P)15 and P32 (Debnath et al., 2018). The percentage of *Ctsk*+ cells in the suture was remarkably higher (31.1%) than the percentage of *Ctsk*+ cells in the calvarial periosteum (4.37%), demonstrating an enrichment of *Ctsk*+ cells in the undifferentiated suture.

1.7 Suture development during embryogenesis

To underpin the genetic aetiology of craniosynostosis, it is important to first understand how sutures are formed during embryogenesis. The maintenance of the undifferentiated sutural mesenchyme alongside the neighbouring proliferative bone fronts is a tightly regulated process, coordinated by various transcription factors including RUNX2 and Osterix (SP7) (Lana-Elola et al., 2007). This crosstalk of signalling between the suture and the developing bone coordinates the specific patterning of each suture. Broadly, there are five main processes in suture formation: stem cell specification and migration, lineage commitment, boundary formation and integrity, osteogenic proliferation and differentiation, and, resorption and homeostasis (Twigg and Wilkie, 2015).

1.7.1 Stem cell specification and migration

Studies at early developmental stages (E7.5 and E8.5) aimed to establish the origin of mesenchymal stem cells during development. In the case of the coronal suture, tracing experiments identified that precursors of cells expressing *Gli1* were localised to the cephalic paraxial mesoderm, immediately flanking the neural tube, because of Sonic hedgehog (*Shh*) accumulation in the ventral notochord and prechordal plate. At E8.5–

9.5, these Shh-responsive cells accumulated within a specific niche above the developing eye, termed the supraorbital regulatory centre (SRC), prior to giving rise to the future coronal suture and flanking ossified tissue (Deckelbaum et al., 2012).

1.7.2 Lineage commitment

At E9.5, following the accumulation of stem cells at the SRC, stem cells become committed to a lineage. The physical divide of cells with osteogenic potential, either mesodermal-derived or neural crest-derived, results in the specific patterning of the future coronal suture (Yoshida et al., 2008). In coordination with the growth of the developing brain, at E11.5 the cells from the SRC extend apically to overlay the brain surface resulting in the neural crest-derived cells giving rise to the future frontal bone, while the mesodermal-derived cells form the overlying parietal bone (Deckelbaum et al., 2012). It has been noted that a significant proportion of the parietal bone is made up of the paraxial progenitors, *Gli1*+ cells, which subsequently adopt an osteogenic fate coordinated by the expression of transcription factors such as *En1* (Engrailed1), *Twist1*, and *Msx2*. This sequence of events promotes expression of the master regulators of osteogenic differentiation, *RUNX2* and *SP7*.

1.7.3 Boundary formation and integrity

The coronal suture is of particular significance in the maintenance of the physical divide between bones of distinct embryological origin: the neural-crest derived frontal bone and the mesodermal-derived parietal bone. This boundary means that the metopic suture is entirely neural crest derived and, as such, perturbations of neural crest development are the most likely cause of metopic synostosis (Liu et al., 2014). The remaining sutures (coronal, sagittal, and lambdoid) have dual neural crest and mesodermal origins, but it is only the coronal suture for which this is true across the entire length (Morriss-Kay and Wilkie, 2005). Maintenance of these bone fronts are coordinated by *En1*, *Twist1*, and

Msx2, together with members of the *Ephrin/Eph* receptor and *Jagged/Notch* families (Yen et al., 2010, Ting et al., 2009, Merrill et al., 2006). Notably, mouse studies have shown that *Efnb1* (a member of the Ephrin family of transmembrane ligands for Eph receptor tyrosine kinases) demarcates the position of the coronal suture (Twigg et al., 2006, Babbs et al., 2011); ephrin receptors are also documented to mediate repulsive signalling between distinct cell types (for example, neural-crest and mesodermal-derived cells), thus controlling cell fate (Niethamer and Bush, 2019, Wilkinson, 2014). In humans, heterozygous loss-of-function variants in *EFNB1* cause craniofrontonasal syndrome (CFNS) (Twigg et al., 2004, Twigg et al., 2006), characterised by severe hypertelorism, a central nasal groove associated with unilateral or bilateral coronal synostosis, alongside other features such as mild cutaneous syndactyly, cleft lip and palate, diaphragmatic hernia, and agenesis of the corpus callosum (Bukowska-Olech et al., 2021).

1.7.4 Osteogenic proliferation/differentiation

Of equal importance to maintaining the undifferentiated sutural mesenchyme, is the establishment of the growing bone fronts flanking the suture. Positive differentiation signals emanating from osteoid (a collagenous unmineralized matrix produced by osteoblasts), alongside provision of growth factors from the dura mater underlying the suture (Opperman, 2000, Levi et al., 2012), stimulate osteogenesis of the surrounding bone fronts. Two key pathways coordinating this process of cell proliferation, differentiation, and cell fate, are hedgehog and FGF signalling.

1.7.4.1 The hedgehog signalling pathway

The hedgehog pathway is activated by ligand binding (including SHH, Indian hedgehog [IHH], and Desert hedgehog [DHH]) to a 12-pass transmembrane protein, Patched-1 (PTCH1), which releases repression on a second transmembrane G-protein coupled

receptor known as smoothened (SMO). This culminates in the translocation of Gli family proteins to the nucleus to initiate transcription of target genes (Cohen et al., 2015). These targets include *PTCH1* and *GLI1*, providing signals for negative and positive feedback control, respectively. As previously remarked, *Shh* expression at E7.5 is an early signal in the establishment of the coronal suture (section 1.7.1) alongside increased expression of *Gli1*, a marker of stemness. *Gli1*+ cells are later identified in the postnatal suture, becoming active upon binding of Ihh (Zhao et al., 2015). As a result, mice null for *Ihh*$^{-/-}$ show a 50% reduction in parietal and frontal bone formation, alongside wider regions of undifferentiated mesenchyme; this phenotype is believed to be due to a defect in differentiation rather than proliferation (Lenton et al., 2011). Conversely, overexpression of *Ihh* leads to craniosynostosis from increased chondrocyte differentiation, and subsequent matrix mineralisation (Amano et al., 2014, Klopocki et al., 2011).

Variants in genes affecting the hedgehog pathway, including *SMO* (Twigg et al., 2016), *RAB23* (Jenkins et al., 2007), and *MEGF8* (Twigg et al., 2012), have been reported in humans to cause a very similar phenotypic spectrum of craniosynostosis (either multi-suture or coronal synostosis), polysyndactyly, and tumorigenesis. Studies of mice deficient in *Rab23* (a negative regulator of hedgehog signalling) have shown that these phenotypic features result from aberrant osteoprogenitor proliferation and elevated osteogenesis in the suture due to increased *Gli1* expression (Hasan et al., 2020). Additionally, mice with a *Gli3*-null allele exhibited bilateral lambdoid synostosis between E17.5–18.5 and widened interfrontal and sagittal sutures, coordinated by an osteoprogenitor proliferation and differentiation defect (Rice et al., 2010). In the lambdoid sutures, *Gli3*$^{-/-}$ mice displayed reduced expression of the transcription factor *Twist1* alongside elevated expression of a marker for osteogenesis, *Runx2* (Rice et al., 2010). Notably, Twist1 has been identified in the established coronal suture (E16) (Teng et al., 2018) to directly antagonise *Runx2* expression, providing a mechanism for preventing osteogenesis (Bialek et al., 2004). Adding a further layer to this inhibition, Twist1 displays

inhibitory activity towards bone sialoprotein and osteocalcin, which are downstream targets of Runx2 (Connerney et al., 2006).

1.7.4.2 The FGF/FGFR signalling axis

Additionally, the family of FGFs and their receptors (FGFRs) are essential for bone development. In humans, there are 23 FGF ligands (FGF1–23) which predominantly exert their pleiotropic effects by binding and activating tyrosine kinase receptors, which are encoded by four genes (*FGFR1*, *FGFR2, FGFR3*, and *FGFR4*) (Xie et al., 2020). In the mouse coronal suture, *Fgfr1–3* are shown to be highly expressed in the proliferative osteogenic cells in and around the suture, demarcating the positions of the surrounding bone fronts (Johnson et al., 2000). Consistent with sites of expression, *FGFR2* has been shown to be a key driver of a shift in status towards cell differentiation, and accordingly, defects in *FGFR2* result in the most common of the craniosynostosis syndromes, Apert syndrome (Wilkie et al., 1995, Wilkie et al., 2017). Dysregulation of the FGF/FGFR signalling axis is observed in multiple human disorders, including craniosynostosis (Wilkie et al., 1995, Kan et al., 2002, Wilkie et al., 2010), dwarfism syndromes (Bonaventure et al., 1996), chronic kidney disease (Walker et al., 2016, Hacohen et al., 1998), insulin resistance (Fan et al., 2019), and cancer (Krook et al., 2021). Binding of canonical FGFs to their receptor induces phosphorylation of intracellular tyrosine residues on the tyrosine receptor. The activated receptor is coupled to multiple downstream signalling pathways including RAS-MAPK, PI3K-AKT, PLCγ and STAT. To prevent consistent activation of the pathway several mechanisms have evolved to inhibit FGF signalling, including internalisation and degradation of the receptor (Sorokin et al., 1994), moderation of kinase activity by phosphatases (Camps et al., 1998, Li et al., 2007), and negative regulation by protein interaction with downstream signalling molecules (Faedo et al., 2010, Ornitz and Marie, 2015). For example, sprouty proteins are believed to sequester the adapter protein GRB2, thus preventing activation of the

PI3K-AKT and RAS-MAPK signalling pathways (Yusoff et al., 2002, Vaquero et al., 2019, Hanafusa et al., 2002) (figure 1.5).

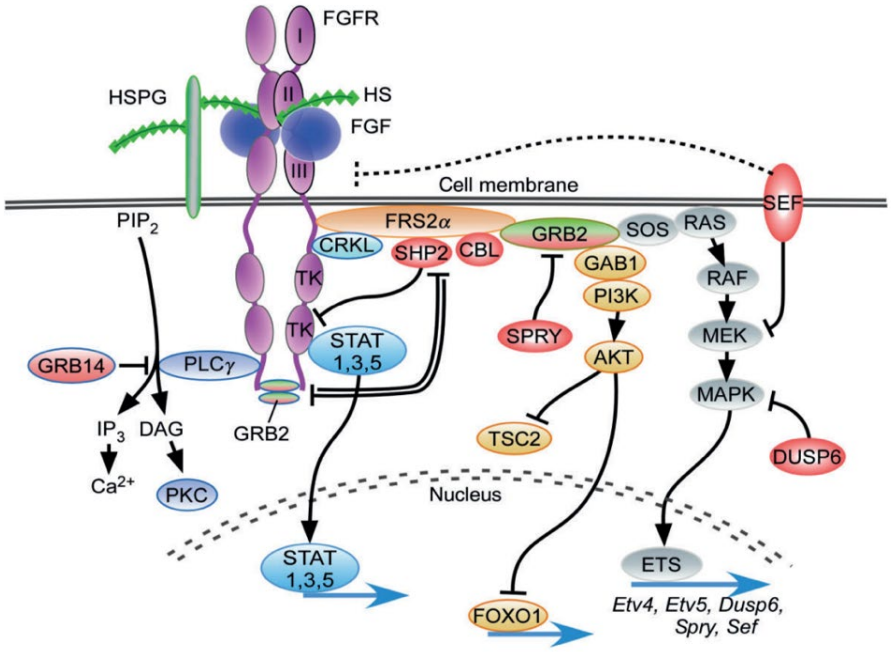


Figure 1.5 Schematic of the FGF/FGFR signalling axis. Binding of FGF to one of the four FGFRs promotes autophosphorylation of the receptor allowing effector molecules to bind. FGFR activation is coupled to several signalling pathways, including PI3K-AKT, RAS-MAPK, PLCγ, and STAT. This image was reproduced under a Creative Commons Attribution-NonCommerical-NoDerivatives International License (CC BY-NC-ND 4.0) (Ornitz and Marie, 2015).

For intramembranous ossification, key growth factors include FGF9 and FGF18 which are expressed in mesenchymal cells and differentiating osteoblasts (Ornitz and Marie, 2015). Heterozygous missense variants in *FGF9*, including p.(Arg62Gly) and p.(Arg190Thr), are documented in humans and mice to cause multiple synostoses syndrome 3, including craniosynostosis of the sagittal suture, owing to a failure of FGF9 homodimerization and a reduced ability to bind FGFR3 (Rodriguez-Zabala et al., 2017, Tang et al., 2017, Dobson et al., 2022). The patient phenotype is remarkably similar to the elbow-knee-synostosis (Eks) mouse model which displays elbow joint fusions, knee joint dysplasia, and craniosynostosis (Murakami et al., 2002, Harada et al., 2009, Harada and Akita, 2020). *Fgf18*-deficient mice show delayed suture closure associated with

decreased proliferation and osteogenic differentiation of the mesenchymal cells (Ohbayashi et al., 2002). Concordant with this data, delivery of exogenous FGF18 to the mouse foetal coronal suture accelerates osteogenesis and is associated with increased expression of *Fgfr1*, *Fgfr2*, *Fgfr3* and *Bmp2* (Nagayama et al., 2013). Notably, variants near *BMP2* have recently been suggested to be associated with non-syndromic midline craniosynostosis alongside variants in *SMAD6* (an inhibitor of osteoblast differentiation mediated by BMP) (see section 1.3.1 and 1.3.2) (Timberlake et al., 2016). Knock out of both *Fgf9* and *Fgf18* (*Fgf9*$^{-/-}$;*Fgf18*$^{-/-}$) demonstrates that these proteins display functional redundancy but with relative differences in tissue expression. For example, *Fgf9* predominates in the developing stylopod with *Fgf9*$^{-/-}$ mice displaying shortening of the bones in the upper arm and thighs; this phenotype is exacerbated by loss of *Fgf18*. Similarly, the *Fgf18* null skeletal phenotype is exacerbated throughout the limbs when one copy of *Fgf9* is deleted. Additionally, *Fgf9*$^{-/-}$;*Fgf18*$^{-/-}$ mice showed loss of *Runx2*, *Pthlh* (parathyroid hormone-like hormone), and *Ihh* expression, supporting reduced osteogenesis as the pathophysiological mechanism associated with *Fgf9* and *Fgf18* variants (Hung et al., 2016).

Given that FGF acts early in development and that FGFRs respond to multiple growth factors, it is not surprising that variants affecting these pathways have widespread effects in multiple tissues and affect both intramembranous and endochondral bone. However, it is known that a specific variant in *FGFR3*, associated with Muenke syndrome, p.(Pro250Arg), can be differentiated from other pathogenic variants in this gene because of its specific dysregulation of intramembranous ossification, without additional perturbation of endochondral bone. FGFR3 exists as two major splice variants, IIIb and IIIc, with IIIc predominantly expressed in proliferating chondrocytes. The p.(Pro250Arg) variant enhances affinity of FGFR3IIIc to unnatural ligands, including FGF2 and FGF9, alongside its canonical ligand (FGF1) (Ibrahimi et al., 2004), resulting in tissue specific effects. In contrast, a variant encoding p.(Lys650Cys) in the intracellular domain of

FGFR3 causes aberrant activation of FGFR3 in a ligand independent manner. The variant causes constitutive activation of FGFR3 due to autophosphorylation of key tyrosine residues on the intracellular domain and causes a more severe and lethal skeletal dysplasia phenotype known as thanatophoric dysplasia, type I or II (Ornitz and Marie, 2015).

1.7.5 Resorption and homeostasis

The maintenance of the suture requires a balance between osteogenesis, mediated by osteoblasts, and resorption, which is regulated by osteoclasts. Perturbations of interleukin 11 (*IL11*) signalling, caused by biallelic loss-of-function variants in the co-receptor *IL11RA* are associated with anomalies of osteoclast differentiation and, subsequently, craniosynostosis (see section 1.4.3). Of note, absence of the localised resorption of jawbone, a prerequisite for secondary dentition, delayed onset of dentition in homozygous mutants of *Il11ra*$^{-/-}$ (Nieminen et al., 2011) further corroborating a defect in osteoclast differentiation as the main pathological mechanism. Members of the cytokine family (including IL6, IL11, IL27, and leukaemia inhibitory factor [LIF]) are known to signal via the common GP130 (Glycoprotein 130) cytokine receptor, and the JAK/STAT pathway. Loss of *Gp130* in mouse models results in lethality, with mice displaying multiple skeletal abnormalities due to a failure in osteoblast and osteoclast function (Yoshida et al., 1996, Shin et al., 2004). In line with these early mouse models, human homozygous variants in *IL6ST* (IL6-signal transducer), which encodes GP130, results in skeletal abnormalities (including craniosynostosis) alongside a wider immunodeficiency defect due to the binding of different GP130 ligands (Schwerd et al., 2017).

1.8 Genes and pathways implicated in craniosynostosis

1.8.1 Developmental genes and their role in craniosynostosis

Maintenance of the sutural mesenchyme and the developing bone is an intrinsically complex and tightly regulated process, requiring the interplay of genes involved in stem cell renewal and differentiation, boundary integrity, osteogenesis, and resorption. These processes are coordinated by genes expressed early in development and therefore perturbations in these genes often disrupt signalling in multiple developmental pathways. This largely explains why syndromic craniosynostosis (which is characterised by defects in multiple tissues) more likely has a genetic cause. Figure 1.6 shows a broad, but non-exclusive, diagram of genes commonly mutated in patients with craniosynostosis, and the developmental pathways implicated for each gene. The fundamental role of these genes in organogenesis and vertebrate patterning implies that many of the pathogenic variants that might give rise to craniosynostosis would also affect multiple organs and subsequently cause a lethal phenotype. This makes it hard to predict precisely which genes, that when mutated, would cause craniosynostosis. However, paradoxically, mutations in a number of these genes are known to specifically cause craniosynostosis. For example, the documented gain-of-function variants within the FGF receptor family (Ibrahimi et al., 2004) that, given the universal importance of FGF signalling for organogenesis, may be expected to cause more widespread phenotypes. One possible explanation is that the cytotoxic effects of gain-of-function mutations within other cells and tissues are mitigated by tight negative feedback regulation, thus suggesting that sutures are more sensitive and tolerant to mutation than the development of early organs.

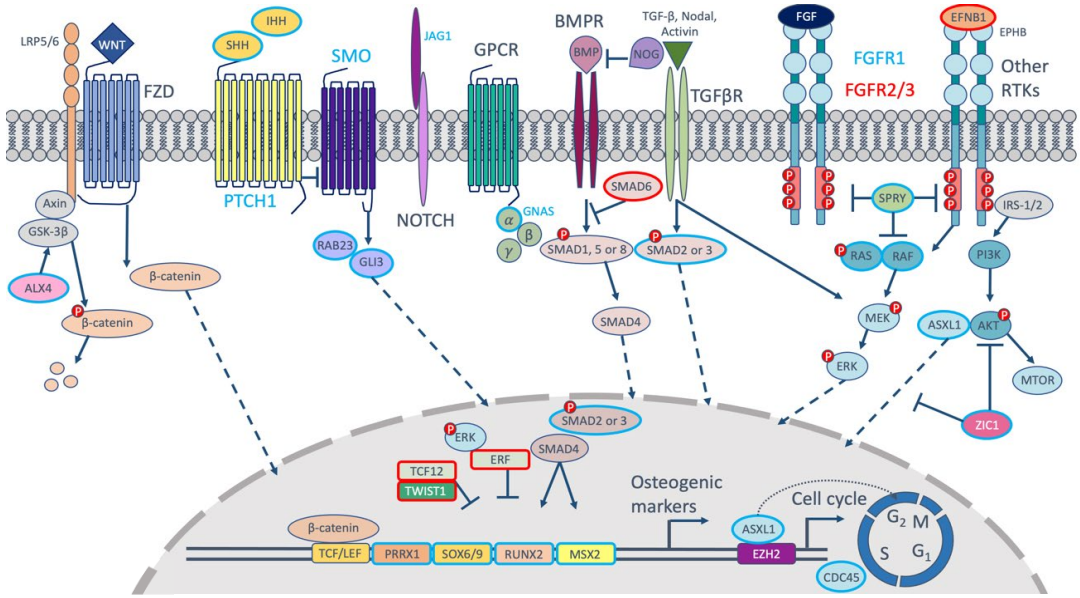


Figure 1.6 A schematic of the main pathways involved in suture development. Genes outlined in red indicate the 7 most commonly mutated genes in patients with craniosynostosis. Genes in blue lettering or are outlined in blue represent a subset of those identified on PanelApp by Genomics England (https://panelapp.genomicsengland.co.uk/panels/168/) as recurrently mutated in patients with craniosynostosis.

1.8.2 The long tail of rarer genetic diagnoses

While ~25% of diagnoses are made by identifying mutations in just seven genes (see section 1.1; figure 1.6), craniosynostosis is characterised by a long tail of rarer genetic diagnoses which affect over 60 genes (Twigg and Wilkie, 2015, Goos and Mathijssen, 2019). In an Oxford survey of 666 individuals with craniosynostosis, pathogenic variants were identified in 20 more rarely mutated genes in 23/666 individuals (3.5%) (Wilkie et al., 2017). This observation was supported by a recent exome sequencing study of patients in Norway (Tønne et al., 2020). The identification of such rare causative variants is inevitably challenging in a clinical diagnostic setting, where the need for high sensitivity (recall), which minimises false negative calls but involves intensive effort, has to be balanced against the need for high precision (positive predictive value), which minimises false positive calls (Hyder et al., 2021). The use of diagnostic gene panels has therefore been adopted as a means to address this problem.

One example of using panel-based approaches for diagnostic screening is provided by the UK's 100,000 Genomes Project (100kGP), delivered by Genomics England Limited (GE) (Smedley et al., 2021). This initiative aimed to facilitate whole genome sequencing (WGS) for 100,000 National Health Service (NHS) patients or relatives with rare disease or cancer, with the primary aim being to return information to participants on variants with sufficient evidence for diagnostic reporting related to their primary condition (Smedley et al., 2021). To facilitate this, GE curated PanelApp (https://panelapp.genomicsengland.co.uk), a publicly available resource containing crowdsourced and disease-focused gene panels for all rare disease groupings within the 100kGP (Martin et al., 2019). Genes on each panel are traffic-light coded based on level of confidence for diagnostic reporting in a given rare disease. Green genes are those in which there are plausible disease-causing variants (*de novo* or rare variants that are fully-penetrant) that affect a functional region of the gene (open reading frame for protein-coding genes) and have been identified in three or more unrelated families with congruent phenotypes, or within 2 or more unrelated families with strong additional functional data. Scrutiny of variants identified in Green genes aids prioritisation of likely pathogenic variants and decreases the number of candidates that diagnostic laboratories are required to screen. Genes that do not meet these criteria are listed as Amber or Red, corresponding to moderate or insufficient evidence for gene-disease association, respectively (Martin et al., 2019).

Based on this classification, there are 65 genes listed as Green in PanelApp for craniosynostosis (version 3; accessed 12th December 2022). Of those, 24 (37%) may be considered "core" craniosynostosis-associated genes, in which variants of a particular molecular type are associated with craniosynostosis in over 50% of the individuals and are considered to perturb fundamental components in the biology of cranial suturogenesis. The remaining genes are those in which craniosynostosis is a phenotype less frequently associated with pathogenic variants within that gene, although likely

causally associated. This latter grouping includes some examples in which there is a wider disturbance in osteogenesis, owing to aberrant osteoblast or osteoclast activity. Whilst the use of these panels greatly facilities workflow through complex genomic datasets, their full utility is critically dependent on regular updating to take account of recent research findings, thus maximising diagnostic sensitivity. This is particularly pertinent for heterogenous disorders like craniosynostosis which is characterised by rare mutations in over 60 genes. Notably, the last systematic update to PanelApp was completed in 2019 and therefore variants in a number of genes may now be considered diagnostic but are not routinely analysed as the gene is either considered Amber or Red or is currently not detailed in PanelApp.

1.9 Next generation sequencing

As is the case for genetic disorders in general, starting in 2010 the ability to sequence whole exomes and genomes has revolutionised the investigation of the genetic causes of craniosynostosis (Ng et al., 2010, Vissers et al., 2010, Gilissen et al., 2014, Taylor et al., 2015). Patients with craniosynostosis who present to the NHS England are screened for SNVs and copy number variations (CNVs) in the most commonly mutated genes, yet this will typically only identify a genetic cause in ~22% of individuals (Wilkie et al., 2017). There remains a considerable number of patients who may harbour variants in genes which are less frequently mutated in craniosynostosis, or perhaps harbour a variant in a novel gene (not yet documented in the pathogenesis of craniosynostosis). As such, whole exome/genome sequencing (WES/WGS) offers a huge advantage for routine diagnostics but also for the continued identification of new disease genes.

Next generation sequencing (NGS) provides a method of high-throughput sequencing and has transformed the speed and capacity of genetic testing, holding the potential for whole genomes to be sequenced for under ~£1000 (Schwarze et al., 2020). Traditional

methods of sequencing, including Sanger dideoxy-sequencing (Sanger et al., 1977) are still widely used but are more suitable for sequencing short fragments of amplified DNA (up to 1000 bp), for example when analysing a known variant in a DNA sample or verifying plasmid constructs. Following a need for increased capacity and lower cost of sequencing, a range of NGS technologies were brought to market. Often referred to as ‘second-generation’ NGS methods, the first of these technologies relied on much shorter DNA fragments (300–500 bp) and generally showed a higher error rate compared to Sanger sequencing. Given the high error rare, these sequencing methods rely on high sequence coverage (massively parallel sequencing) in order to obtain a consensus DNA sequence to distinguish real variants from noise (Slatko et al., 2018).

Although there are differences in the technologies, in general these methods rely on individual DNA molecules being tethered to a platform. The DNA is then amplified, and DNA synthesis occurs using chemical or labelled nucleotides which can be detected or imaged by the platform. For example, Ion Torrent technology relies on the sequential washing of wells containing amplified regions of DNA with one of the four nucleotides. Incorporation of a complementary nucleotide releases a hydrogen ion which changes the pH of the solution; the change in pH is detected and recorded as a voltage spike (Rothberg et al., 2011). In contrast, Illumina sequencing uses a method called bridge amplification whereby DNA molecules, anchored with adapters on both ends, are first amplified to form clusters on a solid glass support containing oligonucleotides complementary to the 5' and 3' adapters. During subsequent sequencing, fluorescently labelled nucleotides are sequentially incorporated and detected. There are a variety of Illumina sequencing protocols, including MiniSeq, MiSeq, NextSeq, NovaSeq and HiSeq, depending on the DNA fragment size, number of reads required and output (Goodwin et al., 2016). For example, MiSeq can perform 25 million reads per run for an output of 15 Gb (2x 300 bp reads), while NextSeq can provide 130 Gb with 400 million reads at 2x 150 bp read length. These second-generation sequencing methods are suitable for

exome and genome sequencing, but the short-read paired-end sequencing does not perform well in long repetitive elements, copy number alterations, and structural variants owing to the limited size of the initial DNA fragment. As such, they are better suited to detecting SNVs. To overcome these limitations, 'third generation' NGS employ long-read sequencing (Goodwin et al., 2016) approaches using a variety of technologies (including, Nanopore and PacBio).

1.10 Strategies for analysing WGS data

Due to the complexity and scale of WGS data, elucidation of likely pathogenic variants for monogenic human disease often results from systematic mining (Gilissen et al., 2014), prioritising variants that: (a) are moderate to high consequence protein coding, (b) display a likely Mendelian MOI (dominant or recessive), and (c) are consistent with available pedigrees (for example, *de novo* or inherited). For dominant disorders, the variant should be absent (or extremely rare if non-penetrance is suspected) from known databases of genomic variation, including gnomAD (Karczewski et al., 2019) and UK Biobank (Karczewski et al., 2022); for recessive disorders the likely pathogenic variant may display a higher allele frequency (typically below <0.001), but with no reported homozygous variants.

1.10.1 Protein coding variants

Approximately 85% of the known genetic causes for Mendelian disorders are protein coding. Given that the majority of WGS analyses have focussed on this ~1–2% of the human genome (Botstein and Risch, 2003), it is possible that this enrichment of protein coding disease genes results from ascertainment bias, however several important reasons remain as to why protein coding variants are prioritised. First, the search space is reduced from ~4 million variants to <25,000 for follow-up (Gilissen et al., 2012),

reducing the pool of variants for screening by almost 160x. Second, there exist several large population cohorts, including gnomAD (Karczewski et al., 2020), which are available to examine the rarity of a given variant within a healthy population, permitting variant filtering based on allele frequency (AF) for a given mode of inheritance. Third, extensive research has facilitated the development of a considerable number of predictive tools for variant pathogenicity, including Combined Annotation-Dependent Depletion (CADD) (Rentzsch et al., 2019), Sorting Intolerant From Tolerant (SIFT) (Ng and Henikoff, 2003), and Polymorphism Phenotyping 2 (PolyPhen2) (Adzhubei et al., 2010). These tools can provide an overall deleterious score to each candidate variant based on information such as location within the protein (for example, within a functional domain), impact (missense or more damaging), rarity, and conservation across species.

1.10.2 Segregation patterns

Another consideration when attempting to identify a disease-causing allele is the family pedigree. In families with more than one affected individual it is possible to reduce the search criteria to shared variation; this strategy is even more powerful if it is possible to sequence the most distantly related affected family member (for example, two affected siblings are expected to share 50% DNA, while 25% of DNA is shared with a grandparent) (Ng et al., 2010, Krawitz et al., 2010). Additionally, in families where consanguinity is expected, it is anticipated that the variant is more likely to be homozygous in the affected proband, within a large region of homozygosity, and inherited from both parents. Indeed, one study showed that in a consanguineous family, 17 out of 318 non-synonymous variants in the proband were autosomal homozygous (Becker et al., 2011). Of these 17 variants, only three were identified in large regions of homozygosity (defined as any region with 500 or more consecutive homozygous single nucleotide polymorphisms [SNPs]) with the identified pathogenic variant residing in the largest of these regions. A more systematic approach investigating 73 individuals from

54 consanguineous families confirmed that the causative gene was positioned within a region of homozygosity larger than 2 Mb in 93% of cases (Hildebrandt et al., 2009). Additionally, for trio sequencing, joint calling permits phase information about two variants, allowing interpretation as to whether two heterozygous variants in a proband are in *cis* or *trans* (Koboldt, 2020). This permits the identification of compound heterozygous variants with increased power if a recessive disorder is anticipated.

1.10.3 Identifying *de novo* variation

A further benefit of trio sequencing is the potential to identify *de novo* mutations (DNMs). Current estimates suggest that the average DNM rate is between 1.0–1.8 x 10^{-8} per nucleotide per generation (Jonsson et al., 2017) which equates to ~60 DNMs in the entire genome and less than 2 affecting the coding sequence (Deciphering Developmental Disorders Study, 2017). This figure is dependent on factors like paternal age, with an additional 2 mutations per year of paternal age (Kong et al., 2012, Rahbari et al., 2016, Kessler et al., 2020, Jonsson et al., 2017) and, to a lesser extent, the maternal age with an additional 0.5 mutations per year of age (Wong et al., 2016, Jonsson et al., 2017). Additionally, estimates suggest that DNMs account for almost half of the cases of patients with a severe developmental disorder, with a roughly equal balance of loss-of-function and altered protein function (Deciphering Developmental Disorders Study, 2017, Kaplanis et al., 2020). This is in stark contrast to the estimated 3.6% of patients with a developmental disorder and European ancestry harbouring a recessive variant (Martin et al., 2018), although this figure is elevated to 31% in a Pakistani background because of frequent autozygosity. Nevertheless, it has been shown that in ~6% of individuals with autozygosity, suggestive of a consanguineous union, there was a plausible pathogenic DNM, highlighting the importance of screening DNM regardless of family history (Deciphering Developmental Disorders Study, 2017).

1.11 Research aims

The aim of my thesis has been to identify previously unrecognised pathogenic SNVs in patients with craniosynostosis and functionally validate their consequence. To address this, I have mined WES or WGS data from the UK 100kGP and our in-house database of patients, known as 'CranioVar'. Analysis of the 100kGP (chapter 3) identified missed genetic diagnoses within genes with a known involvement in craniosynostosis, and further highlighted candidate genes for review. Candidate variants identified from the literature or the 100kGP were included in a resequencing panel consisting of 42 genes (chapter 4), aiming to identify further cases from 617 unsolved patients with craniosynostosis to corroborate findings. A variant in *SPRY1* identified through screening recessive candidates from the 100kGP was functionally validated in chapter 5. Additionally, the functional consequence of variants in *PRRX1* (chapter 6) and *PTCH1* (chapter 7) were assessed, consolidating variants within these genes, which were previously not described in the context of craniosynostosis, as pathogenic. Overall, this work has served to delineate new genetic causes of craniosynostosis, further improving diagnostic pathways and providing a key evidence base to further understand pathophysiological mechanisms associated with craniosynostosis.

Chapter 2

General methodology

The following methods have been used across multiple chapters. Specific primers or conditions used are denoted in the methods section of the appropriate chapter.

2.1 DNA Extraction

DNA was extracted from patient samples (blood, saliva, buccal swabs, hair and lymphoblastoid cells) using protocols and reagents detailed in the Zymo Midiprep Kit (D4075). Saliva (2 ml) was mixed with equal volume of BioFluid and Cell Buffer and 70 µl of proteinase K (20 mg/ml). Samples were mixed thoroughly before being incubated at 55°C for 2 hours. Hair samples were incubated at 55°C for 2 hours with 360 µl of nuclease-free water, 360 µl of solid tissue buffer, 40 µl of DTT (1M), and 40 µl of 20 mg/ml proteinase K. Buccal swabs were rinsed in 200 µl of 'BioFluid and Cell Buffer' before adding 20 µl of proteinase K. Swabs were incubated at 55°C for 10 min before continuing with the extraction protocol (as per manufacturer's protocol). DNA was quantified using Qubit dsDNA broad range assay (Q32850, Life Technologies) and the quality was checked by running 3 µl of DNA on a 1% agarose gel.

2.2 Dideoxy-sequencing

For plasmids, DNA (200 ng) was mixed with 1.75 µl of BigDye Terminator (3.1) buffer (5x), 0.5 µl of BigDye (Applied Biosystems, 4337455), 1 µl of 10 µM primer, and nuclease-free water (to a final volume of 10 µl). Samples were incubated in a thermal cycler at 95°C for 5 min, followed by 95°C for 15 s, 66°C for 10 s and 60°C for 4 min for 24 cycles. For PCR products, 2 µl of product was added to 0.5 µl of Exonuclease I (Exo1, 10 U/µl) and 0.5 µl FastAP Thermosensitive Alkaline Phosphatase (1 U/µl, Thermo Scientific™), adding 2 µl nuclease-free water to a final volume of 5 µl. The total 5 µl reaction was incubated in a thermal cycler at 37°C for 30 min, then 85°C for 15 min. The ExoSAP product was added to 1.75 µl nuclease-free water, 1.75 µl BigDye Terminator buffer (5x), 1 µl of primer, and 0.5 µl BigDye. For both plasmid DNA and PCR product,

DNA was precipitated using 80% EtOH, and washed (70% EtOH), before resuspension in HiDi formamide and sending for dideoxy-sequencing.

2.3 MiSeq Next Generation Sequencing

DNA extracted from blood and saliva samples was analysed by next generation sequencing (MiSeq). Briefly, 100 ng of genomic DNA (gDNA) was amplified using a Q5 enzyme protocol (2 µl of 5x Q5 buffer, 0.5 µl of each forward and reverse primer containing CS tags [10 µM of forward and reverse], 0.2 µl of dNTPs [10 mM each], 0.1 µl of Q5 high-fidelity DNA polymerase and up to 10 µl of nuclease-free water) and thermocycler conditions: 98°C for 30 s, followed by 30 cycles of 98°C for 10 s, 66–70°C (depending on primer) for 30 s, and 72°C for 30 s, and a final extension of 72°C for 8 min. PCR products were diluted 1:100 in nuclease-free water before barcoding. Diluted products (1 µl) were mixed with 5 µl of 2x iProof high-fidelity master mix (BIO-RAD, 172-5310) and 2 µl of water, then 2 µl of Fluidigm Access Array CS barcodes (2 µM). Samples were placed in a thermocycler at 98°C for 2 min, followed by 8 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 30 s, and a final extension of 72°C for 2 min. Barcoded products were pooled and quantified (Qubit dsDNA High Sensitivity protocol). The purified pool was checked using a 2200 TapeStation instrument (Agilent High Sensitivity D1000 ScreenTape) prior to sequencing using Miseq. Data was analysed, including alignment, variant calling, and annotation using Amplimap software (Koelling et al., 2019).

2.4 Whole Genome Sequencing

Blood DNA was sent to NovoGene Company LTD (Cambridge) for WGS using a NovaSeq instrument. Sequencing data were processed and incorporated into CranioVar for analysis.

2.5 Centrifugation

All centrifugation steps requiring volumes of 15 ml to 50 ml were carried out using an Eppendorf 5810 R bench centrifuge. All centrifuge steps requiring smaller volumes in Eppendorf tubes (up to 2 ml) were undertaken using a VWR Micro Star 17 centrifuge.

2.6 Cell Culture

2.6.1 Lymphoblastoid Cell Culture

Blood (1.3 ml) was mixed with an equal volume of Phosphate Buffered Saline (PBS) and added to 3 ml of Sigma Histopaque®-1077 before centrifugation at 580 *g* for 20 min at room temperature (RT). Peripheral blood mononuclear cells (PBMCs) were collected, PBS washed, and centrifuged (RT, 300 *g*, 5 min). The supernatant was discarded, and the cell pellet resuspended in PBS. Epstein-Barr virus (EBV, B95-9) was added to the PBMCs and incubated at 37°C for 1.5 hours. Cells were added to phytohemagglutinin M-form (Gibco, 10576-015) and maintained in culture medium (RPMI Medium 1648 [1x] supplemented with 15% Fetal Bovine Serum [FBS, Gibco], 1% penicillin-streptomycin [Gibco, 15140-122] and 1% L-glutamine [Gibco, 25030-081]), incubated at 37°C, 5% CO_2. Cells were routinely screened for mycoplasma infection using a MycoAlert™ PLUS Mycoplasma Detection Kit (Lonzo, LT07-710) and luminometer.

2.6.2 Fibroblast Cells

Fresh skin samples were obtained from the scalp after surgery and washed three times in sterile culture media (Dulbecco's Modified Eagle Medium [DMEM] with 10% FBS) and 10x penicillin-streptomycin. The sample was rinsed a further three times in sterile culture media containing 1x penicillin-streptomycin before being sliced into 1 mm^2 pieces. The tissue was transferred to a T25 cell culture flask and evenly distributed across the bottom of the flask before being covered in media. Once fibroblast cells were 50% confluent, the

cells were harvested using 0.25% Trypsin-EDTA (Sigma-Aldrich, T4049) and seeded to further flasks.

2.6.3 Adherent Cell Culture

All adherent cells (COS7, U2OS, HEK293, C2C12, HT1080 and fibroblasts) were maintained and grown in DMEM supplemented with 10% FBS, 1% penicillin-streptomycin and 1% L-glutamine, and were stored in a humidified incubator at 37°C, 5% CO_2. Cells were split once they reached ~80% confluency using Trypsin-EDTA.

2.7 Western Blot

Fibroblast cells were seeded to reach 90% confluency after 24 hours and lymphoblastoid cells were given fresh media 24 hours prior to harvest by centrifugation (300 *g*, 5 min). Cellytic MT Cell lysis buffer (Sigma-Aldrich, C3228), containing proteases (cOmplete; Sigma-Aldrich, 4693116001) and phosphatase (PhosStop; Sigma-Aldrich, 4906845001) was added to each sample and left on ice for 30 min. The lysed samples were centrifuged at 18,000 *g* for 15 min at 4°C and the supernatant collected. Total protein concentration was quantified using a BCA Quantification Kit (ThermoScientific, 23225) and absorbance read at 562 nm (Spectra Max M2E). Samples were loaded onto polyacrylamide gels (4–20% Mini-PROTEAN® TGX™ Precast Gels; BIO-RAD, 4561094), containing 20–50 µg of lymphoblastoid or fibroblast sample, 6x loading dye, and nuclease-free water. Samples were run for 1.5 hours at 150 V in 1x TGS (Tris-Glycine-SDS) buffer before being transferred to a membrane (Immobilion®-P PVDF) for 1 hour at 100 V in 1x transfer buffer (25 mM Tris, 192 mM glycine [pH 8.3] and 20% methanol [vol/vol]). The membrane was blocked with 5% non-fat blotting-grade milk (Bio-Rad), diluted in 1x Tween Tris-buffered saline [T-TBS]), and left shaking gently for 1 hour at RT before incubating with primary and secondary antibodies diluted in 5% no-fat milk and T-TBS.

SuperSignal West Pico Plus Chemiluminescent Substrate Kit (Thermo Scientific, 34579) was used to develop the signal for 5 min before imaging.

2.8 Reverse transcriptase PCR (RT-PCR)

Lymphoblastoid cells were cultured and harvested by centrifugation (300 *g* for 5 mins). RNA was extracted from cell pellets using the RNeasy Mini Kit (Roche) following the manufacturer's protocol. DNase I (5 μl of 1U/ μl, [Sigma-Aldrich, AMPD1]) was added to 5 μl of reaction buffer (R6273) and 50 μl of eluted RNA and incubated at RT for 15 min. Stop Solution (1 μl of 50 mM EDTA) was added to DNase-treated RNA and incubated at 70°C for 10 min, before reverse transcription following manufacturers guidelines for 1 μg of RNA (RevertAid First Strand cDNA Synthesis Kit [ThermoScientific, K1622]). A 1 μl sample of each random hexamer and oligo $(dT)_{18}$ primers were used in each cDNA synthesis reaction (ThermoScientific, K1622). The product of the first strand cDNA synthesis reaction was used directly in PCR or quantitative real-time PCR.

2.9 Quantitative real-time PCR (qPCR)

A 1 in 10 dilution of cDNA or 50 ng of gDNA was mixed with 5 µl of PowerUp SYBR Green Master Mix (Thermo Scientific, A25741), 0.5 µl of each forward and reverse primer (10 µM) and up to 10 µl of nuclease-free water. Each sample and primer pair were tested in triplicate, including a no template control (nuclease-free water). The Ct values were calculated using the QuantStudio™ Flex Real-Time PCR System (96-well) following manufacturer's recommended standard cycling protocol for SYBR reagents (adjusting the annealing temperature dependent on primer) and default dissociation steps. Primer efficiency was measured using serial dilutions of cDNA samples from 10–0.15625 ng/µl and calculated following: $(10^{(-1/slope)} - 1) \times 100$.

Chapter 3

Increasing the diagnostic yield in craniosynostosis through systematic mining of the 100,000 Genomes Project Data

3.1 Introduction

Over the past decade, NGS has transformed the capacity, speed, and cost of sequencing (Goodwin et al., 2016, Schwarze et al., 2020). Previously, identification of SNVs was limited to targeted analyses of gene fragments, however NGS now offers the possibility to sequence millions of reads simultaneously, allowing sequencing of whole exomes or genomes at a much faster rate. This increase in power has incentivised the transformation of clinical care, allowing the potential for NGS to be implemented in routine diagnostic practice.

3.1.1 The UK 100,000 Genomes Project (100kGP)

Initiated in 2012, the 100kGP, delivered by Genomics England Limited (GE), aimed to capitalise on NGS technology to sequence 100,000 NHS patients with cancer or rare disease (Smedley et al., 2021). The primary aim of this initiative was to return information to participants on variants with sufficient evidence for diagnostic reporting related to their primary condition. A genetic diagnosis is important for several reasons: (1) identifying a genetic cause may be essential for correct disease classification, highlighting specific complications to screen for or manage (Splinter et al., 2018), (2) it enables an accurate recurrence risk to be given to the parents and offspring which may facilitate prenatal genetic testing, (3) it ends the diagnostic odyssey and provides a 'label' which may open doors to additional support (Wu et al., 2020), and (4) it advances knowledge on human biology and pathophysiology.

However, a difficulty when faced with WGS data is being able to identify potentially disease-causing alleles amidst ~4 million sequencing variants (Auton et al., 2015) which largely consist of non-pathogenic SNPs, sequencing errors, and variation within the non-coding regions. This is particularly challenging for disorders which display substantial

locus heterogeneity, such as craniosynostosis for which over 60 genes are known to be recurrently mutated in patients (Twigg and Wilkie, 2015). To capture this variation, GE use two variant prioritisation pipelines: (a) Tiering and (b) Exomiser, aiming to identify SNVs and insertions or deletions (indels) < 50 base pairs (bp) (Hyder et al., 2021, Smedley et al., 2021) for a given rare disease. Variants that filter through these pipelines are evaluated by the genomic medicine centre (GMC) at which that patient was recruited; any variants that are considered diagnostic are reported back to the families. In addition to diagnostic reporting by the GMCs, the 100kGP dataset is available to researchers through the Genomics England Clinical Interpretation Partnership (GeCIP). Primarily, this provides a large dataset in which to identify novel disease genes related to a given condition and potentially facilitates the development of new therapies. Through this alliance, researchers can also identify missed diagnoses and report them back to the corresponding GMC for review.

3.1.2 Aims of this chapter

Through analysis of data from the 100kGP, I aimed to identify novel disease-associated genes for craniosynostosis and corroborate these findings through a resequencing analysis of unsolved craniosynostosis cases from our in-house database (discussed in chapter 4). Systematic mining of the data revealed diagnostic variants that were missed by the GE/GMC bioinformatics pipelines, which validated that my methods were able to identify clinically relevant variants. Therefore, first I aimed to discount known pathogenic variants in craniosynostosis genes that had been missed by the triaged bioinformatics pipelines, prior to the identification of novel disease-associated genes.

3.2 Methods

All analyses were performed in the GE Research Environment considering individuals annotated under the normalised disease grouping of ‘Craniosynostosis syndromes’, unless otherwise stated. For all approaches, analyses were performed on patient data aligned to hg19 and hg38 independently, prior to final grouping of all variant information where appropriate.

3.2.1 Craniosynostosis disease cohort

The clinical protocol for the 100kGP was approved by East of England-Cambridge South Research Ethics Committee (14/EE/1112). Written informed consent to obtain samples for genetics research was obtained. Patients are recruited through GMCs whereby tissue samples (mostly blood) were collected for DNA extraction before WGS using Illumina instruments. Patients were recruited under the normalised disease grouping of ‘Craniosynostosis syndromes’ if the patient displayed multiple suture fusions and/or had additional clinical features (syndromic) or a positive family history. All patients had undergone previous genetic screening for common causes of craniosynostosis and, if syndromic, they displayed a normal chromosomal microarray. In 51 families recruited, written informed consent had previously been obtained by researchers in the CGG, Oxford, to investigate genetic causes of craniosynostosis (Oxfordshire Research Ethics Committee B [C02.143] and London-Riverside Research Ethics Committee [09/H0706/20]) (Hyder et al., 2021). As it was not generally permitted to export patient phenotype information from the GE Research Environment, this is provided only where either (1) the patient had been independently recruited to the Oxford study or (2) the use of phenotypic information was approved for the purpose of publication in Hyder et al., 2021.

3.2.2 The GE Interpretation Pipeline

3.2.2.1 Tiering Data

For the GE interpretation pipeline, SNVs and indels are annotated as Tier 1, Tier 2, Tier 3, or Tier null. In brief, the 'Tiering' algorithm constitutes a diagnostic grade gene panel (accessed through PanelApp: https://panelapp.genomicsengland.co.uk/) which details known disease-causing genes for a given pathology using a traffic light system. Green genes display high level evidence for disease pathology and can be used confidently in genome interpretation, whereas Amber and Red genes display moderate to insufficient evidence for gene-disease association, respectively. Genomes are interrogated as family units, where available, with additional covariates used in the algorithm including allele frequency, MOI, segregation patterns, and intersection with a protein coding element. Tier 1 variants display high impact (for example, loss-of-function) or are *de novo* moderate impact missense variants; these variants must display appropriate segregation based on reported human phenotype ontology (HPO) terms detailed for each family member. GMCs were mandated to review Tier 1 and Tier 2 variants and had the discretion to review Tier 3 if deemed necessary. Further information on the Tiering pipeline is detailed in Hyder et al., 2021.

3.2.2.2 Exomiser

Part-way through the program (Data Release V7, 25/07/19) Exomiser (Smedley et al., 2015) was incorporated into the GE interpretation pipeline. Exomiser was configured to remove non-coding and low-quality variants, and for all MOI considered (autosomal dominant, autosomal recessive, X-linked dominant, X-linked recessive and mitochondrial) variants were excluded with a minor allele frequency (MAF) greater than 0.1% (or 2% for compound-heterozygotes). The following reference databases were used to ascertain the MAF: 100,000 Genomes Project reference samples, 1000 Genomes, ESP, TOPMed, UK10K, ExAC, and gnomAD (excluding the Ashkenazi

Jewish Population). Exomiser scores each variant by rarity, using the above frequency databases as a source. A score for rarity is combined with a pathogenicity score generated by Polyphen2 (Adzhubei et al., 2010), SIFT (Ng and Henikoff, 2003) and MutationTaster (Schwarz et al., 2014), which culminates in an overall rank for each candidate. In parallel, a score is also produced based on the phenotypic similarity of the patient, as reported by assigned HPO terms, to phenotypes listed in OMIM and Orphanet rare diseases known to be associated with the gene, mouse and zebrafish models, and phenotypes associated with neighbouring genes in the StringDB database. As with analysis of Tier 3 data, it was not mandated for GMCs to review the Exomiser hits prior to January 2021.

3.2.3 Independent Bioinformatic Analyses

All filtering strategies used in the following analyses are detailed in figure 3.1. In all analyses, candidate variants were intersected with variants identified in CranioVar (our in-house database), reported variants identified through the Deciphering Developmental Disorder (DDD) study and a list of *de novo* variants provided by a WGS analysis, as detailed previously (Reijnders et al., 2018, Calpena et al., 2019).

3.2.3.1 Recessive/ X-linked Analyses

All variants annotated as homozygous or hemizygous in the participant variant call format (VCF) (V10 Data Release; 03/09/2020) files were called and annotated with Ensembl 'Variant Effect Predictor' (VEP) (McLaren et al., 2016). An allele-frequency based filter was used, considering an overall gnomAD frequency of less than 0.005. Remaining variants were investigated manually, considering factors such as pathogenicity scores (reported by Polyphen, SIFT, and CADD scores), functional consequence, position within a functional domain, amino acid conservation in

vertebrates and proteins with related domains (for missense variants), and similarity of patient phenotype to reported animal models.

3.2.3.2 Compound heterozygous variants

Firstly, variants from any gene listed in Ensembl were called, using Platypus variant caller, from unsolved duos and trios enrolled in the craniosynostosis cohort of the V11 (18/12/2020) GE data release (230 individuals, 80 families). Variants annotated as 'PASS' in the VCF file were selected and subjected to decomposition and normalisation (Tan et al., 2015), aiming to expand multiallelic variants into multiple rows of annotation and left align all annotations representing the same variant. Second, pass-filtered variants were annotated with SnpEff (Cingolani et al., 2012) and Vcfanno (Pedersen et al., 2016) and indexed prior to segregation analysis with Slivar (Pedersen et al., 2021). Compound heterozygous variants were extracted, annotated with VEP, and filtered by a CADD (Rentzsch et al., 2019) score greater than 20, or missing CADD, and protein-coding consequence.

3.2.3.3 Families with more than one affected individual

Families with more than one affected individual (9 unsolved families) enrolled in the V9 data release (02/04/2020) were screened for segregating and non-segregating variants following the workflow described for the compound heterozygous analysis, with the addition of filtering also by *de novo* or recessive variants.

3.2.3.4 DNMs

DNMs not ranked Tier 1 or Tier 2 (and therefore Tier 3) were found in the tiering data (V10 Data Release; 02/04/2020) by filtering on DNM annotation. These calls were intersected with genes listed in the developmental disorder genotype-to-phenotype (DDG2P) database (accessed: http://www.ebi.ac.uk/gene2phenotype/) and/or reported

in recent literature, to generate a shortlist of DNM calls for individual scrutiny. Second to this approach, multi-sample VCF files for families enrolled as trios were passed through the Platypus 'bayesianDeNovoFilter.py' script under default parameters. The modified multi-sample VCF was annotated with base and stringent filters (see https://research-help.genomicsengland.co.uk/display/GERE/De+novo+variant+research+dataset) and protein-coding variants annotated as 'PASS' were extracted. All pass-variants were annotated with VEP and filtered. Non-coding, intronic (excluding canonical splice-site variants), synonymous, and variants annotated as benign in ClinVar were excluded from the dataset, leaving the remaining variants to be investigated manually.

3.2.3.5 Exomiser

The top 20 Exomiser hits in the V9 data release (02/04/2020) were screened for each affected individual in the craniosynostosis cohort.

3.2.3.6 Indels

Insertion, deletion, and frameshift variants were called from the V11 data release (18/12/2020) of the unsolved craniosynostosis cohort participant VCF files. These variants were annotated with VEP and intersected with genes listed in the DDG2P database (Thormann et al., 2019). Non-coding, intronic (excluding variants affecting the canonical splice-site), and variants annotated as benign were removed before filtering by absence from gnomAD (v2.1.1). Indels identified in more than three unrelated families were removed, aiming to increase identification of rare, and likely pathogenic, variants. Variants were manually assessed and removed from analyses if: (1) they intersected a region of low complexity, (2) were flagged in the common dbSNP track on UCSC, (3) were within a poorly conserved region, or (4) affected an alternatively spliced exon. Variants that adhered to any of these categories but intersected a Green gene on PanelApp were retained.

3.2.3.7 Cryptic splice site variants

Variants were called in unsolved participant VCF files (V11; 18/12/2020), selected for 'PASS' variants and intersected with DDG2P genes. SpliceAI (Jaganathan et al., 2019) was run on the remaining variants and filtered by a probability of 70% or greater that the variant caused an increase or decrease in the activity of the donor or acceptor splice site. The remaining variants were merged with probability of loss-of-function intolerance (pLI) and loss-of-function observed/expected upper bound fraction (LOEUF) scores and retained if the LOEUF value was below, or equal to, 0.25. Finally, variants were removed if they were present in gnomAD.

3.2.3.8 UTR variants

Variants intersecting DDG2P genes and that were annotated as a untranslated region (UTR) variant in participant VCF files enrolled in the V11 GE data release (18/12/2020) were annotated with VEP, including the 'UTRannotator' plugin (Zhang et al., 2020). Variants were subsequently filtered by UTR consequence, CADD score greater than 20 or missing, and screened manually.

3.2.3.9 Burden Analysis

First, all variants intersecting with genes listed on Ensembl were called from participant VCF files in probands with an ethnicity listed as White-British, or not stated (73 families), and a normalised disease grouping of 'Craniosynostosis syndromes'. Variants were filtered by gnomAD AF less than 0.000045, based on previously reported frameworks for a maximum credible AF for a causative variant in craniosynostosis (Calpena et al., 2020, Whiffin et al., 2017), and a CADD score of $\geq$ 20 or missing. For any gene with more than one variant identified aligning to both filters, variants were considered for further scrutiny, including functional consequence, amino acid conservation of missense variants, rarity, evidence in animal models, similarity of patient phenotypes (as reported by HPO terms),

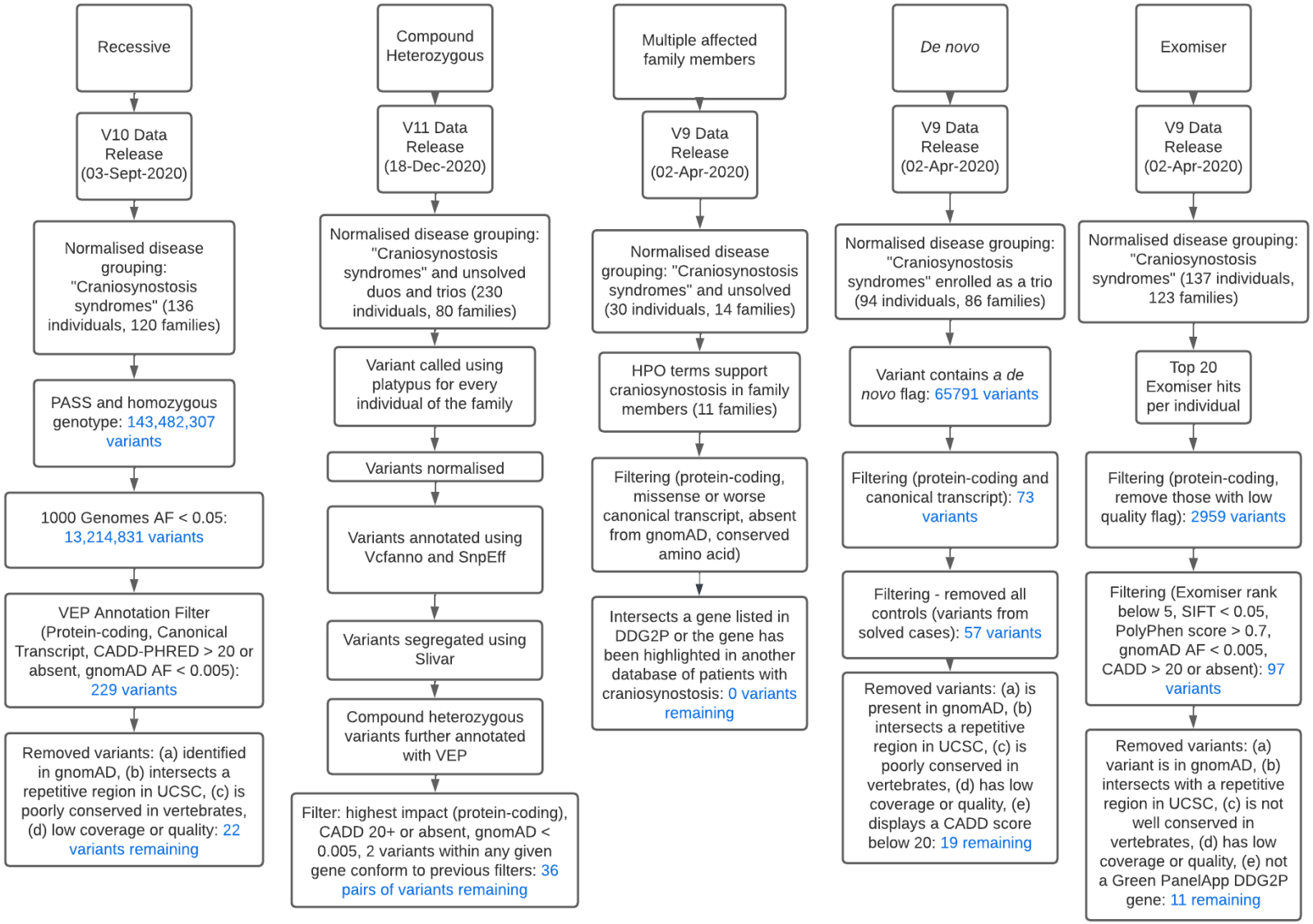
and gene function. Candidate genes were screened against the 100kGP rare disease cohort for incidence of 'likely pathogenic' variants across normalised disease groupings. Burden was calculated as the number of individuals within a normalised disease category with likely pathogenic variants identified in the gene of interest, divided by the total number of probands enrolled under that normalised disease grouping.

3.2.4 Validation of a variant in *BCL11B*

3.2.4.1 Deep Sequencing

DNA extracted from blood and saliva samples were analysed by NGS using a MiSeq instrument. Further experimental details are described in chapter 2, section 2.3. Primers containing CS tags were used to amplify a 213 bp fragment of *BCL11B*: forward (5'-ACACTGACGACATGGTTCTACAGATGAGCCTTCCAGCTACATT-3'), reverse: (5'-TACGGTAGCAGAGACTTGGTCTCTGTCGCCCAGGAAATTCA-3').

A


Recessive
V10 Data Release (03-Sept-2020)
Normalised disease grouping: "Craniosynostosis syndromes" (136 individuals, 120 families)
PASS and homozygous genotype: 143,482,307 variants
1000 Genomes AF < 0.05: 13,214,831 variants
VEP Annotation Filter (Protein-coding, Canonical Transcript, CADD-PHRED > 20 or absent, gnomAD AF < 0.005): 229 variants
Removed variants: (a) identified in gnomAD, (b) intersects a repetitive region in UCSC, (c) is poorly conserved in vertebrates, (d) low coverage or quality: 22 variants remaining
Compound Heterozygous
V11 Data Release (18-Dec-2020)
Normalised disease grouping: "Craniosynostosis syndromes" and unsolved duos and trios (230 individuals, 80 families)
Variant called using platypus for every individual of the family
Variants normalised
Variants annotated using Vcfanno and SnpEff
Variants segregated using Slivar
Compound heterozygous variants further annotated with VEP
Filter: highest impact (protein-coding), CADD 20+ or absent, gnomAD < 0.005, 2 variants within any given gene conform to previous filters: 36 pairs of variants remaining
Multiple affected family members
V9 Data Release (02-Apr-2020)
Normalised disease grouping: "Craniosynostosis syndromes" and unsolved (30 individuals, 14 families)
HPO terms support craniosynostosis in family members (11 families)
Filtering (protein-coding, missense or worse canonical transcript, absent from gnomAD, conserved amino acid)
Intersects a gene listed in DDG2P or the gene has been highlighted in another database of patients with craniosynostosis: 0 variants remaining
De novo
V9 Data Release (02-Apr-2020)
Normalised disease grouping: "Craniosynostosis syndromes" enrolled as a trio (94 individuals, 86 families)
Variant contains a de novo flag: 65791 variants
Filtering (protein-coding and canonical transcript): 73 variants
Filtering - removed all controls (variants from solved cases): 57 variants
Removed variants: (a) is present in gnomAD, (b) intersects a repetitive region in UCSC, (c) is poorly conserved in vertebrates, (d) has low coverage or quality, (e) displays a CADD score below 20: 19 remaining
Exomiser
V9 Data Release (02-Apr-2020)
Normalised disease grouping: "Craniosynostosis syndromes" (137 individuals, 123 families)
Top 20 Exomiser hits per individual
Filtering (protein-coding, remove those with low quality flag): 2959 variants
Filtering (Exomiser rank below 5, SIFT < 0.05, PolyPhen score > 0.7, gnomAD AF < 0.005, CADD > 20 or absent): 97 variants
Removed variants: (a) variant is in gnomAD, (b) intersects with a repetitive region in UCSC, (c) is not well conserved in vertebrates, (d) has low coverage or quality, (e) not a Green PanelApp DDG2P gene: 11 remaining

B

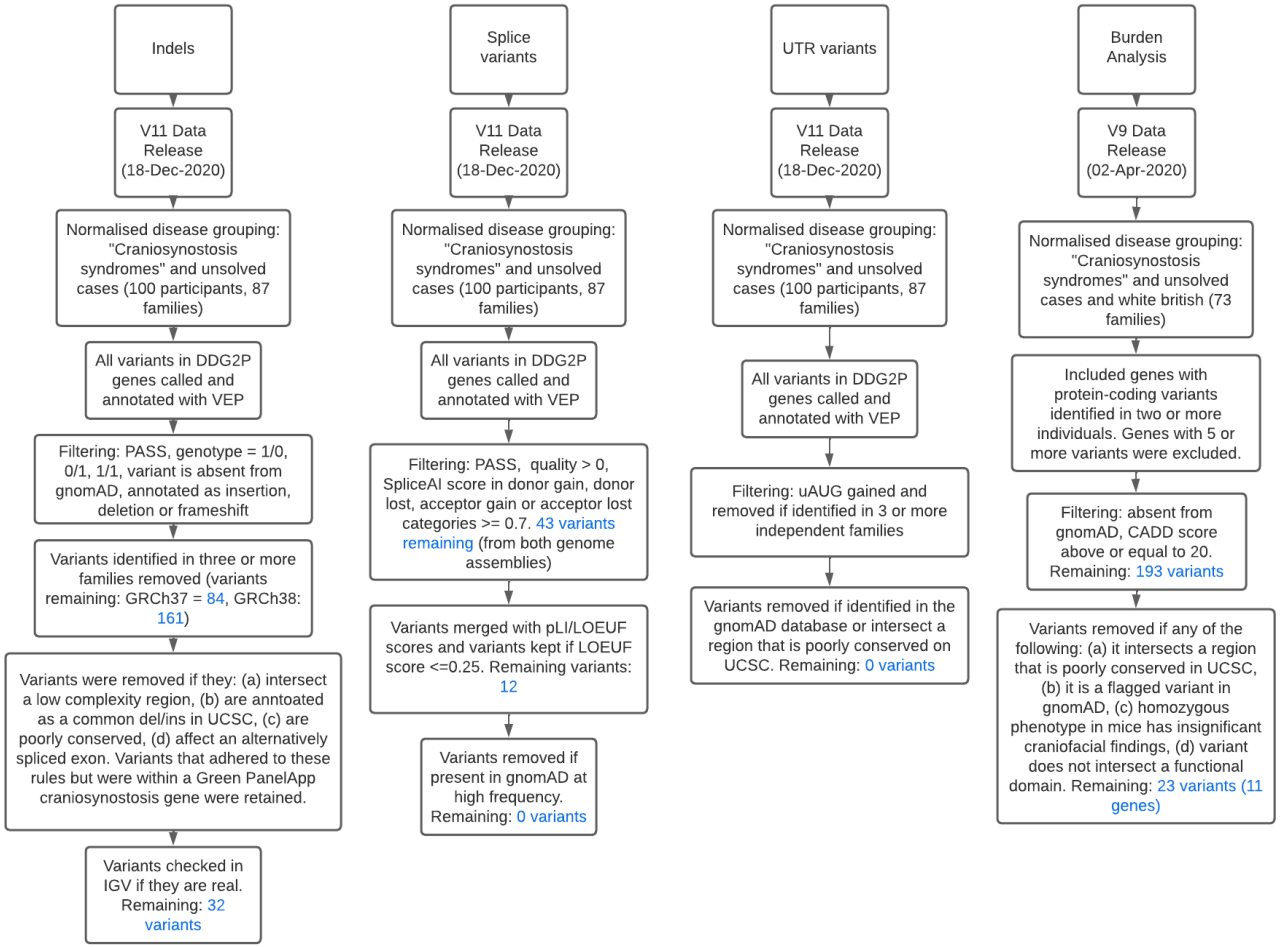


Figure 3.1 Filtering strategies and workflows used in bioinformatic analyses. For each workflow, the filtering criteria has been detailed, alongside which version of the GE data-release was used. The number of variants remaining after filtering are annotated in blue and correspond to tables A3.1–6. (A) Analysis of recessive, compound heterozygous, segregating, and *de novo* variants, alongside the top 20 Exomiser hits. (B) Analysis of indels, splice and UTR variants, and the workflow for the burden analysis.

3.3 Results

All variants considered diagnostic, or which displayed reasonable evidence for pathogenicity and were therefore considered as candidates, are detailed in table 3.1 and 3.2 respectively. The summary of analyses is provided in the same order as outlined in the methods (sections 3.2.3.1–9) and figure 3.1.

3.3.1 Homozygosity/hemizygosity: A homozygous variant in *TMCO1* in a family with reported consanguinity

Following filtering, 229 variants remained which were manually screened (figure 3.1). Variants were removed if homozygotes were reported in gnomAD with the alternate allele or displayed a low-confidence annotation flag, or if a missense variant intersected a region that was poorly conserved. This left 22 variants remaining (table A3.1); a previously identified homozygous variant in *SPRY1* (encoding p.(Leu27*), see chapter 5) was used as a positive control during filtering. This analysis revealed a homozygous nonsense variant in *TMCO1*, encoding p.(Trp78*) (table 3.1), in a patient annotated under the normalised disease grouping of 'Craniosynostosis syndromes'. This gene was missed by the GE triaged bioinformatics pipelines as *TMCO1* was not Green in PanelApp at the time the patient was recruited. The variant was reported as a researcher-identified potential diagnosis (RIPD) and later classified by the GMC as pathogenic. However, despite this patient being enrolled within the primary cohort considered in this screen, the proband lacked HPO terms for craniosynostosis.

3.3.1.1 *TMCO1*

Transmembrane coiled-coil domains 1 (*TMCO1*) is a channel regulator. It functions to permit the flow of calcium ions out of the endoplasmic reticulum and mediates tail-anchored protein insertion (Sharkia et al., 2019). Loss-of-function alleles in *TMCO1* are

associated with craniofacial dysmorphism, skeletal anomalies, and mental retardation in a documented Amish population with multiple affected siblings from a consanguineous union (Xin et al., 2010). Later reports identified similar loss-of-function variants in consanguineous families (Caglayan et al., 2013, Pehlivan et al., 2014) with consistent clinical features: low set ears, synophrys, hypertelorism, and craniosynostosis. Tall stature (above the 99th percentile) was documented in the original report (Xin et al., 2010) and has since been noted in a further three individuals with loss-of-function variants in *TMCO1* (Sharkia et al., 2019).

3.3.2 Compound heterozygous variants in *SPTBN1*

Analysis of compound heterozygous variants identified 36 pairs of variants requiring manual screening (figure 3.1, table A3.2). Variants were assessed manually considering the AF reported in gnomAD, known animal models with significant craniofacial findings, tolerance of either amino acid to missense variants and constraint of neighbouring amino acids, and information reported in ClinVar. Following this, compound heterozygous variants were highlighted in *SPTBN1*, encoding p.(His111Gln) and p.(Ser1895Cys), as candidates requiring further investigation (table 3.2; for further information see section 3.4.1.2).

3.3.3 Analysis of segregating variants in multiply affected families

No additional likely pathogenic variant was identified through analysis of segregating variants in families in which an additional family member presented with craniosynostosis. Variants that were found to segregate were discounted for multiple reasons, including: (a) being located in a non-coding region, (b) affecting a gene with insufficient evidence on gene function and no other supporting variant identified in another patient or animal model, or (c) having a predicted low impact consequence. In total, very few diagnoses were made in families with more than one affected individual.

Indeed, only 28.6% and 16.7% of families with an affected parent (families enrolled as a duo or trio) or affected sibling (enrolled as a quartet), respectively, received a likely genetic diagnosis (Hyder et al., 2021). Analysis of family units may have not been powerful in this instance owing to no large pedigrees being enrolled to the 100kGP or perhaps owing to additional interactors (for example, polygenic or environmental).

3.3.4 DNM analysis: four missing diagnoses and one candidate disease gene

After filtering, the remaining 19 variants (figure 3.1, table A3.3) were screened manually, assessing the consequence (missense or worse), the conservation of the affected position amongst vertebrates (not considered for frameshifts, indels, or stop-gains), supporting animal models or human disease associations, and severity of the consequence. This highlighted a heterozygous frameshift in *PTCH1*, encoding p.(Asn97Thrfs*20), and a missense substitution in *TCF12*, p.(Leu624Phe) (table 3.1), which were considered diagnostic. Two further DNMs were identified in the indel analysis (*ARID1B* and *BCL11B*; section 3.3.6) but were filtered out of the initial DNM screen owing to assigned quality filters within the GE *de novo* pipeline. In addition, a heterozygous missense variant affecting *ANKRD28* was highlighted as a candidate (table 3.2; further information in section 3.4.2.3) because of its role in the inhibition of phosphatase activity of *PPP1C* (Kwiek et al., 2006a, Stefansson et al., 2008). As such, *ANKRD28* was included in a resequencing panel of unsolved cases of craniosynostosis in downstream analyses (see chapter 4). Other variants which were considered likely damaging based on the consequence, position within the protein, and amino acid conservation within vertebrates (*CALD1*, *FGD6*, and *EPHB3*) but for which there was limited supporting data were discounted from further experimental analyses.

3.3.4.1 *PTCH1*

The membrane-bound receptor PTCH1 is a regulator of the hedgehog signalling pathway (see chapter 1, section 1.7.4.1). Upon ligand binding by SHH (Cohen et al., 2015), PTCH1 releases its repression on the G-protein coupled receptor, SMO, thus increasing downstream hedgehog signalling. The hedgehog signalling pathway is essential for mesodermal tissue patterning and differentiation during early development. Dysregulation of this pathway results in craniofacial abnormalities, including holoprosencephaly, hypotelorism, and craniosynostosis (Twigg et al., 2012, Pan et al., 2013, Twigg et al., 2016) (see chapter 7 for further information). Considering the role of this protein in a known developmental pathway, and the absence of any other convincing pathogenic variant, this variant was considered diagnostic (see section 3.4.2.1 for more information).

3.3.4.2 *TCF12*

In addition, the DNM analysis identified a heterozygous missense variant in *TCF12*, encoding p.(Leu624Phe) (table 3.1), one of the 7 most frequently mutated genes in craniosynostosis. Of note, this variant had been identified by the diagnostic lab involved in the case prior to enrolment within the 100kGP but had been missed by pipelines employed by GE/GMC owing to incomplete penetrance within the family (see section 3.4.2.2 for more information).

3.3.5 Analysis of the top Exomiser hits identifies a variant in *SMAD2*

Analysis of the top 20 Exomiser hits identified a variant in mothers against decapentaplegic homolog 2 (*SMAD2*), encoding p.(Leu408Pro) (table 3.1, table A3.4), which was submitted as a RIPD. Additionally, this method highlighted the *PTCH1* and *ANKRD28* variants as the top ranked candidates for those individuals.

3.3.5.1 *SMAD2*

SMAD2 is a member of the SMAD family of proteins, which are signal transducers and transcriptional moderators of multiple signalling pathways, namely involving TGFβ signalling. SMAD2 is recruited to the TGFβ receptor by its MH2 domain through an interaction with the SMAD anchor for receptor activation (SARA) protein. Following phosphorylation, SMAD2 dissociates from SMAD4 allowing translocation into the nucleus to bind to target promoters (Hata and Chen, 2016). Variants affecting TGFβ signalling (including those in *TGFBR2*, *TGFBR3*, *SMAD2* and *SMAD3*) are associated with Loeys-Dietz syndrome (LDS), a connective tissue disorder affecting the cardiovascular, skeletal, and ocular systems. LDS patients often present with aortic aneurysms, hypertelorism and bifid/broad uvula or cleft palate, with craniosynostosis described in a subset of patients with LDS (Schepers et al., 2018).

3.3.6 Indels are frequently missed from the GE/GMC triaged bioinformatics pipelines

Following filtering, the remaining 32 variants (figure 3.1, table A3.5) were checked manually in IGV to ascertain if they were real or an artifact of sequencing. Variants in *ARID1B, BCL11B, MEGF8 and MMP21* were identified as likely pathogenic with an additional variant in tousled-like kinase 2 (*TLK2*) (NM_006852.3:c.1316_1331del; p.(Arg439Ilefs*17)) found in a patient with craniosynostosis classified within the 'Intellectual Disability' cohort of the 100kGP. Of the four variants identified in the primary craniosynostosis cohort, two of these were *de novo* (*ARID1B* and *BCL11B*; table 3.1) and two identified alongside missense variants as biallelic (*MEGF8* and *MMP21*; table 3.1). Of note, the *BCL11B* variant was identified in only 4/39 reads indicating potential mosaicism. The level of variation was quantified using deep sequencing of DNA obtained from blood and saliva samples; data from triplicate technical analyses identified 20%

(3215/16320) of variant reads in the blood sample and 34% (4114/12240) in saliva, confirming mosaicism.

3.3.6.1 *ARID1B*

AT-rich interactive domain-containing protein 1B (ARID1B) is a DNA-binding subunit of the Brahma-associated factor (BAF) chromatin remodelling complexes. ARID1B functions to regulate differentiation of stem cells exiting the cell cycle in postmitotic neurones (Sim et al., 2015). Variants in *ARID1B* are the most common cause of intellectual disability and are also associated with Coffin-Siris syndrome (Mignot et al., 2016), with defining features including aplasia or hypoplasia of the distal phalanx, developmental delay, and distinctive craniofacial features. Craniosynostosis has been described in two previous cases of Coffin-Siris syndrome: one with a variant in *ARID1B* (Mignot et al., 2016), and a second with a 2p25 deletion encompassing *SOX11* (Hempel et al., 2016).

3.3.6.2 *BCL11B*

The BAF chromatin remodelling complex subunit (*BCL11B*) gene is a key transcription factor with multiple roles, including regulating early development of the central nervous, immune, and cardiac systems (Lennon et al., 2017). Previously, knock-out studies of *Bcl11b* in mice demonstrated severe coronal craniosynostosis from embryonic day E16.5 (Kyrylkova et al., 2016) with loss of *Bcl11b* resulting in ectopic expression of *Runx2* and *Fgfr2* and downregulation of *Twist1*; this pattern of gene expression is characteristic of early osteogenic differentiation (Kronenberg, 2004). Following these reports, a patient was identified with a heterozygous *de novo* variant (c.7C>A; p.(Arg3Ser)) affecting a critical residue involved in binding the RBBP4-MTA1 complex. A murine model of the p.(Arg3Ser) variant at post-natal day 0, also displayed coronal craniosynostosis and reduced anterior fontanelles (Goos et al., 2019).

3.3.6.3 *MEGF8*

MEGF8 is a negative regulator of hedgehog signalling. The transmembrane protein assembles with MGRN1, a RING superfamily E3 ligase, to form a complex which catalyses the ubiquitination and degradation of SMO (Kong et al., 2020). As such, loss-of-function variants in *MEGF8* result in increased hedgehog signalling. While mice heterozygous for loss-of-function variants in *Megf8* are phenotypically normal, homozygous mutants display preaxial digit duplication, heterotaxy, and congenital heart defects (Kong et al., 2020). Variants in *MEGF8* have been described in humans to be associated with Carpenter syndrome, a recessive disorder characterised by multi-suture craniosynostosis, polysyndactyly of the hand and feet, and obesity, with the cause being determined as increased hedgehog signalling (Twigg et al., 2012). However, a key feature of the human *MEGF8* Carpenter syndrome-like phenotype is the identification of defective lateralisation, with three individuals reported with complete situs inversus, dextrocardia, and transposition of the great arteries (Twigg et al., 2012), consistent with the reported homozygous phenotype in mice. This phenotype was also documented in an ENU-induced mouse model with a missense variant in *Megf8* (p.(Cys193Arg)) (Zhang et al., 2009).

3.3.6.4 *MMP21*

Matrix metallopeptidase 21 (*MMP21*) is involved in the breakdown of extracellular matrix during embryonic development and tissue remodelling. A recessive screen in mice identified a role for *MMP21* in congenital heart disease, owing to abnormal ciliary function (Li et al., 2015). Following this report, recessive variants in *MMP21* were identified in nine independent cases of heterotaxy (the failure to establish normal left-right asymmetry) (Guimier et al., 2015) but, to date, there are no known cases of variants in *MMP21* and craniosynostosis.

3.3.7 No likely pathogenic UTR or splicing variants were identified

After filtering, no variants remained with reasonable evidence for pathogenicity.

3.3.8 Two candidate disease genes were identified through a burden analysis of the craniosynostosis cohort and the wider rare disease cohort

Analysis of two or more variants within a given gene in independent families resulted in 23 variants in 11 genes remaining (figure 3.1, table A3.6), which were uniquely absent from gnomAD, resided within or would disrupt a functional domain, and were identified within genes for which there were convincing animal models to support abnormal craniofacial phenotypes. This analysis identified two variants in transcription factor CP2 (*TFCP2*): a missense variant and a stop-gain in independent families (table 3.2; see section 3.4.5.2 for more information). Further scrutiny of the CGG in-house database (CranioVar) identified one further family harbouring an inherited heterozygous variant affecting a canonical splice site in *TFCP2*. In addition, two variants in heat shock protein 90 alpha family class A member 1 (*HSP90AA1*) (table 3.2; see section 3.4.5.1 for more information) were considered as likely pathogenic owing to correct segregation of the variants within affected members of the same family, in addition to meeting the filtering criteria described. Both genes were considered further in a resequencing analysis of unsolved patients with craniosynostosis (see chapter 4).

Table 3.1 Diagnostic variants identified within the 100kGP

Screen	Gene	Amino acid change	cDNA change	GT[a]	gnomAD AF (v2.1.1)	CADD
De novo	*PTCH1*	p.(Asn97Thrfs*20)	NM_000264.3:c.290del	0/1	-	-
De novo	*TCF12*	p.(Leu624Phe)	NM_207036.1:c.1870C>T	0/1	-	26.5
Exomiser	*SMAD2*	p.(Leu408Pro)	NM_005901.5:c.1223T>C	0/1	-	30
Indel	*ARID1B*	p.(Gly1199Profs*14)	NM_020732.3:c.3594delinsCCCCCA	0/1	-	-
Indel	*BCL11B*	p.(Gly270Alafs*256)	NM_138576.3:c.808_809insCGAGCG GCGGCGGGATGGTGAGCCGCGG	0/1	-	-
Indel	*MMP21*	p.([Val224Glyfs*29]); ([His259Asp])	NM_147191.1:c.[671_684del]; [775C>G]	0/1 1/0	-	-;26.4
Indel	*MEGF8*	p.([Arg1499His]); ([Phe2589del])	NM_001410.2:c.[4496G>A]; [7766_7768delTCT]	0/1 1/0	-	27;-
Recessive	*TMCO1*	p.(Trp78*)	NM_019026.4:c.233G>A	1/1	-	40

Table 3.2 Candidate genes and the associated variants identified through independent analyses of the 100kGP craniosynostosis cohort/ CranioDB

Screen	Gene	Amino acid change	cDNA change	GT[a]	gnomAD AF (v.2.1.1)	CADD
De novo	*ANKRD28*	p.(Asn696Tyr)	NM_001195099.1:c.2086A>T	0/1	-	28.6
Burden	*HSP90AA1*	p.(Glu15Glyfs*4)	NM_005348.3:c.43dup	0/1	-	-
Burden	*HSP90AA1*	p.(Met119Thr)	NM_005348.3:c.356T>C	0/1	-	29.8
Burden	*TFCP2*	p.(Lys236Arg)	NM_005653.4:c.707A>G	0/1	-	29.8
Burden	*TFCP2*	p.(Glu110*)	NM_005653.4:c.328G>T	0/1	-	39
Burden	*TFCP2*	p.(?)	NC_000012.12:c.966+3A>T	0/1	-	23.4
Compound heterozygous	*SPTBN1*	p.([His111Gln]); ([Ser1895Cys])	NM_003128.2:c.[333C>G];[5684C>G]	0/1	-	25.3; 24.2

[a]GT = Genotype; 0/1 = heterozygous, 1/1 = homozygous for the alternative allele

3.4 Discussion

In parallel analyses, several bioinformatic pipelines were employed to identify likely pathogenic variants within the craniosynostosis cohort enrolled in the 100kGP, aiming to identify candidate genes involved in the pathogenesis of craniosynostosis. Through these methods, eight families (with a previously unsolved genetic basis to their craniosynostosis) received a genetic diagnosis, and a further four genes (eight variants) were highlighted as candidates for further scrutiny. Although identification of diagnostic variants was not the primary outcome of this analysis, scrutiny of this cohort highlighted distinct categories of variants which were missed from the GE triaged bioinformatics pipelines, primarily *de novo* variants (outside of Green PanelApp genes) and indels, highlighting systematic errors with the clinically implemented pipelines. Of note, although the contributions of my colleagues made an important uplift to the total diagnoses of the 100kGP craniosynostosis cohort, the focus here is on the variants identified through my own methods (*ARID1B, BCL11B, MEGF8, MMP21, PTCH1, SMAD2, TCF12* and *TMCO1*).

3.4.1 Recessive analysis

Recessive variants make up a small percentage (3–4%) of known diagnoses for developmental disorders (Martin et al., 2018); although autosomal recessive inheritance is documented in seven out of 20 (35%) core genes for which specific types of variants are associated with craniosynostosis in more than half of affected individuals (data taken from an Oxford cohort across a 10-year period) (Twigg and Wilkie, 2015), presentation of patients in clinic with variants in these genes is very rare. It is therefore not surprising that very few diagnoses were made through screening the 100kGP craniosynostosis cohort for recessive variants.

Within the 100kGP primary craniosynostosis cohort (v9 data release, 02/04/2022), there were 12/115 families where consanguinity was confidently reported or suspected; of these, 3/12 (25%) had an identified likely-causative variant (*IFT80*, *HUWE1*, *SPRY1*), leaving 8 families undiagnosed (see chapter 5 for further information on *SPRY1*). Despite efforts to screen large regions of homozygosity (ROH), defined as any region greater than 2 Mb (Hildebrandt et al., 2009), only one further candidate was identified – a stop-gain in *TMCO1* (p.(Trp78*)).

Given the possibility that recessive variants could exist outside of the consanguineous cohort, albeit at lower frequency, a further screen of the wider population was conducted. This aimed to identify homozygous or compound heterozygous variants within the primary craniosynostosis cohort. Perhaps in line with figures indicating the low proportion of developmental disorders which are attributable to recessive variants (Martin et al., 2018), no further clinical-grade diagnoses were identified. However, compound heterozygous variants were flagged for further review in *SPTBN1:* p.([His111Gln]);([Ser1895Cys]).

3.4.1.1 *TMCO1*

Notably, while *TMCO1* did not appear in the Tiering data, it was ranked in Exomiser at position 1 despite the individual displaying no HPO terms for craniosynostosis. Nevertheless, the patient displayed consistent clinical features previously attributed to *TMCO1* loss-of-function variants (section 3.3.1.1). The stop-gain was present in gnomAD (v2.1.1) in three alleles out of 250,000 but no homozygous p.(Trp78*) variants were reported at this position. Additionally, there are no reported homozygous variants in gnomAD at any amino acid within the gene, thus supporting selection against homozygosity in the general population. Given the high concordance of the phenotype described in the literature (section 3.3.1.1) and the documented clinical features of the

participant in the 100kGP, alongside the absence of any other likely candidate, this variant was reported to the GMC and considered pathogenic.

3.4.1.2 *SPTBN1*

The βII-spectrin gene (*SPTBN1*) is a member of the spectrin superfamily of proteins which connect cytoskeletal elements to the cell membrane and the nucleus. The classifying feature of these proteins is the presence of 2 to 74 tandem arrays known as spectrin repeats (SR) (between 106–122 residues in length) (Liem, 2016). In addition, some members of the spectrin family contain two tandem calponin-homology domains (CH1 and CH2) that bind to actin. To date, a spectrum of *SPTBN1* heterozygous missense and loss-of-function variants have been reported in patients with neurodevelopmental disorders (Cousin et al., 2021, Rosenfeld et al., 2021), encompassing variants in both CH domains and multiple SR domains, but craniosynostosis has not been described. Nevertheless, homozygous knockout of *Sptbn1* in mouse models results in lethality, with embryos (E9.5) displaying a reduced head size and absence of a branching network of vessels in the yolk sac (Tang et al., 2003).

Compound heterozygous variants were identified in a patient with craniosynostosis (table 3.2) enrolled in the 100kGP. The p.(His111Gln) variant resides within the CH1 domain, which unlike CH2, is capable of binding actin alone (Gimona and Winder, 1998); the p.(Ser1895Cys) variant resides within SR15 which mediates ankyrin binding (Djinovic-Carugoa et al., 2002). Abrogation of the SR15 domain destroys the ability to bind ankyrin, yet this interaction is entirely phenocopied by substitution of a key tyrosine residue within SR15 (p.(Tyr1874Ala)) (Cousin et al., 2021). A documented nearby heterozygous missense variant within SR15 (p.(Glu1886Gln)) was shown to cause abnormal cell membrane protrusions in induced pluripotent stem cells (Cousin et al.,

2021) but did not affect the ability to bind ankyrin-B. Although craniosynostosis has not been described in neurodevelopmental disorder patients with variants in *SPTBN1*, the spectrum of variants described to date have all been monoallelic. Biallelic variants in murine models result in abnormalities of the head with the authors noting the similarity of the phenotype of *Sptbn1*$^{-/-}$ mice to *Smad2*$^{+/-}$/*Smad3*$^{+/-}$ mice (Tang et al., 2003). In addition, deficiency of *Sptbn1* was associated with the mis-localisation of *Smad3* and *Smad4* and loss of the TGF-β dependent transcriptional response, a known pathway in the pathogenesis of craniosynostosis (Twigg and Wilkie, 2015). Taken together, variants within this gene were considered as candidates in further analyses (see chapter 4).

3.4.2 *De novo* variant analysis

In accordance with current literature indicating that almost 50% of developmental disorders are attributable to DNMs (Kaplanis et al., 2020, Deciphering Developmental Disorders Study, 2017), analysis of *de novo* variants was the single best way to increase diagnostic yield within the 100kGP craniosynostosis cohort, accounting for 26/34 (76%) of the total diagnoses (researcher-identified and GE/GMC) (Hyder et al., 2021). Given that there are on average fewer than 2 protein coding DNM in each genome (Deciphering Developmental Disorders Study, 2017) analysis of this dataset offered high sensitivity with minimal time burden.

Of the identified DNMs, 10 (29%) of the single-nucleotide or indel-associated DNMs were missed by the GE/GMC bioinformatics pipelines. Five DNMs were identified within Green genes on panels applied to those individuals but were missed as the gene was not Green when the patient was enrolled (*SMAD6*, *TRAF7*, *HNRNPK*, *FBXO11, PTCH1*). At the time, no process existed for GMCs to routinely reanalyse unsolved cases following panel updates. An additional two DNMs were missed by the GE/GMC pipelines (*ARID1B* and *TCF12*) because of insufficient calling of indels and reliance on HPO terms for assigning

segregation patterns within families, respectively. Of note, both *ARID1B* and *TCF12* are Green genes on panels applied to those individuals. Details of the *PTCH1*, *TCF12 and ARID1B* variants shall follow as these were identified through my analyses.

3.4.2.1 A pathogenic heterozygous frameshift variant in *PTCH1*

Analysis of the 100kGP identified a *de novo* frameshift variant in *PTCH1* (p.(Asn97Thrfs*20)) in a patient with metopic craniosynostosis, which was absent from gnomAD and predicted to abolish all functional domains. Given that aberrations in hedgehog signalling are frequently associated with craniosynostosis (further details in chapter 7), this variant was considered likely pathogenic. Additionally, the phenotype of this patient was consistent with previous literature of a frameshift variant (encoding p.(Val502Glyfs*13)) in a patient with metopic craniosynostosis (Beltrami et al., 2020), alongside patients with deletions of 9q22.3 (which includes *PTCH1*) that classically present with metopic synostosis (Yamada et al., 2020). A further two patients were detailed in an abstract at the 2021 European Society of Human Genetics (ESHG) conference with loss-of-function variants in *PTCH1* (a *de novo* frameshift and an inherited stop-gain from an affected parent), both with metopic craniosynostosis (accessed: https://www.abstractsonline.com/pp8/#!/9102/presentation/1801). Importantly, heterozygous variants in *PTCH1* are a known cause of nevoid basal cell carcinoma syndrome (NBCCS) (Hong et al., 2016, Martinez et al., 2019, Johnson et al., 1996). Although NBCCS was not described in the proband recruited to the 100kGP with craniosynostosis, this condition typically presents from 35 years of age (Gorlin, 1995) and should therefore be monitored as the patient gets older. Importantly, this diagnosis has immediate implications for managing the risk of developing complications associated with NBCCS, including wearing protective sunscreen, avoiding excess sun exposure and ultraviolet (UV) protection on sunglasses (Lo Muzio, 2008).

3.4.2.2 A *de novo TCF12* variant in one member of a multiply affected family

A variant in *TCF12*, encoding p.(Leu624Phe), was identified in one member of a family comprised of three male siblings with differing cranial phenotypes. All siblings were listed under the normalised disease grouping of 'Craniosynostosis syndromes', despite one of the siblings displaying a much milder phenotype of wormian bones. The Tiering pipeline assumes that individuals within the same family unit, and who are assigned the same disease grouping, are likely to have inherited the same variant and, as such, filters out any unshared variation. Incidentally, variants at Leu624 are already reported in the literature to cause coronal craniosynostosis (Sharma et al., 2013) and therefore, despite the lack of segregation in the male siblings, this variant was considered pathogenic. The variant affects a highly conserved residue within the bHLH DNA-binding/dimerisation domain of *TCF12.* Concordant with this observation, it was shown that variants within the bHLH domain resulted in an impairment of the ability of the protein to form homo- or heterodimers, thus reducing overall protein activity by almost 75% (Sharma et al., 2013). Further reports of patients with coronal craniosynostosis and variants coinciding with the bHLH domain of *TCF12* support these conclusions (Di Rocco et al., 2014). It's important to note however that based on the assumption of shared phenotypic variation within families, it is logical for pipelines to filter out any non-segregating variants (Ng et al., 2010), however, this strategy likely performs better in developmental disorders with more discrete phenotypes, rather than more heterogenous conditions like craniosynostosis.

3.4.2.3 *ANKRD28* was highlighted as a candidate disease gene for further scrutiny

Additionally, a *de novo* missense variant, encoding p.(Asn696Tyr), within ankyrin repeat domain 28 (*ANKRD28*) was identified as a compelling novel candidate in a patient with metopic craniosynostosis and cardiac abnormalities. The variant was absent from gnomAD and led to substitution of a highly conserved residue within one of the ankyrin repeat domains. *ANKRD28* functions as a putative regulatory factor of protein phosphatase 6. It has been shown to selectively inhibit the phosphatase activity of

protein phosphatase 1 catalytic subunit (*PPP1C)* which in turn moderates the phosphorylation of the transcriptional regulator, heterogenous nuclear ribonucleoprotein K (*HNRNPK*) (Kwiek et al., 2006b, Stefansson et al., 2008). Of note, variants in *PPP1CB* are associated with Noonan syndrome (Bertola et al., 2017, Huckstadt et al., 2021), in which craniosynostosis is a feature. In addition, several individuals with loss-of-function variants in *HNRNPK* (reviewed in Goos and Mathijssen, 2019) display clinical features consistent with Au-Kline syndrome (intellectual disability, unique craniofacial phenotype, cardiac anomalies, and in some cases craniosynostosis). *ANKRD28* has been shown to be significantly upregulated (1.9-fold) in left ventricular hypertrophied tissue (Eleftheriadou et al., 2017) but further studies are required to define its role in cardiac anomalies. While there are no known human disease associations of variants in *ANKRD28*, to date, two out of the 34 diagnoses made in the 100kGP primary craniosynostosis cohort are attributable to variants in *PPP1CB* and *HNRNPK*.

3.4.3 Analysis of pathogenic indels

Second to *de novo* variants, a large proportion of missing diagnoses were a result of the failure to identify indels which were filtered out early in the GE pipelines. The initial discovery of a missing indel in *MEGF8* highlighted a systematic error in the GE bioinformatics workflow for indel identification. As such, I developed a pipeline for indel calling from the GE dataset, described in section 3.2.3.6, which facilitated the identification of a further three missing indels in *ARID1B, BCL11B* and *MMP21*. Of note, two out of the four missing indels were also DNM (*ARID1B* and *BCL11B*) yet these had not initially been called in the DNM screen due to low quality scores.

3.4.3.1 *MEGF8*

The first case of a missing indel was identified in a patient with clinical features suggestive of Carpenter syndrome. Although variants in three genes have been

described in Carpenter syndrome (*RAB23, MEGF8* and, more rarely, *GLI3*) variation in *MEGF8* was scrutinised further. It was possible that a variant in *RAB23* could have been missed, but this gene should have been routinely screened prior to the patient's enrolment in the 100kGP and therefore a variant in *MEGF8* was considered more likely. Similar to the *TCF12* case, a heterozygous missense variant in *MEGF8*, encoding p.(Arg1499His) (table 3.1), was identified at a position previously reported to be pathogenic for craniosynostosis, specifically in Carpenter syndrome (Twigg et al., 2012). Twigg and colleagues showed that morpholino knockdown of *MEGF8* in zebrafish embryos caused epiboly defects – a phenotype which was largely rescued by injecting *MEGF8* RNA (12% embryos with epiboly as opposed to 83% with morpholino alone). However, injection of RNA containing the p.(Arg1499His) missense variant was unable to entirely rescue the phenotype with 60% of embryos still displaying epiboly defects, indicating defective early gastrulation (Twigg et al., 2012). Despite this previous functional evidence and the proband displaying the defining features of Carpenter syndrome, this missense variant had not been tiered and was ranked in Exomiser at position 96 as *MEGF8* is a recessive disease gene and only one heterozygous variant had been identified. Further scrutiny of variation in *MEGF8*, primarily removing all assigned quality filters used in the GE bioinformatics pipelines, highlighted a heterozygous deletion of one amino acid, p.(Phe2589del), which was absent from gnomAD and highly conserved, but coincided with a repetitive region of the DNA sequence. Although this patient was recruited as a singleton, subsequent sequencing of the parental samples confirmed that the two variants were in *trans*. Despite the indel residing within a low complexity region, the strong phenotypic features suggestive of Carpenter syndrome and the identification of two variants in *MEGF8*, alongside functional data supporting the pathogenicity of p.(Arg1499His), was deemed sufficient to classify these variants as diagnostic. Future algorithms should incorporate published variant level data, particularly those in Green panel genes, as a means of retaining variants that may otherwise be filtered out on factors like segregation or MOI.

3.4.3.2 *ARID1B*

ARID1B is the most frequently mutated gene in intellectual disability (ID) (Bogershausen and Wollnik, 2018, Ben-Salem et al., 2016, Wieczorek et al., 2013, Yu et al., 2015) yet its association with craniosynostosis is far less frequent (Mignot et al., 2016). Despite being listed under the normalised disease grouping of 'Craniosynostosis syndromes', a defining feature of the individual in which the *de novo* frameshift was identified (p.(Gly1199Profs*14)) was in fact ID and, therefore, this variant was classified as pathogenic by the GMC. Given that ID was likely the primary phenotype in this case, it shall not be discussed further.

3.4.3.3 *BCL11B*

There is mounting evidence for the role of *BCL11B* variants in coronal craniosynostosis (see section 3.3.6.2) (Goos et al., 2019) and, as such, the identified frameshift (p.(Gly270Alafs*256)) was considered likely damaging. However, only 4/39 (10%) of the reads were attributable to this variant, raising the suspicion of mosaicism. To ascertain if this was true, tissue samples were obtained from two independent embryonic lineages (blood and saliva), which were found on deep sequencing to contain 20% and 34% of variant reads, respectively, thus confirming mosaicism. Of note, while there is substantial evidence building for the involvement of *BCL11B* in craniosynostosis, this gene remains unlisted on PanelApp. If mosaicism is identified to be a frequent cause of variation within this gene, filters considering the number of reads permitted for the alternative allele will need to be made less stringent.

3.4.3.4 *MMP21*

The remaining missing indel was identified in *MMP21* which is a Green panel gene for heterotaxy. Similar to the *MEGF8* case, the indel was identified in a recessive disease gene; in both cases, the second variant (a heterozygous missense) was not specifically

flagged by GE/GMC owing to the assumption of a biallelic MOI and no second hit identified post filtering. A variant in *MMP21*, encoding p.(Val224Glyfs*29), was identified in a patient harbouring a heterozygous missense variant (p.(His259Asp)) on the opposite allele (table 3.1). *MMP21* displays a biallelic MOI with human variants associated with a limited spectrum of clinical features, namely heterotaxy and cardiac anomalies (Guimier et al., 2015). Although craniosynostosis is not a defining feature of *MMP21* variants, the additional clinical features of this proband, particularly the heterotaxy, was sufficient to assign diagnostic status to this variant. Interestingly, this family was one of the 12 reported to be consanguineous in which I failed to identify a compelling homozygous variant contributing to the phenotype; indeed, *MMP21* did not feature in any ROH called by Illumina, thus highlighting the importance of screening variants outside of large ROH (> 2 Mb) within consanguineous families.

3.4.3.5 Difficulties of identifying indel variants within the 100kGP cohort

A difficulty when analysing indels is deciphering the real variants from sequencing noise and therefore a threshold was set to retain variants identified in three or fewer independent families. This threshold was set as recent literature indicates that *SMAD6* variants are the 5th most common cause of craniosynostosis (Calpena et al., 2020) and yet only two cases were identified in the 100kGP. Therefore, it was assumed that it would be unlikely to identify pathogenic indels in more than three cases and instead these would more likely represent common SNPs or sequencing errors. Despite this being one of the largest cohorts of patients with craniosynostosis who have received WGS, it is likely these data are still relatively underpowered, and a much larger cohort would be required to identify two or more novel variants in any given gene.

3.4.4 Exomiser

Perhaps unsurprisingly, the top candidates identified in the Exomiser screen were also identified in the parallel DNM analysis, including variants in *PTCH1* and *ANKRD28* described previously, and one hit in *SMAD2* (p.(Leu408Pro)). The *SMAD2* variant resided within the MH2 domain, a documented 'hot spot' for variants within this protein (Zaidi et al., 2013, Micha et al., 2015, Zhang et al., 2017, Granadillo et al., 2018, Schepers et al., 2018, Cannaerts et al., 2019). While craniosynostosis has never been described in association with *SMAD2* variants, it has been documented in association with variants in other closely related members of the SMAD family, namely *SMAD3* (Schepers et al., 2018, Baskin et al., 2020) and *SMAD6* (Calpena et al., 2020). However, given that there are no other supporting cases of patients with craniosynostosis and variants in *SMAD2*, the GMC reviewed and assigned this variant as a 'hot' variant of unknown significance. Importantly, variants in *SMAD2* have been shown to increase the risk of aortic aneurysms (Cannaerts et al., 2019), and consequently this patient is now receiving follow up clinical assessments.

The Exomiser algorithm performed remarkably well in identifying diagnostic-grade variants within the 100kGP solved craniosynostosis cohort and correctly returned 81% of diagnostic SNVs within the top 10 candidates listed (Hyder et al., 2021). Variants that were not detected through this approach include the *TCF12* variant displaying non-penetrance (section 3.3.4.2) and the described missing indels (section 3.3.6: *MEGF8*, *ARID1B*, *BCL11B*, and *MMP21*). In addition, the identified diagnostic SNVs were annotated as the top ranked candidate in 48% of cases and within the top 3 candidates in 70% of cases. Clearly, my analysis of the top 20 Exomiser hits for all individuals is time intensive and would not be feasible in a routine diagnostic setting. Nevertheless, analysis of the top 3 Exomiser hits would capture many of these diagnoses and, indeed, new NHS England guidelines now stipulate that GMCs must screen the top three

Exomiser hits for every participant; this requirement was not present when I first screened this data.

3.4.5 Burden Analysis

After removing any outstanding diagnostic variants, the remaining unsolved cases were screened for enrichment of pathogenic variants in any given gene, highlighting *HSP90AA1* and *TFCP2* as candidate genes. Variants were retained if they were identified in an unaffected parent and were only removed if they were identified in the gnomAD v2.1.1 database, displayed a CADD score below 20 (except if CADD was unreported in indels), and if only one variant was identified in each gene. To date, there are no documented reports of craniosynostosis and putative pathogenic variants in either *HSP90AA1* or *TFCP2*, but both genes are implicated in pathways with a known role in development.

3.4.5.1 *HSP90AA1*

Heat shock protein 90 alpha family class A member 1 (*HSP90AA1*) encodes a molecular chaperone that aids in the correct folding of specific target proteins. The precise function of *HSP90AA1* is unknown, but it has been implicated in the regulation of autophagy (Yang et al., 2022), glucose homeostasis (Abarkan et al., 2019), cell proliferation in skeletal muscle (Chen et al., 2022b), and recently was identified as a candidate gene in craniofacial disorders (Berenguer et al., 2018). Indeed, *HSP90AA1* was upregulated in mouse embryos exposed to prenatal retinoic acid (RA) (Berenguer et al., 2018). RA has been shown to increase osteogenesis within the sutural mesenchyme (James et al., 2010) and, in support, variants in known regulators of RA degradation (*POR* and *CYP26B1*) have been shown to cause craniosynostosis (Flück et al., 2004, Laue et al., 2011) (see chapter 1, section 1.4.3). Additionally, the variant was considered of

additional interest owing to the segregation of the p.(Met241Thr) missense variant in the affected father with sagittal craniosynostosis.

3.4.5.2 *TFCP2*

TFCP2 is a member of the grainyhead (Grh)/CP2 family of transcriptions factors. The first member of this family was identified in *Drosophila* in mutant embryos displaying fragile cuticles alongside a 'grainy' and discontinuous head skeleton (Bray and Kafatos, 1991). Since these early reports, members of the Grh/CP2 family have emerged as important regulators of craniofacial fate. Loss of *Grhl3* in mice results in neural tube and palatal defects, and premature apposition of the frontal and parietal bones (Ting et al., 2003, Goldie et al., 2016, Peyrard-Janvid et al., 2014), while complete loss of *Grhl2* is associated with developmental delay, facial clefting, exencephaly and mid-gestational embryonic lethality (E11.5) (de Vries et al., 2020). A human phenotype of variants in *TFCP2* has not yet been delineated. From the 100kGP dataset, I identified two heterozygous variants in *TFCP2* in independent families, a stop-gain (p.(Glu110*)) and a missense substitution (p.(Lys236Arg)). Importantly, variation at Lys236, alongside a second substitution at Gln234, have been shown to prevent *TFCP2* from binding DNA (Ming et al., 2018), implicating this amino acid as a key binding residue. In addition, a further patient was identified within our in-house database, CranioVar, to harbour a splicing variant in *TFCP2* (c.966+3A>T; table 3.1). SpliceAI (Jaganathan et al., 2019) predicted a 93% increase in the likelihood that the variant would introduce a new donor acceptor site. Despite this, non-penetrance was evident in all three identified families and therefore further cases would be required to corroborate findings. As a result, *TFCP2* (alongside *HSP90AA1*) was analysed in the resequencing of 617 unsolved patients with craniosynostosis (see chapter 4).

3.4.6 Diagnostic yield

The final diagnostic rate for craniosynostosis within the 100kGP was 29.8% (34/114) with higher success rates for syndromic conditions (39%) compared to non-syndromic (6.25%). Out of these diagnoses, 16 were uniquely identified by the CGG, representing a significant uplift in the number of diagnoses that can be achieved from researcher-identified diagnoses, compared to the GE/GMC pipelines alone (Hyder et al., 2021). Despite this, a large proportion of the craniosynostosis cohort (~70%) remain without a genetic diagnosis. Possible explanations for this include novel disease genes for which there was only one likely pathogenic variant identified within the cohort but for which additional cases would be required to validate the findings. Craniosynostosis has also undergone extensive genetic screening and therefore many of the indisputably pathogenic variants should already have been detected in routine analysis. The more complex intronic and structural variants, which are characteristically harder to identify, and which do not lend themselves to Illumina sequencing, may remain. In addition, we may detect more genes displaying non-penetrance which would require additional cases to be confident of their pathogenicity. As such, the 114 craniosynostosis families within the 100kGP would be underpowered to detect many of these variants. Nevertheless, this work has made a positive contribution towards improving the current GE triaged bioinformatics pipelines as set out in a recent update on the NHS England guidelines for rare disease analysis (McMullan et al., 2021). First, GE has incorporated improved indel calling aiming to retain variants which were previously screened out on quality. Interestingly, this would still filter out the indel variant in *MEGF8* because this intersects a region of low complexity. Second, *de novo* variants are now screened regardless of whether the variant is identified in a Green gene on PanelApp and third, the top three Exomiser hits are analysed for all participants. Incorporating these measures would achieve an 85% diagnostic sensitivity (an increase of 38%) (Hyder et al., 2021).

Chapter 4

Panel-based sequencing for the identification of SNVs and CNVs in a cohort of 617 patients with craniosynostosis

4.1 Introduction

While analysis of the craniosynostosis cohort in the 100kGP proved successful in increasing the diagnostic rate, identifying novel genetic causes was less fruitful. Despite the 100kGP being the largest cohort of unsolved cases who have received WGS, it is still relatively underpowered to identify novel disease genes. For example, variants in *SMAD6* are the 5th most common cause of craniosynostosis (Calpena et al., 2020), yet only two individuals out of 115 families enrolled in the programme were found to have *SMAD6* variants (Hyder et al., 2021). Variants in *SMAD6* were considered a useful measure of the incidence of pathogenic variants associated with incomplete penetrance within the 100kGP cohort because, at the point of recruitment, this gene was not considered a diagnostic-grade gene. Indeed, *SMAD6* variants were first described to be causative for craniosynostosis in 2016 (Timberlake et al., 2016, Di Rocco et al., 2023), four years after patients were initially recruited. Since, further evidence has consolidated *SMAD6* as a Green PanelApp gene allowing any patient recruited after 2018 to be screened for *SMAD6* variants. Similarly, two variants in independent families with craniosynostosis recruited to the 100kGP were identified in *SOX6* (which in 2019 was considered a Red PanelApp gene); these variants contributed to the first report of 19 individuals with variants in *SOX6* associated with a neurodevelopmental syndrome, craniosynostosis, and osteochondromas (Tolchin et al., 2020). Currently, *SOX6* is considered an Amber PanelApp gene, meaning that newly recruited patients will not be automatically screened for variants in this gene. Given that only two individuals harboured variants in *SMAD6* or *SOX6* within this cohort, it is unlikely that two or more likely damaging variants would be found in any one novel gene causally related to craniosynostosis within this cohort. Therefore, I aimed to corroborate findings from candidate disease genes identified from the 100kGP with additional cases identified in the Oxford cohort of unsolved craniosynostosis patients (consisting of 1276 unsolved families [23/11/22]).

From the initial 100kGP screen, 7 candidate disease genes (table A4.1) were selected for further analysis of any additional variants in patients with an unsolved genetic cause to their craniosynostosis: *ANKRD28, HSP90AA1, PTCH1, SMAD2, SPRY1, SPTBN1,* and *TFCP2* (see chapter 3 for variant information; for *SPRY1* see chapter 5). This consisted of three *de novo* variants (*ANKRD28, PTCH1*, and *SMAD2*), one recessive research gene (*SPRY1*), one compound heterozygous variant (*SPTBN1*) and two genes identified as candidates from the burden analysis (*HSP90AA1* and *TFCP2*) (table 4.1; table A4.1).

Including the candidate genes identified within the 100kGP, a panel of 42 genes was curated considering genes under two other broad categories. A subset of 12 genes were selected owing to evidence in mouse or zebrafish models that complete loss or constitutive activation resulted in craniosynostosis (table 4.1; table A4.1) (Lee et al., 2019). Additionally, 23 genes were selected due to reports of more than one individual with craniosynostosis and a putatively pathogenic variant (table 4.1; table A4.1). Identification of variants in multiple independent families included analysis of *de novo* variants in patients with non-syndromic midline craniosynostosis (Timberlake et al., 2017a), variants identified in our in-house database (CranioVar or CranioDB), knowledge of patients attending the Oxford clinic with known genetic variants (identified through panel screening), variants identified through collaborators or craniofacial geneticists, and *de novo* variants identified through the Deciphering Developmental Disorders (DDD) project. This group of 'multiple case' genes can be differentiated into two categories: those with a known role in craniosynostosis, where I aimed to ascertain the frequency of pathogenic variants in a given gene within a cohort of unsolved patients, and those which are considered candidate genes for craniosynostosis (table 4.1).

A minor exception to the above mentioned three categories (100kGP candidate gene, available supporting animal model, or multiple variants reported in independent families)

is *SMAD3*, one of four genes associated with Loeys-Dietz syndrome (LDS, OMIM: #609192). LDS is characterised by tortuous arteries and aortic aneurysms, hypertelorism, and a bifid uvula or cleft palate (Iqbal et al., 2021). Patients with LDS often display other skeletal abnormalities, including craniosynostosis (Loeys et al., 2005, Schepers et al., 2018). Of the four LDS-associated genes (*SMAD2, SMAD3, TGFBR1*, and *TGFBR2*), craniosynostosis is described in patients with heterozygous missense or loss-of-function alleles in *TGFBR1* and *TGFBR2* (Loeys et al., 2005). Additionally, a gain-of-function variant in a mouse model, p.(Met318Arg), provided further evidence that variants in *TGFBR1* are associated with craniosynostosis (Gallo et al., 2014); however, the genotype-phenotype correlation for *SMAD2* and *SMAD3* is less certain. *SMAD2* was included in the panel owing to the identification of a *de novo* variant in the 100kGP (see chapter 3, section 3.3.5). However, to date, there have only been two reports of *SMAD3* variants associated with craniosynostosis. In the first, the authors failed to provide any further variant level information or link to a supporting publication (Schepers et al., 2018); a second report detailed a patient with dolichocephaly and biallelic variants in *SMAD3* (Baskin et al., 2020). Owing to this uncertainty, all LDS-associated genes were included in the panel to assess the frequency of craniosynostosis in association with variants in these genes, and to identify any further supporting cases of variants in *SMAD2* and *SMAD3*, to underpin the role of these genes in its aetiology.

Table 4.1 Category of genes included in the resequencing panel

	Multiple Cases	100kGP Candidate	Animal Model with Craniosynostosis
Candidate Genes	*IL11, NAA25, SLC12A2, SMAD3, SMURF1, SOX11*	*ANKRD28, HSP90AA1, PTCH1, SMAD2, SPRY1, SPTBN1, TFCP2*	*AXIN2, DUSP6, EPHA4, FGF3, FGF4, GDF6, LMX1B, NELL1, PDGFRA, PTHLH, SIX2*
Known Disease Genes	*ALPL, ALX4, BCOR, DHRS3, ERF, GNAS, MSX2, NFIA, PITX2, PRRX1, RNU12, SOX6, TGFBR1, TGFBR2, TFAP2B, ZIC1, ZNF462*		*FGF9*[a]

[a]Recent evidence suggests that *FGF9* should be considered a disease gene for craniosynostosis (Tooze et al., 2023)

4.2 Methods

4.2.1 Probe design

A pool of probes (*n* = 2054; 120 bp in length) were designed and synthesised by IDT to cover all coding regions of the canonical transcript for each gene; each region was covered by at least two probes (2x tiling) which extended into the introns or the 3' and 5' UTRs by at least 60 bp (up to 90 bp in some cases), depending on sequence compatibility for probe design.

4.2.2 Unsolved craniosynostosis cases

A total of 617 samples without a genetic diagnosis were selected for analysis (table 4.2). As the non-syndromic (NS) sagittal and NS metopic samples were the most abundant, a maximum of 100 samples for each of these categories was set. Of the NS sagittal and metopic cases, samples were prioritised if parental samples were available or if they were included in a previous targeted sequencing analysis of *PRRX1* but displayed incomplete coverage (due to additional interest as a research gene; see chapter 6). For the remaining diagnoses, all available unsolved samples were included regardless of availability of parental samples or inclusion in prior genetic screens (table 4.2).

4.2.3 Targeted capture next-generation sequencing

Analysis of 617 samples for variants in 42 genes was undertaken using IDT's hybridisation and capture protocol (version 4, May 2019). A subset of 297 sample libraries were previously prepared for capture-based resequencing using a 'Twist Bioscience' kit by Dr Eduardo Calpena (Calpena Samples, table 4.2). A further 320 sample libraries were prepared using the 'xGenTM DNA Library Prep EZ Kit' and protocol (Integrated DNA Technologies [IDT], protocol version 1 [December 2021]) (Tooze samples, table 4.2). DNA was fragmented to ~200 bp, before adding adapters and indexing primers, and subsequently analysed using broad-range Qubit and D1000 TapeStation reagents (average fragment size of 330 bp). The prepared libraries were pooled to a total of 6 μg of DNA and between 32–40 samples per hybridisation capture reaction. The hybridisation reactions were carried out at 65°C for 16 hours. After hybridisation, the pooled libraries were washed and post-capture PCR was performed following the manufacturer's protocol (IDT xGen hybridisation capture of DNA libraries for NGS target enrichment, 7 PCR cycles were used) for a panel containing 2054 probes. The amplified capture reactions were washed with beads before quantification and validation using high-sensitivity (HS) Qubit and HS D1000 TapeStation reagents. Next-generation sequencing analysis was performed using a Miseq v3 kit (150-cycle) and data were analysed using amplimap software under default parameters (Koelling et al., 2019) (including mapping, coverage analysis, and variant calling with GATK). Variants with a low number of reads or quality were retained under the default calling parameters and filtered out in analysis.

4.2.4 Identification of SNVs

Heterozygous SNVs were filtered on the basis of rarity (gnomAD v3.1.2 allele frequency below 0.000045). This figure was selected as the maximum credible population allele frequency for any causative variant in craniosynostosis, as described in previous reports

(Calpena et al., 2020, Whiffin et al., 2017). This figure considers that the largest proportion of cases attributable to a single nucleotide substitution in craniosynostosis is a missense variant in *FGFR2* (c.749C>G; p.(Pro250Arg)), accounting for 24/666 cases (Wilkie et al., 2017), and assumes a penetrance of 0.2. Given these assumptions, the maximum expected population allele frequency is 1/2000 (frequency of craniosynostosis) x 24/1332 (considering number of alleles) x 1/0.2 (penetrance) = 0.000045. Variants were subsequently filtered by CADD score (≥ 20, or not reported) and likely consequence (missense or worse) and screened manually. Recessive variant analysis was conducted using an allele frequency threshold of 0.01. All identified likely damaging variants were confirmed by dideoxy-sequencing (chapter 2, section 2.2; primers detailed in table A4.2).

4.2.5 Coverage analysis

In a subset of patients (*n* = 479) included in the targeted sequencing analysis, coverage was assessed. Normalised coverage was calculated by dividing the coverage for each gene by the mean coverage of all genes (*n* = 41; *BCOR* was excluded as the only gene on the X-chromosome) for a given sample. This value was normalised to the average coverage for all samples in a given gene. The value corresponding to the gene with the greatest deviation from 1 was removed from each sample before the mean and standard deviation of coverages per sample for all genes was ascertained. The standard deviation was used to assign a threshold for sample cut off (figure A4.1a–b). After ranking each sample by standard deviation (SD) (figure A4.1a), the error behaviour was calculated as $(SD_b - SD_a)/(b - a)$ where a and b correspond to ranked samples (figure A4.1d). Data quality noticeably deteriorated at *n* = 479 which corresponded to a standard deviation threshold of 0.045. Normalised exon coverage and/or coverage for non-overlapping probe regions was considered for known haploinsufficiency genes (*ALX4, ERF, PRRX1, SOX6, ZIC1*) or for any gene which showed an abnormal signal in at least one sample

in the whole gene coverage analysis. For exon and probe analyses, each sample was analysed for deletions or duplications, considering two or more consecutive exons/probes displaying a normalised coverage below 0.6 (heterozygous deletion) or above 1.4 (heterozygous duplication). Identified likely damaging CNVs were confirmed independently by array. Breakpoints of any deletions were ascertained by a combination of information from an array, qPCR, and long amplification PCR (see chapter 6, section 6.2.2).

4.2.6 RT-PCR

RNA from patient lymphoblastoid cell lines was used to generate cDNA for RT-PCR analysis (methods in chapter 2, section 2.8). Primers were designed to span exons 1–4 of *SMAD3* for amplification of cDNA (primers detailed in table A4.3).

4.3 Results

4.3.1 Sequencing metrics

The average coverage per bp for all samples screened was 134, with the minimum coverage value in a sample identified as 10, and the maximum value being 788. The average coverage per bp was analysed for each exon of the 42 genes screened across all samples; 0.8% of samples had a coverage value below 10 (at which reliable calling becomes doubtful), and 48% and 79% of samples displayed an average coverage value above 100 and 50, respectively.

4.3.2 Six likely pathogenic SNVs: *ALX4*, *MSX2*, *PRRX1*, *SMAD3*, and two *SOX6*

Filtering by absence from gnomAD and a CADD score above or equal to 20 left 76 variants remaining for individual scrutiny (table A4.4); compound heterozygous variants

in *SPTBN1* (c.333C>G; p.(His111Gln) and c.5684C>G; p.(Ser1895Cys)) and a heterozygous stop-gain in *TFCP2* (c.328G>T; p.(Glu110*)) were originally identified in patients recruited to the 100kGP (see chapter 3; table 3.2) and were included as controls. Data were also screened under a recessive disease model (section 4.2.4), but no further candidates were identified. Identified variants were analysed in IGV to retain those with true calls (56/76, 74%); all false positives showed below 20% of the reads with the variant allele, although mosaicism was considered in all cases. For all true calls, the conservation amongst vertebrates (as documented on the UCSC web browser), effect on a functional domain, consequence (missense or worse), MOI (dominant or recessive; for any gene with a documented biallelic mode of inheritance, a monoallelic cause was also considered), constraint metrics for missense variation (Samocha et al., 2017), and current literature were used to ascertain likely pathogenicity. Dideoxy-sequencing confirmed the presence of any likely damaging variants and segregation patterns were assessed in families in which additional family member samples were available.

After removing those that were not confirmed by dideoxy-sequencing, inherited from unaffected family members (unless non-penetrance was known for a given gene), or did not conform to the likely molecular pathophysiological mechanism associated with variants in that gene, there remained five variants considered diagnostic owing to known reports of similar variants in patients with craniosynostosis (table 4.3). These variants consisted of heterozygous missense variants in *ALX4* (NM_021926.4:c.817G>A; p.(Glu273Lys)) and *MSX2* (NM_002449.5:c.442C>A; p.(Pro148Thr)), a loss of the start-codon in *PRRX1* (NM_022716.4:c.2T>G; p.(?)), and two truncating variants in *SOX6* (NM_001367873.1:c.436_445+13del; p.(?) and NM_001367873.1:c.1624G>T; p.(Glu542*)) (figure 4.1). Importantly, parental samples were available for four of these families and confirmed that the *MSX2* missense and the *SOX6* stop-gain variants were *de novo*, the *ALX4* variant was inherited from an affected father with sagittal synostosis, and the *PRRX1* variant was inherited from an unaffected parent (figure 4.1). Despite

non-penetrance being evident for *PRRX1*, work conducted in parallel suggests that heterozygous loss-of-function variants in this gene are a relatively frequent cause of craniosynostosis (see chapter 6), and it was therefore retained in the analysis. For the remaining *SOX6* frameshift variant (c.436_445+13del; p.(?)), the patient was enrolled into the Craniofacial study as a singleton and therefore it was not possible to assess variant segregation.

Additionally, a donor splice-site variant in *SMAD3* (NM_005902.4:c.206+1G>A; p.(?)) was identified in a patient with non-syndromic uni-coronal synostosis (table 4.3). While variants in *SMAD3* are described in patients with LDS (see section 4.1) the association with craniosynostosis is less clear, and therefore further functional validation and *in silico* prediction was considered necessary. Predictive software (SpliceAI) suggested a 98% chance of donor loss and 33% chance of donor gain, predicting loss of one functional copy of the gene. Although exon 1 is alternatively spliced in *SMAD3*, the variant affects the predominant transcript and would be predicted to disrupt the MH1 domain which is essential for DNA binding (BabuRajendran et al., 2010). Additionally, previous studies have also reported loss-of-function variants affecting this exon (Haller et al., 2015, Fitzgerald et al., 2014) and a patient was described with LDS in ClinVar with a heterozygous variant affecting the same nucleotide (NM_005902.4:c.206+1G>C), which was considered likely pathogenic. To investigate this variant further, I obtained RNA from patient-derived lymphoblastoid cells and two control samples for RT-PCR. A region spanning exons 1–2 was amplified from each cDNA sample and showed no difference between the proband and the control samples (figure 4.2a). Additionally, dideoxy-sequencing of the product spanning exons 1–2 showed no sequencing abnormalities in the cDNA. To ascertain if this data was representative of both alleles or just the wildtype, I screened the sequencing data for the presence of any informative SNPs and identified one heterozygous SNP in exon 2 (rs1065080; A>G); this finding was confirmed by dideoxy-sequencing (figure 4.2b). Analysis of cDNA at the same position identified a

homozygous variant of the alternate allele (figure 4.2b), thus confirming the presence of only one allele in the cDNA analysis and loss of one functional copy of *SMAD3*. However, given that the variant in exon 2 is ~100 kb away from the missense substitution, it was not possible to determine the phase of these variants. Nevertheless, combining this functional data with the previous reports of patients with LDS and loss-of-function variants in exon 1 was sufficient to consider this variant pathogenic.

Table 4.2 Samples included in the resequencing analysis by diagnosis

Total Samples Sequenced by CRS diagnosis[a]	Calpena Samples (TWIST libraries)	Tooze Samples (IDT libraries)	Total Samples by diagnosis	Identified diagnostic SNV/ CNV
NSUC	161	0	161	2 (1.2%)
NSBC	29	1	30	
NSS	0	100	100	
NSUL	3	7	10	
NSBL	2	0	2	
NSM	0	100	100	
NSMultiple	34	17	51	1 (2.0%)
SUC	16	3	19	
SBC	5	2	7	1 (14.3%)
SS	3	45	48	1 (2.1%)
SUL	2	0	2	
SBL	1	0	1	
SM	0	30	30	
SMultiple	25	15	40	3 (7.5%)
Other	16	0	16	
Total	**297**	**320**	**617**	**8 (1.3%)**

[a]NS = non-syndromic, S = syndromic, UC = uni-coronal, BC = bi-coronal, S = sagittal, UL = uni-lambdoid, BL = bi-lambdoid, M = metopic

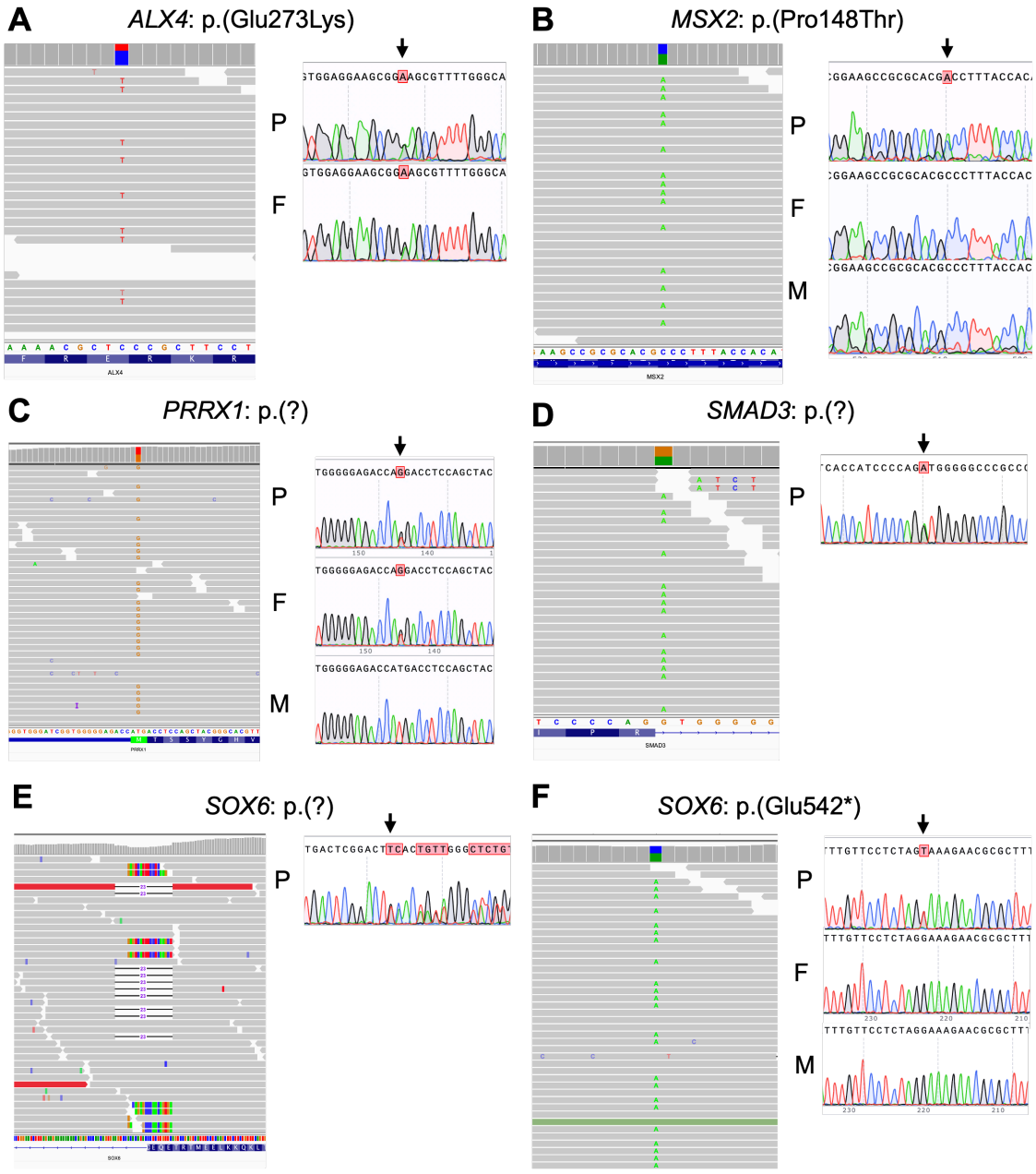


Figure 4.1 Sequencing confirmation from IGV and dideoxy sequencing. (A–F) Left, IGV screenshot of the identified variant after targeted capture next-generation sequencing. Right, chromatogram from dideoxy-sequencing of gDNA from the affected proband and parents, if available: P = proband, F = father, M = mother. In all dideoxy-sequencing traces, the black arrow denotes the position of the variant allele and the red box highlights nucleotides that differ from the consensus sequence. Confirmation of the (A) *ALX4* familial variant, (B) *MSX2 de novo* variant, (C) *PRRX1* inherited loss of the start codon, (D) *SMAD3* donor splice site substitution, (E) *SOX6* frameshift variant, and (F) *SOX6 de novo* stop-gain.

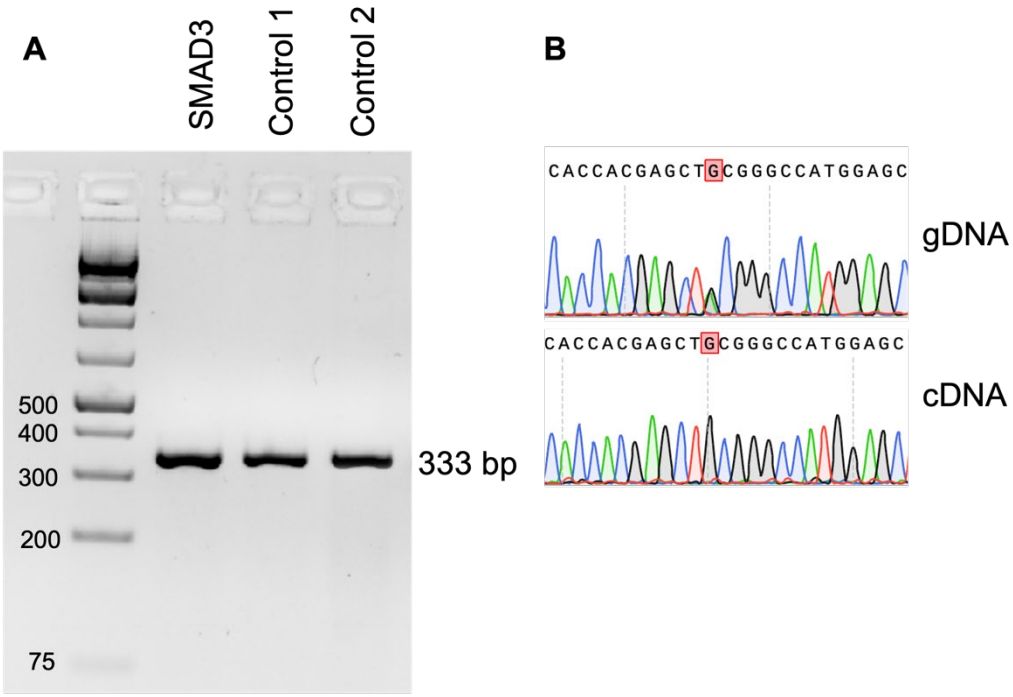


Figure 4.2 RT-PCR analysis of an aberrant splicing defect in *SMAD3*. (A) RNA was extracted from patient-derived lymphoblastoid cells from the proband and two controls before reverse transcription. For each sample, cDNA was amplified using primers in exon 1 and exon 2 (forward: 5'-GCTGGAAGAAGGGCGAGCAGAA-3'; reverse: 5'-GGTAGTGGTAGGGATTCACGCAGAC-3'), generating a product of 333 bp. (B) Top, a dideoxy-sequencing trace from gDNA amplified using primers in intron 1 and exon 2 (forward: 5'-GGACCTCTGGATCTGGGAGAAATGAG-3'; reverse: 5'-GGTAGTGGTAGGGATTCACGCAGAC-3') identifying a heterozygous SNP (rs1065080; A>G). Bottom, a dideoxy-sequencing trace of cDNA amplified between exon 1 and 2 (using primers as in panel A) identifying a homozygous SNP at the same position as the trace above. In both traces, the variant position is highlighted by the red box.

Table 4.3 Diagnostic SNVs from 617 samples in the resequencing analysis

Gene	cDNA change	Protein change	Segregation	Variant Zygosity[a]	Del Score	CADD	Comments	Craniosynostosis/ clinical information
ALX4	NM_021926.4:c.817G >A	p.(Glu273Lys)	Paternally inherited (familial)	Het	5	29	The variant affects the 60th position of the homeodomain. Substitution at the adjacent position, p.(Arg272Pro), is reported as pathogenic for parietal foramina in ClinVar.	Proband and father have sagittal synostosis.
MSX2	NM_002449.5:c.442C >A	p.(Pro148Thr)	*De novo*	Het	5	24.7	The variant is within the homeobox domain. Substitutions at Pro148 are pathogenic for craniosynostosis (Jabs et al., 1993, Florisson et al., 2013).	Non-syndromic multiple (bi-coronal and sagittal synostosis). Transverse palmar crease, glue ear. Maternal methotrexate.
PRRX1	NM_022716.4:c.2T>G	p.(?)	Paternally inherited	Het	4	24.3	Loss of the start codon in a functionally relevant gene (see chapter 6).	Syndromic multiple (sagittal and right lambdoid).
SMAD3	NM_005902.4:c.206+1 G>A	p.(?)	Singleton	Het	6	36	SpliceAI predicts a 98% chance of donor loss and 33% chance of donor gain. One patient is reported in ClinVar with LDS and a variant affecting the same nucleotide.	Non-syndromic uni-coronal
SOX6	NM_001367873.1: c.436_445+13del	p.(?)	Singleton	Het	6	-	The variant intersects the exon 3–intron 3 boundary and is real on IGV. The variant was confirmed by WES data in CranioDB and dideoxy sequencing. Variants in *SOX6* are a known cause of	Syndromic multiple. Low anterior hairline, synophrys, small mouth. Visual loss secondary to raised intercranial pressure. Hydrocephalus.

							craniosynostosis (Tolchin et al., 2020).	
SOX6	NM_001367873.1:c.1624G>T	p.(Glu542*)	*De novo*	Het	6	44	Confirmed by dideoxy sequencing, loss-of-function variants in *SOX6* are associated with craniosynostosis (Tolchin et al., 2020).	Syndromic multiple. Hyperphosphatemic rickets, inflammatory bowel disease, eczema, delay in walking.

[a]Het = Heterozygous; Del Score = Deleterious Score (0 = likely benign, 6 = likely pathogenic)

4.3.3 Candidate variants

Having discounted those variants which were (a) identified outside of relevant functional domains, or (b) for which there are available functional analyses that dispute variant pathogenicity, or (c) for which there was evidence that the specific molecular mechanism associated with variants of that type (for example, missense) are not pathogenic, there were 12 candidate variants remaining (see table A4.4 for further information on reasons for discounting variants). To identify additional cases to support gene-disease association of candidate genes, rare variants were screened assuming an allele frequency < 0.000045 (Whiffin et al., 2017, Calpena et al., 2020)) (table A4.5). This added three variants (two in *ALPL* and one in *ALX4*), bringing the total number of candidate variants to 15 (table 4.4). Although some of these candidates may be contributing to the patient phenotype, most of these variants would be considered a variant of unknown significance (VUS) owing to evidence of segregation of the variant allele in an unaffected parent or insufficient evidence on the likely consequence of the variant (table 4.4). One method of increasing confidence in gene-disease association is the identification of more than one patient with congruent phenotypes and a potentially disease-causing allele in the same gene. Therefore, any gene for which more than one individual was identified with a diagnostic (table 4.3) or candidate variant (table 4.4) from the resequencing analysis, or prior analysis of the 100kGP dataset is discussed further. Additionally, through-out my DPhil, I have focussed attention on variants in homeobox transcription factors as a result of the functional characterisation of variants within the *PRRX1* homeodomain (see chapter 6), and therefore two variants in *LMX1B* (identified in one patient) are also discussed.

4.3.3.1 *ALPL*

Variants in *ALPL*, which encodes the tissue nonspecific alkaline phosphatase enzyme (TNSALP), cause a mineralisation defect known as hypophosphatasia. This is a

heterogenous disorder, ranging from severe bone mineralisation defects, seizures, and hypercalcemia to milder phenotypes including teeth anomalies and susceptibility to bone fractures (Fenn et al., 2021). The more severe phenotypes demonstrate recessive inheritance, while the moderate to mild presentations display both dominant and recessively inherited variants (Mornet et al., 2020). Variants are identified across the length of the protein, with no specific phenotypic correlation based on the position of the variant. Within this screen, analysis of variants in *ALPL* identified six heterozygous variants in individuals with craniosynostosis. For the three *ALPL* variants which were absent from gnomAD, functional assays indicate no effect on the activity of TNSALP (Mornet et al., 2020) and they were subsequently discounted. Of the three rare heterozygous missense variants, two were identified in the ALPL mutation database (p.(Gln76Arg) and p.(Arg223Gln)) and were shown to have dominant negative effects compared to the wildtype protein (Mornet et al., 2020). In both cases, the variant was inherited from an unaffected family member and would therefore be considered a VUS. Despite this, it may be expected that non-penetrance would be frequent for *ALPL* variants given the variability in phenotype and genotype documented in patients with *ALPL* variants (Mornet et al., 2020); however, in a cohort of 72 patients with hypophosphatasia the parents were not systematically sequenced, and therefore a current estimate of penetrance is not clear (Jandl et al., 2021).

4.3.3.2 *ALX4*

Aristaless-like homeobox 4 (*ALX4*) is member of the homeobox class of transcription factors which regulate several early developmental pathways. *ALX4* has been shown to couple with other transcription factors, including *MSX2*, to control lineage commitment of the mesenchymal progenitor cells (Stains and Civitelli, 2003, Antonopoulou et al., 2004). Haploinsufficiency of *ALX4* causes parietal foramina, characterised by enlarged openings of the parietal bones (Mavrogiannis et al., 2006), while gain-of-function variants

are believed to be causative of non-syndromic craniosynostosis with reduced penetrance (Yagnik et al., 2012, Fonteles et al., 2021). Six heterozygous missense variants in *ALX4* were identified within the resequencing cohort and three of these affect highly conserved residues within the homeodomain. The p.(Glu273Lys) variant locates at the 60th residue of the homeodomain; variants at the prior residue are reported to be pathogenic for parietal foramina. The individual displays non-syndromic sagittal synostosis, and the variant was also found to segregate from the patient's affected father (figure 4.1). A rare variant affecting the 44th residue of the homeodomain, p.(Arg257His), was identified in two families within this resequencing cohort. In both cases, the variant was inherited from an unaffected family member, although for one family there were two affected siblings with non-syndromic bi-coronal synostosis in which the variant segregated. Variants affecting the same residue have been described in a patient with non-syndromic craniosynostosis, p.(Arg257Cys), also displaying non-penetrance (Fonteles et al., 2021). Consistent with the documented non-penetrance, there are 5/~250,000 reported heterozygous p.(Arg257His) alleles and 8/~280,000 alleles for p.(Arg257Cys) in gnomAD v2.1.1.

4.3.3.3 *LMX1B*

LIM homeobox transcription factor 1 (*LMX1B*) plays an essential role in dorso-ventral patterning of the vertebrate limb (Isojima et al., 2014). Loss-of-function heterozygous variants within the LIM domain or homeodomain of *LMX1B* are associated with nail-patella syndrome, characterised by dysplastic nails, absent or hypoplastic patellae, elbow dysplasia, and renal anomalies (Harita et al., 2017). I identified one individual within the resequencing cohort with two potentially damaging heterozygous variants in *LMX1B*. One variant within the second LIM domain, p.(Val128Met), was confirmed *de novo* and displayed a high CADD score (29.8); the second variant affects the 41st residue, p.(Ser259Thr), of the homeodomain and segregated from an unaffected parent. The residue within the homeodomain locates to the recognition helix (helix 3) which is

responsible for binding the major groove of DNA, although it is not one of the key residues mediating this interaction (Noyes et al., 2008). Indeed, DNA-binding assays investigating the effect of a substitution at the same residue in HOXD13 showed no difference in protein activity compared to the wildtype protein (Johnson et al., 2003). The patient displays no phenotypes suggestive of nail-patella syndrome, and instead shows only non-syndromic sagittal synostosis. While these variants could be contributing towards the phenotype, the non-penetrance of the variant in the more functionally relevant domain, alongside the lack of functional evidence makes analysis complicated. The second *de novo* variant may be sufficient to significantly reduce the levels of *LMX1B* resulting in loss-of-function and craniosynostosis, although further studies and cases would be required to confirm this hypothesis.

4.3.3.4 *MSX2*

Similar to *ALX4*, heterozygous loss-of-function variants in *MSX2* are associated with parietal foramina. However, a specific gain-of-function variant affecting proline-148 within the homeodomain (7th residue of the homeodomain) causes craniosynostosis (see chapter 1, section 1.4.1, for further information) (Florisson et al., 2013). In this work, I identified one patient with bi-coronal synostosis and a *de novo* variant affecting proline-148, encoding p.(Pro148Thr), explaining the phenotype in this individual. A second patient with uni-coronal synostosis was found to harbour a rare variant affecting the 18th position of the homeodomain, inherited from their unaffected father, resulting in a change of arginine to histidine at this position (p.(Arg159His)). The variant resides within helix I of the homeodomain, believed to be essential for protein stability (Wilkie et al., 2000), and is conserved in both *MSX1* and *MSX2*. Previously, a patient with parietal foramina was described with a deletion of Arg159–Lys160; binding assays revealed that deletion of these two residues bound a target sequence with an affinity of 3% that of the wildtype protein, supporting loss-of-function of the variant allele (consistent with parietal foramina)

(Wilkie et al., 2000). The predicted molecular mechanism was destabilisation of the protein, owing to loss of the salt bridge formed between lysine at position 19 and glutamate at position 30. Therefore, the phenotype is likely attributable to deletion of amino acid 19 and, to a lesser extent, deletion of the 18th residue (Arg159). Moreover, substitution of the 18th residue in LMX1B (p.(Ala213Pro)) showed no difference to the wildtype protein in a transactivation assay (Dreyer et al., 2000), while variants at this position in HESX1 (p.(Asn125Ser)) are a known benign polymorphism (Brickman et al., 2001). Of further note, this position is variable in other members of the homeodomain family, indicating low conservation within the family and tolerance to a variety of amino acids at this position. As such, this variant was not considered further in my analyses.

4.3.3.5 *NFIA*

NFIA encodes the nuclear factor 1A transcription factor which plays a key role in activating various embryonic differentiation pathways. Heterozygous truncating variants or intragenic deletion of *NFIA* are associated with brain abnormalities, intellectual disability, and occasionally, urinary tract defects (Senaratne and Quintero-Rivera, 1993, Uehara et al., 2021). More recently, five patients were described with heterozygous loss-of-function variants in *NFIA* and craniosynostosis (Yoon et al., 2020, Tønne et al., 2021, Tønne et al., 2022, Chen et al., 2022a, Tooze et al., 2023); two of these variants encoded frameshifts and two encoded deletions (13 Mb and 8 kb, both disrupting *NFIA*). Studies of *Nfia* homozygous knockout in mice demonstrate that over 95% of neonates die within 2-weeks of birth, displaying an absence of the corpus callosum and signs of early hydrocephalus (das Neves et al., 1999). NFIA heterozygous littermates (*Nfia*$^{+/-}$) show an absence of the corpus callosum alongside renal abnormalities (Lu et al., 2007), supporting a key role for this gene during early vertebrate development. A defect in cell differentiation of astroglia cells of the midline has been proposed as the likely pathophysiological mechanism behind NFI-related disorders (Ogura et al., 2022, Uehara et al., 2021). Within this resequencing cohort, three patients were found with novel

missense variants in *NFIA*, with two of these affecting the CTF-NFI transcription modulation domain. While missense variants are reported within the N-terminal MH1 DNA-binding domain to be likely pathogenic for chromosome 1p32-p31 syndrome, currently there is little evidence to suggest that missense variants within the CTF-NFI domain would be damaging and therefore without the presence of additional functional work, these variants are considered VUS'.

4.3.3.6 *PDGFRA*

PDGFRA encodes a receptor tyrosine kinase which responds to platelet-derived growth factor (PDGF); variants within this gene are typically associated with gastrointestinal stromal tumours (GIST). Mutational hotspots are identified in three domains: the activation loop in exon 18, the juxtamembrane domain in exon 12, and the kinase domain in exon 14 (Guérit et al., 2021). A specific substitution within the activation loop, p.(Asp842Val), in mice causes coronal synostosis which can be detected at P15. I identified two *bona fide* variants in *PDGFRA* in two families, predicting a stop gain, p.(Gln736*), and a substitution in exon 18, p.(Ser851Leu). The truncating variant was identified in a patient with Saethre-Chotzen syndrome, usually characterised by loss-of-function variants in *TWIST1* (Howard et al., 1997, Zhou et al., 2018, Twigg and Wilkie, 2015). TWIST1 has been shown to induce *PDGFRA* expression, suggesting that variants in PDGFRA could also present similar phenotypes (Eckert et al., 2011, Yeeravalli et al., 2021, Sepporta et al., 2022). Nevertheless, there are 22 high-confidence loss-of-function calls reported in gnomAD v2.1.1 from ~150,000 exomes or genomes which, in contradiction to its pLI score of 1, indicates some tolerance to loss-of-function variation (Karczewski et al., 2019). The missense substitution lies close to the described mutation associated with coronal synostosis in the mouse model and within a region constrained for missense variation (observed/expected = 0.18) (Karczewski et al., 2020). However, functional studies show that substitution of other nearby residues within exon 18 display very weak gain-of-function activity of the variant compared to wildtype (Ip et al., 2018),

therefore, the effect of p.(Ser851Leu) is hard to predict. Interestingly, consistent with the phenotype of the individual with the stop-gain variant, this patient also displays bi-coronal synostosis alongside features suggestive of Saethre-Chotzen syndrome (ptosis, broad thumbs, and toes, and bi-lateral 2–3 toe syndactyly). The confounding factor for both cases is the segregation of the variant allele from an unaffected parent, and therefore more cases are required to consolidate the role of *PDGFRA* in craniosynostosis.

4.3.3.7 *SMAD3*

In addition to the splicing variant in *SMAD3* (see section 4.3.2), a missense substitution of the linker region of the protein (c.608G>A; p.(Gly203Asp)) was identified in patient with non-syndromic uni-coronal synostosis. This phenotype is consistent with the patient reported with the donor splice site substitution. However, while missense substitutions are described in both the MH1 and MH2 domains of SMAD3, no heterozygous missense variants are documented within the linker region within patients with LDS (Schepers et al., 2018). Interestingly, a variant at the adjacent position (p.(Ala202Thr)) was identified in a patient with colorectal cancer and was considered benign (Fleming et al., 2013). However, unlike residue 202, glycine-203 is conserved in the closely related protein SMAD2, but this position is variable in other SMAD proteins including SMAD1, SMAD4, SMAD5 and SMAD6 which are more distantly related. Parental samples were unavailable for both individuals with *SMAD3* variants, and it is therefore unknown whether they were *de novo* (which would increase confidence in variant pathogenicity) or inherited. As such, the missense variant would require further functional validation before being considered likely pathogenic.

4.3.3.8 *TFCP2*

Previous analysis of the 100kGP dataset and CranioVar identified three individuals with variants in *TFCP2*, encoding two loss-of-function variants and a heterozygous missense

(see chapter 3, section 3.3.8). Here, I identified one further individual with a stop-gain in *TFCP2* and non-syndromic sagittal synostosis, alongside a documented heart murmur. The previous three individuals displayed bi-coronal synostosis *(n* = 2) or uni-coronal with additional fusion of the metopic suture; developmental delay was documented in two patients. These variants are considered as candidates for two reasons. First, in all four reported cases, the variant was inherited from an unaffected parent. Second, the gene displays a pLI score of 0, suggesting tolerance to loss-of-function variation. Consistent with this, there are 51 loss-of-function variants reported in gnomAD v2.1.1 (encoding stop-gains, frameshifts, or splicing variants), out of ~240,000 alleles. Within this resequencing study, I identified one individual with a stop-gain in *TFCP2* out of 990 alleles (including only samples with complete coverage of *TFCP2* and excluding both control patients already recruited to the 100kGP). This figure is in line with the number of loss-of-function variants that would be expected to be identified in a cohort of 497 individuals, given the number of loss-of-function variants reported in the gnomAD data (X^2; *p*-value = 0.087). As such, this variant was considered an incidental finding, and further cases would be required to support haploinsufficiency of *TFCP2* in the aetiology of craniosynostosis.

Table 4.4 Candidate SNVs identified in the resequencing analysis which were confirmed by dideoxy-sequencing

Gene	cDNA change	Protein change	Segregation	AF	Variant zygosity[a]	Del Score	CADD	Comments
ALPL	NM_000478.6:c.227A>G	p.(Gln76Arg)	Maternally inherited (unaffected)	7.95E-06	Het	6	27.6	Identified in one individual with non-syndromic sagittal synostosis. Compound heterozygous variants, including p.(Gln76Arg) are documented as pathogenic for hypophosphatasia (Mornet et al., 2020). Mutation at this position results in a marked reduction in TNSALP activity compared to the wildtype (0.03) (Angel et al., 2020).
ALPL	NM_000478.6:c.668G>A	p.(Arg233Gln)	Paternally inherited (unaffected)	1.99E-05	Het	5	32	A heterozygous variant at this position was shown to display a dominant negative effect compared to the wildtype protein (Angel et al., 2020). There is one reported pathogenic heterozygous variant in the Angel cohort and seven other compound heterozygotes at this position.
ALX4	2 x NM_021926.4:c.770G>A	2 x p.(Arg257His)	Maternally inherited (unaffected); segregates in affected sibling and also in carrier mother (unaffected).	1.99E-05	Het	5	27.4	This variant is identified in two individuals (non-syndromic sagittal and non-syndromic uni-coronal); in one family this variant is found to segregate in an affected sibling. There are five reported alleles in gnomAD v2.1.1, alongside eight reported alleles corresponding to p.(Arg257Cys). A child with coronal synostosis has been described with an p.(Arg257Cys) missense variant which was inherited from their unaffected mother (Fonteles et al., 2021). This variant affects the 44th position of the homeodomain.
EPHA4	NM_004438.5:c.2464G>C	p.(Glu822Gln)	Singleton	0	Het	6	26.2	Highly conserved residue within the protein tyrosine domain. There is little evidence from crystal structures to suggest that this is a critical residue (Bowden et al., 2009, Xu et al., 2013). This individual has non-syndromic metopic synostosis.

LMX1B	NM_001174146.2:[c.382G>A; c.776G>C]	p.([Val128Met]); p.([Ser259Thr])	*De novo*; paternally inherited (unaffected)	0	Het	6; 3	29.8; 23.1	Both variants have 1 reported heterozygote in gnomAD but no homozygous variants. Val128 is outside of the homeodomain and is *de novo* (confirmed by dideoxy sequencing) while the p.(Ser259Thr) variant is within the homeodomain and paternally inherited.
MSX2	NM_002449.5:c.476G>A	p.(Arg159His)	Paternally inherited (unaffected)	1.19E-05	Het	6	31	Arg159 encodes the 18th position of the homeodomain; lysine is the consensus amino acid at this position for this class of homeobox transcription factor. An indel at this position is associated with parietal foramina.
NELL1	NM_006157.5:c.2399G>T	p.(Arg332*)	Paternally inherited (unaffected)	0	Het	6	36	The variant encodes a stop-gain towards the N-terminus of the protein which would result in nonsense-mediated decay; the gene has a pLI score of 0 suggesting that loss-of-function variants are well tolerated.
NFIA	NM_001134673.4:c.1123A>G	p.(Ile375Val)	Paternally inherited (unaffected)	0	Het	4	23.8	Variant is within the CTF-NFI domain (transactivation domain); it is 20 amino acids away from a described p.(Thr395Met) variant in a patient with intellectual disability (Ogura et al., 2022). The amino acid at this position is well conserved in vertebrates.
PDGFRA	NM_006206.6:c.2206C>T	p.(Gln736*)	Paternally inherited (unaffected)	0	Het	6	40	The gene has a pLI score of 1 (suggesting intolerance to loss-of-function variants). Both missense variants and loss-of-function are associated with gastrointestinal tumours (Guérit et al., 2021).
PDGFRA	NM_006206.6:c.2552C>T	p.(Ser851Leu)	Maternally inherited (unaffected)	0	Het	5	27.4	Listed as a VUS in ClinVar for gastrointestinal stroma tumour (GIST). The variant is nearby to a hotspot of missense mutations for GIST within the kinase domain (Guérit et al., 2021). There is one reported allele in gnomAD v2.1.1.
SLC12A2	NM_001046.3:c.2956A>G	p.(Thr986Ala)	Singleton	0	Het	4	23.5	There are two heterozygous variants in gnomAD v2.2.1. Variants at nearby residues (p.(Pro988Ser), p.(Glu980Lys),

								p.(Glu979Lys)) (McNeill et al., 2020, Mutai et al., 2020, Adadey et al., 2022) are considered likely pathogenic for hearing abnormalities.
SMAD3	NM_005902.4:c.608G>A	p.(Gly203Asp)	Singleton	0	Het	4	25.3	The variant intersects the linker region of the protein. Missense variants are described for both MH1 and MH2 domains but only loss-of-function is described as pathogenic in the linker region (Schepers et al., 2018). The variant resides near to a splice site and SpliceAI suggests a 0.24 probability of acceptor loss and 0.23 probability of donor loss. p.(Gly203Asp) is described as a VUS for familial thoracic aortic aneurysm and aortic dissection in ClinVar.
SMURF1	NM_181349.3:c.2036G>A	p.(Arg679Gln)	Maternally inherited (unaffected)	0	Het	6	26	Variant coincides with a functional HECT domain (ubiquitin-transferase) and is within a region constrained for missense variation (missense constraint = 0.47 [67 observed, 142 expected]). The variant is not described in COSMIC.
SPRY1	NM_001258038.2:c.770_771del	p.(Tyr257Serfs*20)	Does not segregate in the affected sibling	0	Het	6	-	As Spry-1 only contains one coding exon, this variant would be predicted to escape nonsense mediated decay and disrupt the cysteine-rich domain. Available WES data on this individual confirmed that this variant is present in a very low number of reads but dideoxy sequencing was unable to identify the variant (suggestive of a mosaic variant). This variant belongs to a family with three affected siblings born to a consanguineous union; therefore, there is a likely recessive cause to their craniosynostosis. Incidentally, this patient also harboured a copy number gain in *SPRY1* (see table 4.5).

TFCP2	NM_005653.5:c.1259C>G	p.(Ser420*)	Paternally inherited (unaffected)	0	Het	6	38	The variant encodes a stop-gain in a gene with a low probability of loss-of-function intolerance score (pLI = 0; suggesting loss-of-function is well tolerated). The stop-gain occurs after any functional domain but would be predicted to cause haploinsufficiency of the variant allele.

[a]Het = Heterozygous
AF = allele frequency (gnomAD v3.1.2)
Del Score = Deleterious Score (0 = likely benign, 6 = likely pathogenic)

4.3.4 Four CNVs were identified: Two *PRRX1* deletions, and a *PDGFRA* and *SPRY1* copy number gain

Analysis of whole gene coverages (figure 4.3) identified a heterozygous whole-gene deletion in *PRRX1*, a heterozygous partial duplication in *PDGFRA* and a copy number gain in *SPRY1* (figure 4.3, table 4.5). Breaking down the coverage of *PRRX1* to an exon (figure 4.4a) and probe analysis (figure 4.4b–c), revealed one further patient with a deletion of exons 1–2. The whole gene and partial gene deletions in *PRRX1* were independently verified by an array. The individual with the whole gene deletion was found to harbour a ~10.5 Mb deletion on chromosome 1, which included *PRRX1*; the individual with the partial deletion of exons 1–2 was found on breakpoint analysis to have inherited the deletion from an unaffected parent (table 4.5 and see chapter 6). Genes with known haploinsufficiency (*ALX4, ERF, SOX6*, and *ZIC1*) were screened in depth looking at normalised coverage of each exon and non-overlapping probe region to identify any partial gene deletions, however this revealed no further convincing CNVs (figures A4.2–3).

The two identified copy number gains in *PDGFRA* and *SPRY1* (table 4.5) were further verified through analysis of coverage of non-overlapping probe regions (figure 4.5 and figure 4.6, respectively). Whole gene and partial gene duplications are reported in *PDGFRA* in the database for genomic variation (DGV) and, as such, this was considered likely incidental. Whole and partial gene gains in *SPRY1* are also documented in DGV, however this variant was primarily discounted from further analyses owing to lack of segregation in affected siblings. The patient belongs to a family with three affected siblings from a consanguineous union (Blair et al., 2000), and therefore a recessive variant was anticipated. The analysis of WES data in CranioVar from one of the affected siblings did not identify the *SPRY1* gain. Given known consanguinity, regions of

homozygosity larger than 1 Mb were also analysed in the index patient, but these did not intersect *SPRY1*. Additionally, the normalised coverage value for *SPRY1* in the index patient was shown to be ~2.5, potentially indicating a triplication on one allele. Collectively, this suggests that there is likely an alternative genetic explanation for the phenotype in these three siblings. This patient was also found to harbour a mosaic frameshift variant in *SPRY1*, encoding p.(Tyr257Ser*20) (table 4.4), which was considered incidental owing to recent evidence that indicates heterozygous loss-of-function variants in *SPRY1* are likely not disease-causing (see chapter 5) (Tooze et al., 2022a).

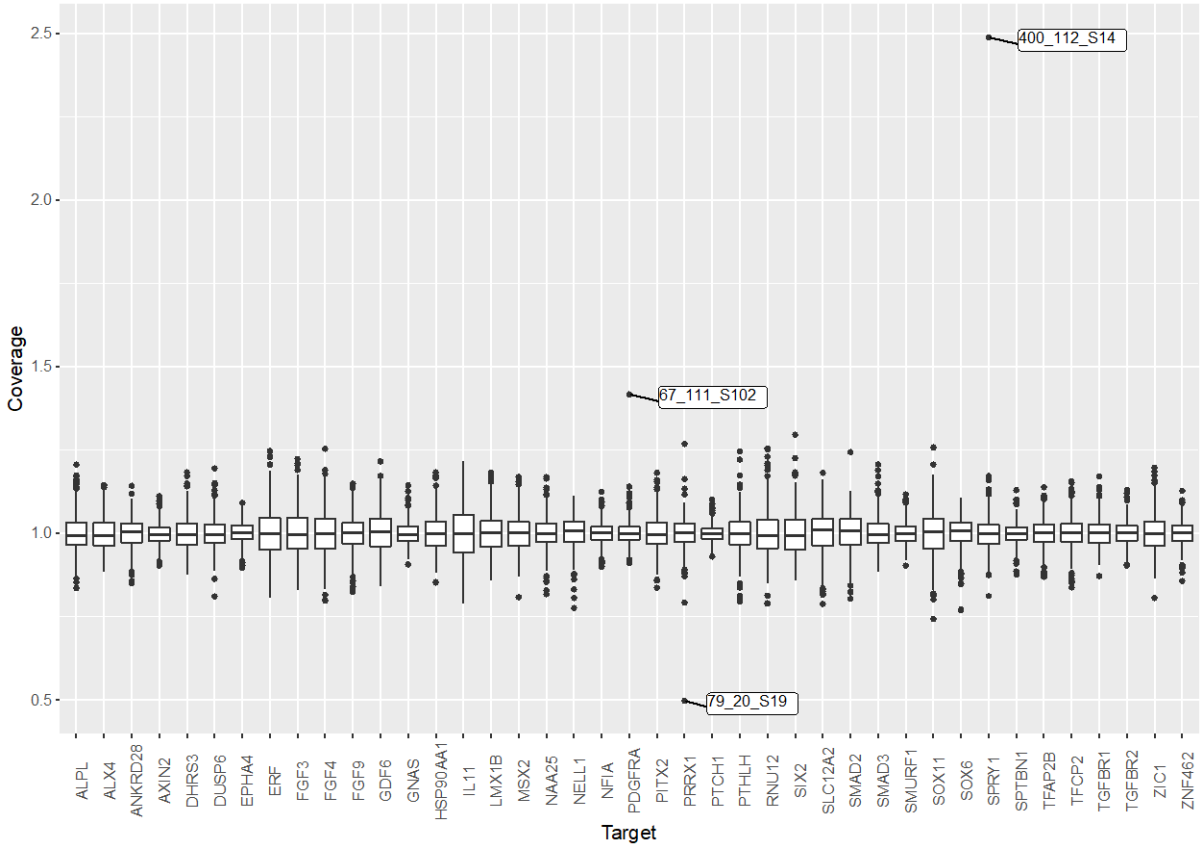


Figure 4.3 Boxplot of normalised gene coverage for included samples (*n* = 479). For each sample, the mean coverage for a given gene (*n* = 41) was normalised to the mean coverage for all genes screened in that sample; this value was normalised to the mean coverage for a given gene assayed across all samples (*n* = 479). RStudio was used to display a boxplot of values (the R-code used to generate this plot is detailed in figure A4.4), plotting the mean coverage for each gene, the upper (Q3) and lower (Q1) interquartile range (IQR) and outliers (Q1 – 1.5 x IQR; Q3 + 1.5 x IQR). Samples ID's that are below 0.6 or above 1.4 have been annotated as labels.

Table 4.5 Identified copy number variants from 479 samples

Gene	CNV	Segregation	Variant Zygosity[a]	Del Score	CADD	Comments	Craniosynostosis/ clinical information
PRRX1	Partial gene deletion (exons 1–2)	Maternally inherited	Het	-	-	Confirmed independently by array and breakpoint PCR (Chr1:170649418–170725330 [GRCh38])	Non-syndromic uni-coronal (left). Teratogenic exposure.
PRRX1	Whole gene deletion	Not confirmed	Het	-	-	Confirmed by array (GRCh37): 1q24.2q25.2 (Chr1:167699849–178212109)	Syndromic bi-coronal.
PDGFRA	Duplication of exons 3–22	Not tested	Het	-	-	Large duplications are reported in DGV (*n* = 3) that intersect this region	Non-syndromic multiple (sagittal, left lambdoid). Bilateral squamosal synostosis also present.
SPRY1	Copy number gain in exon 3	The variant is absent in the affected sibling	Suspected triplication	-	-	There is known consanguinity within this family and the CNV was not found to segregate in the affected siblings or intersect a known ROH. As such, there is likely an alternative genetic explanation for this family.	Syndromic sagittal.

[a]Het = Heterozygous
Del Score = Deleterious Score (0 = likely benign, 6 = likely pathogenic)

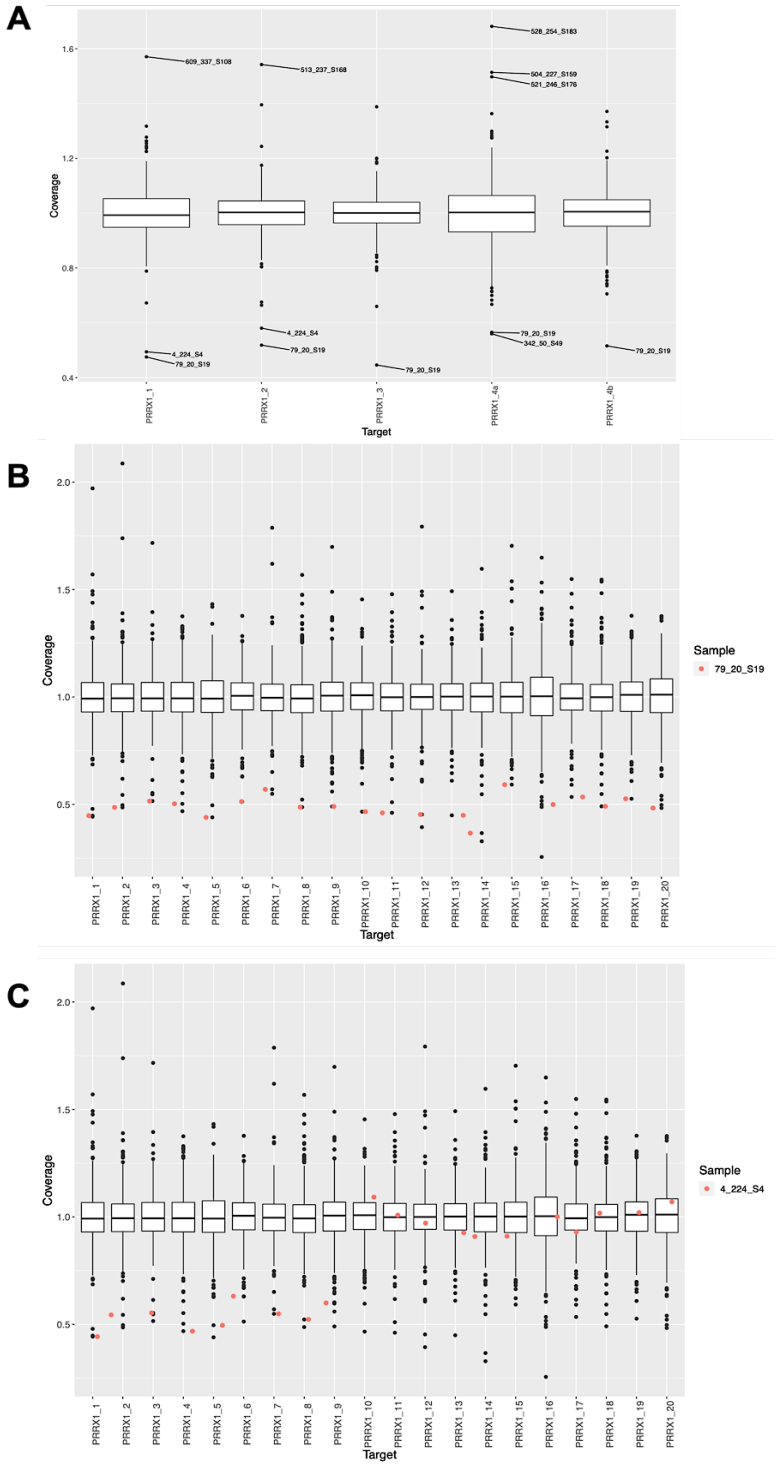


Figure 4.4 Boxplots of the normalised exon and probe coverage for *PRRX1*. (A–C) Boxplots show the normalised mean coverage, upper (Q3) and lower (Q1) interquartile range (IQR), and outliers (Q1 – 1.5 x IQR; Q3 + 1.5 x IQR) for *PRRX1*. (A) Samples with a normalised exon coverage above 1.4 or below 0.6 are labelled. (B) Normalised non-overlapping probe coverage for all regions are shown in red for sample 79_20_S19, illustrating a 50% reduction in coverage for all regions (heterozygous whole gene deletion). (C) Normalised non-overlapping probe coverage for all probe regions are shown in red for sample 4_224_S4 showing a heterozygous deletion from probes 1–10, corresponding to exons 1–2 (consistent with panel A).

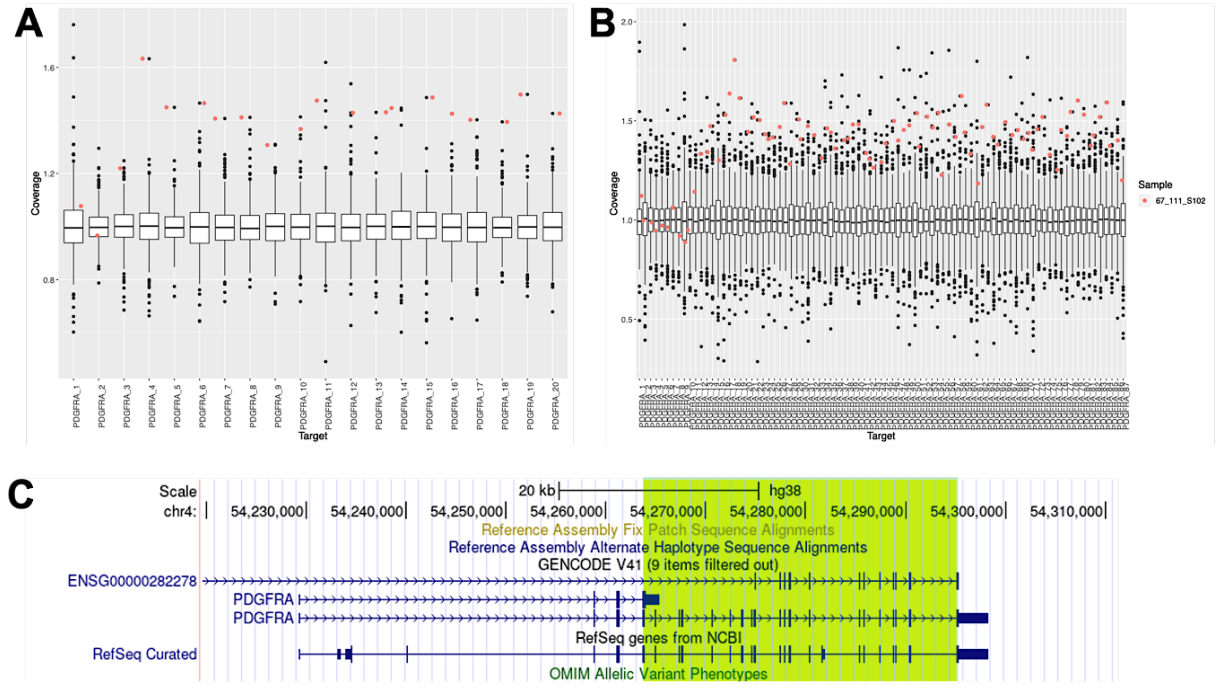


Figure 4.5 Boxplot of the normalised exon and probe coverage for *PDGFRA*. (A–B) The boxplots show the normalised mean coverage for each exon or non-overlapping probe region in *PDGFRA*, the upper (Q3) and lower (Q1) interquartile range (IQR) and outliers (Q1 – 1.5 x IQR; Q3 + 1.5 x IQR). The normalised exon (A) or probe (B) coverage for sample 67_111_S102 is shown in red, illustrating a heterozygous duplication from exons 3–20. (C) A screenshot from UCSC displaying the region affected by the duplication in green.

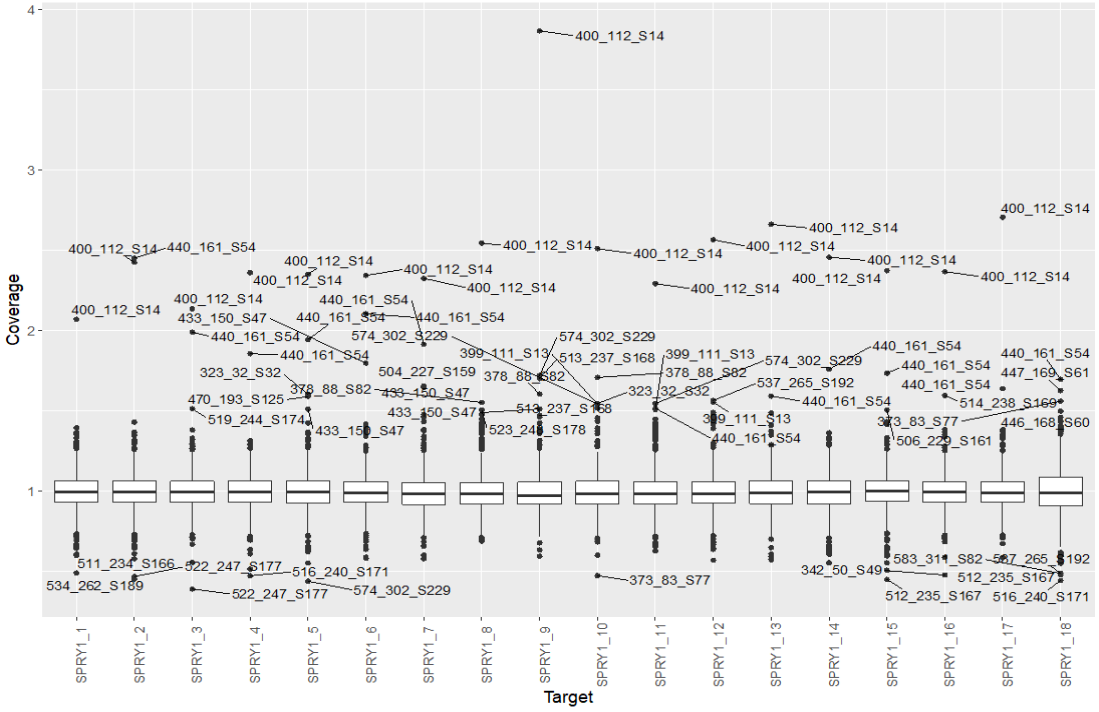


Figure 4.6 Boxplot of normalised probe coverage for *SPRY1*. The boxplot shows the mean coverage for each non-overlapping probe region in *SPRY1*, the upper (Q3) and lower (Q1) interquartile range (IQR) and outliers (Q1 – 1.5 x IQR; Q3 + 1.5 x IQR). Samples ID's that are below 0.5 or above 1.5 have been annotated. Sample 400_112_S14 shows a consistent increase in mean coverage across all probes and corresponds to the *SPRY1* sample identified in figure 4.3.

4.4 Discussion

Targeted resequencing of 617 unsolved patients with craniosynostosis for variants in 42 genes achieved a diagnostic success rate of 8/617 (1.3%). The gene with the greatest diagnostic success was *PRRX1*, for which this analysis identified three patients with SNV or CNVs that contributed to an experimental cohort (see chapter 6).

4.4.1 *PRRX1*

Variants in *PRRX1* made up 38% of the total diagnoses (table 4.3 and 4.5) found through this screen. This in part likely represents an ascertainment bias and inflated estimate of the total proportion of diagnoses attributable to *PRRX1*, compared to other genes within this screen. This is because non-syndromic sagittal and metopic samples were selected if they had not previously been screened for variants in *PRRX1* or had low coverage of this gene, owing to wider interest of this gene as a candidate (see chapter 6). Additionally, given other experimental evidence that penetrance of craniosynostosis is only ~12.5% in families with *PRRX1* variants (see chapter 6), any variant found to be inherited from an unaffected parent in this gene was not discounted (for example, the variant causing loss of the start codon and the partial gene deletion of exons 1–2). In contrast, variants in other genes within this screen found on dideoxy-sequencing to be inherited from unaffected family members, without a known human *de novo* variant reported in the literature or other database for that gene, were not prioritised in this analysis (table A4.4).

4.4.2 *SOX6*

SOX6 variants were identified in two patients, increasing the total number of cases of patients with craniosynostosis (*n* = 7) and *SOX6* variants by 40% (Tolchin et al., 2020,

Chen et al., 2022a, Tagariello et al., 2006, Tooze et al., 2023). In line with previous reports of variants in *SOX6* (Tolchin et al., 2020), both variants within this screen were loss-of-function and predicted to abrogate *SOX6* activity, supporting haploinsufficiency as the likely mechanism. Interestingly, one of the patients had previously undergone WES by the Clinical Genetics Group (Miller et al., 2017), prior to the identification of *SOX6* variants being associated with neurodevelopmental delay and craniosynostosis. In the WES analysis, the frameshift variant in *SOX6* had been missed as the variant spanned the exon 3–intron 3 boundary; at the time, the scoring algorithm used to analyse the WES data assigned the variant as intronic, possibly due to the reverse orientation of this gene, and it was therefore given a low predictive score for pathogenicity. Additionally, the WES was originally aligned using Bowtie2, which was unable to clearly align the deletion variant; later alignment using Burrows-Wheeler Aligner (BWA) clearly showed the presence of this variant and has since been used in future analyses of unsolved cases. The *de novo* stop-gain (p.(Glu542*); table 4.3) was identified in a patient who was enrolled into the 100kGP pilot scheme with neurodevelopmental delay and craniosynostosis. Patients recruited in the pilot scheme were not available for independent analysis by the GeCIPs and I was therefore unable to screen this patient data through this approach. Additionally, the pilot scheme was analysed prior to the *SOX6* publication (Tolchin et al., 2020) and therefore any variant in *SOX6* would not have been prioritised by the Tiering algorithms (Smedley et al., 2021). Importantly, *SOX6* is currently an Amber PanelApp gene meaning that 100kGP patients would not be automatically screened for *SOX6* variants. However, there are now seven patients known with craniosynostosis and loss-of-function variants in *SOX6*, highlighting a need to upgrade its status to Green (Tolchin et al., 2020, Hyder et al., 2021, Chen et al., 2022a, Tooze et al., 2023).

4.4.3 *ALX4*

In this work, a heterozygous missense variant in *ALX4*, encoding p.(Glu273Lys), was identified in a patient and their father with non-syndromic sagittal synostosis. This variant was confirmed independently in both family members by an NHS diagnostic laboratory. However, a currently unanswered question is how to reconcile findings of heterozygous missense variants in craniosynostosis with those identified in parietal foramina, since heterozygous missense variants are reported across the protein in both disorders. Notably, there are no apparent genotype-phenotype correlations based on the distribution of missense variants across the homeodomain, or the protein as a whole. One potential explanation is that similar to *MSX2*, parietal foramina is a result of loss-of-function variation whereas craniosynostosis is a phenotype of gain-of-function. In support, several truncating variants are described in patients with parietal foramina (Mavrogiannis et al., 2006) and functional analyses using reporter assays previously suggested that two rare missense variants in patients with craniosynostosis (p.(Val7Phe) and p.(Lys211Glu)) caused a gain-of-function effect (Yagnik et al., 2012). However, this hypothesis is hard to explain in light of evidence that individuals with variants affecting the same amino acid are described in both disorders. Indeed, five individuals are reported with a variant affecting Arg257: two in this study (p.(Arg257His)), one previous report (p.(Arg257Cys)) (Fonteles et al., 2021), one unaffected mother of a child with craniosynostosis harbouring a p.(Arg257Ala) variant (unpublished), and a further individual with parietal foramina described with a missense variant in Decipher (p.(Arg257Pro)). It is currently not clear whether these patients with craniosynostosis also have phenotypes suggestive of parietal foramina.

Discounting variants outside of the functionally relevant homeodomain, there are eight families with craniosynostosis and rare variants, indicating the variants in *ALX4* could be

enriched in patients with craniosynostosis. To provide a comparator, I examined data from the gnomAD database (v2.1.1). In ~250,000 alleles in gnomAD, there are 90 missense substitutions that potentially disrupt the homeodomain. Including patients screened within this resequencing analysis (*n* = 554 patients with complete coverage of *ALX4*), two previous exome or genome sequencing studies (*n* = 558) (Timberlake et al., 2017a, Miller et al., 2017) and the UK 100kGP cohort of patients with craniosynostosis (*n* = 142), I identified 3 variants in 2508 alleles. These data indicate a marginal over-representation (3-fold; Fisher exact test: $p < 0.03$) of variants associated with craniosynostosis compared to gnomAD, consistent with a weak causal contribution to craniosynostosis.

For four of the families with craniosynostosis, the variants could be considered as likely pathogenic/ pathogenic due to segregation of the variant in affected family members (2x p.(Arg215Trp), and p.(Glu273Lys)) or because of the identification of a *de novo* variant (p.(Ala267Thr)) (Fonteles et al., 2021) (figure 4.7). Additionally, evidence suggests that variants affecting the second residue of the homeodomain (Arg215 in ALX4) are likely damaging as they disrupt a key DNA-binding residue (Noyes et al., 2008). For the remaining variants in patients with craniosynostosis, there is documented non-penetrance meaning that these variants would be considered a VUS without additional functional evidence. For example, a report of a patient with craniosynostosis and a variant affecting the 3rd position of the ALX4 homeodomain (p.(Arg216Trp)), inherited from their unaffected mother (Di Rocco et al., 2023). Although there are no specific reports for ALX4, there are functional studies of variants affecting the same residue in related homeodomain proteins. Evidence in PRRX1 (a related homeobox transcription factor) suggests that variants affecting the 3rd residue cause abnormal nuclear localisation and are associated with incomplete penetrance (see chapter 6). Substitutions of the 44th position of the homeodomain (Arg257 in ALX4) in other

homeobox transcription factors (including SIX3, PHOX2B, and HNF1A) have been shown to disrupt protein stability and function, or prevent DNA-binding activity (Laflamme et al., 2004, Reiff et al., 2010, Najmi et al., 2017, Chi, 2005). Given these data, there is sufficient evidence to suggest that variants in *ALX4* are associated with craniosynostosis with reduced penetrance, which is possibly due to a gain-of-function effect.

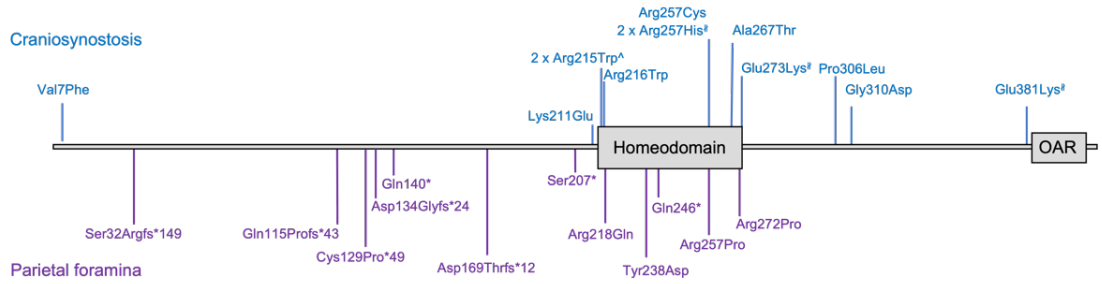


Figure 4.7 *ALX4* variants identified in patients with craniosynostosis and parietal foramina. In blue, variants identified in patients with craniosynostosis, highlighting variants identified in this resequencing analysis (‡) and CranioDB (^). Rare variants reported in previous literature are included in the schematic (Fonteles et al., 2021, Yagnik et al., 2012, Di Rocco et al., 2023). Below, variants identified in patients with parietal foramina are annotated in purple (Mavrogiannis et al., 2006, Lee et al., 2018, Walters et al., 2021).

4.4.4 *MSX2*

Additionally, a missense variant in *MSX2* (encoding p.(Pro148Thr)), the first gene to be implicated as a genetic cause of craniosynostosis (Jabs et al., 1993), was identified. Loss-of-function variants in *MSX2* are typically associated with enlarged parietal foramina (Wilkie et al., 2000), however, a specific gain-of-function variant at proline-148 has been shown to cause a dominant craniosynostosis phenotype by increasing the affinity of MSX2 to its targets (Ma et al., 1996). Variants at this position are not described in gnomAD and affect a highly conserved position of the homeodomain. Proline to histidine (Jabs et al., 1993), proline to leucine (Florisson et al., 2013, Janssen et al., 2013), and proline to serine (Timberlake et al., 2017a) have all been previously described, and here I identified a fourth amino acid change at this position – proline to threonine. Interestingly, it was known that the patient's mother had taken methotrexate (a folic acid antagonist used to treat autoimmune diseases) during the pregnancy. Craniosynostosis is a known rare complication of foetal methotrexate exposure, with patients typically displaying multi-suture synostosis of the coronal and/or lambdoid sutures (Zarella et al., 2016). For this individual, it was counselled that the craniosynostosis was likely attributable to the teratogenic exposure, yet the identification of the *de novo MSX2* mutation disputes this original hypothesis and is able to explain the multi-suture synostosis in this individual (bi-coronal and sagittal synostosis) as having a genetic origin. This will be of paramount importance when the affected individual is an adult, as the offspring risk is increased from zero to 50%.

Notably, a second rare variant (gnomAD v2.1.1 allele frequency = 1.19E-05) in *MSX2*, encoding p.(Arg159His), was identified in a patient with coronal synostosis. Arginine to leucine or histidine are also reported rarely in gnomAD at this position (7.07E-06 and 3.98E-06, respectively), indicating a hyper mutable position and, indeed, the variant affects a CG dinucleotide. Importantly, non-penetrance has been documented in other

members of the homeobox class of transcription factors, including *PRRX1* (chapter 6) but also *ALX4* as described above (Fonteles et al., 2021), and therefore the presence of this variant at low frequency in the general population does not in itself provide sufficient evidence to discount pathogenicity of the p.(Arg159His) variant. However, there is no convincing evidence to suggest that the 18th position of the homeodomain is essential for binding DNA, or that it stabilises the protein structure, and therefore further functional studies would be required to ascertain any likely pathophysiological mechanism.

4.4.5 *SMAD3*

The last variant identified to be diagnostic was a missense substitution at a canonical splice site in *SMAD3* (c.206+1G>A; p.(?)), resulting in loss of one functional copy of *SMAD3* which was confirmed by the analysis of patient cDNA. The variant would be predicted to affect the MH1 domain for which loss-of-function variants are known to cause LDS (Schepers et al., 2018). Importantly, a patient with LDS is described in ClinVar to harbour a variant at the same position, albeit a different nucleotide change (c.206+1G>C). The patient described here displayed non-syndromic uni-coronal synostosis, but it was unknown if they also had additional features suggestive of LDS. However, a known complication of LDS is the development of aortic aneurysms and, interestingly, there is a documented family history of mitral valve prolapse which has been identified in 28% of a cohort of patients with *SMAD3* variants (Chesneau et al., 2020). It would therefore be important that this patient is monitored for this in future, and it would be interesting to assess segregation of the *SMAD3* variant in the wider family.

4.4.6 Candidate variants

Importantly, this study highlighted several further candidate variants which may be of future research interest, namely the two rare variants discussed in *ALX4* affecting the same residue in different families (p.(Arg257His)). However, a real challenge in assigning

pathogenic status to many of the variants identified in this study, is that the majority display non-penetrance. Nevertheless, the most recent discoveries of genes in craniosynostosis (*SMAD6* and *PRRX1*) were both associated with a marked reduction in penetrance (16% and 12.5%, respectively) (Calpena et al., 2020) (see chapter 6), and therefore analysis of variants displaying non-penetrance is an important avenue to pursue to enable further genetic discoveries.

4.4.7 Success of targeted resequencing for genetic discovery

The overall diagnostic success rate was only 1.3%; although this figure is considerably lower than what could be achieved through whole exome or genome sequencing, there are several reasons why panel-based sequencing might be preferred. Primarily, for the analysis of multiple samples in parallel, targeted enrichment is considerably cheaper per sample. In my resequencing, total costs were ~£40/ sample including library preparation and target enrichment of all samples. The libraries from a subset of 297 samples included in this panel were prepared using a TWIST Bioscience kit (Calpena samples, table 4.2) before I started this project, however it would have been possible to reduce the total cost of the resequencing project further to ~£30/ sample using IDT's library preparation kit. Comparatively, estimates of whole exome (£382–3592 per sample) or genome (£1312–17243 per sample) sequencing (Schwarze et al., 2018) are considerably higher. Nevertheless, if WES had been achievable for all of this cohort, the diagnostic success rate would likely be elevated.

A further consideration is that this resequencing analysis consisted of a large proportion of non-syndromic sagittal and metopic individuals (which accounted for 32% [200/617] of the cohort in total), for which a genetic diagnosis is far less frequently identified. In a previous cohort analysis a positive genetic diagnosis was identified in 1 patient out of 232 (0.4%) with non-syndromic sagittal synostosis and 0/97 individuals with non-

syndromic metopic synostosis (Wilkie et al., 2017). Consistent with previous reports (Wilkie et al., 2017, Tønne et al., 2020), syndromic bi-coronal and multi-suture craniosynostosis displayed the highest diagnostic success rate (14.3% and 7.5%, respectively). Out of the diagnoses achieved in this resequencing study, 5/8 (62.5%) were made in patients with multi-suture synostosis (either syndromic or non-syndromic), giving an overall diagnostic success rate of 3.8% (5/131) for multi-suture synostosis. The remaining diagnoses were one patient with syndromic sagittal synostosis, and two patients with non-syndromic uni-coronal synostosis (table 4.2), likely reflecting the greater genetic component associated with the unique pathophysiology of the coronal suture.

Last, while the short-read sequencing used in this analysis does not lend itself to the identification of more complex CNVs which may affect regulatory elements, introns, or overall gene structure (for example, inversions), I was successful in identifying four CNVs in three genes which affected coding regions. Importantly, two of the heterozygous CNVs added to the experimental *PRRX1* cohort (see chapter 6), providing further evidence to support that haploinsufficiency of *PRRX1* is a likely cause of craniosynostosis. Nevertheless, the limit of this technology is such that CNVs in repetitive elements or variants that affected gene regulation through topologically associating domains would not have been identified and should be investigated through independent analyses.

Chapter 5

Complete loss of Spry-1 (sprouty homolog 1) function in a child with craniosynostosis, inner ear, and renal anomalies

5.1 Introduction

Studies in *Drosophila* first identified sprouty as an important negative regulator of FGFRs and other receptor tyrosine kinase (RTK) signalling pathways (Hacohen et al., 1998). In *Drosophila*, sprouty mutants displayed ectopic branching of the airways alongside increased FGF signalling (Hacohen et al., 1998). Following this early work, four mammalian sprouty homologues were described (encoded by *SPRY1–SPRY4*), that in humans range in size from 288–319 amino acids (aa) (figure A5.1), and share three regions of homology: a canonical Cbl-tyrosine kinase-binding domain (PTB; aa 51–57), a serine-rich domain (SRD; aa 112–131) and a cysteine-rich domain (CRD; aa 181–306; numbering refers to Spry-1) (Guy et al., 2009). The CRD is well conserved between members of the Spry family (figure A5.1) and is hypothesised to be important for homo- and hetero-oligomerisation of the Spry family members and protein translocation to the plasma membrane (Vaquero et al., 2019). Outside of the CRD domain, there is little conservation between family members, except for a small region of the SRD and a conserved tyrosine residue within the PTB (aa 53 of human Spry-1).

The molecular mode of action of sprouty proteins are not well understood (Vaquero et al., 2019). Although sprouty proteins lack enzymatic activity, they have been shown to decrease activation of the RAS-MAPK pathway, potentially by binding and sequestering the growth factor receptor-bound protein 2 (GRB2) adapter protein (Hanafusa et al., 2002), following sprouty phosphorylation by RTKs at key tyrosine residues (for example, aa 53 in Spry-1 described above) (Vaquero et al., 2019). This in turn would prevent membrane localisation of son of sevenless (SOS), a RAS guanine exchange factor (Lorenzo and McCormick, 2020), and block activation of the signalling pathway (figure 5.1). However, studies on Spry-2 indicate that inhibition of the RAS/MAPK pathway occurs after the point of receptor and RAS activation, suggesting that sequestering of GRB2 is unlikely to be the inhibitory mechanism. Follow-up analyses showed that Spry-

1 and Spry-2 were able to inhibit phosphorylation of extracellular signal-regulated protein kinase 2 (ERK2) downstream of active RAS (Yusoff et al., 2002). Phosphorylated ERK (pERK) signals entry of the protein into the nucleus, which in turn enables pERK to phosphorylate nuclear substrates and regulate downstream transcriptional processes (Maik-Rachline et al., 2019).

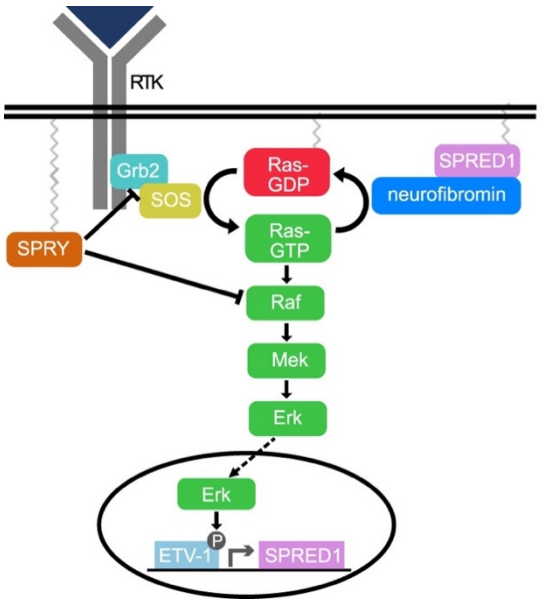


Figure 5.1 Inhibitory mechanism of Sprouty proteins on the RAS-MAPK pathway. Sprouty proteins are believed to inhibit RTK-mediated signalling by competing with GRB2 for binding with SOS, a RAS guanine exchange factor. Sequestrating of GRB2 by Spry would prevent downstream activation of the MAPK signalling pathway. Spry proteins are down-regulated by dephosphorylation of a key phosphotyrosine residue by Shp2 phosphatase or through interaction with c-Cbl, an E3 ubiquitin-ligase, which targets the protein for proteasomal degradation. This image is reproduced from Lorenzo and McCormick (2020) under a Creative Common License (Attribution-NonCommerical 4.0 International).

To date, there are no known human disorders attributable to mutations in *SPRY1*, but variants in *SPRY2* and *SPRY4* have been described in association with IgA nephropathy (IgAN) and hypogonadotropic hypogonadism, respectively. A variant in *SPRY2*, encoding p.(Arg119Trp), was identified to segregate in four generations of a large family (*N* = 31) with 9 affected individuals with autosomal dominant IgAN (Milillo et al., 2015). Functional studies showed a reduction in pERK1 and pERK2 in all *SPRY2* mutation carriers, compared to healthy controls (Milillo et al., 2015). Given that Spry-2 is a negative regulator of RAS-MAPK signalling, potentially through preventing ERK

phosphorylation, this result is somewhat surprising as it would be anticipated that loss of Spry-2 function would increase pERK. Additionally, two individuals with sporadic IgAN were also shown to have dysfunctional ERK1/2 but did not harbour a variant in *SPRY2*, implicating other members of the pathway in the aetiology of IgAN; causative variants in other members of the MAPK-ERK pathway are yet to be identified (Milillo et al., 2018). Additionally, one heterozygous variant at p.(Arg119Trp) has since been reported in gnomAD (v2.1.1), alongside two other variants at the same amino acid encoding glutamine and proline. Given that no further mutations have been described in *SPRY2* associated with IgAN, this variant has since been annotated in ClinVar as a risk factor for IgAN.

Variants in *SPRY4* have been reported in patients with congenital hypogonadotropic hypogonadism (CHH) (Miraoui et al., 2013) and its anosmia-associated form, Kallmann syndrome (KS), which is associated with skeletal abnormalities, midline malformations, and hearing loss. In 10% of cases, a mutation in *FGFR1* is responsible for KS (Hu and Bouloux, 2010). Early zebrafish and mouse models showed that *Spry4* was widely expressed during early development at sites consistent with *Fgf3* (Furthauer et al., 2002, Maximy et al., 1999). Loss of Spry-4 in mice (*Spry4*$^{-/-}$) resulted in death shortly after birth owing to severe defects of the mandible; polysyndactyly was also present in all *Spry4*$^{-/-}$ offspring (Taniguchi et al., 2007). Following this, *SPRY4* was tested as a candidate gene in the pathogenesis of CHH (classified as KS or normosmic idiopathic hypogonadotropic hypogonadism) in 386 unrelated individuals. In this cohort, 14 patients harboured heterozygous missense variants in *SPRY4* (Miraoui et al., 2013). However, 8/14 individuals were found to have either a p.(Lys177Arg) or p.(Ser241Tyr) variant, which are reported in 454 and 1279 alleles out of ~280,000 alleles, respectively, in gnomAD (v2.1.1); the frequency of these variants is too high to be considered causative. In three

of the 14 individuals, a second rare variant was also found in *DUSP6* and *FGFR1* which is likely to contribute to their phenotype (Miraoui et al., 2013).

While bona fide pathogenic variants in human *SPRY1* have not been reported (two heterozygous variants are reported in the literature; see section 5.4), studies in mice have shown that homozygous deletion of the entire coding sequence of *Spry1* (orthologous to human *SPRY1*; figure A5.1) affects multiple developmental processes. In contrast, heterozygous mice (*Spry1*$^{+/-}$) are phenotypically normal. In the original report (Basson et al., 2005), 95% of homozygous neonates (mixed genetic background) displayed unilateral or bilateral ureter and kidney malformations, and 71% died within 5 months. These abnormalities were principally attributable to loss of negative feedback regulation of RET-RTK in the Wolffian duct (Basson et al., 2005), and could be phenocopied by homozygous substitution of a critical tyrosine residue, p.(Tyr53Ala), within the Cbl-tyrosine kinase-binding domain (Vaquero et al., 2019). Notably, crossing onto a different mouse strain background (FVB) was associated with reduced penetrance of urogenital malformations in *Spry1*$^{-/-}$ animals (Yang et al., 2021).

Subsequently, detailed studies of *Spry1*$^{-/-}$ mice have revealed additional defects in multiple organs and cell types, including rostral cortex (Faedo et al., 2010), mammary gland (Koledova et al., 2016), mesenchymal stem cells (Urs et al., 2010) and muscle satellite cells (Shea et al., 2010). These phenotypes are exacerbated by adding loss-of-function alleles of other *Spry* genes (namely *Spry2*), indicating partial functional redundancy between Spry proteins (Ching et al., 2014). Depending on strain background, both increased and decreased skull size were documented in *Spry1*$^{-/-}$ mice, but craniosynostosis was not reported (Percival et al., 2017).

Bona fide disease-causing variants of human *SPRY1* have not been reported and therefore the consequence of complete loss-of-function of the human orthologue was

previously unknown. Here I describe a homozygous *SPRY1* nonsense variant, identified in a patient with syndromic craniosynostosis, severe defects of the inner ear, and mild kidney abnormalities, as a result of complete loss of Spry-1.

5.1.1 Case report

The index patient (II-1) is the first-born male child of consanguineous parents (half-first cousins) of Indian ancestry. Apart from familial hypercholesterolaemia in the mother (I-2), both parents are healthy and phenotypically normal.

Antenatal scans had shown increased nuchal translucency at 12 weeks and dolichocephaly from 20 weeks. Array comparative genomic hybridisation (aCGH) on an amniocentesis sample was normal. The male child was born at 39 weeks' gestation and noted at birth to have scaphocephaly, turricephaly, posterior ridging of the sagittal suture and bitemporal narrowing, with tense, symmetrically positioned cystic lesions (3 x 1 cm) over the parietal bones bilaterally. At 12 weeks of age, his length and weight were around the 50th centile; his head circumference could not be accurately measured owing to the cystic lesions. Distinctive facial features (figure 5.2a) included hypertelorism, a broad base to his nose, large ears with earlobe creases and naevus flammeus over his philtrum and central forehead. There was a capillary malformation on his upper back. In his extremities there was soft skin on his hands with deep palmar and plantar creases, and a proximally placed overriding 2nd toe on the right. Computerised tomography (CT) scan of his head (aged 10 weeks) showed sagittal synostosis and bilateral subcutaneous cystic lesions over the lambdoid sutures (figure 5.2b). Aged 4 months he had spring-assisted cranioplasty, subsequently removed aged 8 months, and removal of the scalp lesions, which were shown histologically to be dermoid cysts.

He failed his new-born hearing screen and was subsequently found to have bilateral profound sensorineural hearing loss. Magnetic resonance imaging (MRI) scan of his inner ear (aged 8 months) demonstrated bilateral inner ear dysplasia. On the right, there was cochlear hypoplasia with anterior off-set, markedly hypoplastic apical turn (cochlear hypoplasia 4), similar to the unwound cochlea found in branchio-oto-renal syndrome. The vestibule and semi-circular canals were relatively normal. On the left there was an incomplete partition type 1 malformation of the cochlea with abnormal vestibule and semi-circular canals (figure 5.2d,e). The VIIIth nerves were normal bilaterally. Aged 15 and 18 months, he underwent bilateral cochlear implantation in a two-stage procedure (left then right). MRI of the brain showed a small corpus callosum, which was dysmorphic posteriorly (figure 5.2f).

A renal ultrasound and renogram showed a simple left renal parapelvic cyst, while the right kidney appeared normal. He had normal renal function and was normotensive with no microalbuminuria; testes and external genitalia were normal. Echocardiogram at 3 months of age showed a small atrial septal defect (which had resolved spontaneously by 4 years) and mild supravalvular pulmonary stenosis of no clinical significance. His speech and language development were delayed, consistent with his hearing loss. However, his motor and cognitive development has been normal. At 4 years of age, he used 3-word sentences and followed complex instructions. Ophthalmology review was normal, with no concerns regarding his vision.

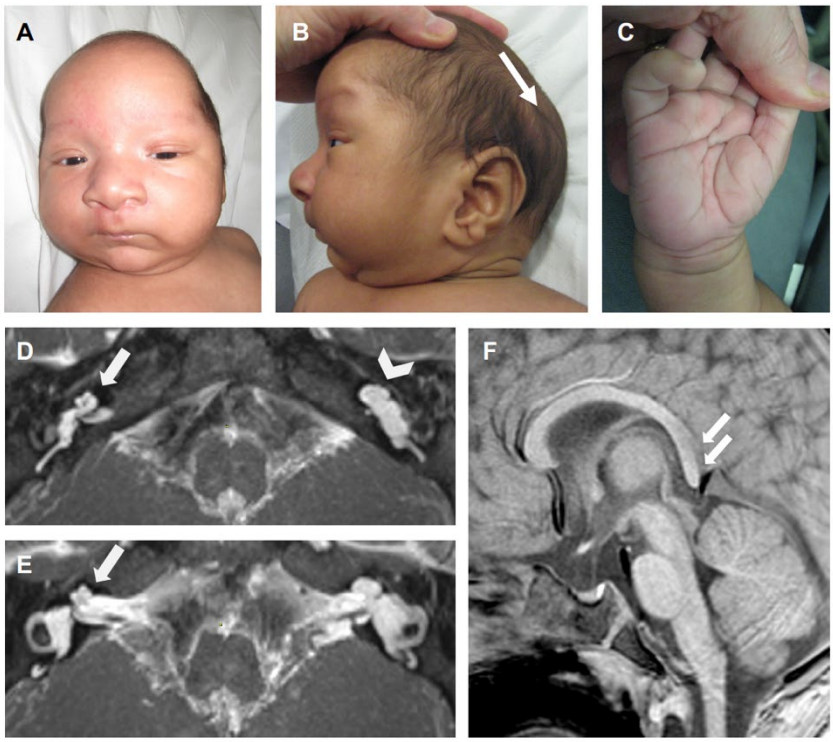


Figure 5.2 Clinical features of proband II-1 homozygous for a *SPRY1* nonsense variant. (A-C) Photographs of the index patient aged 11-weeks showing scaphocephaly, turricephaly, bitemporal narrowing, hypertelorism, earlobe creases, cystic lesions overlying the lambdoid suture (B, arrow) and deep palmar crease (C). (D-F) Axial maximum intensity projection 3D high resolution T2-weighted images aged 8 months. (D-E) On his right side, he displays cochlear hypoplasia type 4 (arrows) with reasonably normal semi-circular canals. On the left, he has an incomplete partition type 1 cochlea (arrowhead), dilated vestibule with dysplastic lateral semi-circular canals. (F) Sagittal T1-weighted images showing small corpus callosum with hypoplastic splenium (double arrows).

5.2 Methods

5.2.1 Genetic investigation and variant filtering of whole-genome sequencing data

Following normal routine genetic investigations, the family was enrolled into the Genomics England 100kGP, and Illumina WGS was performed on the proband and both parents (Smedley et al., 2021). No clinically reportable variants were identified. The participant VCF files were examined in the GE Research Environment, as previously described (Hyder et al., 2021) (see chapter 3, section 3.2.3.1). Following annotation with Ensembl VEP (McLaren et al., 2016), the variants were filtered using a gnomAD (v2.1.1) allele frequency of ≤0.01 (Karczewski et al., 2020), CADD score ≥20 (Rentzsch et al., 2019), and no intersection with low complexity regions. Given the reported consanguinity, variants within homozygous regions identified using Illumina's ROHcaller were prioritised.

5.2.2 Functional analysis of cDNA and protein

Blood DNA was isolated from the proband and both parents to prepare gDNA and lymphoblastoid cell lines as detailed in chapter 2 (section 2.1 and 2.6.1). Dideoxy-sequencing of gDNA was used to confirm the *SPRY1* variants (section 2.2). RNA was obtained from lymphoblastoid cells and, following which, cDNA was analysed (including deep sequencing) by RT-PCR; protein quantification of Spry-1 was undertaken by western blotting. All primers used in analyses are detailed in table A5.1.

5.2.2.1 RT-PCR

Lymphoblastoid cells were cultured for RNA extraction and cDNA synthesis as described in chapter 2 (section 2.8). A 2 μl sample of cDNA was diluted in a master mix containing 5 μl 5x Q5 reaction buffer, 1.25 μl each of 10 μM forward (5'-

GAAAAGGGATTTCAGATGCATGCCAGG-3') and reverse primer (5'-CATGCTTTTCTTGTCTTGGTGCTGTC-3'), 0.5 μl of dNTPs (10 mM), 5 μl of 5x Q5 GC enhancer, 0.25 μl of Q5 high fidelity enzyme and made up to a final volume of 25 μl with water. Samples were placed in a thermocycler under the following conditions: 98°C for 30 s, followed by 35 cycles of 98°C for 10 s, 70°C for 20 s, 72°C for 30 s, and a final elongation of 72°C for 2 min. Samples underwent dideoxy-sequencing by the MRC Weatherall Institute of Molecular Medicine sequencing facility.

5.2.2.2 Next Generation Sequencing (MiSeq)

A 50 ng sample of each cDNA sample was prepared for deep sequencing as detailed in the methods chapter (section 2.3). An annealing temperature of 70°C for 30 s was used in the Q5 protocol.

5.2.2.3 Western blot

A protocol for analysis of protein by western blot was followed as detailed in the methods chapter (section 2.7). Anti-Spry-1 (rabbit monoclonal antibody [mAb] D9V6P #13013, Cell Signaling Technology), diluted 1:1000, was added to the membrane and left shaking at 4°C overnight. After 16 hours, the membrane was washed with 1x T-TBS, and secondary donkey anti-rabbit-HRP (diluted 1:10,000; Abcam, ab97085) was incubated with the membrane for 1 hour at RT, shaking gently. Following imaging, the primary and secondary antibodies were removed using the restore western blot stripping buffer (Thermo Scientific) before being stained with anti-GAPDH-HRP (Cell Signaling Technology, 14C10) diluted 1:10,000 at RT for 1.5 hours.

5.2.3 Targeted resequencing of *SPRY1*

A search for additional variants in *SPRY1* was undertaken by targeted resequencing of 617 individuals with previously undiagnosed craniosynostosis (see chapter 4; primers in table A5.2).

5.3 Results

5.3.1 Molecular analysis

Initial genetic investigations of the proband included normal postnatal aCGH and craniosynostosis gene panel testing. Analysis of WGS data from the 100kGP showed five regions of homozygosity ≥2 Mb (figure 5.3b); after filtering, three homozygous variants remained, including one stop-gain (in *SPRY1*) and two predicted missense variants (in *CCDC80* and *RASSF6*). Further evaluation of the *CCDC80* and *RASSF6* homozygous variants, as well as two heterozygous *de novo* missense variants (in *ATXN10* and *POMP*), did not support pathogenicity of these variants (table 5.1), leaving the *SPRY1* variant as the sole candidate to be prioritised for further analysis. Dideoxy-sequencing confirmed the homozygous nonsense variant (c.80T>A, encoding p.(Leu27*)) in *SPRY1*, which was heterozygous in both parents (figure 5.3c).

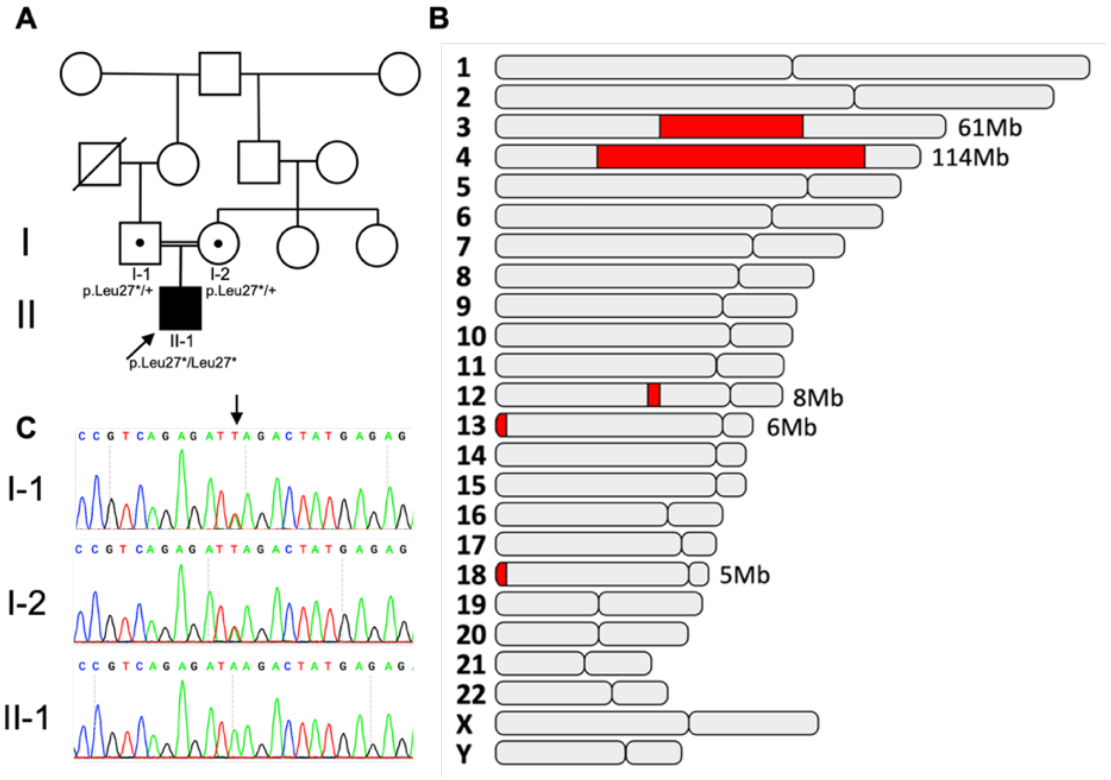


Figure 5.3 Family pedigree and homozygosity analysis of WGS data from the proband II-1. (A) Pedigree figure illustrating both heterozygous parents (I-1 and I-2) with known consanguinity and the index patient (II-1), homozygous for p.(Leu27*) in Spry-1. (B) The output from ROHcaller is depicted in a schematic showing each chromosome, with regions of homozygosity larger than 2 Mb highlighted in red. *SPRY1* resides within the largest region of homozygosity (114 Mb) located on chromosome 4. (C) Dideoxy sequencing traces for each member of the pedigree. The arrow indicates the position of the c.80T>A nonsense variant in *SPRY1*.

Table 5.1 Variants prioritised after filtering by allele frequency and CADD score

Gene	Chr	Pos[a]	Nucleotide and amino acid change	Mode of inheritance	Allele Frequency (AF)[b]	dbSNP	CADD score	Reported Phenotypes	Reason discounted
CCDC80	3	112609989	NM_199511.2: c.2414C>T; p.(Ala805Val)	Recessive	0.008783	rs56683778	31	Colorectal carcinoma	60 reported homozygous p.(Ala805Val) variants in gnomAD v2.1.1
RASSF6	4	73587902	NM_177532.5: c.320A>G; p.(Tyr107Cys)	Recessive	-	rs776608439	25.5	Multiple cancers	Multiple amino acids tolerated at this position.
SPRY1	4	123401671	NM_001258038.2 : c.80T>A; p.(Leu27*)	Recessive	-	-	35	Craniofacial malformation Kidney and urinary tract malformation	-
POMP	13	28664562	NM_015932.6: c.155A>G; p.(Glu52Gly)	*De novo*	-	-	31	Keratosis linearis with ichthyosis congenita and sclerosing keratoderma; proteasome-associated autoinflammatory syndrome 2	Associated disorders arise from loss-of-function variants. Described phenotypes are not consistent with our patient (Poli et al., 2018).
ATXN10	22	45700287	NM_013236.4: c.397C>T; p.(Arg133Cys)	*De novo*	0.000024	-	28.1	Spinocerebellar ataxia	There are 6 other reported heterozygous p.(Arg133Cys) variants in gnomAD. *Atxn10* null heterozygous mice show no overt phenotype (Wakamiya et al., 2006).

[a]Genome GRCh38

[b]gnomAD AF taken from v2.1.1 (combined exome and genome data). AF for *de novo* variants was set at below 0.001, whereas AF for recessive variants was set at less than 0.01

5.3.2 Functional validation of the escape from nonsense-mediate decay

The premature termination codon, p.(Leu27*), arises in the last and only coding exon of *SPRY1* (figure 5.4a) and is predicted to affect all transcripts (figure 5.4b). As the variant resides within the last coding exon, it is predicted to escape nonsense-mediated decay (NMD) (Lindeboom et al., 2019). To confirm escape of NMD, I undertook an RT-PCR analysis of RNA extracted from lymphoblastoid cells and, in support, identified RNA products in all three family members (figure 5.5a). The presence of the variant in the cDNA in approximately equal quantity to the normal allele was confirmed by deep sequencing of RT-PCR products from the parental samples, consistent with normal production of mutant RNA transcript and escape from NMD in lymphoblastoid cells (table 5.2).

In addition, there are no other further methionine residues prior to Leu27 in an alternative reading frame that would allow for read-through of the stop-gain and protein to be produced (figure 5.4b). Therefore, to assess the effect of the variant on Spry-1 protein synthesis, I undertook western blotting and identified a band of ~35 kDa in all healthy controls and both parents, corresponding to the expected size of Spry-1 (figure 5.5b). Of note, the antibody used was raised to an epitope comprising residues surrounding aa 70 of human Spry-1, located C-terminal of Leu27 (figure 5.4a). Accordingly, the predicted truncated protein appeared entirely absent in the proband and the band intensity was considerably weaker in both heterozygous parents compared to controls, consistent with loss of one copy of full length Spry-1. Results of replicate analyses and additional controls are shown in figure A5.2.

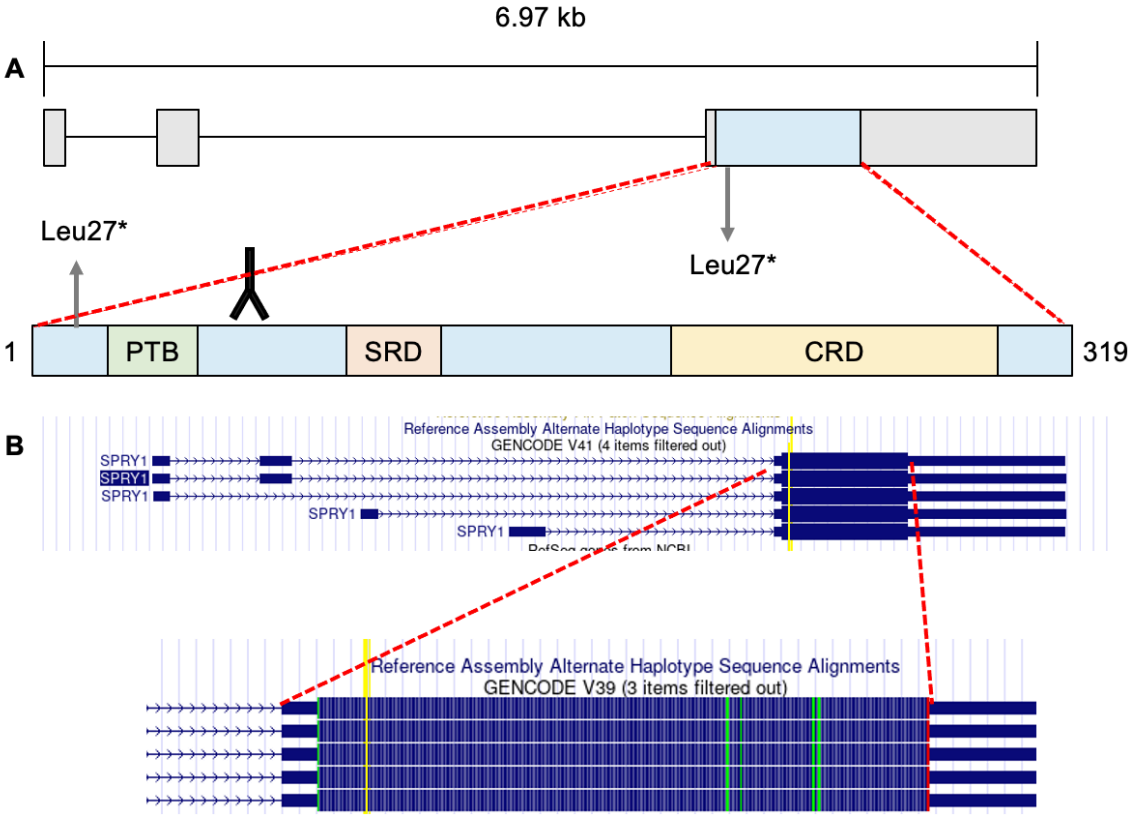


Figure 5.4 *SPRY1* exon structure and predicted consequence of p.(Leu27)* nonsense variant. (A) Schematic of *SPRY1* with the coding region marked in blue and the non-coding exons in grey. Below, encoded protein showing conserved domains (PTB, phosphotyrosine-binding domain; SRD, serine-rich domain; CTD, cysteine-rich domain). The position of the antibody in the translated protein is indicated (inverted Y). (B) Screenshot from the UCSC web browser (https://genome.ucsc.edu/) showing that all transcripts of *SPRY1* would be affected by the p.(Leu27*) variant (position marked by the yellow line). Below, the coding exon has been zoomed in to identify any further methionine residues (highlighted in green), which all reside downstream of p.(Leu27).

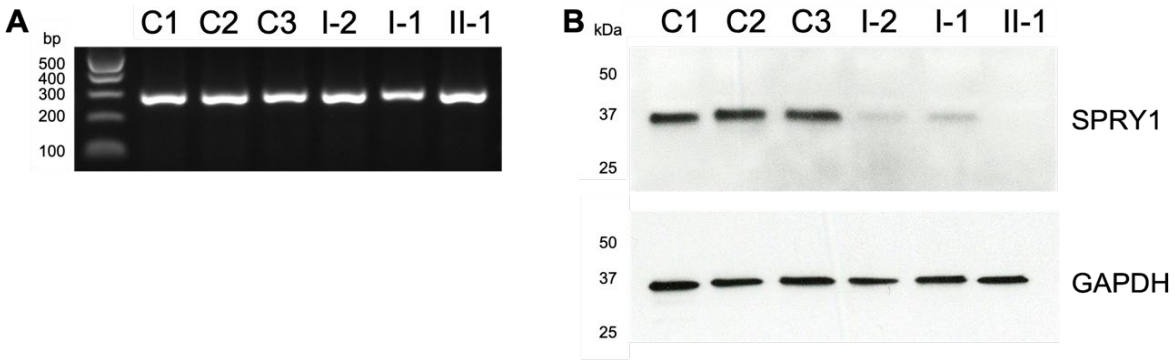


Figure 5.5 Identification of *SPRY1* nonsense variant and effect on expression of RNA and protein. (A) RT-PCR of *SPRY1* including three unaffected controls (C1–3), the mother (I-2), father (I-1), and proband (II-1). (B) Western blot using an antibody against Spry-1, targeting an epitope downstream of the stop-gain; lanes are labelled as in part B. The blots were stripped and re-probed using an antibody against GAPDH as a positive control.

Table 5.2 Deep sequencing of cDNA from both parents and the affected child

Family Member	Reference Allele	Alternative Allele
I-1	11841/22180 = 53%	10337/22180 = 47%
I-2	6126/12692 = 48%	6565/12692 = 52%
II-1	5/21937 = 0.02%	21930/21937 = 100%

5.4 Discussion

To the best of my knowledge, this is the first description of a patient harbouring a homozygous nonsense variant in *SPRY1*. This variant is predicted to affect all known *SPRY1* splice-forms and to lead to a severely truncated Spry-1 protein that has lost all conserved features. The complete loss of full-length functional Spry-1 protein in the proband was corroborated by western blot analysis (figure 5.5b). The parents, who were phenotypically normal, were heterozygous for the variant and exhibited diminished amounts of full-length Spry-1 protein.

Major points of interest arising from this work are the delineation of human *SPRY1* loss-of-function phenotypes in homozygous and heterozygous states, and comparison with the mouse null mutant. The patient shares many phenotypes previously reported in mouse *Spry1*$^{-/-}$ mutants (see section 5.1), but with apparent differences in relative severity. Renal and ureteric anomalies, which are a major feature of the murine *Spry1*$^{-/-}$ phenotype and often lead to reduced viability, were manifested in the patient only as a unilateral simple parapelvic cyst, which did not compromise renal function. By comparison, the patient exhibited major developmental abnormalities of the inner ear, affecting both the cochlea and vestibular systems; whereas inner ears of *Spry1*$^{-/-}$ mice are normal (Zhang et al., 2014) but manifest similar anatomical features when crossed with *Spry2* loss-of-function alleles (Wright et al., 2015). Given current evidence that the renal and inner ear abnormalities are predominantly driven by dysregulated signalling through different growth factor/RTK axes (GDNF/RET and FGF/FGFR, respectively), it is possible that these signalling pathways differ in their relative sensitivity to Spry-1 inhibition in the mouse and human. A failure to repress FGF/FGFR signalling is consistent with the presence of sagittal synostosis in the patient, since dysregulated FGF/FGFR signalling is a well-documented cause of this phenotype (Twigg and Wilkie, 2015). The patient manifested several additional phenotypes including dysmorphic facial

features, bilateral dermoid cysts over the scalp, and minor brain anomalies; the identification of additional cases is required to determine whether these fit into a consistent pattern.

There are two previous reports of putatively pathogenic heterozygous *SPRY1* variants: a frameshift (p.(Gln6fs)) (Timberlake et al., 2017b) in a patient with sagittal synostosis, and a stop-gain (p.(Glu79*)) (Chen et al., 2014) in a patient with Noonan syndrome (OMIM #609942). Two arguments suggest that these variants alone are unlikely to account for the associated phenotypes reported. First, the gnomAD dataset (v2.1.1) identifies 13 individuals out of ~140,000 sequenced, including normal controls, heterozygous for loss-of-function variants in *SPRY1*. This relatively high frequency conflicts with such alleles having highly penetrant pathogenic effects and accords with the low pLI score (pLI = 0). Second, in the current study, both parents were heterozygous for the loss-of-function variant in *SPRY1* but displayed no clinical phenotype. Importantly, no homozygous loss-of-function variants are recorded in the gnomAD v2.1.1 or v3.1.1 datasets.

Based on similarities of the patient's phenotype with the mouse *Spry1*$^{-/-}$ mutant and absence of other credible pathogenic variants in the patient, it is likely that complete loss of Spry-1 accounts for the abnormal phenotypes described. Further corroboration of genotype-phenotype correlations will require the identification of additional individuals harbouring homozygous variants in *SPRY1*; however, attempts to identify such individuals through GeneMatcher or by the sequencing of 617 patients with undiagnosed craniosynostosis (chapter 4) were unsuccessful.

In conclusion, I present the first evidence of a Spry-1 knockout in humans. Although these observations are based on a single case study, there are substantial overlaps with the reported phenotypes in *Spry1*$^{-/-}$ mutants in mice. Together with functional evidence

showing loss of full-length protein within the proband, and lack of any other identified genetic pathology, I propose that the *SPRY1* variant is likely causative.

Chapter 6

Pathogenic variants in the paired-related homeobox 1 (*PRRX1*) gene cause craniosynostosis

6.1 Introduction

The mammalian paired-related homeobox family comprises two genes, *PRRX1* and *PRRX2*, which classify within the PRD class of homeobox transcription factors (Holland et al., 2007). The orthologous genes in mice (*Prrx1*/*Prrx2*) were previously named *MHox*/*Prx1* and *Prx2*, respectively. The homeobox class contains a highly conserved 60 aa domain known as the homeodomain, comprising an N-terminal arm and three alpha-helices which mediate DNA-binding (figure 6.1b). Sequence-specific recognition is mediated by the third helix (recognition helix) docking in the major groove of the DNA while the N-terminal arm docks in the minor groove (Chu et al., 2012); this specificity is defined by residues at positions 2, 3, and 5–8 of the N-terminal arm, alongside residues at positions 47, 50, 51, 54 and 55 of the recognition helix (Noyes et al., 2008) of the homeodomain (figure 6.1c–d). *PRRX1* consists of 5 coding exons; alternative splicing of exon 4 generates two distinct protein isoforms, *PRRX1a* (NM_022716.4; 245 aa in humans) and *PRRX1b* (NM_006902.5; 217 aa) (figure 6.1a). *PRRX1a* and *PRRX2* (which is not subject to alternative splicing) contain a highly conserved 14 aa C-terminal otp, aristaless, and rax domain (OAR motif), which is absent in *PRRX1b* (Norris and Kern, 2001b). The OAR domain is present in over 40 other paired-type homeodomain containing proteins, but *PRRX1* and *SHOX* are the only two genes which display alternative splicing to remove the OAR domain in one isoform (Norris et al., 2000). There is conflicting evidence with respect to the function of the OAR domain, with early reports suggesting it functions to repress transactivation (Norris and Kern, 2001b, Norris and Kern, 2001a), similar to the function of the OAR domain in the homeodomain-containing proteins, PITX2 (Footz et al., 2009) and CART1 (Brouwer et al., 2003). In contrast, other reports suggest that the OAR domain in *PRRX1* increases gene transactivation and subsequent loss of this domain results in decreased DNA-binding potential (Jones et al., 2001).

Prrx1 has been shown to be widely expressed within the mouse coronal suture at E15.5 and E17.5 and is a marker of stem cells in the postnatal sutural mesenchyme (Farmer et al., 2021, Takarada et al., 2016, Bragdon et al., 2022). Postnatal skeletal stem cells expressing *Prrx1* are present in the periosteum of long bones where they contribute to bone development and regeneration (Bassir et al., 2019). Additionally, these cells have been shown to reside exclusively within the postnatal calvarial suture, respond to WNT signalling by differentiating into osteoblasts, and are able to regenerate bone upon heterotopic transplantation (Takarada et al., 2016, Wilk et al., 2017). *Prrx1* and *Prrx2* exhibit overlapping expression patterns in undifferentiated mesenchymal cells of the head, limb buds, axial mesoderm, and branchial arches, although there are differences particularly in the heart and brain (Leussink et al., 1995, Lu et al., 1999). Null mutants in mice demonstrated partially redundant roles of *Prrx1* and *Prrx2* affecting craniofacial development (Martin et al., 1995). Whereas $Prrx1^{+/-}$ heterozygotes were normal, $Prrx1^{-/-}$ null mice died at birth with cleft palate and defects of multiple bones of the face and lateral skull and absent supraoccipital bone; additional anomalies were present in the limbs (short, thickened endochondral bones) and vertebrae (incomplete vertebral arches) (Martin et al., 1995). Although $Prrx2^{-/-}$ null mutants were phenotypically normal, $Prrx1^{-/-}$;$Prrx2^{-/-}$ double knockouts exhibited a marked exacerbation of the $Prrx1^{-/-}$ phenotype, including a severely reduced lower jaw (micrognathia), pointed snout, absence of external ears, and a cleft mandible and tongue which obstructed the airways (ten Berge et al., 1998). The double mutants ($Prrx1^{-/-}$;$Prrx2^{-/-}$) died within an hour after birth, while $Prrx1^{-/-}$ null mutants survived for 24 hours.

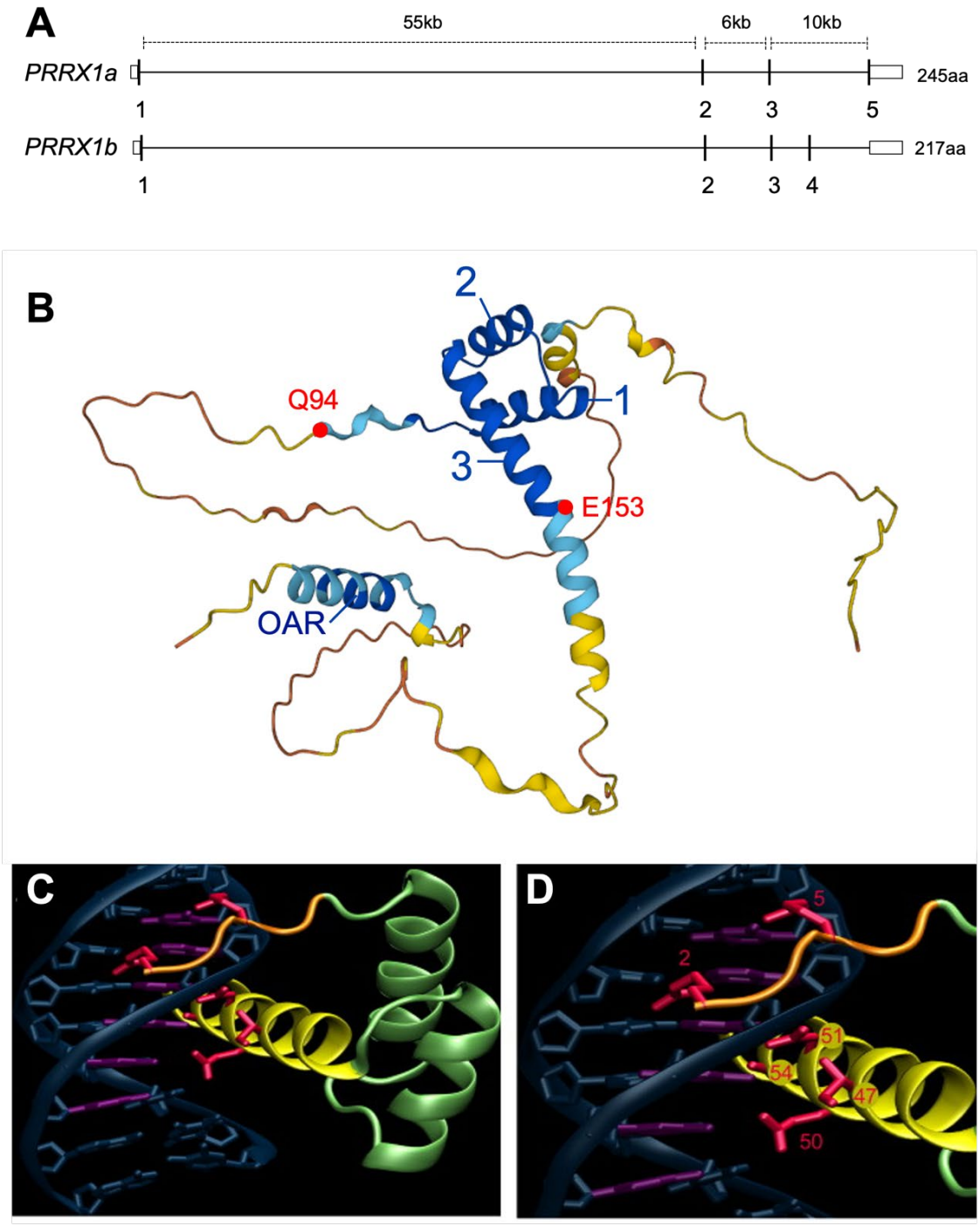


Figure 6.1 Structure of *PRRX1* and binding of homeodomain proteins to DNA. (A) Schematic of the exon structure (exons 1–5) of human *PRRX1a* (245 aa) and *PRRX1b* (217 aa). Alternative splicing of exon 4 results in an OAR domain present in PRRX1a, which is absent in PRRX1b. (B) 3D structure of PRRX1a (P54821) obtained from AlphaFoldDB: https://alphafold.ebi.ac.uk/entry/P54821. Residues shown in dark blue indicate a very high per residue confidence score of the structure (>90), light blue shows good confidence (70–90), and yellow and orange predict low (50–70) and very low (<50) confidence in the proposed structure, respectively. The start and end of the homeodomain are marked by Q94 and E153, respectively. The three alpha helices within the homeodomain are shown in dark blue and numbered (1–3) and the OAR domain (in PRRX1a) is annotated. (C–D) Diagram of Msx-1 (a representative homeodomain protein) binding to DNA via helix 3 and the N-terminal arm of the homeodomain. Images re-used from Noyes et al., 2008. (D) The key residue positions within the homeodomain are noted (2, 5, 47, 50, 51, 54).

Influenced by this early mouse work, *PRRX1* (located at human chromosome 1q24.2), was tested as a candidate disease gene in infants with agnathia-otocephaly (OMIM #202650), a severe craniofacial malformation resulting in reduction or absence of the mandible, alongside microstomia, hypo- or a-glossia and ventromedial auricular malposition or fusion (Faye-Petersen et al., 2006). Following the identification of a heterozygous variant in *PRRX1* (NM_022716.4:c.337T>C; p.(Phe113Leu) [incorrectly documented as p.(Phe113Ser) in the original report]) (Sergi and Kamnasaran, 2011), three other reports of homozygous (NM_022716.4:c.691G>C; p.(Ala231Pro)) (Celik et al., 2012) or heterozygous variants have been published, the latter comprising two *de novo* frameshift variants: NM_022716.4:c.266_269dup; p.(Arg92Glufs*8), incorrectly reported as p.(Arg92Glufs*98), and NM_022716.4:c.269del; p.(Lys90Argfs*42), incorrectly reported p.(Lys90Argfs*131) (Dasouki et al., 2013, Donnelly et al., 2012). Since then, no further patients with agnathia-otocephaly and variants in *PRRX1* have been reported.

Alongside variants in *PRRX1*, genetic testing of patients with agnathia-otocephaly have identified heterozygous variants in the homeobox gene *OTX2*, including a frameshifting variant (NM_001270525.2:c.316delC; p.(Gln106Asnfs*11)) in a large family consisting of 17 individuals with micro/anophthalmia (Chassaing et al., 2012). In this family, two affected cousins with the *OTX2* variant gave birth to children with agnathia-otocephaly but owing to the severity of the condition DNA was unavailable from both otocephalic children (Chassaing et al., 2012). Of note, several members of the family displaying micro/anophthalmia were negative for mutations in *OTX2*. However, the authors reported a second family with a child displaying agnathia-otocephaly who was found to harbour a *de novo* frameshifting variant in *OTX2* (NM_001270525.2:c.130delC; p.(Arg44Glyfs*15)) (Chassaing et al., 2012). Both reported frameshifting variants were predicted to abrogate the *OTX2* homeodomain. An additional loss-of-function variant affecting the homeodomain (Jones et al., 2017), a whole gene deletion (Patat et al., 2013), and an

intragenic duplication have since been described (Sergouniotis et al., 2015) in association with agnathia-otocephaly. In support, heterozygous (*Otx2*$^{+/-}$) mouse mutants show craniofacial malformations consistent with otocephaly (Matsuo et al., 1995), although the severity of the phenotype is dependent on genetic background (Hide et al., 2002). Given this strain-dependent variability, and evidence of non-segregation of the *OTX2* variant in the family with micro/anophthalmia, additional genetic factors have been hypothesised as modifiers of the phenotype. Molecular testing in humans and linkage analysis in mice have highlighted potential modifier-genes, including *PRRX1*, alongside *ALX4*, *MSX1*, *PGAP1* and *TWSG1* (Hide et al., 2002, Chassaing et al., 2012). However, variants in these genes are yet to be identified as modifiers in the aetiology of agnathia-otocephaly.

More recently, an unpublished WGS study from my supervisors' laboratory using Complete Genomics technology, performed prior to the period of my DPhil research, had revealed a *de novo* mutation in *PRRX1* (NM_022716.4:c.52delC; p.(Arg18Alafs*23)) in a patient with craniosynostosis, a much milder craniofacial abnormality. In the same WGS analysis, which included nine families with sporadically affected children with bi-coronal or multi-suture synostosis, an additional individual harbouring a heterozygous missense variant within the *PRRX1* homeodomain (NM_022716.4:c.449G>A; p.(Arg150His)) was identified (Complete Genomics Trios, table 6.1). This patient was included in a previous whole exome analysis (Family 22) (Miller et al., 2017), but the missense variant had not been flagged as it was inherited from an unaffected father. At that time, additional exome and targeted re-sequencing of *PRRX1a* (Miller and Timberlake Resequencing and WES cohort, table 6.1) had highlighted a further six inherited heterozygous variants in patients with craniosynostosis, as well as unaffected relatives: c.161A>C; p.(Asp54Ala), c.283C>T; p.(Arg95*), 2 x c.285G>T; p.(Arg96Met), c.371G>A; p.(Arg124Gln), and c.439G>A; p.(Ala147Thr). Through collaboration, one further missense variant within the homeodomain was identified in a family with multiple

affected individuals (c.432C>G; p.(Asn144Lys)). In total, 8 variants in 9 independent families with individuals with craniosynostosis were identified. This suggested that variants in *PRRX1* were enriched within the craniosynostosis cohort, and that non-penetrance was frequent. Given previous evidence in mice that loss-of-function variants in *PRRX1* result in craniofacial abnormalities (Martin et al., 1995, ten Berge et al., 1998) and that *PRRX1* is expressed in the major cranial sutures, identifying a subset of cells with characteristics of osteoprogenitors (Takarada et al., 2016, Wilk et al., 2017), it was deemed that the loss-of-function mutations were likely damaging. However, the pathogenic mechanism for each missense variant was less clear, particularly given that a heterozygous missense variant has been previously reported in association with agnathia-otocephaly (p.(Phe113Leu)) (Sergi and Kamnasaran, 2011). Therefore, I aimed to establish the functional consequence of missense variants in the PRRX1 homeodomain and identify further patients with variants in *PRRX1* and craniosynostosis.

6.2 Methods

6.2.1 Targeted sequencing of *PRRX1*

Prior to the period of my DPhil research, PCR had been used to screen 388 unsolved samples from patients with craniosynostosis for variants in *PRRX1*. I screened an additional 541 samples (Tooze Cohort, table 6.1) for variants in *PRRX1* using IDT's hybridisation and capture protocol (details within chapter 4, section 4.2). Probes were designed to ensure that all coding regions of *PRRX1* were captured by at least two probes (table A6.1). Sequencing data were analysed using amplimap software (Koelling et al., 2019) (including mapping, coverage analysis, and variant calling) and variants were filtered on the basis of rarity (gnomAD v2.1.1 (Karczewski et al., 2020); allele frequency below 0.000045) (Calpena et al., 2020, Whiffin et al., 2017), CADD score (≥ 20, or not reported), and likely consequence (missense or worse). The coverage of

PRRX1 was assessed in 479 samples from patients with craniosynostosis (as detailed in chapter 4, section 4.2.5). The normalised coverage for 20 non-overlapping probe regions (covering all exons of *PRRX1*) were analysed for deletions or duplications, considering two or more consecutive probes displaying a normalised coverage below 0.6 (heterozygous deletion) or above 1.4 (heterozygous duplication).

6.2.2 Isolation of breakpoints in two patients with deletions in *PRRX1*

The breakpoints were narrowed down starting with available information from array and exome depth analysis, by a combination of qPCR and long amplification PCR (available array/exome information in table A6.2, primers are detailed in table A6.3). For long amplification (up to 15 kb), a 25 ng sample of gDNA was mixed with 10 μl of 2x LongAmp hot start master mix (NEB, M0533S), 0.5 μl of each forward and reverse primer (10 μM) (table A6.3) and up to 20 μl total volume of nuclease-free water. The sample was amplified in a thermocycler at 95°C for 2 min, followed by 35 cycles of 95°C for 30 s, 58°C for 45 s, and 65°C for 16 min and a final extension of 65°C for 10 min. Once a unique product had been identified by gel electrophoresis in the affected individuals, the region was narrowed down to a product of <1 kb following a Q5 enzyme protocol. A 25 ng sample of gDNA was mixed with 12.5 μl of 2x Q5 high-fidelity master mix (NEB, M0492S), 1.25 μl of each forward and reverse primer (table A6.3) (10 μM) and up to 25 μl total volume of nuclease-free water. The sample was amplified in a thermocycler at 98°C for 30 s, followed by 35 cycles of 95°C for 10 s, 66°C for 30 s, and 72°C for 30 s and a final extension of 72°C for 2 min. A control primer pair amplifying a region of *SYNJ2* (SYNJ2_FW: 5'-CCACTGTGTTGAGCTGATGA-3', SYNJ2_RV: 5'-CAACGGGAAATGCTGCAAAG-3') was used to confirm the presence of a band in a healthy control sample. The PCR product corresponding to the *PRRX1*-deletion was analysed by dideoxy-sequencing to establish the precise breakpoints.

Table 6.1 Phenotypic composition of cohorts of unsolved CRS probands

	Complete Genomics Trios[b]		**Miller and Timberlake Resequencing and WES Cohort**		**Tooze Cohort**		**Totals**		
Phenotype[a]	*PRRX1*-negative	*PRRX1*-positive	*PRRX1*-negative	*PRRX1*-Positive	*PRRX1*-negative	*PRRX1*-Positive	PRRX1-negative	PRRX1-positive	Percent (%) positive
NS metopic			253		68		321		
NS unicoronal			19	1	159	1	178	2	1.11
NS sagittal			531	1	79		610	1	0.16
NS unilambdoid			19	1	9		28	1	3.45
NS multiple			17	2	50	1	67	3	4.29
NS bicoronal	1	1	3		28		32	1	3.03
NS bilambdoid			2		2		4		
S metopic			16	1	29		45	1	2.17
S unicoronal			4		20		24		
S sagittal	1		15		41		57		
S unilambdoid			1		2		3		
S multiple	1	1	9		35		45	1	2.17
S bicoronal	1		2		7	1	10	1	9.09
S bilambdoid					1		1		
Carpenter			1				1		
Crouzon			4				4		
Pfeiffer			3				3		
Saethre Chotzen			2		8		10		
Other			1				1		
Total	**4**	**2**	**902**	**6**	**538**	**3**	**1444**	**11**	**0.76%**

[a]NS = non-syndromic, S = syndromic

[b]9 cases were originally sequenced but 3 have since been solved and removed from the total numbers in this table

6.2.3 Plasmid construction

6.2.3.1 *PRRX1a* missense constructs

Two vectors (pcDNA3.1) containing either *Prrx1a* or *Prrx1aΔOAR* were originally obtained as a gift from Michael Kern (Norris et al., 2000). Dr Kerry Miller derived missense variants p.(Asp54Ala), p.(Arg96Met), and p.(Phe113Leu) in the *Prrx1a* construct; Associate Professor Stephen Twigg derived the p.(Arg150His) construct. Using site-directed mutagenesis (SDM), I created the p.(Arg124Gln) and p.(Asn144Lys) constructs in the *Prrx1a* vector. Initially, 40 ng of plasmid DNA was added to 1.5 µl of each 10 µM primer (table A6.4), 1 µl of 10 mM of each dNTP, 10 µl of Q5 buffer and enhancer, 1 µl of Q5 enzyme and up to 50 µl of nuclease-free water. Reactions were placed in a thermal cycler (BIORAD T100^{TM}) at 95^{o}C for 2 min, followed by 95^{o}C for 30 s, 55^{o}C for 1 min and 72^{o}C for 10 min, for 12 cycles. Amplicons were digested with DpnI (NEB, #R0176) and CutSmart Buffer for 4 hours at 37^{o}C. Digested products were transformed into competent *E. coli* cells (NEB 5-alpha, C2987) using the manufacturers 'High Efficiency Transformation Protocol'. Plasmid DNA was extracted using a miniprep kit (QIAprep Spin 27106, QIAGEN) and sent for dideoxy-sequencing (chapter 2, section 2.2).

Subsequently, *Prrx1* was amplified with primers containing EcoRI and HindIII (table A6.5; figure A6.1a) restriction sites at the 5' and 3' ends, respectively, before subcloning into a hemagglutinin (HA) and FLAG-tagged vector obtained from Addgene (FLAG-HA-pcDNA3.1, #52535). *Prrx1* cDNA was amplified by a Q5 protocol (1 µg of plasmid DNA [*Prrx1a*], 12.5 µl Q5 hot start 2x master mix, 10 µM of each forward and reverse primer [table A6.5] and up to 25 µl nuclease-free water). Products were amplified in a thermocycler at 98^{o}C for 30 s, followed by 35 cycles of 98^{o}C for 10 s, 70^{o}C for 20 s and 72^{o}C for 30 s, with a final extension of 72^{o}C for 2 min. The HA and FLAG tagged plasmid was digested, alongside the amplified PCR products (1 µg of plasmid or PCR product),

using 1 μl of each HindIII-HF and EcoRI-HF, and 1x CutSmart Buffer. The products were incubated with enzyme at 37°C for 15 min before adding 1 μl of alkaline phosphatase for a further 15 min. Digested products were run on a 1% agarose gel and gel purified (NEB, T1020 gel extraction kit). Briefly, purified *Prrx1a* inserts were ligated into the cut FLAG-HA-pcDNA3.1 plasmid using a quick ligation protocol (NEB, M2200) and a 4:1 ratio of insert DNA to vector (μl). Ligated products were transformed into chemically competent *E. coli* cells following the manufacturer's protocol. Colonies were screened by colony PCR and successfully cloned colonies were cultured in LB, supplemented with ampicillin (50 μg/ml). The plasmids were purified using a maxiprep kit (Qiagen, 12162) before sequence confirmation by dideoxy sequencing (figure A6.1b).

The p.(Ser104Gly) and p.(Arg115Trp) variants were constructed directly in the HA/FLAG-tagged vector using the Q5 SDM kit (NEB, E0554S): 12.5 μl Q5 hot start high-fidelity 2x master mix, 0.5 μM of each forward and reverse primer (table A6.4), 25 ng of template DNA and up to 25 μl reaction of nuclease-free water. The reactions were placed in a thermocycler under the following conditions: 98°C for 30 s, followed by 25 cycles of 98°C for 10 s, 65/70°C (for p.(Ser104Gly) and p.(Arg115Trp), respectively) for 20 s and 72°C for 5 min; a final extension of 72°C for 2 min was performed. Following SDM, 1 μl of product was incubated with 5 μl of kinase, ligase, and DpnI (KLD) 2x reaction buffer, 1 μl of 10x KLD enzyme mix (NEB, M0554), and 3 μl of nuclease-free water for 5 min at RT. Finally, 5 μl of the KLD reaction was transformed into 50 μl of chemically competent cells following the manufacturer's guidelines; sequences were confirmed by dideoxy-sequencing (figure A6.1b).

6.2.3.2 pGL2-Luciferase-TN-C construct

Human and mouse gDNA were used as templates to amplify the tenascin-c (TN-C) promoter sequence. A 100 ng sample was mixed with 5 μl of 5x Q5 reaction buffer, 5 μl

of Q5 enhancer, 1 μl of each 10 mM dNTP, 1 μl of each 10 μM forward and reverse primer (table A6.4), 0.5 μl of Q5 high fidelity polymerase, 0.25 μl of Taq polymerase and up to 25 μl of nuclease-free water. Each product was amplified in a thermocycler under the following conditions: 98°C for 3 min, followed by 35 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 30 s and a final extension of 72°C for 5 min and 4°C hold. A 4 μl sample of amplified product was subcloned into 0.5 μl of pGEM®-T easy vector (Promega, A1360) using 5 μl of 10x T4 ligase buffer (Promega, M1801) and 1 μl T4 DNA ligase. The ligation reaction was left overnight at RT. The following day, the ligated product was transformed into chemically competent *E. coli* cells following manufacturer's guidelines. Restriction digest using EcoRI (15 min at 37°C) was used to confirm presence of the insert and positive clones were confirmed by dideoxy-sequencing. Following confirmation of the correct TN-C sequence (human or mouse), the DNA was cut out of the pGEM vector using XmaI and HindIII enzymes and gel purified (NEB Monarch DNA extraction kit). The pGL2-eNOSPromoter-Luciferase plasmid was obtained from Addgene (Plasmid #19297) and digested with XmaI and HindIII for 15 mins at 37°C, followed by 15 min incubation with 1 μl alkaline phosphatase. Digested and purified TN-C insert, and pGL2-Luciferase, were ligated following previous methods and transformed into chemically competent *E. coli* cells. The pGL2-Luciferase-TN-C constructs (mouse and human derived TN-C) were purified using a miniprep kit (Qiagen, 27104).

6.2.4 Dual luciferase reporter assay

Cells (C2C12/HEK293/COS7/HT1080) were seeded to reach 70% confluency at the point of transfection. Plasmids (FLAG-HA-pcDNA3.1-*Prrx1a* constructs, Renilla-luciferase driven by a TK promoter [RL-TK], and a plasmid containing the TN-C promoter upstream of luciferase [pGL2-Luciferase-TN-C]) were co-transfected to a final concentration of 1 μg, alongside 3 μl of Lipofectamine-2000 transfection reagent. After 24 hours, luciferase activity was measured (Promega GloMax Luminometer) using a

Dual Luciferase Reporter Assay® (Promega, E1910), following the manufacturer's instructions. The cell lysate used for input to the dual luciferase assay was then used as input for a western blot to assess transfection efficiency.

6.2.5 Western Blot

After 24 hours of transfection, the cell lysate was harvested by centrifugation (300 *g* for 5 mins) and a protocol for analysis of protein by western blot was followed as detailed in chapter 2 (section 2.7). A 1/1,000 dilution of anti-HA (rabbit mAb; Cell Signalling Technology, #C29F4) or anti-FLAG (mouse mAb [clone M2]; Sigma-Aldrich, #F1804) was used to detect each *Prrx1* construct, and a 1/10,000 dilution of anti-GAPDH-HRP (Cell Signalling Technology, #14C10) was used a control. Both primary antibodies (HA/FLAG) were left at room temperature for an hour before staining with secondary IgG antibodies conjugated to horse radish peroxidase (HRP): donkey polyclonal antibody (pAb) against rabbit IgG (secondary stain for HA; Abcam, ab97085), or donkey pAb against mouse IgG (secondary stain for FLAG; Abcam, ab98799), diluted 1/10,000.

6.2.6 Immunofluorescence

COS7 cells were seeded onto coverslips to reach 60% confluency after 24 hr in a 6-well plate, before transfection with 0.5 µg of plasmid and 3 µl of Lipofectamine-2000 (ThermoFisher Scientific). After 18 hr, cells were fixed using paraformaldehyde (4%, pH 7.4) and permeabilised using 0.2% Triton X-100, before blocking with 2% bovine serum albumin (BSA). Anti-HA (HA-Tag [C29F4], rabbit mAb; Cell Signaling, #3724) was diluted 1:100 in SignalBoost Immunoreaction Enhancer Solution 1 (Calbiochem, KP31812) and added to coverslips to incubate for 1 hour at RT. Coverslips were washed with 1x PBS before adding secondary antibody (Alexa Fluor 488 donkey anti-rabbit IgG [Life Technologies, #A-21206]), diluted 1:500 in SignalBoost Immunoreaction Enhancer Solution 2 (Calbiochem, KP31855). Cells were incubated with secondary antibody for 45

min at RT and counter-stained with Alexa Fluor 647 Phalloidin diluted 1:20 in 1x PBS (Cell Signaling, #8940) for 15 min, followed by staining with 4',6-diamidino-2-phenylindole (DAPI, 2 mg/ml) diluted 1:2,000 in 1x PBS for 10 min. Coverslips were washed (1x PBS) and mounted to slides using VECTASHIELD Antifade Mounting Medium H-1000 (Vector Laboratories), and imaged using Zeiss 880 Inverted Confocal and ScanR microscopes. The ScanR microscope imaged 25 random sections of each slide (in technical triplicates) and cells were counted for distribution of protein within the nucleus or cytoplasm, blinded to the transfected construct used. A minimum of 698 cells were counted for each variant.

6.3 Results

6.3.1 Identification of rare heterozygous single nucleotide or indel variants of *PRRX1* in craniosynostosis

6.3.1.1 Resequencing analysis

Prior analysis of WGS, exome, and targeted resequencing data revealed 6 variants (from 7 families) predicting loss-of-function or substitution of the homeodomain (table 6.2; figures 6.2–3). Additionally, a missense variant at a highly conserved residue N-terminal to the homeodomain was also identified (c.161A>C; p.(Asp54Ala), Family 3) (table 6.2). Based on these preliminary findings I screened a further cohort of patients with undiagnosed craniosynostosis (*n* = 541), considering a variety of sutures (Tooze cohort table 6.1). This revealed 1 further unrelated subject with a heterozygous loss of the start codon (NM_022716.4:c.2T>G; p.(?), Family 1) (table 6.2; figure 6.2–3). To provide a comparator I examined data from the gnomAD database (v2.1.1) and UK Biobank. In ~250,000 alleles in gnomAD, there are four loss-of-function alleles and 34 missense substitutions (excluding a known benign substitution, p.(Ser104Gly)) that potentially disrupt the homeodomain. In ~735,000 alleles in the UK Biobank, there are three loss-

of-function variants, and 110 missense variants (excluding p.(Ser104Gly)) predicted to disrupt the homeodomain. This indicates a significant over-representation (16-fold; Fisher exact test: $p < 0.000001$) in the resequencing studies (two loss-of-function and five homeodomain missense variants in 2898 alleles; table 6.2) compared to gnomAD and the UK Biobank, consistent with a causal contribution in the affected individuals.

6.3.1.2 Coverage analysis

In a subset of 479 individuals in the resequencing study who were analysed using a targeted capture approach, I developed a robust method to screen for whole exon or gene deletions in *PRRX1* (chapter 4, section 4.2.5). Initially, analysis of whole gene coverage revealed a suspected whole gene deletion in one individual with syndromic bi-coronal synostosis (Family 15). Additional investigation of the normalised coverage of non-overlapping probe regions showed a 50% reduction of probes covering all exons, consistent with a heterozygous whole gene deletion (figure 6.4b). Results from a microarray completed in 2013 identified a 10.5 Mb deletion, including *PRRX1* (table A6.2). The large deletion likely explains the additional learning disability documented in this patient. Notably, the craniosynostosis could be a feature of the deletion of any gene within this region; however, given evidence gathered in this project it is likely that the craniosynostosis is owing to haploinsufficiency of *PRRX1*. Further scrutiny of non-overlapping probe coverages showed a 50% decrease in values corresponding to probes spanning exons 1–2 (figure 6.4a) in a second individual with non-syndromic uni-coronal synostosis (Family 14). Results from a genome-wide array supported these findings and were used to narrow down the breakpoints of the deletion (figure 6.5; right hand panels). The breakpoint analysis revealed that the variant segregated in the unaffected mother.

6.3.1.3 Additional cases through collaboration

Prior to my work on *PRRX1*, one patient was identified through collaboration to harbour a missense variant within the PRRX1 homeodomain: c.432C>G; p.(Asn144Lys) (see section 6.1). However, during the course of the project a further three families (5 affected individuals) were identified by collaboration with the community of craniofacial geneticists. These comprised one frameshift which segregated in two affected cousins (NM_022716.4:c.291delT; p.(Asn97Lysfs*35), Family 7), a substitution of the homeodomain in two affected cousins (NM_022716.4:c.343C>T; p.(Arg115Trp), Family 8), and a partial gene deletion (exons 2–5, Family 13) (table 6.2; figure 6.3a). The breakpoints of the deletion were narrowed down using information gained from an array (table A6.2). The PCR product corresponding to the deletion in exons 2–5 was confirmed by dideoxy sequencing (figure 6.5; left hand panels). Given the additional cases identified through my resequencing screen and those presented from collaborators, the number of variants identified in PRRX1 in families with craniosynostosis increased from 9 to 15 (as summarised in column 1 of table 6.2), providing a substantial uplift to the evidence base.

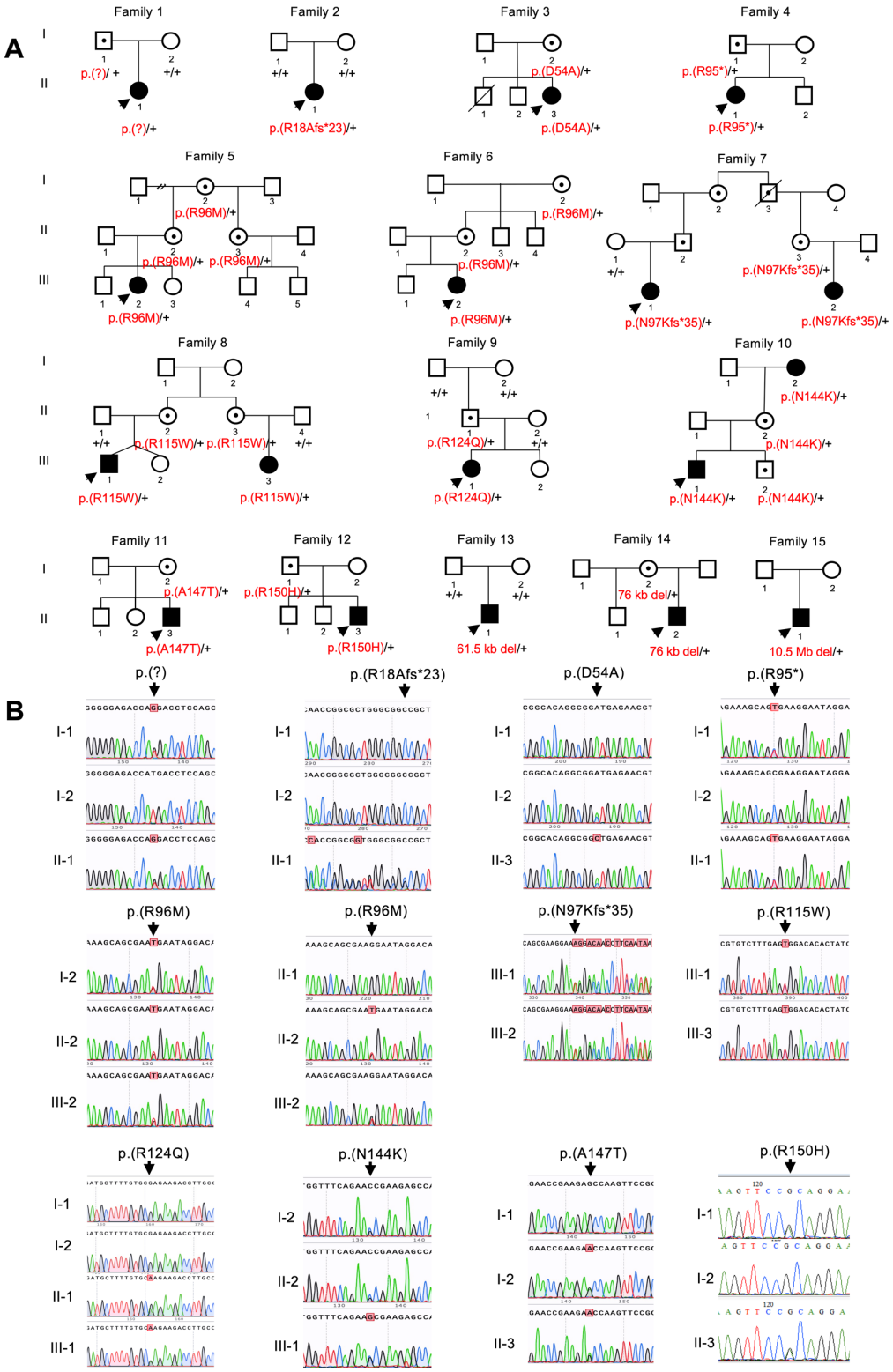


Figure 6.2 Pedigree figures and dideoxy sequencing confirmation of each variant. (A) Pedigrees of all families in which a small nucleotide variant or deletion was identified in *PRRX1*. The index patient is denoted by the arrowhead. If available, inheritance of the variant from family members have been confirmed by sequencing and documented on the pedigree (red lettering). A dot in the middle of the pedigree symbol indicates a heterozygous carrier. (B) Confirmation of small nucleotide variants (by dideoxy-sequencing); sanger traces show a representative result from forward or reverse sequencing and arrows indicate the position of each variant.

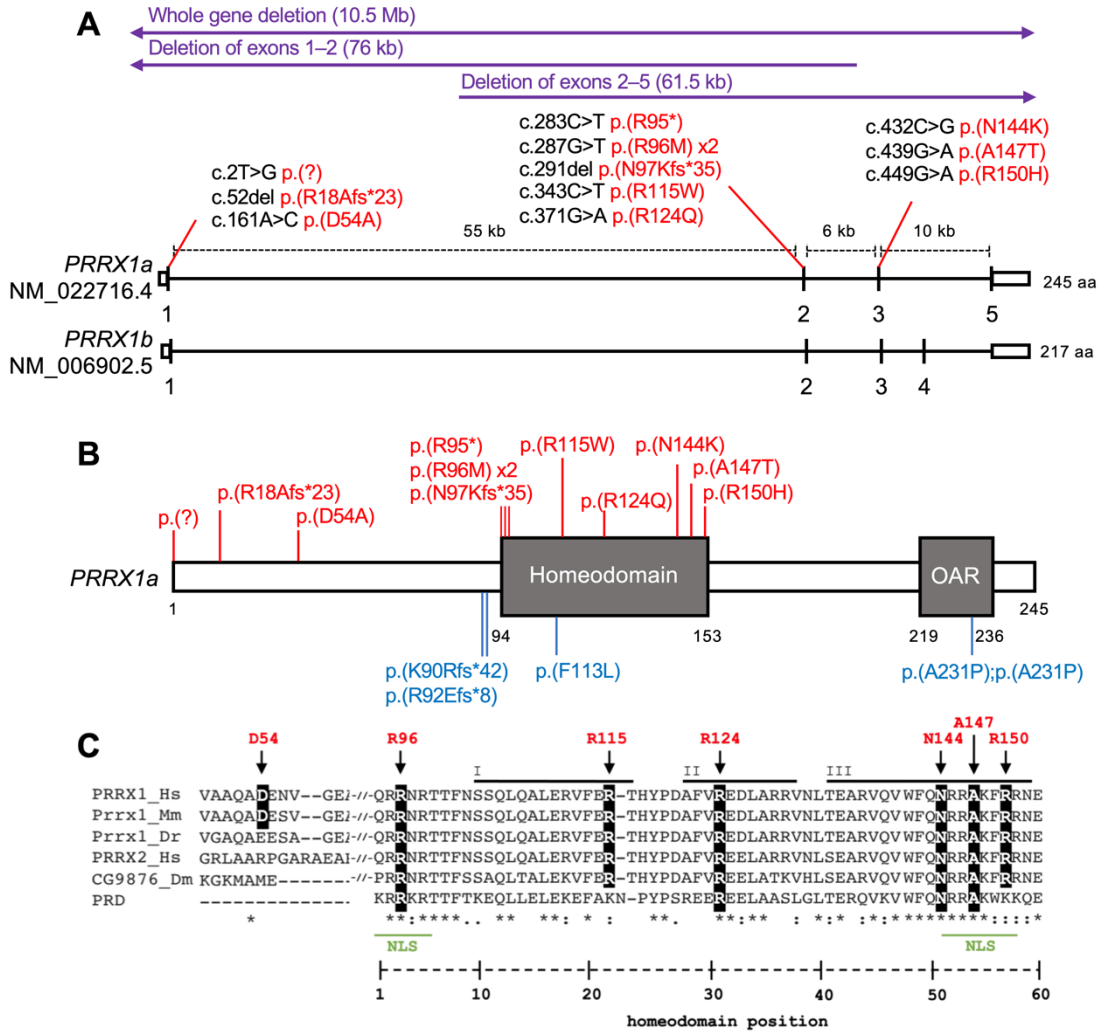


Figure 6.3 Structure, conservation, and variants in *PRRX1*. (A) Variants identified in patients with craniosynostosis (combining work undertaken during the course of this thesis with previous observations) are highlighted in red with the purple lines above indicating the three deletions identified in independent patients. Arrows indicate that the deletion extends beyond the *PRRX1* gene. (B) A schematic representation of the PRRX1a protein (P54821) showing the position of the homeodomain and the OAR domain. Variants identified in craniosynostosis are highlighted in red, variants reported in patients with agnathia-otocephaly are shown in blue. The only homozygous variant reported is p.(Ala231Pro) identified in a patient with agnathia-otocephaly. (C) Conservation of the amino acids surrounding p.Asp54 (left) and the 60 amino acids of the homeodomain (right). Sequences correspond to human PRRX1 and PRRX2 (PRRX1_Hs and PRRX2_Hs, respectively), mouse (Prrx1_Mm), zebrafish (Prrx1_Dr), drosophila (CG9876_Dm), and the consensus sequence for the PRD class of homeodomain-containing proteins. The arrows indicate the position of the missense substitutions identified in this study. The three alpha helices of the homeodomain are also highlighted (I, II, III). Two predicted nuclear localisation sequences (NLS) are annotated in green. Asterisks (*) represent complete conservation across all alignments, a colon (:) represents alignment residues with similar biochemical properties, and the period (.) denotes conservation between groups of weakly similar properties.

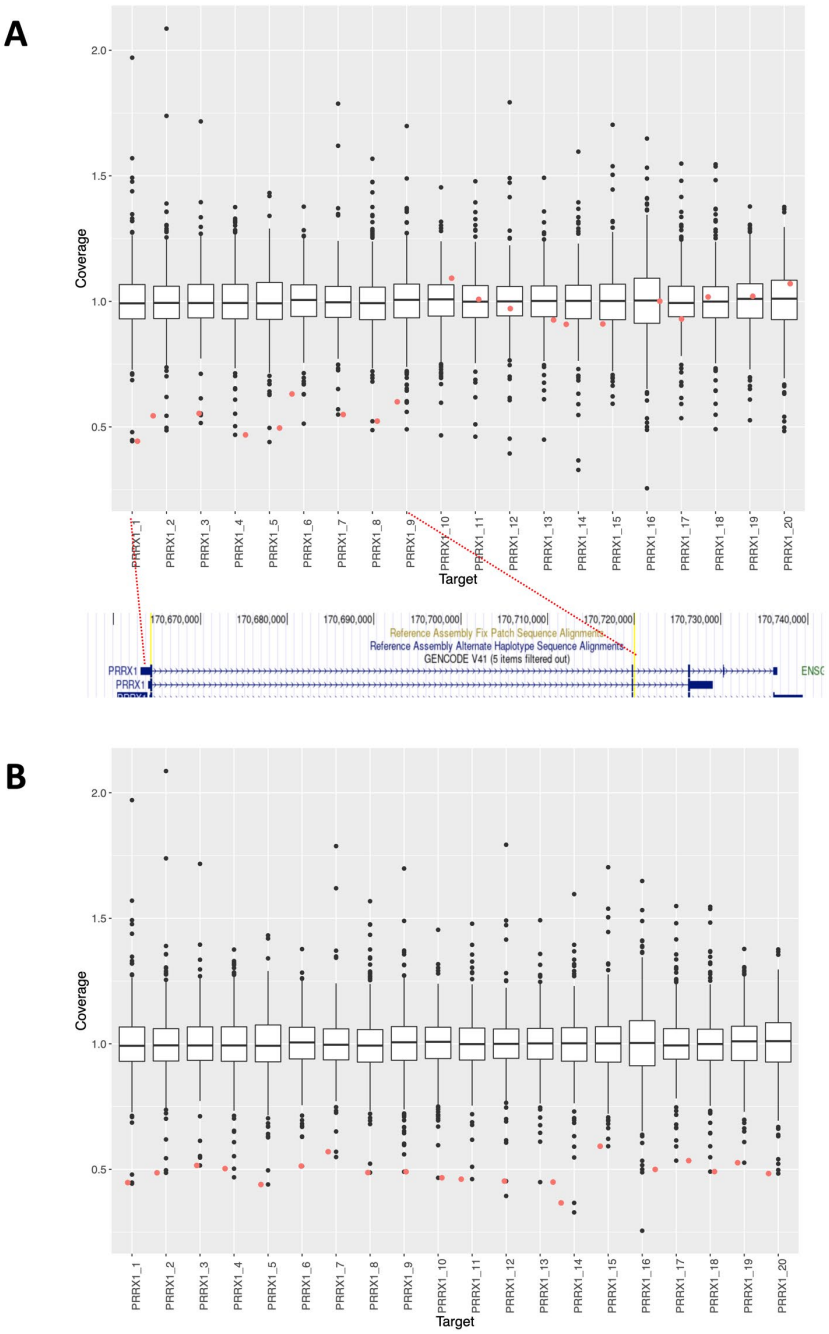


Figure 6.4 Analysis of coverage of *PRRX1* in a screen of 479 unsolved craniosynostosis patients. (A–B) Images show a boxplot of the normalised probe coverage for 20 non-overlapping regions in PRRX1. For each target, the horizontal centre line in the box marks the mean normalised coverage considering 479 samples, the top and bottom line of the box mark the upper (75th percentile; Q3) and lower (25th percentile; Q1) interquartile ranges (IQR). For each box plot, the vertical line extends to the maximum and minimum values (Q3 + 1.5*IQR and Q1 – 1.5*IQR, respectively). A normalised probe coverage value above or below the maximum and minimum value are annotated as a dot. (A) Partial deletion of *PRRX1* (sample II-2 from Family 14; red dots), resulting in reduced normalised coverage of probes 1–9 (below 0.6), corresponding to the deletion of exons 1–2 as denoted by the yellow lines in the below UCSC screenshot (Chr1: 170664249–170719963 [GRCh38]). (B) Whole gene deletion of *PRRX1* (sample II-1 from Family 15; red dots) shown by a drop in coverage (below 0.6) across all probes (spanning exons 1–5; Chr1: 170664249 – 170736267 [GRCh38]).

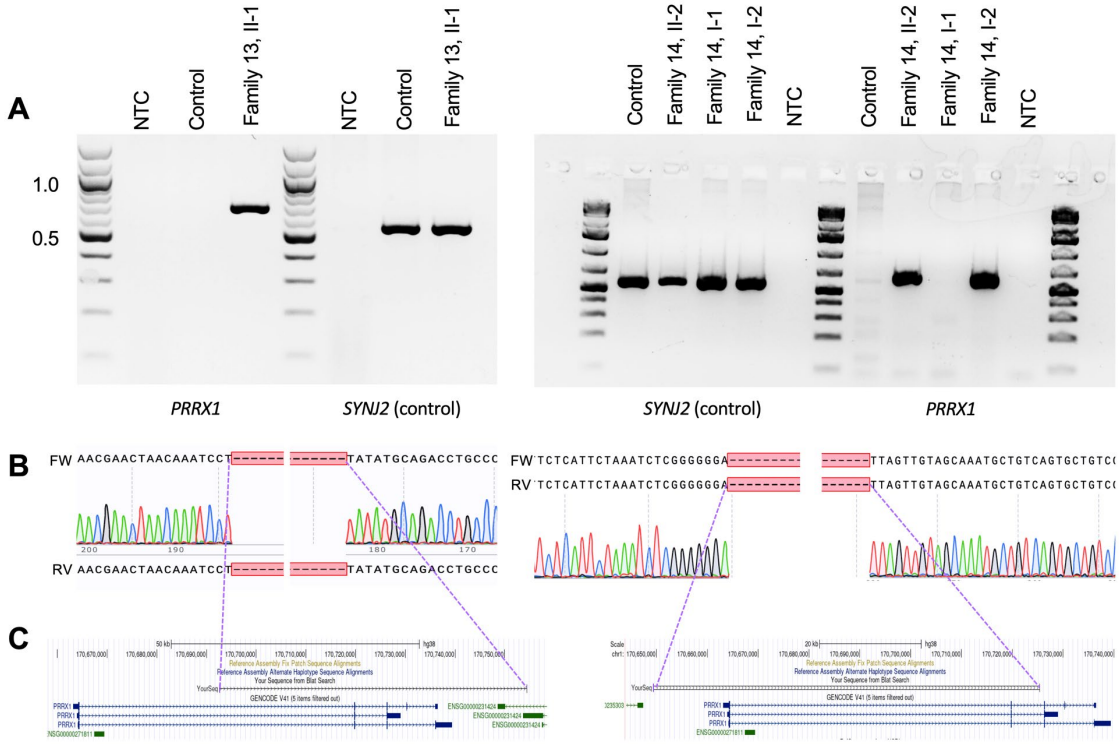


Figure 6.5 Analysis of the breakpoints of partial deletions including *PRRX1*. (A–B) Left, breakpoints in Family 13; right, breakpoints in Family 14. (A) Breakpoint PCR of the affected sample, available family members, and a healthy control showing presence of the amplified product using *PRRX1* primers in both families with the deletion, but no product in the healthy control. Primers amplifying a region of *SYNJ2* (552 bp) were used to ensure that the healthy sample could be amplified. NTC = no template control. (B) Forward (FW) and reverse (RV) dideoxy sequencing traces showing the region of the respective 3' and 5' breakpoints. (C) A screenshot from UCSC genome browser indicating the region of *PRRX1* where the two breakpoints align (GRCh38: left = Chr1:170692716_170754397del, right = Chr1:170649419_1707253330del).

Table 6.2 Rare heterozygous *PRRX1* variants identified in this and previous studies

Family #	Individuals with variants[a] (* = index patient)	Suture fusion[b]	Syndromic (s)/non-syndromic (ns)	Exon(s) #	cDNA change	Amino acid change	CADD score	gnomAD prevalence (v2.1.1)	*De novo*?	% nuclear localisation (mean±SD)
1†	I-1, **II-1***	S+LL	ns	1	c.2T>G	p.(?)	24.3	0		-
2	**II-1***	BC	ns	1	c.52delC	p.(Arg18Alafs*23)	-	0	Y	-
3	I-2, **II-3***	Me	s	1	c.161A>C	p.(Asp54Ala)	23.3	0		78.7±4.1
4	I-1, **II-1***	S+RC	ns	2	c.283C>T	p.(Arg95*)	36	0		-
5	I-2, II-2, II-3, **III-2***	S	ns	2	c.287G>T	p.(Arg96Met)	29.5	0		39.1±1.8
6	I-2, II-2, **III-2***	RC	ns	2	c.287G>T	p.(Arg96Met)	29.5	0		39.1±1.8
7‡	II-3, **III-1*, III-2**	BC; P	s; ns	2	c.291delT	p.(Asn97Lysfs*35)	-	0		-
8‡	II-2, II-3, **III-1*, III-3**	S; S	s; s	2	c.343C>T	p.(Arg115Trp)	23.4	1.06e-5		50.4±6.6
9	II-1, **III-1***	RL	ns	2	c.371G>A	p.(Arg124Gln)	31	0	Y (in II-1)	49.0±0.4
10	**I-2**, II-2, **III-1***, III-2	BC; BC	s; s	3	c.432C>G	p.(Asn144Lys)	25.2	0		49.5±3.7
11	I-2, **II-3***	S+BC	ns	3	c.439G>A	p.(Ala147Thr)	28.9	0		56.3±5.1
12	I-1, **II-3***	S+BC	s	3	c.449G>A	p.(Arg150His)	31	0		49.3±8.4
13‡	**II-1***	Mu	ns	2–5	-	61.5 kb deletion	-	0	Y	-
14†	I-2, **II-2***	LC	ns	1–2	-	76 kb deletion	-	0		-
15†	**II-1***	BC	s	1–5	-	10.5 Mb deletion	-	0		-

[a]Individuals with craniosynostosis indicated with bold lettering
[b]BC, bicoronal; LC, left coronal; Me, metopic; Mu, multi-suture synostosis; P, pansynostosis; RC, right coronal; RL, right lambdoid; S, sagittal.
† Affected individuals/families identified by resequencing of 541 samples (section 6.2.1).
‡ Additional affected individuals/families identified by collaboration during the current study.

6.3.2 A dual luciferase assay was unable to establish a pathogenic mechanism for any missense variant

Initial experiments using a green fluorescent protein (GFP) reporter plasmid as a proxy for the *Prrx1a* construct obtained from the Kern lab indicated that a 6 (μl):1 (μg) ratio of lipofectamine-2000 to DNA in a 6-well plate would yield the highest transfection efficiency but at a cost of reduced cell viability (figure A6.2). Transfection of a reduced quantity of lipofectamine (3 μl) resulted in a marginal decrease in transfection efficiency (-5.9%) but a notable improvement on cell morphology and viability. At this same dose of lipofectamine, increasing the DNA concentration to 2 μg resulted in a marginal improvement in transfection efficiency by 3.1%.

Given the minimal loss in transfection efficiency using 1 μg of DNA compared to 2 μg, further experiments were optimised with 3:1 ratio of lipofectamine to DNA. Initial experiments were conducted using the constructs provided by the Kern lab, alongside a construct containing the sequence for the mouse tenascin-c promoter (TN-C) (reported as -257 to +198) (Norris and Kern, 2001b). In an attempt to replicate their experimental findings, I used C2C12 cells and transfected the expression and reporter plasmids at 40% of the final DNA concentration, and the internal control (RL-TK) at 20% of the total DNA concentration (Norris and Kern, 2001a). Given previous experimental findings from this group that the OAR domain was a repressor of transactivation (Norris and Kern, 2001b), the *Prrx1ΔOAR* construct was included as a positive control. However, the relative luciferase activity (compared to RL-TK) was notably low in all constructs; moreover, the *Prrx1ΔOAR* construct showed a reduction in relative luciferase activity compared to the wildtype Prrx1 (figure 6.6a), potentially disputing Kern's original results. Two mutants were included in the experiment (p.(Arg150His) and p.(Ala147Thr)); the p.(Arg150His) mutant was selected as previous data investigating a homologous position in SHOX (p.(Arg173Cys) and p.(Arg173His)) (Sabherwal et al., 2004, Schneider et al.,

2005), a member of the PRD class of homeobox transcription factors, showed abnormal nuclear localisation of the protein. Schneider and colleagues also demonstrated that a position homologous to p.(Ala147Thr), p.(Arg170Pro) in SHOX, resulted in a reduced ability of the protein to form a homodimer and weak DNA-binding potential (Schneider et al., 2005). Although these mutants showed a reduction in relative luciferase activity compared to the wildtype protein, there was no difference in relative luciferase activity of the wildtype protein compared to an empty vector, suggesting that the detected signal was owing to background luciferase activity or failure of the experiment itself (figure 6.6a). To confirm that the dual luciferase system (reagents and luminometer) were able to detect an increase in activity and that my transfection conditions were working, I included a constitutively active BMPR1A (BMPR1A c.a.) control, as used in a recent paper by our group to assay the effect of *SMAD6* variants (Calpena et al., 2020). This showed a substantial increase (20x) in luciferase activity compared to the Prrx1 system (figure 6.6b).

To ascertain why the Prrx1 system had such a low luciferase reading, I sequenced the TN-C construct which was received from Michael Kern prior to my engagement in the project. While the insert sequence correctly aligned to the mouse TN-C promoter of multiple strains, I identified three errors in the sequence. First, a single nucleotide deletion (position -89) of an adenine in a poly-A repeat (figure A6.3a); previous data suggest that the dA.dT homopolymer is important for inducing DNA bending and excluding nucleosomes to increase accessibility to transcription machinery (Copertino et al., 1995). Second, a single nucleotide substitution (C>T; position +94); although this nucleotide is well conserved in other mouse strains, its conservation within vertebrates is less striking (figure A6.3b). Both errors were consistent across multiple mouse strains, suggesting a sequence error rather than strain difference. Third, the insert size was reported to be 455 bp (-257 to +198) (Norris and Kern, 2001b) while the construct contained only 444 bp. The missing 11 bp corresponded to a KROX-20 GC box, which

is reported to inhibit the constitutive proximal promoter in some cell types (Copertino et al., 1995). Of note, other similar assays have employed constructs without the KROX-20 sequence (for example, -247 to +121 of the mouse TN-C (Jones et al., 2001)). Given the inconsistences in the mouse TN-C reporter construct inherited from the Kern lab, I re-constructed the mouse TN-C luciferase reporter, ensuring that all reported regions were included (-287 to +230) (Norris and Kern, 2001b) and also generated a human (-527 to +197) TN-C construct to assess any differences, following previous literature (Dasouki et al., 2013). Both human and mouse constructs were made in a pGL2-luciferase vector obtained from Addgene.

The mouse and human TN-C constructs were tested in a dual luciferase assay using the *Prrx1*Δ*OAR*, wildtype, and p.(Ala147Thr) constructs, alongside the BMPR1A c.a. control. However, the results again showed an increased activity with the BMPR1A c.a. control and no difference with the wildtype construct compared to an un-transfected control (figure 6.7a). In Dasouki's paper (using the human TN-C construct), 18 μl of transfection reagent was used to transfect 6.15 μg of DNA (3:1 ratio; 3 μg of each expression and reporter plasmid, 0.15 μg RL-TK) into 1 x 10^4 cells. Previous optimisation experiments using a higher lipofectamine dose and DNA concentration had resulted in a marked reduction in cell viability; however, in an attempt to increase luciferase signal, I increased the amount of DNA to 4 μg. This resulted in a subtle increase in luciferase activity using the wildtype protein compared to an un-transfected control (figure 6.7b). Of note, in both experiments the wildtype protein demonstrated a higher luciferase activity compared to the ΔOAR construct (figure 6.7a,b). Additionally, the mouse TN-C construct consistently out-performed the human TN-C construct (figure 6.7a,b). In the original paper by Kern, the authors demonstrated that different cell types augmented luciferase activity (Norris and Kern, 2001b). As such, COS7 cells (used in Dasouki et al., 2013) were tested alongside C2C12; results showed a subtle difference between cell types (figure 6.7c) but

indicated elevated activity using the C2C12 cells. HT1080 cells were also tested as used in previous reports with human TN-C constructs (Yeo et al., 2018), however there were no clear differences in luciferase activity compared to other cell types (figure A6.4).

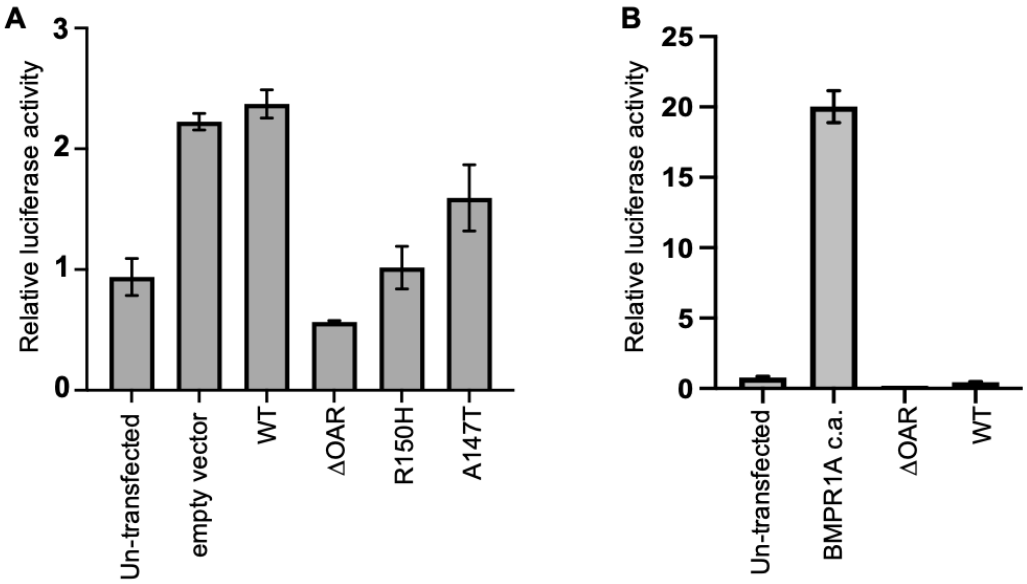


Figure 6.6 Dual luciferase assay using C2C12 cells transfected with wildtype and mutant *Prrx1* constructs and a TN-C luciferase reporter plasmid (A) An initial experiment using Michael Kern's *Prrx1* plasmids (wildtype [WT], ΔOAR deletion, and two mutants: p.(Arg150His) [R150H], p.(Ala147Thr) [A147T]) alongside transfection of an empty vector and an un-transfected control. Cells were transfected at a 3:1 ratio (400 ng of expression plasmid, 400 ng of TN-C reporter and 200 ng RL-TK). (B) Assay conducted with the addition of a positive control (BMPR1A c.a.). (A–B) Results are from triplicate technical analyses and error bars show ± standard deviation from the mean.

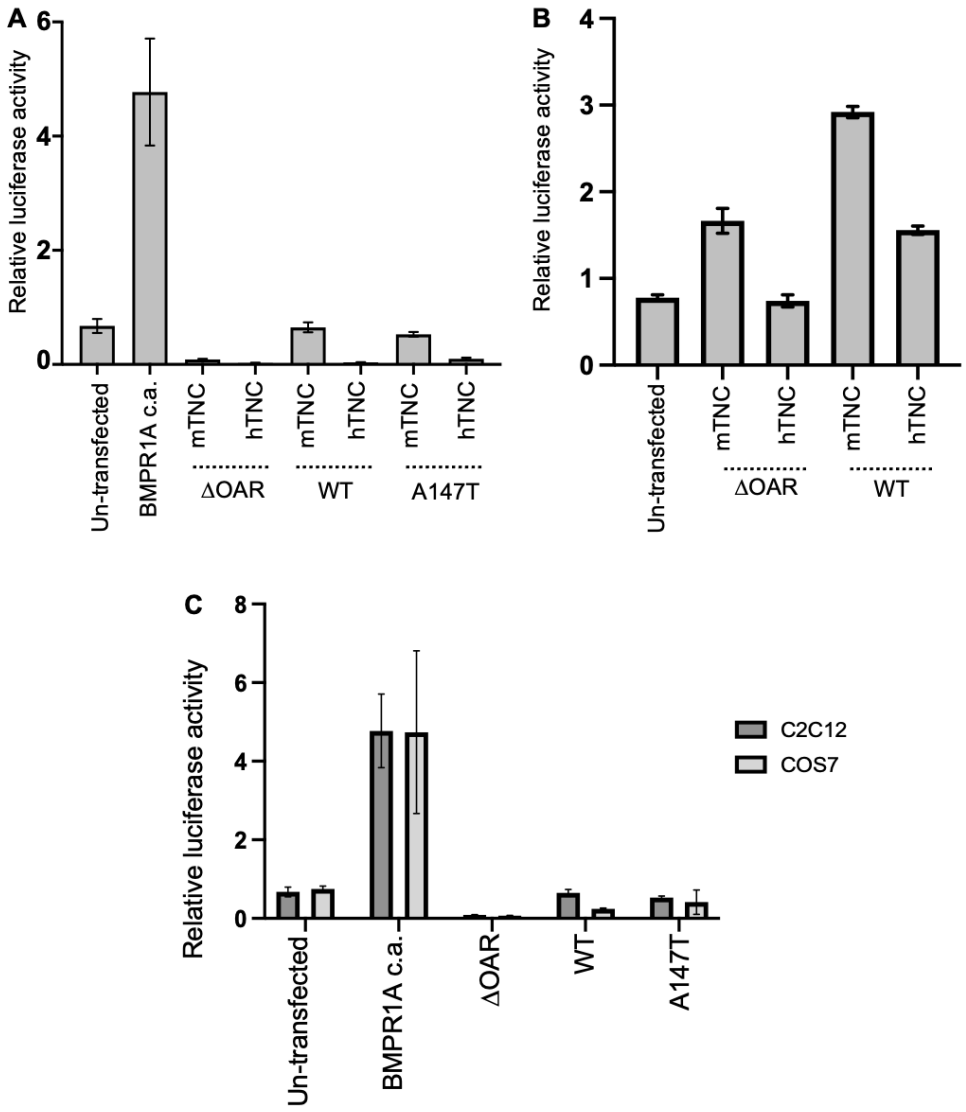


Figure 6.7 Optimisation of dual luciferase assays using new TN-C constructs (A) C2C12 cells were transfected with *Prrx1* constructs (ΔOAR deletion, wildtype [WT], and one mutant (p.(Ala147Thr) [A147T]) and new mouse (mTNC) or human (hTNC) TN-C constructs at a 3:1 ratio of transfection reagent to DNA (400 ng of each expression and reporter plasmid and 200 ng of RL-TK). (B) Experiment conducted as in panel A but using 4 µg of DNA (2 µg of each expression and reporter plasmid and 0.2 µg of RL-TK) and 8 µl of lipofectamine (2:1 ratio). (C) Experiment conducted as in panel B but using two cell types (C2C12 and COS7) and the mouse TNC reporter. (A–C) Results are from triplicate technical analyses and error bars show ± standard deviation from the mean.

To investigate if any of the missense substitutions in PRRX1 would cause a significantly different luciferase signal to the wildtype protein, all mutagenised constructs were tested for relative luciferase activity (figure 6.8a). The mutants (including ΔOAR) displayed an attenuated luciferase activity compared to the wildtype but the wildtype protein was no different to the empty vector. As such, it was difficult to establish if the signal was owing to any real interaction, or background luciferase activity. To test this, the mouse TN-C construct was transfected alone (figure 6.8b) (i.e., without the transcriptional activator, Prrx1) which showed elevated luciferase activity, thus indicating that the signal was predominately owing to the luciferase construct alone.

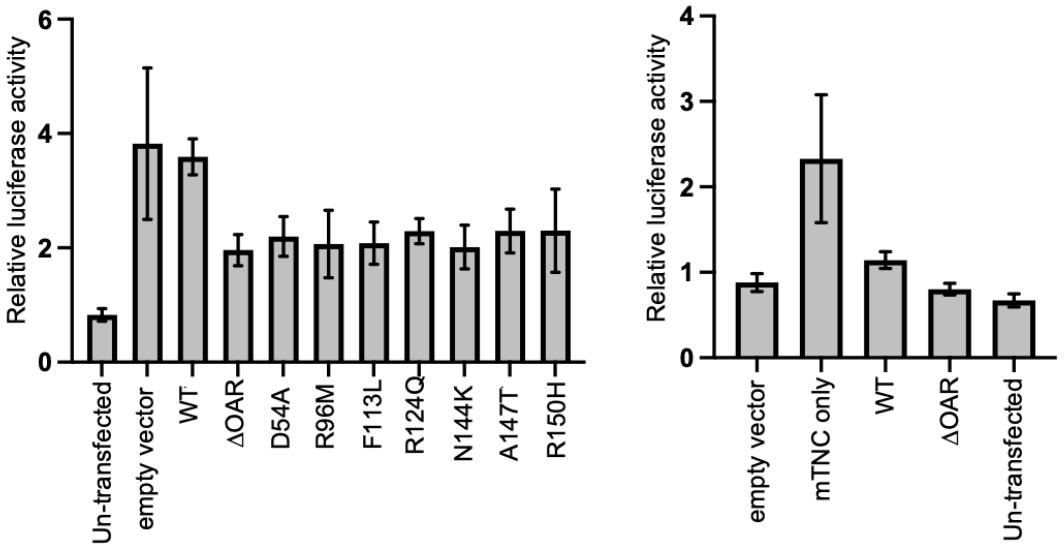


Figure 6.8 Relative luciferase activity of Prrx1 constructs. (A) C2C12 cells were transfected with *Prrx1* constructs, an empty pcDNA3.1 vector or no plasmid, considering a total DNA concentration of 4 µg. (B) Experiments were conducted as in panel A, but included the transfection of mouse TN-C and the RL-TK vector, but no *Prrx1* expression plasmid. (A–B) Results are from triplicate technical analyses and error bars show ± standard deviation from the mean.

Next, it was important to establish that the *Prrx1* constructs were successfully being transfected into the cells. However, the construct (pcDNA3.1) did not contain a suitable tag. Although it was possible to detect the construct using an antibody against Prrx1, using this strategy would have meant I was unable to distinguish endogenous versus overexpressed protein. Given this, all *Prrx1* constructs were sub-cloned into a pcDNA3.1 plasmid which contained an N-terminal HA and FLAG tag. Western blot analyses (figure 6.9a) confirmed that the plasmids were transfected into the cell, but the subsequent dual luciferase assay showed no difference compared to prior experiments (figure 6.9b).

Attempts to contact Michael Kern to trouble-shoot the protocol proved unsuccessful. Nevertheless, having confirmed that (a) the luciferase assay protocol was working by using the BMPR1A c.a. control, (b) the plasmids were being transfected into the cell and being expressed using a western blot analysis, and (c) that the TN-C luciferase construct contained no sequencing errors, it seemed unlikely that I would be able use a luciferase assay to investigate the effect of missense variants within the PRRX1 homeodomain.

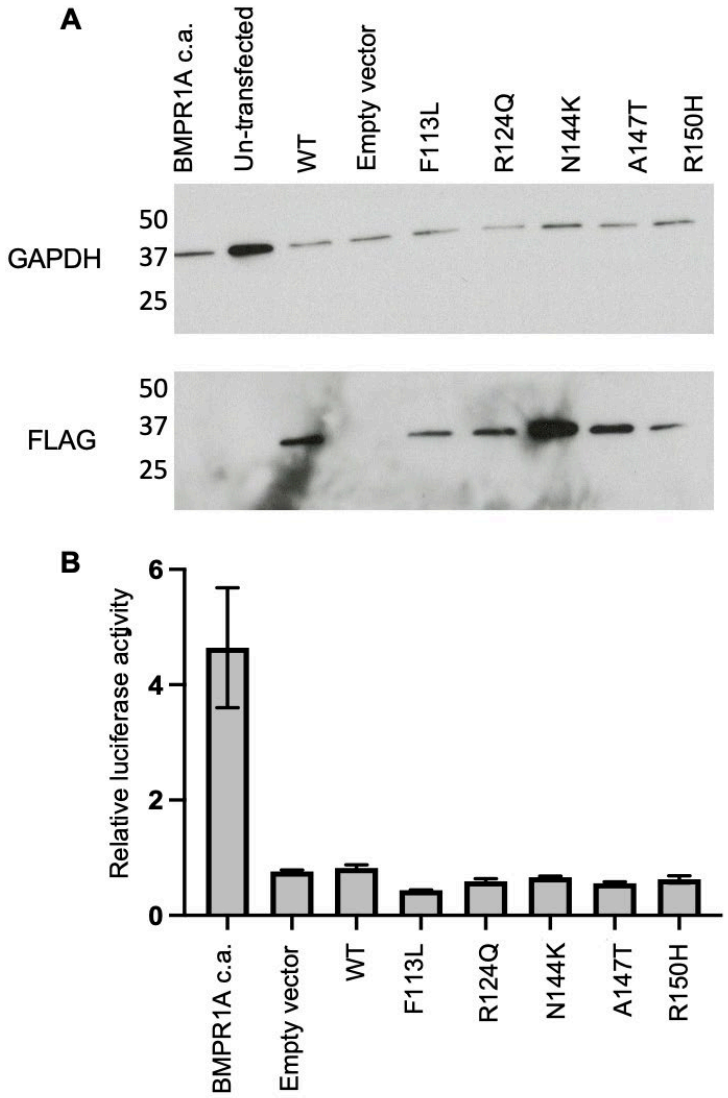


Figure 6.9 Protein levels and relative luciferase activity of wildtype and mutant *Prrx1* constructs. C2C12 cells were transfected with flag-tagged constructs considering a 3:1 ratio of transfection reagent to DNA. (A) Western blot analysis detecting GAPDH (positive control; 36 kDa) and FLAG (specific for the *Prrx1* constructs; expected 32 kDa) for all transfected constructs. (B) Luciferase assay comparing the activity of firefly-luciferase (an indication of transactivation of TN-C) to renilla-luciferase (driven by the RL-TK plasmid). Results are from triplicate technical analyses and error bars show ± standard deviation from the mean.

6.3.3 PRRX1 missense variants affect nuclear localisation

To investigate if missense variants in PRRX1 affected the ability of the protein to localise to the nucleus, I transfected tagged constructs into COS7 cells and undertook immunofluorescence analyses (figure 6.10a). This pathogenic mechanism was previously implicated for a missense variant in SHOX at the homologous position to the first homeodomain missense variant identified in *PRRX1* in a patient with craniosynostosis, p.(Arg150His). In addition, homeodomain proteins are predicted to have two nuclear localisation sequences (NLS) (figure 6.3c) at the N- and C-terminus of the homeodomain: aa 94–101 and aa 143–150 (Do et al., 2007, Moede et al., 1999), intersecting several of the identified variants (p.(Arg96Met), p.(Asn144Lys), p.(Ala147Thr), p.(Arg150His)).

Nuclear localisation was assayed for the wildtype protein (Prrx1), a homeodomain SNP (p.(Ser104Gly)) with an allele frequency of 0.0011 in gnomAD v2.1.1 that is classified as benign (incidentally, 479 heterozygous p.(Ser104Gly) substitutions were also identified in the 100kGP rare disease cohort [*n* = 74,348; allele frequency = 0.0032]), the seven missense variants identified in patients with craniosynostosis, and a heterozygous missense variant in the homeodomain (p.(Phe113Leu)) reported in an individual with agnathia-otocephaly (Sergi and Kamnasaran, 2011). Slides were initially analysed by confocal microscopy and subsequent images were automatically screened using a ScanR microscope (698–1187 cells counted for each variant). Cells were counted blindly and assigned as having either a cytoplasmic or nuclear distribution of Prrx1 (figure 6.10b). When transfecting cells using the wildtype Prrx1 construct, the encoded protein localised within the nucleus in 74% of cells (averaged across three repeats) counted (figure 6.10b); quantitatively similar results were obtained for the p.(Ser104Gly) polymorphism (67%) and the p.(Asp54Ala) variant (79%; this being the only missense variant outside of the homeodomain). In contrast, all six of the craniosynostosis-

associated homeodomain missense variants demonstrated abnormal localisation of Prrx1, with the protein localising to the cytoplasm in the majority of cells. The most severe quantitative defect was obtained for cells transfected with the p.(Arg96Met) variant, for which Prrx1 localised to the cytoplasm in 61% of cells (a 35% reduction in nuclear localisation compared to wildtype Prrx1); the other homeodomain missense substitutions displayed a 17–25% reduction in cells displaying nuclear localisation compared to the wildtype construct. Interestingly, for the agnathia-otocephaly variant (p.(Phe113Leu)) (Sergi and Kamnasaran, 2011) 82% of cells displayed a nuclear phenotype (similar to the wildtype protein), suggesting a different pathogenic mechanism(s) for this variant (figure A6.5).

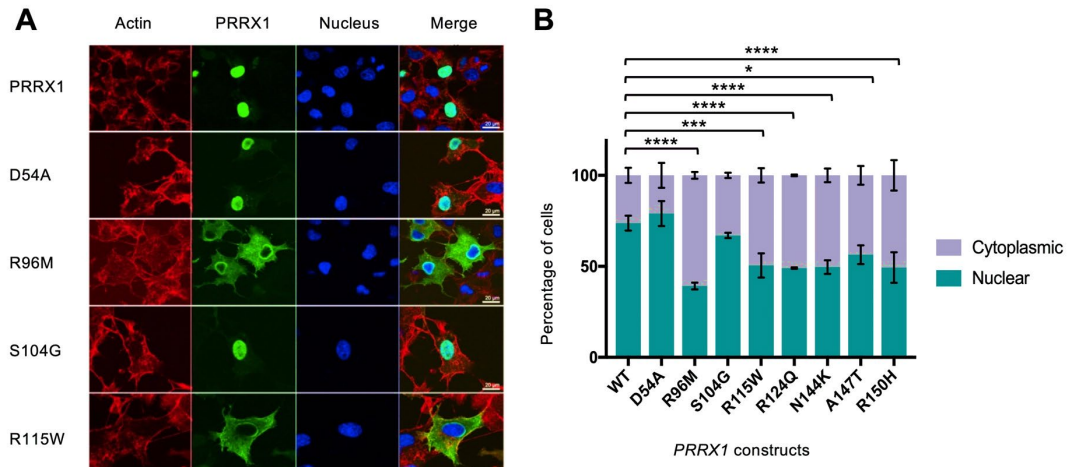


Figure 6.10 Immunofluorescence analysis of PRRX1 homeodomain missense substitutions. (A) Confocal microscopy was used to analyse the distribution of Prrx1 in the nucleus or cytoplasm of COS7 cells transfected with plasmids containing wildtype Prrx1 or mutant (D54A, R96M, S104G, R115W, R124Q, N144K, A147T, R150H). From left to right images show cytoplasmic staining of actin using Phallodin-647 (red), staining of HA-tagged Prrx1 (green), nuclear staining using DAPI (blue) and an overlay of the three channels. Images are representative of the major phenotype assessed over three technical replicates. (B) Results were quantified using a ScanR microscope that randomly took images of 25 sections of each slide (previously analysed using confocal microscopy). Blinded cell counts were scored as either cytoplasmic or nuclear considering a total of >698 cells per construct. Error bars represent ± SD. A two-way ANOVA (with Bonferroni correction) was used to assess the difference between nuclear distribution of cells expressing the wildtype and mutant proteins (P = * 0.0156, *** < 0.005, **** ≤ 0.0001).

6.3.4 Clinical phenotype in individuals with heterozygous *PRRX1* loss-of-function variants

The pattern of craniosynostosis, and associated clinical features, in 17 affected individuals from the 14 families for which there was evidence for pathogenicity (i.e., excluding the p.(Asp54Ala) variant) was analysed. The most frequent presentation (table 6.3, table A6.6) was synostosis of two or more sutures, either pure bicoronal (5/17), or other suture combinations (6/17). Together this represents a ~3-fold excess of multi-suture presentations (11/17; 65%) compared with ~22% for the craniosynostosis population as a whole (Wilkie et al., 2017); single suture synostoses were sagittal (3/17), uni-coronal (2/17) and uni-lambdoid (1/17). Six individuals had required a second major craniofacial procedure. The associated phenotype is relatively non-specific (figure 6.11), with 10/17 individuals classified clinically as non-syndromic (table A6.6). Those considered to be syndromic had a variety of additional features. Family 10, in which the proband and his grandmother had bicoronal synostosis and the grandmother additionally required surgery for unilateral ptosis, had a clinical diagnosis of Saethre-Chotzen syndrome. One affected individual from Family 7, also with bicoronal synostosis, was diagnosed with Pfeiffer syndrome owing to broad thumbs and halluces; in most individuals, the hands and feet were normal. Recurrently noted dysmorphic features included small, posteriorly rotated, or low-set ears (7 subjects) and midface hypoplasia (4 subjects). Cognitive ability was usually in the normal range, except for two individuals (III-1 in Family 10 and II-1 in Family 15) with documented chromosomal abnormalities (respectively coincidental to, and causative of, the identified *PRRX1* variant). Syndromic diagnoses were otherwise based on minor facial dysmorphic features such as low frontal hairline, eyelid ptosis, prominent orbits, mild midface hypoplasia and small and/or low-set ears with or without external auditory canal stenosis, which did not amount to a recognisable pattern. Overall, this analysis suggests that the prognosis for affected individuals is good, provided that the consequences of craniosynostosis and any

associated intracranial hypertension are addressed, and coincident chromosomal abnormalities are excluded.

To examine the inheritance pattern of the rare *PRRX1* variants, parents and, where available, additional family members were tested. Except for Family 2 in which the original *de novo* mutation was identified, the *de novo* variant arising in the father of Family 9, and the *de novo* 61.5 kb deletion in Family 13, most variants were inherited from parents without craniosynostosis (pedigrees shown in figure 6.2). Of note, the p.(Asn144Lys) variant (Family 10) was inherited from an affected grandmother, but craniosynostosis was not evident in the carrier mother. The p.(Arg95*) variant (Family 4) was inherited from a clinically syndromic father, noted to have an atypical skull shape, but who did not present with craniosynostosis. Overall, after excluding the proband in each family, the penetrance of craniosynostosis in other family members proven or deduced to be *PRRX1* carriers was 3/24 (12.5%); hence, it is apparent that many individuals heterozygous for these variants are non-penetrant or only mildly clinically affected.

Table 6.3 Subjects with craniosynostosis analysed for rare, deleterious *PRRX1* variants in three screens[a]

	Nonsyndromic		Syndromic		Combined	
	Total	**PRRX1 positive**	**Total**	**PRRX1 positive**	**Total**	**PRRX1 positive**
Metopic	321		46	1	367	1 (0.27%)
Sagittal	611	1	57		668	1 (0.15%)
Unilateral coronal	181	2	24		205	2 (0.98%)
Bilateral coronal	33	1	11	1	44	2 (4.55%)
Uni- or bilateral lambdoid	33	1	4		37	1 (2.70%)
Other multisuture	69	3	46	1	115	4 (3.48%)
Suture not specified			19		19	0 (0.00%)
Combined	**1248**	**8**	**207**	**3**	**1455**	**11 (0.76%)**

[a]See table 6.1 for further details.

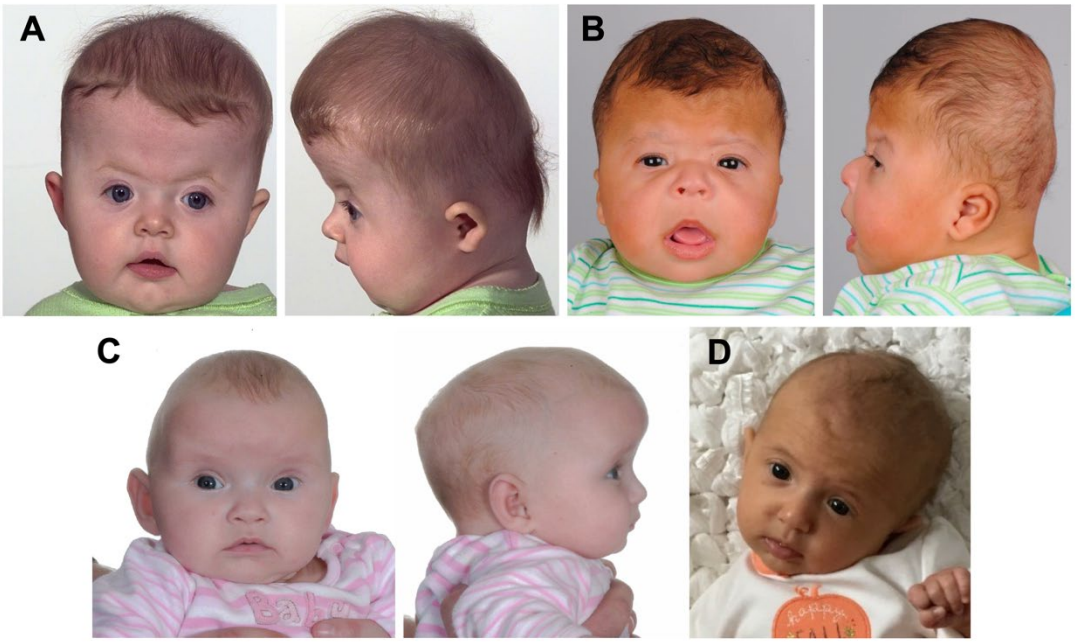


Figure 6.11 Pre-operative clinical presentation of craniosynostosis in patients with *PRRX1* missense or loss-of-function variants. (A) Individual II-1 from Family 2 aged 5 months with non-syndromic (NS) bicoronal synostosis (p.(Arg18Alafs*23)). (B) Individual III-1 from Family 10 aged 3 months with syndromic bicoronal synostosis (p.(Asn144Lys)); this individual also harbors a maternally inherited del(19)(p13.11). (C) Individual III-2 from Family 5 aged 3 months with NS sagittal synostosis (p.(Arg96Met)). (D) Individual III-1 from Family 9 aged 2 months with NS right lambdoid synostosis (p.(Arg124Gln)).

6.4 Discussion

In this work I describe the identification of heterozygous loss-of-function and deleterious missense variants of *PRRX1* associated with craniosynostosis in 14 unrelated families. In total, there was one loss of the start codon, three variants (nonsense or frameshift) that introduce a premature stop codon, six missense variants (in seven families) that affect highly conserved residues within the homeodomain, and three complete or partial deletions of *PRRX1*. The six homeodomain variants were all shown to reduce nuclear localisation of PRRX1 in a cellular assay and were considered pathogenic; in addition, one patient with syndromic metopic synostosis and a rare variant outside the homeodomain (p.(Asp54Ala)) was identified. Notably, Asp54 intersects a putative PRX domain (aa 28–45) which was previously shown to have a subtle inhibitory effect on protein transactivation (Norris and Kern, 2001a). Nevertheless, p.(Asp54Ala) did not abrogate nuclear localization in this study and therefore the pathogenic significance of this variant remains uncertain. Overall, these observations suggest partial or complete loss-of-function of the variant *PRRX1* allele, consistent with haploinsufficiency as the likely pathogenic mechanism. Interestingly, a small number of individuals have been described with heterozygous deletions of 1q24.2, associated with syndromic learning disability with short stature, brachydactyly and facial dysmorphism as additional features (Burkardt et al., 2011, Ashraf et al., 2015, Chatron et al., 2015). Although *PRRX1* falls outside of the shortest region of overlap for the full syndrome, a pair of monozygotic twins has been described with a ~2.6 Mb deletion including *PRRX1*, both of whom required surgery for craniosynostosis (Ashraf et al., 2015).

Despite previous literature showing transactivation of the TN-C promoter by PRRX1, I was unable to replicate these findings in order to analyse the effect of missense variation within *PRRX1*. I identified several contradictions to previously published experiments including the hypothesis that the OAR domain acts to repress TN-C transactivation

(Norris and Kern, 2001a). Concordant with other studies indicating that the OAR domain instead acts as a transcriptional activator (Jones et al., 2001), in my experiments, loss of the OAR domain resulted in reduced luciferase activity. The result was observed when using the original constructs from Michael Kern, alongside my corrected mouse and human TN-C constructs. Additionally, the identification of three errors in the original TN-C construct provided another level of uncertainty (Norris and Kern, 2001a). Later studies investigating agnathia-otocephaly using a human TN-C reporter (Dasouki et al., 2013) were also difficult to replicate. Of note, all papers reporting patients with variants in *PRRX1* and agnathia-otocephaly showed incorrect variant annotation, incomplete methodology, and, in some cases, the absence of any evidence supporting a variant allele in the affected individual. Indeed, RNA sequencing data obtained from Dr Majed Dasouki failed to support the documented variant owing to an average reported read depth of 2. Incidentally, all reported agnathia-otocephaly papers contain the same lead author (Dr Deepak Kamnasaran) and likely follow the identical or similar experimental protocols. As such, it draws into question the validity of previous reports using these transactivation assays to assess the effect of *PRRX1* variants, particularly in the context of agnathia-otocephaly.

However, consistent with Kern's early report, I did find a subtle difference in output from the transactivation assay dependent on cell type and, in support of his claims, C2C12 yielded the strongest signal. Nevertheless, the luciferase activity, and thus transactivation of TN-C, was not consistently high enough nor above background levels to draw conclusions. Despite this, the process enabled optimisation of cell type and transfection conditions for the immunofluorescence analyses. While C2C12 cells appeared to yield a marginally stronger signal in the transactivation assay compared COS7 (figure 6.7c), COS7 cells were selected for the immunofluorescence analyses owing to faster growth rates (permitting screening of more transfected cells after 24 hours) and visually improved cell morphology at higher magnifications. Additionally,

initial flow cytometry analyses (using a GFP plasmid as a proxy for the PRRX1 construct) demonstrated that a 6:1 ratio of lipofectamine (μl) to DNA (μg) yielded the highest transfection efficiency (figure A6.2), but at the cost of reduced cell viability. The transactivation assays were therefore largely conducted using a total of 1 μg of DNA and 3 μl of lipofectamine per well of a 6-well plate. However, this was adjusted to a 6:1 ratio (using a total of 500 ng of DNA) for the immunofluorescence analyses, which successfully prevented visible differences to cell morphology owing to transfection but maintained transfection efficiency. Transfection reagents were left to incubate with cells over-night as there was little difference observed between a 6-hour or over-night incubation (figure A6.4).

The immunofluorescence analyses suggested a reduced ability of the mutant PRRX1 protein to translocate into the nucleus as a contributing pathogenic mechanism. Four of the six missense substitutions reported in this study coincide with two predicted NLS within PRRX1 (Bernhofer et al., 2021) at identical positions to those reported in other members of the homeobox superfamily (Do et al., 2007, Wang et al., 2021, Sabherwal et al., 2004). In support, missense variants at either the N- or C-terminal NLS in PITX2B (another PRD homeobox transcription factor) also result in abnormal nuclear localisation (Wang et al., 2021, Schneider et al., 2005). The importance of these two regions of the homeodomain is also well established for DNA binding. Residues at position 2, 3 and 5–8 on the N-terminal arm, and residues 47, 50, 51, 54, and 55 of the homeodomain have all been shown to contribute to DNA-binding specificity (Noyes et al., 2008); in this study, variants affecting three of these critical residues were identified (3: p.(Arg96Met), 51: p.(Asn144Lys), and 54: p.(Ala147Thr)) (figure 6.1d, figure 6.3c). Furthermore, in the wider homeobox superfamily literature, variants that correspond in position to the PRRX1 homeodomain missense substitutions have been reported as pathogenic in multiple disorders (table A6.7).

Surprisingly, I also found abnormal nuclear localisation of the p.(Arg115Trp) and p.(Arg124Gln) variants. Substitutions at these positions are predicted to effect DNA-binding (https://predictprotein.org/), and the equivalent position to Arg124 in the homeobox protein PROP1 was reported to disrupt this function (table A6.7). Both residues reside within the major alpha helices in PRRX1 (figure 6.3c), with residue 31 of the homeodomain (Arg124) interacting with Glu135 in the recognition helix and capable of forming homeodomain-DNA salt bridges (Zhou et al., 2012). Changes at residue 22 (Arg115) have been hypothesised to affect protein interaction with co-activators (Dawson et al., 1998), possibly reducing efficiency of nuclear import. Although I did not assay DNA-binding activity, an impaired ability to enter the nucleus would be sufficient to abrogate PRRX1-mediated gene regulation. In this study I prioritized analysis of variants identified in patients with craniosynostosis, however, there are several other rare homeodomain missense variants reported in population databases such as gnomAD and UK Biobank that could be disease-causing. Future clinical and functional studies could explore further the effect of rare missense variation across the homeodomain to establish any hidden genotype-phenotype correlation. Additionally, while variants outside of the homeodomain could disrupt protein transactivation (as hypothesised for the p.(Asp54Ala) variant), the 3.5-fold enrichment of rare variants within the homeodomain (based on the resequencing survey), combined with the low disease penetrance of pathogenic variants, indicates that establishing disease causality for variants outside this well-studied domain would be challenging.

The phenotype associated with *PRRX1* pathogenic variants and craniosynostosis appears relatively non-specific, with a variety of sutures fused and no diagnostic syndromic features. However, there was a clear tendency to fusion of multiple sutures – especially bicoronal and other multi-suture presentations, which in combination were 3-fold over-represented compared to craniosynostosis as a whole. The requirement for

repeated craniofacial procedures (5/17 individuals) was also high. Together these considerations indicate that a diagnosis of *PRRX1*-related craniosynostosis serves as a prognostic marker of greater severity and suggests that children harbouring pathogenic *PRRX1* variants should be monitored throughout their growth for possible late development of raised intracranial pressure. Based on the aggregated data from the three craniosynostosis cohorts that we screened, the prevalence of pathogenic *PRRX1* variants was 0.76% (11/1455). However, this figure is subject to potential bias, for example the mix of subjects screened was not based on a cross-sectional population sample, and only a subset of the total (*n* = 479) was screened for deletions. A lower prevalence of 0.3% (3/981) was determined from a prospective UK 18-year birth cohort of 981 affected individuals (unpublished), however the 95% confidence interval is wide (0.06–0.9%). Hence, I anticipate that the screening of other large population cohorts should reveal further affected individuals with *PRRX1* variants, alongside unaffected family members. An important consideration for variant interpretation is that the penetrance of pathogenic variants in the heterozygous state is low (estimated as 12.5%; 3 of 24 relatives of the proband who are heterozygous for the pathogenic variant are affected). The loss-of-function intolerance score (pLI) and loss-of-function observed/expected upper bound fraction (LOEUF) are 0.24 and 0.66, respectively (Karczewski et al., 2020). These moderate constraint values are concordant with the deduction that haploinsufficiency can cause significant congenital disease, but with reduced penetrance. Hence, transmission of a *PRRX1* variant from an unaffected parent does not exclude a causal connection in an individual with craniosynostosis. The variable phenotype is likely attributable to a complex combination of genetic factors (as discussed below) and environmental differences, including fetal head constraint (Sanchez-Lara et al., 2010).

An important and currently unresolved question is how these observations can be reconciled with previous reports of three heterozygous *PRRX1* variants (Sergi and

Kamnasaran, 2011, Donnelly et al., 2012, Dasouki et al., 2013) and, less problematically, one homozygous variant (Celik et al., 2012) associated with the much more severe phenotype of agnathia-otocephaly. The homozygous variant (c.691G>C; p.(Ala231Pro)) could be consistent with substantially reduced function of PRRX1, leading, as in the mouse (ten Berge et al., 1998), to severe deficiency of facial structures. The association of the heterozygous variants are harder to explain in the light of our new data, particularly given that two of the three variants predict frameshifts (c.266_269dup; p.(Arg92Glufs*8) (Dasouki et al., 2013) and c.269del; p.(Lys90Argfs*42) (Donnelly et al., 2012)) of the protein occurring very close to the nonsense variant p.(Arg95*) reported here. Of note, no individual in whom the deleted region included *PRRX1* had the more severe craniofacial features of agnathia-otocephaly, excluding haploinsufficiency as the mechanism of this more severe phenotype. Moreover, I found no convincing evidence that the p.(Phe113Leu) missense variant (identified in a patient with agnathia-otocephaly) affected nuclear import, although this does not exclude other pathogenic mechanisms. Differences in genetic background might contribute to the discrepant phenotypes, which could be either allelic differences (for example, relative expressivity of mutant and wildtype *PRRX1* allele, as observed in the case of *RBM8A* variants in thrombocytopenia-absent radius [TAR] syndrome (Albers et al., 2013)), or non-allelic differences. Notably, given prior evidence from the mouse of functional redundancy between *Prrx1* and *Prrx2* (ten Berge et al., 1998), and the suspected genetic modifiers reported in association with *OTX2* mutations (Hide et al., 2002, Chassaing et al., 2012) (see section 6.1), coincident loss-of-function variants in *PRRX2* or other homeobox transcription factors could exacerbate the phenotype. Additional comments on these various explanations are provided in table A6.8. Authors of all the previous reports of agnathia-otocephaly and variants in *PRRX1* were contacted with the aim of undertaking comparative studies. Unfortunately, related to early lethality, a lack of suitable biological materials from the individuals with agnathia-otocephaly hampered a more systematic approach to this problem.

Based both on functional analyses and the observation of phenotypes associated with 1q24.2 deletions (discussed above), I conclude that heterozygous loss-of-function and missense mutations within the homeodomain of *PRRX1* are normally associated with craniosynostosis or milder phenotypes. Independent analyses have shown that *Prrx1* is expressed in the mouse postnatal suture, localising to the intra-sutural mesenchyme but absent from the osteogenic fronts (Ouyang et al., 2014, Takarada et al., 2016, Wilk et al., 2017, Farmer et al., 2021). Fate mapping showed that the cells expressing *Prrx1* later differentiated into osteoblasts lining the bony plate surfaces. It was concluded that *Prrx1* expression characterises a subpopulation of cells spanning the late stem cell (Sca1+) and early osteogenic (Runx2+) stages (Takarada et al., 2016). PRRX1 may link to the BMP-MSX2 signalling axis in the cranial suture, overactivity of which has previously been implicated in craniosynostosis (Jabs et al., 1993, Timberlake et al., 2016). Binding of PRRX1 to a specific sequence within the promoter of *Msx2* was demonstrated *in vitro* (Brugger et al., 2004) and, in the mouse mandible, it has been shown that the requirement for a high level of BMP4 to induce expression of the homeobox gene *Msx2* can be bypassed in a *Prrx1*$^{-/-}$;*Prrx2*$^{-/-}$ genetic background, indicating that PRRX1 acts as a repressor in this context (Liu et al., 2005); hence, if the same activity occurs in the cranial suture, the effect of PRRX1 mutation would be to alleviate this repression, thus predisposing to craniosynostosis. In summary, these new human genetic findings complement these mouse studies, confirming the important role of PRRX1 in cranial suture function by providing the first genetic evidence that ~50% reduction in PRRX1 levels is sufficient, in some individuals, to lead to craniosynostosis.

Chapter 7

A novel Carpenter syndrome-like phenotype associated with a homozygous variant in PTCH1

7.1 Introduction

PTCH1 is a transmembrane (TM) receptor that acts as an antagonist of the hedgehog signalling pathway. Hedgehog signalling was first identified in *Drosophila* as an important regulator of segmental patterning and polarity (Nüsslein-Volhard and Wieschaus, 1980), with later research establishing the essential role of this pathway for embryonic development, through regulation of polarity, organogenesis, and survival and proliferation of multiple tissue progenitor and stem cell populations (Ingham and McMahon, 2001, Briscoe and Therond, 2013). The first hedgehog receptor identified in *Drosophila* was Patched (*Ptc*); loss of *Ptc* in embryos resulted in cells in the middle half of each segment expressing pattern elements appropriate to the margins of each segment (Hooper and Scott, 1989, Nakano et al., 1989), identifying a role for hedgehog signalling in developmental patterning. Following this, a human homologue of the *Drosophila* Ptc was identified on chromosome 9, containing 23 coding exons and spanning approximately 34 kb (Hahn et al., 1996). The structure of the 1447-residue human PTCH1 protein has since been resolved by cryogenic electron microscopy (Cryo-EM) and has been shown to contain 12 TM helices. Two large extracellular domains (ECDs) are anchored between TM1–2 (ECD1) and TM7–8 (ECD2) (Gong et al., 2018, Kowatsch et al., 2019), which function to bind its ligand. There are three mammalian hedgehog ligands that bind the PTCH1 receptor (see chapter 1, section 1.7.4.1), comprising desert hedgehog (DHH) which is expressed largely in the gonads, Indian hedgehog (IHH) which has been shown to be important for endochondral ossification, and sonic hedgehog (SHH) which is essential for limb patterning and correct development of the nervous system (Carballo et al., 2018). The two extracellular ligand-binding domains are connected to the N and C termini of each TM domain by a flexible linker and neck helix, respectively (Kowatsch et al., 2019).

The hedgehog signal is transmitted through a second TM (7-helices) protein known as smoothened (SMO). Despite PTCH1 and SMO not physically interacting, it has long been understood that PTCH1 inhibits hedgehog signalling through repressing the activity of SMO (Qian et al., 2019). Structural analyses show that PTCH1 binds SHH in its dimeric form (Gong et al., 2018, Qi et al., 2018, Kowatsch et al., 2019). Upon ligand binding, PTCH1 releases its repression on SMO and allows active signalling (Gong et al., 2018) (figure 7.1a). Given the lack of physical interaction, a second messenger was presumed to mediate hedgehog signalling between PTCH1 and SMO, with recent evidence supporting cholesterol molecules as the key chemical mediator between the two TM proteins (Radhakrishnan et al., 2020) (figure 7.1b). This hypothesis is also supported by crystal structures of SMO which show that cholesterol molecules not only bind to, but are essential for, correct function of SMO (Byrne et al., 2016, Huang et al., 2016, Luchetti et al., 2016). Crystal structure analyses of cholesterol-bound SMO have identified cholesterol bound to the hydrophobic groove of the extracellular cysteine-rich domain (CRD), but also at the centre of its TM domain.

The proposed mechanism of action is that PTCH1 uses its transporter-like function to diminish a pool of accessible cholesterol and repress hedgehog signalling; upon PTCH1 inactivation by SHH, cholesterol accessibility increases allowing SMO activation. In support of a transporter-like function, studies show that PTCH1 can reduce the abundance of free cholesterol in both the inner and outer leaflets of the plasma membrane (Kinnebrew et al., 2021, Zhang et al., 2018). Additionally, PTCH1 has been shown to display sequence homology to Niemann-Pick C1 (NPC1) (Radhakrishnan et al., 2020), a lysosomal membrane protein that transports sterol molecules. Both proteins are found to harbour a sterol sensing domain, which is located between TMs 2–6 of PTCH1 (Davies and Ioannou, 2000). However, unlike NPC1, the direction of sterol transport and the putative donor or acceptor protein for PTCH1 remains unknown.

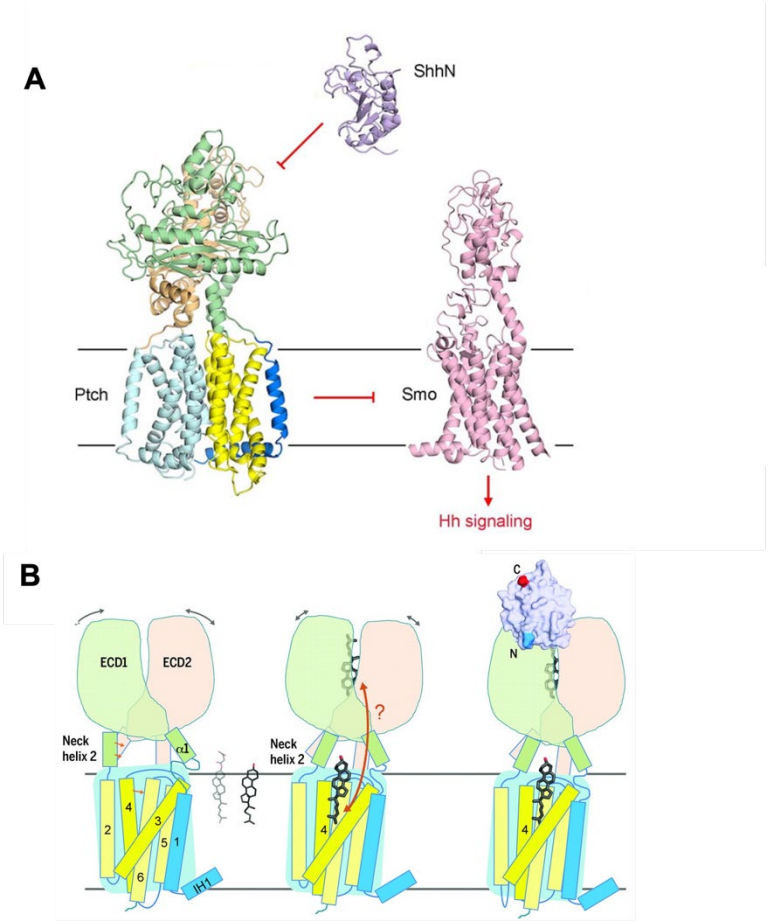


Figure 7.1 Schematic illustration of the interaction of PTCH1 with SHH and SMO. (A) Regulation of the hedgehog signalling pathway by PTCH1. Upon binding of SHH (ShhN: N-terminal domain of human SHH [residues 24 to 197]) to PTCH1, PTCH1 releases its repression of SMO allowing downstream hedgehog signalling. (B) Proposed structural mechanism of the SHH ligand binding PTCH1; ligand binding induces a closed conformational change, potentially forming a tunnel for cholesterol transport. Images obtained from Gong et al., 2018.

The paralogous gene to *PTCH1*, Patched 2 (*PTCH2*), displays significant sequence and structural homology (Motoyama et al., 1998) and has been shown to bind hedgehog ligands, including SHH (Takabatake et al., 1997), and to negatively regulate hedgehog signalling through its interaction with SMO (Carpenter et al., 1998). However, these two genes differ in their expression patterns, with *Ptch1* widely expressed throughout the mouse embryo while *Ptch2* expression is restricted to the skin and spermatocytes (Carpenter et al., 1998). Complete loss of *Ptch1* in mice results in embryonic lethality, while *Ptch2* null mutants are viable and fertile, suggesting functional compensation by Ptch1 (Nieuwenhuis et al., 2006). In line with sites of expression, adult male *Ptch2*

mutants develop skin lesions consisting of alopecia, ulceration, and epidermal hyperplasia (Nieuwenhuis et al., 2006).

While *Ptch1*-null mice die prenatally (E8.5) with neural tube closure defects (Goodrich et al., 1997), there have been reports of mouse models with reduced Ptch1 function that survive until birth. The first model was identified from a screen of spontaneous recessive mutations in mice with mesenchymal dysplasia (herein referred to as the mes model) (Sweet et al., 1996). Later analysis of the mes mutant identified the causative variant to be a 32-bp deletion in the C-terminal cytoplasmic domain of Ptch1 (Makino et al., 2001), which was predicted to introduce a premature termination codon (PTC) 68 amino acids after the last transmembrane domain. The recessive phenotype (*Ptch1*$^{mes/mes}$) was assessed in three different mouse strains (table A7.1) with consistent phenotypes being preaxial polydactyly, skeletal defects, a domed and shortened head, and increased body weight (Sweet et al., 1996, Makino et al., 2001). The second hypomorphic model was identified from a recessive N-ethyl-N-nitrosourea (ENU)-mutagenesis screen (Feng et al., 2013). The model was termed 'dogface-like' (DL) owing to abnormalities in the skull and snout morphology. Homozygous mutants survived until birth but died pre-weaning. Skeletal staining of the DL mutants at E18.5 showed embryos with a similar phenotype to the previously reported mes mutant, including domed skulls, abnormal patterning of the sternum and limbs, and polydactyly of all four limbs; additionally, DL mutants displayed an abnormal snout morphology, synostosis of the lambdoid sutures, and an omphalocele. The causative variant in the *Ptch1*^{DL} mouse was identified as a point mutation at a donor splice site, leading to the production of two abnormal transcripts. Amplification of mutant RNA samples showed two PCR products, a weaker product of similar size to the control (termed *Ptch1*$^{DL\text{-}a}$) and an abundant product which was 119 bp shorter than the wildtype RNA (*Ptch1*$^{DL\text{-}b}$). The weaker product (*Ptch1*$^{DL\text{-}a}$) was shown to contain a 36-nucleotide insertion (12 aa) within the intracellular loop, predicting a new in-frame splice site and a protein of size 1446 aa (the wildtype mouse Ptch1 is 1434 aa).

The second, more abundant product (*Ptch1*$^{DL\text{-}b}$) showed a shorter transcript, owing to skipping of exon 13, which is predicted to introduce a PTC 9 aa into exon 14.

In humans, linkage mapping first identified *PTCH1* as a tumour suppressor gene and causative of nevoid basal cell carcinoma syndrome (NBCCS) (otherwise referred to as Gorlin syndrome [OMIM #109400]) (Hahn et al., 1996). NBCCS is an autosomal dominant disease, characterised by developmental defects including bifid ribs, palmar pits and macrocephaly, and predisposition to tumours (such as basal cell carcinoma) which develop during early adulthood (Reinders et al., 2018). Frameshift, nonsense, missense, and splicing variants, alongside large genomic deletions, and duplications of *PTCH1*, have all been documented in patients with NBCCS. These variants are located in all coding regions, with no hotspot identified (Hong et al., 2016, Reinders et al., 2018, Martinez et al., 2019). Recently, a heterozygous frameshift variant in *PTCH1* was described in a child with overgrowth, dysmorphic facial features, and metopic synostosis. The identified variant (encoding p.(Val502Glyfs*13)) was predicted to alter splicing and affect the sterol-sensing domain (426–616 aa) (Beltrami et al., 2020). The phenotypic features of this patient overlapped with several of those described in patients with a 9q22.3 deletion (which includes *PTCH1*) (Yamada et al., 2020); many of the features associated with a 9q22.3 deletion are also described in Gorlin syndrome, with the exception of craniosynostosis, obstructive hydrocephalus, and intellectual disability (Beltrami et al., 2020). Adding to this initial patient, a further three cases of metopic synostosis with heterozygous loss-of-function variants in *PTCH1* have been described (see chapter 3, section 3.4.2.1), alongside a patient with metopic synostosis and a *de novo PTCH1* frameshift variant identified within the 100kGP data (chapter 3, section 3.3.4.1).

To the best of my knowledge, there are no descriptions in humans of homozygous pathogenic variants in *PTCH1*. Here I identify the first *bona fide* case of two patients

within the same family with a novel homozygous missense variant in *PTCH1*, causing a Carpenter syndrome-like phenotype. I aimed to ascertain the likely pathophysiological mechanism associated with this variant and identify further cases to corroborate findings.

7.2 Methods

7.2.1 Patient-derived lymphoblastoid and fibroblast cells

Blood from the proband and both parents, and a skin sample from the proband were obtained for preparation of EBV-transformed lymphoblastoid and fibroblast cells, respectively, as detailed in chapter 2, section 2.6.1–2.

7.2.2 Variant confirmation from gDNA

Genomic DNA was isolated from fresh blood samples (detailed in section 2.1) from both parents, the proband and their sibling, and amplified using a Q5 protocol with 10 μM of both forward (5'-GCCCTGGAGTAAATACAGCACTCAG-3') and reverse (5'-GATGTTAAGAGCAGTTAAGAAGCAGGAGC-3') primers, and an annealing temperature of 60°C. Products were prepared for dideoxy-sequencing (section 2.2) and sequenced by the WIMM Sequencing Facility.

7.2.3 RT-PCR

Lymphoblastoid cells from both parents and the proband were cultured for RNA analysis by RT-PCR (section 2.8). Primers were designed to amplify fragments which overlap exon boundaries (table A7.2)

7.2.4 WGS

Whole genome sequencing was performed as described in section 2.4. Regions of homozygosity were determined using bcftools (v1.5). Given the known consanguinity,

homozygous variants within any region larger than 2 Mb were extracted. Variants were filtered by gnomAD (v2.1.1) allele frequency ≤0.001, variant reads above 80% of the total, no observation in other individuals in CranioVar, and consequence (missense or worse). Remaining variants were assessed manually for likelihood of being disease causing, considering the number of heterozygotes or homozygotes reported in gnomAD at that position, intersection with reported low complexity regions, evidence of being a SNP, description of protein function, and gene-disease associations.

7.2.5 Rare disease analysis – 100kGP

Recessive variants in *PTCH1* were called from participant VCF files in the rare disease cohort enrolled in the 100kGP and filtered by gnomAD frequency (see section 3.2.3.1 for details).

7.2.6 Western blot

Protein stability in patient-derived lymphoblastoid cells was assayed using a western blot (section 2.7) with primary antibodies staining PTCH1 (rabbit polyclonal antibody; Abcam, ab53715), diluted 1/2000 in 5% no-fat milk and 1x T-TBS, and HRP-conjugated GAPDH (Cell Signaling Technology, 14C10), diluted 1:20,000. Both antibodies were left to incubate with the membrane at 4°C for 16 hours before staining with secondary donkey anti-rabbit-HRP (Abcam; ab97085) for 1 hour at room temperature. Membranes were washed in 1x T-TBS before imaging.

7.2.7 qPCR

Patient lymphoblastoid and fibroblast cells were used to quantify RNA transcripts of hedgehog signalling targets (methods described in section 2.9). Primers used are described in table A7.2.

7.2.8 Plasmid construction

Four lentiviral vectors (pHR-CMV-TetO2) were obtained from the Siebold lab (Wellcome Centre for Human Genetics, University of Oxford): (1) a full-length human *PTCH1* (*PTCH1-FL*) construct, (2) a truncated *PTCH1* mutant (*PTCH1ΔICL3*) containing an in-frame deletion of the intracellular loop 3 (amino acids 630–717), (3) a lentiviral vector with the cloned *ALK* sequence and the AVI-1D4 tag, and (4) a lentiviral vector with the cloned *ALK* sequence and the mVenus-1D4 tag. Initially, 1 μg of the *PTCH1-FL* and *PTCH1ΔICL3* plasmids were amplified by PCR using primers 1–2 and 3–4, (table A7.3) respectively, following a Q5 enzyme protocol (10 μl of 5x Q5 reaction buffer, 10 μl of 5x Q5 high GC enhancer, 2 μl of 10 mM dNTPs, 2 μl of each forward and reverse primer [10 μM], 1 μl of Q5 high fidelity polymerase, 0.5 μl of Taq polymerase, 1 μl DMSO and up to a final volume of 50 μl of nuclease-free water). The reaction was performed in a thermocycler under the following conditions: 98°C for 3 min, followed by 35 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 4 min and a final extension of 72°C for 5 min and 4°C hold. Expected product size was confirmed by gel electrophoresis (1% agarose gel, 150 V for 45 min). Amplified products were ligated into pGEM-T easy vector using 0.5 μl vector, 5 μl T4 ligase buffer, 2 μl *PTCH1* template, 1 μl T4 DNA ligase and up to 10 μl of nuclease-free water. The ligation reaction was left overnight at room temperature before performing a transformation using Invitrogen One Shot™ TOP10 chemically competent cells. Presence of the insert was confirmed by colony PCR and restriction digest of 1 μg of plasmid using EcoRI-HF and KpnI-HF; successful clones were grown overnight in LB media before plasmids were purified using a miniprep protocol (Qiagen).

The NEBaseChanger tool was used to design abutting primers for SDM of the p.(Phe434Ser) and p.(Gln466Leu) variants (primers 5–6, and 7–8; table A7.3) in the pGEM-*PTCH1-FL/ΔICL3* constructs. SDM was performed using recommended conditions for a Q5 protocol (12.5 μl Q5 hot start high-fidelity 2x master mix, 0.5 μM of

each forward and reverse primer, 25 ng of template DNA [pGEM-*PTCH1-FL* or pGEM-*PTCH1ΔICL3*] and up to 25 μl reaction of nuclease-free water) and cycling conditions (98°C for 30 s, followed by 25 cycles of 98°C for 10 s, 66°C/ 72°C for 20 s for each F434S/Q466L primer pair respectively, and 72°C for 5 min; a final extension of 72°C for 2 min was performed). Following SDM, 1 μl of product was incubated with 5 μl of kinase, ligase, and DpnI (KLD; M0554, NEB) 2x reaction buffer, 1 μl of 10x KLD enzyme mix, and 3 μl of nuclease-free water for 5 min at RT. Finally, 5 μl of the KLD reaction was transformed into 50 μl of chemically competent cells (NEB 5-alpha, C2987H) following manufacturer's guidelines. Presence of the correct product was confirmed by colony PCR using primers 9–10 (table A7.3) followed by dideoxy-sequencing.

Following this, 1 μg of the mutagenised pGEM plasmids (pGEM-*PTCH1-FL-F434S/Q466L,* pGEM-*PTCH1ΔICL3-F434S/Q466L*) were digested for 1 hour at 37°C using 1 μl of each EcoRI-HF, KpnI-HF, and BsaI, combined with 5 μl CutSmart buffer (NEB) and up to 50 μl of nuclease-free water. The digested products were run on a 1% agarose gel at 150 V for 45 min, before gel purification (NEB Monarch DNA gel extraction kit). Prior to ligation, the *ALK* insert was removed from pHR-CMV-TetO2 lentiviral vectors, containing the AVI-1D4 or mVenus-1D4 tag, by restriction digest (EcoRI-HF and KpnI-HF; 37°C for 15 min following a NEB protocol). A 1 μl sample of alkaline phosphatase was added to the digestion reaction to prevent re-ligation of the vector before the vector was run on a 1% agarose gel and gel purified. Briefly, rapid ligation was performed following a quick ligation protocol (NEB, M2200S) using a ratio of 1:8 vector to insert (μl) and was followed by transformation of ligated plasmids into one shot Stbl3 chemically competent cells (Invitrogen, C7373-03), using the recommended manufacturer's protocol. Successful transformants were screened by colony PCR; single colonies were resuspended in 10 μl of nuclease-free water and a 1 μl sample was mixed with 4 μl of 5x Phire reaction buffer, 0.4 μl of 10 mM dNTPs, 1 μl of each forward and

reverse primer (10 μM; table A7.3 [primers 1–2 and 3–4]) and 0.4 μl of Phire hot start II DNA polymerase. The PCR reaction was performed under the following conditions: 98°C for 3 min, followed by 35 cycles of 98°C for 10 s, 57°C for 10 s and 72°C for 20 s, with a final extension of 72°C for 1 min. Successful clones were grown overnight in LB containing ampicillin (50 μg/ml) at 37°C, shaking at 220 rpm. Plasmids were purified from *E. coli* by miniprep and digested with KpnI-HF and EcoRI-HF to confirm correct size of the insert. Clones confirmed by digest were sent for dideoxy-sequencing.

7.2.9 CRISPR targeting

7.2.9.1 Construction of the pX458-SpCas9-Ptch1 construct

Initially, 1 μg pX458-SpCas9 (Addgene 48138) plasmid was digested using 1 μl high-fidelity BbsI (NEB R3539S), 1 μl fast thermosensitive alkaline phosphatase (Fast AP 1 U/μl; Thermo Scientific, EF0651), and 2 μl 10x CutSmart Buffer (NEB, B7204S) in a final reaction volume of 20 μl. The digested product was run on a 1% agarose gel and purified using the NEB 'Monarch DNA Gel Extraction Kit' (T1020S). A guide RNA sequence was designed using http://crispor.tefor.net/ considering NGG as the protospacer adjacent motif (PAM) at a maximum distance of 25 bp from the variant site (Ran et al., 2013, Schubert et al., 2021); the selected guide contained a specificity score of 89 and no predicted off-target effect with 1–3 mismatches. Following, the forward and reverse primers corresponding to the guide RNA sequence were modified to contain complementary overhangs for BbsI. Both primers (100 μM of each forward [5'-CACCGTCCCTTCACAACCACGACCC-3'] and reverse [5'-AAACGGGTCGTGGTTGTGAAGGGA-3']) were mixed with 1 μl 10x T4 ligation buffer (NEB, B0202S), 0.5 μl T4 PNK (NEB, M0201L) and up to 10 μl reaction volume of nuclease-free water. The primers were annealed under the following conditions: 37°C for 30 min, followed by 95°C for 5 min, and gradual ramp down to 25°C (-5°C/min). The annealed and phosphorylated primer pair (1 μl) and 50 ng of digested pX458-SpCas9

plasmid were ligated together following manufacturer's protocol for rapid ligation (NEB; Quick Ligation Kit, M2200S). A 5 μl sample of ligated product was transformed into one shot Stbl3 chemically competent cells (Invitrogen, C7373-03), following protocol guidelines, and the cells were plated onto ampicillin agar plates before being left overnight at 37°C. Single colonies were grown overnight at 37 °C, shaking at 220 rpm, in LB supplemented with 50 μg/ml of ampicillin. Plasmid was extracted following an endotoxin-free maxiprep protocol (Qiagen, 12391) and confirmed using dideoxy-sequencing.

7.2.9.2 Culturing mESCs

Mouse embryonic stem cells (mESCs [E14Tg2aIV, 129/Ola]) were grown in Glasgow Minimum Essential Media (GMEM) (BHK-21, Gibco 21710-25) supplemented with 1x glutamine (Gibco, 25030-024), 1x sodium pyruvate (ThermoFisher, 11360-039), 1x MEM non-essential amino acids (Gibco, 11140-035), 1x penicillin-streptomycin (Gibco, 15070-063), 10% Qualified FBS (Gibco, 10370106), 0.1 mM β-mercaptoethanol, and 1200 units of LIF (Cambridge Bioscience, GFM200). The mESCs were seeded onto a 0.1% gelatin-coated 6-well plate at 3 x 10^5 cells per well. After 24 hours (~70% confluency), the cells were co-transfected with 3.3 μg of pX458-SpCas9-Ptch1 and 100 μM single-stranded oligonucleotide donor (5'-CACGGGAGCCCAGAATTACCATCAGTAGGTAGCCGCTGGCCACTCGGATGACAC TGACATCAGAG**G**AGGATTTTAGGATGTCGTC**T**AGGGTCGTGGTTGTGAAGGGAAG CACCTTTTGAGTGGAGT-3'; the bold and underscored G denotes the variant base while the bold and underscored T indicates the site of introduction of a nucleotide to disrupt the PAM site and facilitate variant screening by restriction digest) using 20 μl FuGene HD reagent diluted in Opti-MEM Reduced Serum Medium (31985-062). After 24 hours, the cells were split (0.25% Trypsin-EDTA Solution; Sigma, T4049) and sorted by fluorescence-activated cell sorting (FACS, FACSymphony S6) using GFP as a marker

for transfection. Single cells (GFP+) were sorted into wells of a 96-well flat-bottomed plate and cultured until 70% confluency was achieved, before splitting into three replicates (one plate for DNA extraction, one for storage at -80°C, and one for continued growth).

7.2.9.3 Colony screening

For DNA extraction, colonies growing in 96-well plates were left until confluent before the media was removed and replaced with 100 μl of lysis buffer (Tris 100 mM [pH 8.5], EDTA 5 mM, SDS 0.2%, NaCl 200 mM), containing 10 μg/ml of proteinase K, for 4 hours. The lysed cells were transferred to a V-shaped 96-well plate and gently mixed with 200 μl of ice-cold isopropanol, before centrifugation at 1680 *g* for 10 min. The supernatant was removed by plate inversion before adding 50 μl of 70% EtOH and mixing for a few seconds. The plate was again spun at 1680 *g* for 10 min before removing the ethanol and allowing to air-dry. The DNA pellets were resuspended in 50 μl of Tris-EDTA (TE) buffer and allowed to gently shake for 15 min before screening by PCR (forward: 5'-GGACCTCGTGTTGCTGTATCAGC-3', reverse: 5'-CAACAGGGTGGAGACCACTTCTAA-3'). For each reaction, 10 μl of 2x Phire hot start master mix (ThermoFisher, F125S), 6.5 μl of nuclease-free water, and 1 μl of each 10 μM primer were mixed with 20 ng of DNA. Reactions were placed in the thermocycler at 98°C for 30 s, followed by 35 cycles of 98°C for 5 s, 58°C for 10 s and 72°C for 15 s, and a final extension of 72°C for 1 min. Any clones showing a single PCR amplicon were subjected to restriction digest using BfaI (NEB, R0568S) – 1 μg of product was digested with 1 μl of BfaI, 1x CutSmart buffer and up to 25 μl of nuclease-free water for 1 hour at 37°C. Successful clones identified through PCR and digest were sent for dideoxy-sequencing for variant confirmation. Clones confirmed through dideoxy-sequencing were amplified using primers for a ~1 kb and ~3.5 kb product (1 kb and 3.5 kb forward, respectively: 5'-CCCACACTGGGTGAACAAGAAAAGG-3', 5'-

GGTCTGCCCTTACATAGACTAATGGGA-3'; 1 kb and 3.5 kb reverse, respectively: 5'-GCCAGAGGACAGCAACGAATGG-3', 5'-GCACACAGCGGTGGTTTCTAATACT-3') to assess any local off-target effects. For these primers, the elongation step was increased to 1 min. Following screening, cells were karyotyped by the WIMM Transgenics Facility.

7.2.10 Mouse model

The successful clone (NM_008957.3:c.1259T>C; p.(Phe420Ser)) was injected into blastocysts (C57BL/6) for generation of a chimera, before backcrossing to the C57BL/6J background.

7.3 Results

7.3.1 Case Report (III-1)

The proband (III-1) is the first-born child to parents of consanguineous union (first cousins) of Pakistani descent. Both parents were healthy and there is no known family history of congenital disorders. In particular, no features of NBCCS were elicited, either on direct questioning or upon examination of the skin (including palmar and plantar surfaces) face or teeth, with the exception of a bifid right lateral incisor present in the father. The child displayed a complex head shape suggestive of multi-suture craniosynostosis (figure 7.2a). The posterior fontanelle and left lambdoid sutures were noted to be abnormally positioned, and wide, open bony defects were documented on the right parietal and left coronal regions. Bulging was present above the left coronal suture. An MRI scan revealed an absent corpus callosum and she also displayed ventriculomegaly and hydrocephalus. Additional craniofacial features included pseudoproptosis, nystagmus, low set ears, and cleft palate. She has visual impairment, a normal right audiogram, and a suspected left aural atresia. Developmental delay was documented from a young age. In addition, she showed incomplete simple

polysyndactyly of the 1st, 2nd and 3rd web on both hands and feet (figure 7.2b,c), and postaxial polysyndactyly of both hands (figure 7.2c) alongside increased hair growth on the extensor surfaces. Moreover, a large midline epigastric hernia was observed (figure 7.2d). At 2 months of age, she underwent surgery for insertion of a primary ventriculoperitoneal shunt to relieve intracranial pressure.

7.3.2 Case Report (III-2)

After one year, the family were planning a second child and were counselled that with the known consanguinity the recurrence risk would be 25%. During their second pregnancy, a transabdominal ultrasound scan was performed at 21 weeks' gestation, which showed early indications of craniosynostosis, abnormalities of the brain (severe bilateral ventriculomegaly with dilatation of the third ventricle, hydrocephalus, macrocephaly, and agenesis of the corpus callosum), nuchal oedema, and postaxial polydactyly of the hands and feet. A normal heart, thorax, abdominal wall, gastrointestinal tract, urogenital tract, and spine were noted. The phenotypic features showed striking similarities with the patient's older sister and strongly suggested that they shared the same condition. The parents made an informed decision to continue the pregnancy.



Figure 7.2 Phenotype of the index patient (III-1). (A) Side, front, and aerial profile of the index patient (III-1) at 2 months of age showing a dysmorphic facial appearance, cleft palate, downward slanting palpebral fissures, and low set ears. (B) Right and left foot showing polysyndactyly. (C) Syndactyly of the fingers. (D) Epigastric hernia.

The second child (III-2), also female, was born through a planned Caesarean section at 41 weeks and was noted to have a clover-leaf shaped head, low-set and malformed ears, clouding of the corneas, a broad flattened nasal bridge, syndactyly of both hands and feet, and microphthalmia. An MRI scan showed the presence of an enlarged cranium, dysplastic left plagiocephaly, and a degree of trigonocephaly; she was also shown to have ventriculomegaly with a narrow aqueduct and agenesis of the corpus callosum. A CT scan confirmed the premature fusion of the metopic suture and partial fusion of the left and right coronal sutures. The right lambdoid and squamous sutures were broadly open, and the axis of the sagittal suture was documented to be displaced. Her nasal bridge was flattened, and the nasal passage was narrow posteriorly, but patent. An echocardiogram showed a trace of aortic regurgitation, which was not considered to be clinically significant. She was noted at discharge to have an epigastric hernia and umbilical granuloma. At one month of age, she was diagnosed with pyloric stenosis and a pyloromyotomy was performed. At two months of age, she was documented with

hydrocephalus and underwent an operation to insert a right ventriculoperitoneal shunt. Her weight was on the 25th centile and her head circumference over the 99.6th centile. She displayed twitching seizures which improved after surgery, but she continued to show ongoing periods of stiffness and vacancy and was prescribed Keppra to manage symptoms. At 7 months old her parents reported daily seizures, lasting ~10 seconds or longer, and prolonged periods of drowsiness potentially due to being postictal for a large proportion of the day. A CT scan showed a reduced ventricular size, Chiari 1 malformation, a tight posterior fossa, dysmorphic brain appearance, and low positioned tonsils. She showed limited interaction with toys or her parents, raising concerns over vision, but had no evidence of any hearing abnormalities.

7.3.3 RT-PCR did not identify any splicing defects in *RAB23*

Based on a diagnosis of Carpenter syndrome, initial genetic testing by the diagnostic laboratories focused on identifying a likely pathogenic variant in *RAB23* but did not identify any sequence variation within the coding regions. Subsequently, RNA was obtained from lymphoblastoid cells from two control samples, both parents, and the proband (III-1) for assessment of aberrant splicing defects in *RAB23*. Primers were designed to span multiple exons and the cDNA product was analysed by gel electrophoresis (figure 7.3). Although it was not possible to amplify the 5' UTR and non-coding exon 1, analysis of products from exon 2 to the 3' UTR showed no difference between the controls, parents (II-1 and II-2), and the index patient's sample (III-1).

7.3.4 WGS identified a single homozygous missense variant in PTCH1

Following this, blood DNA from the proband was sent for WGS. After filtering by allele frequency and homozygosity, 20 variants remained with the most plausible candidate being a homozygous missense variant in *PTCH1* (NM_000264.5:c.1301T>C; p.(Phe434Ser)) (table 7.1), which intersected the 7th largest ROH (spanning ~12 Mb)

(figure A7.1). Importantly, both *RAB23* and *MEGF8*, for which recessive variants are known to cause Carpenter syndrome, did not locate in regions of autozygosity. Additionally, the analysis of heterozygous variants in *GLI3* (for which dominant mutations are less frequently associated with Carpenter syndrome) did not identify any likely pathogenic variant. Subsequently, the *PTCH1* variant was confirmed by dideoxy-sequencing of gDNA (figure 7.4b). Cord blood was obtained for analysis of *PTCH1* in the affected sibling (III-2); this showed segregation of the c.1301T>C variant (figure 7.4b), adding further evidence to support a causal role of the homozygous missense variant in PTCH1.

The missense variant in PTCH1 (p.(Phe434Ser)) is unreported in gnomAD v2.1.1 (considering ~250,000 alleles) and the UK Biobank (considering ~735,000 alleles) and is highly conserved over evolutionary time (figure 7.5). Amino acid 434 lies within the neck helix 2 region (figure 7.5) and is anticipated to interact with the palmitoyl moiety of the PTCH1 ligand, SHH, (Qian et al., 2019, Gong et al., 2018, Qi et al., 2018) but also is suspected to be important for cholesterol binding (Radhakrishnan et al., 2020) (figure 7.6). A change from phenylalanine at this position to serine would be predicted to encode a much smaller and polar side chain, potentially disrupting interaction with its ligand and any free cholesterol molecules.

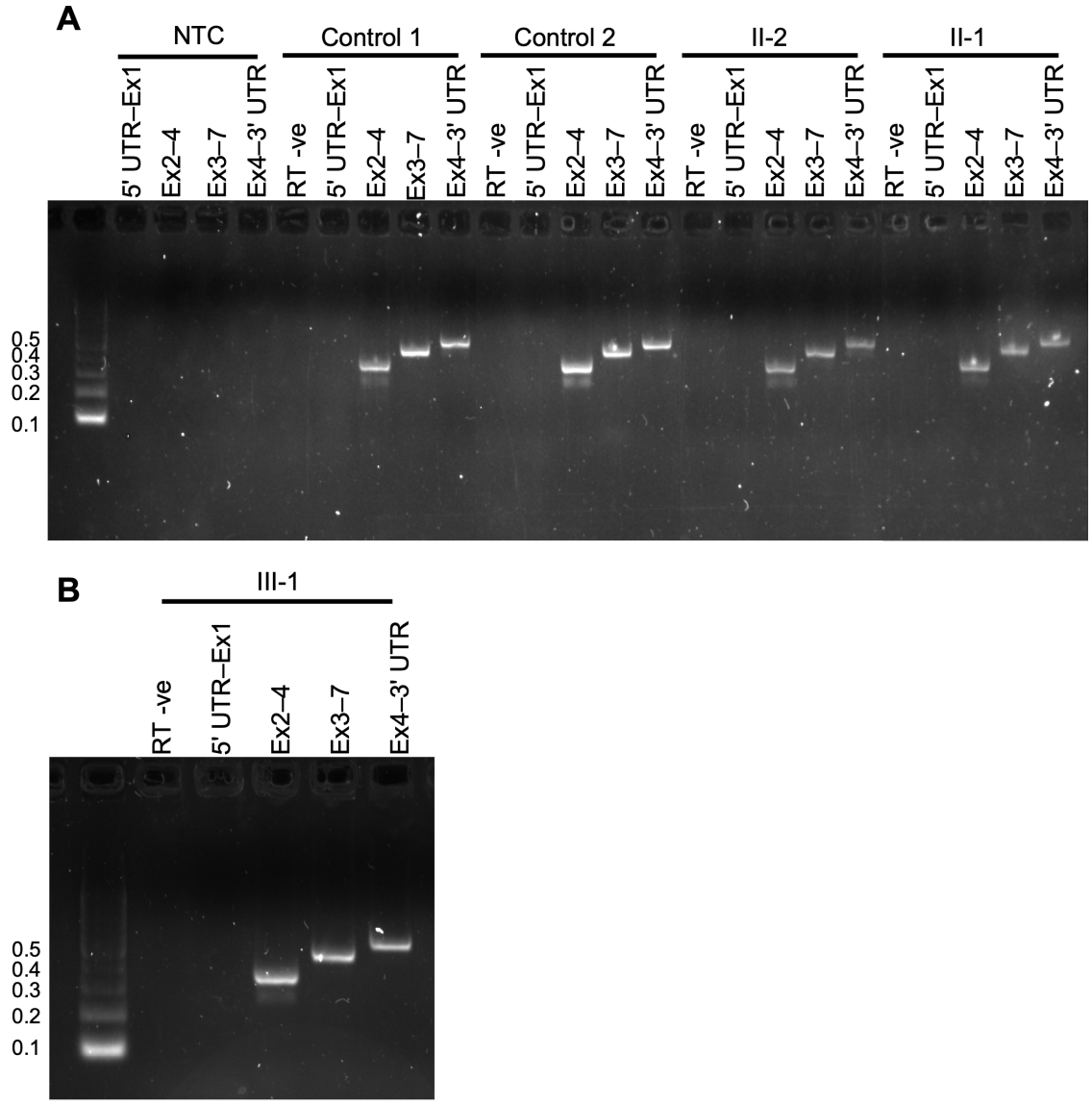


Figure 7.3 RT-PCR analysis of *RAB23*. (A–B) Amplification of RNA products in *RAB23*, considering three regions: exons 2–4 (Ex2–4), exons 3–7 (Ex3–7), and exon 4–3' UTR (Ex4–3' UTR); it was not possible to amplify a product between the 5' UTR and exon 1. (A) cDNA from two control samples, and both parents (II-1 and II-2); NTC = no template control. (B) cDNA from the affected proband (III-1).

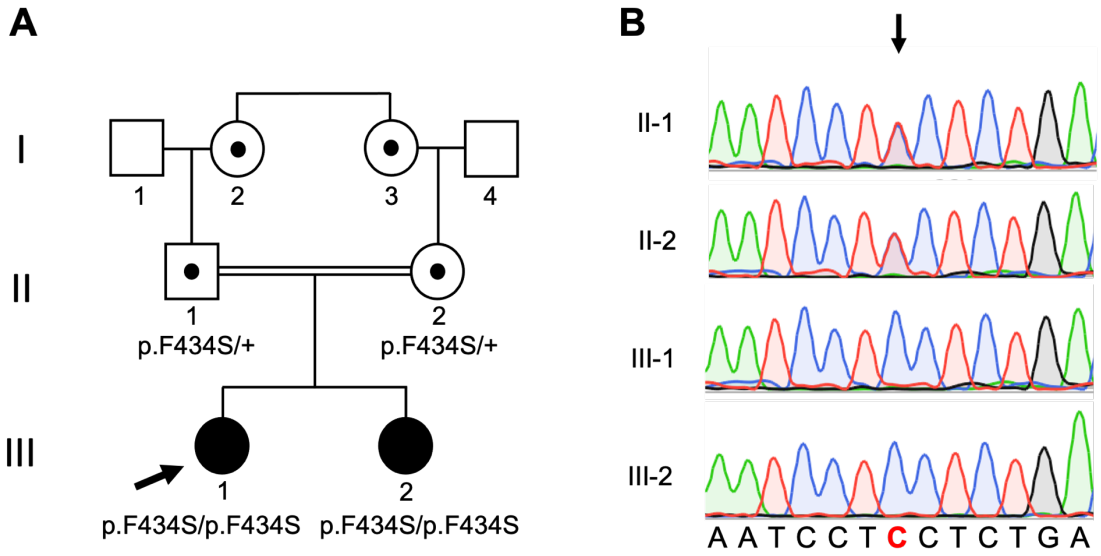


Figure 7.4 Pedigree and dideoxy-sequence confirmation of NM_000264.5:c.1301T>C. (A) Pedigree diagram displaying both heterozygous parents and their affected children. The black arrow indicates the proband. (B) Dideoxy-sequencing traces of both parents and affected children, confirming the presence of the heterozygous c.1301T>C; p.(Phe434Ser) variant in both parents (II-1 and II-2) and evidence of homozygosity in the affected proband (III-1) and her sibling (III-2). The black arrow indicates the position of the nucleotide change and the red lettering highlights the nucleotide that differs from the wildtype sequence. Note, the peak height of the variant allele (C in blue) overlaps the wildtype allele (T in red) in both heterozygous parents.

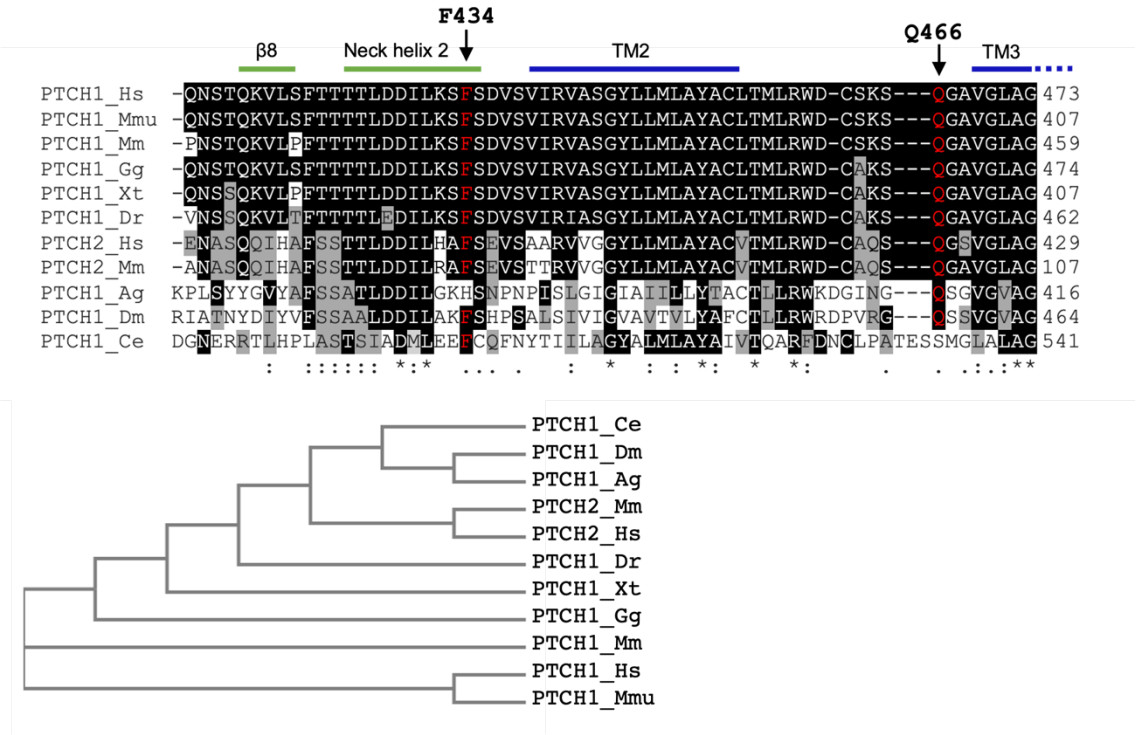


Figure 7.5 Conservation of variant residues. Top, a multiple-protein sequence alignment of PTCH1 and PTCH2. Abbreviations are as follows: Hs = human, Mmu = rhesus monkey, Mm= mouse, Gg = chicken, Xt = tropical clawed frog, Dr = zebrafish, Ag = mosquito, Dm = fruit fly, Ce = nematode worm. The position of the two variants is indicated by the black arrow and above label (F434 or Q466). Secondary structural elements are annotated including β-sheet 8, alongside the neck helix 2 region and two transmembrane (TM) domains (TM2 and TM3). Black or grey shaded amino acids indicated identical or similar residues to the human PTCH1 sequence, respectively. Conservation of amino acids at both variant positions are highlighted in red. For each alignment, an asterisk (*) denotes full conservation across all aligned sequences, a colon (:) indicates conservation of residues with similar biochemical properties, and a period (.) annotates conservation of residues with weakly similar biochemical properties. Below, a phylogenetic tree of the aligned PTCH1 sequences.

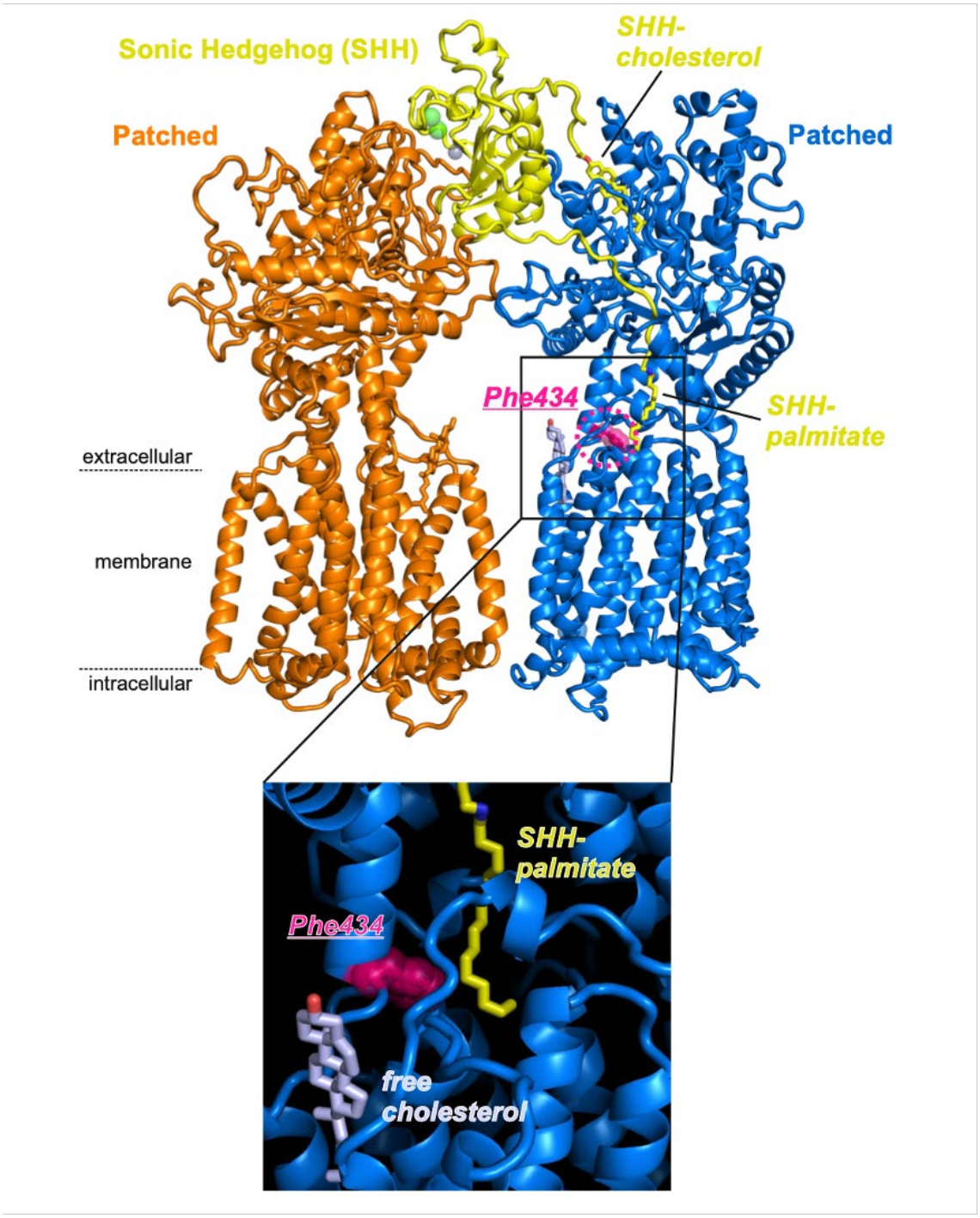


Figure 7.6 Structure and position of PTCH1 Phe434. Top, binding of SHH (yellow) to two PTCH1 molecules. The position of Phe434 in the neck helix is shown in pink. Below, zoomed in 3D structure of Phe434 in proximity to the palmitate moiety in SHH and a free cholesterol molecule. Images obtained from the Siebold lab.

Table 7.1 Variants remaining after filtering WGS data

Gene	Nucleotide change	Protein change	Genomic coordinate (GRCh38)	WT/ variant reads (% variant)	gnomAD AF (v2.1.1)	CADD	Gene-disease association (mode of inheritance)	Segregation (dideoxy sequencing)	Comments
ACY1	NM_000666.3:c.242C>T	p.(Thr81Met)	Chr3:51985443	0/39 (100%)	6.17e-4	16	Aminoacylase 1 Deficiency (biallelic)	-	1 homozygote reported in gnomAD.
ANGEL2	NM_001300755.2:c.658_659insTT	p.(Glu221Leufs*19)	Chr1:213007116	2/20 (91%)	-	-	-	-	Intersects a repeat masker element. Identified as a common SNP.
CCDC180	NM_001348010.4:c.2331_2332insTAGGAG	p.(Ser778*)	Chr9:97330692	0/13 (100%)	-	-	-	-	Intersects a region highlighted as common variation in dbSNP.
CCDC71	NM_022903.4:c.1022A>G	p.(Lys341Arg)	Chr3:49163187	0/39 (100%)	2.36E-4	5.7	-	-	Low CADD score and 59 heterozygotes reported in gnomAD.
ELOVL7	NM_001104558.2:c.17T>C	p.(Leu6Pro)	Chr5:60784102	0/27 (100%)	4.637E-05	18.8	-	-	6 heterozygotes reported in gnomAD. This residue shows low conservation in vertebrates and is predicted benign.
FEZ1	NM_005103.5:c.86G>A	p.(Cys29Tyr)	Chr11:125489692	0/34 (100%)	-	11.1	-	-	Low conservation of this residue in vertebrates and low CADD score.
HEPACAM	NM_001411043.1:c.115C>T	p.(Pro39Ser)	Chr11:124925040	0/37 (100%)	2.15E-4	23.4	Mega-lencephalic leukoencepha lo-pathy with subcortical cysts (biallelic)	-	59 homozygotes in gnomAD.

IL16	NM_001172128.2:c.3150C>G	p.(Asn1050Lys)	Chr15:81301344	0/36 (100%)	2.3E-4	23.7	-	Heterozygous in the affected sibling and both parents; homozygous in the proband	61 heterozygotes reported in gnomAD; the variant displays a CADD score above 20. There is absence of homozygosity in the affected sibling.
PLPP7	NM_001411021.1:c.251C>T	p.(Ser84Phe)	Chr9:131290248	0/32 (100%)	8.00E-6	25.5	Associated with myopathy-related disorders	Homozygous in both affected siblings and the unaffected father; heterozygous in the mother	There are two heterozygotes reported in gnomAD. The variant does not intersect any functional domain but does show a high CADD score. Homozygous knockout in mice shows no craniofacial abnormalities. RNA analysis identifies *PLPP7* predominantly expressed in striated muscle (particularly heart, skeletal muscle, and the tongue). The homozygous variant is identified in the unaffected father.
PTCH1	NM_000264.5:c.1301T>C	p.(Phe434Ser)	Chr9:95478101	0/28 (100%)	-	31	Nevoid Basal Cell Carcinoma Syndrome (monoallelic)	Heterozygous in both parents; homozygous in both affected children	Repressor of hedgehog signalling. This variant is absent from gnomAD and within a functionally relevant binding pocket for cholesterol and the SHH ligand. PTCH1 is widely expressed in all tissues and the null phenotype in mice shows abnormal craniofacial development.
RGS7BP	NM_001029875.3:c.277A>G	p.(Thr93Ala)	Chr5:64507822	0/28 (100%)	1.99e E-5	24.3	-	Heterozygous in the affected sibling and both unaffected parents; homozygous in the proband	There are five heterozygotes reported in gnomAD. The null phenotype in mice (MGI) shows no phenotypic abnormality (including morphology or neurological). Tissue expression analysis predominantly identifies *RGS7BP* expressed in the brain, indicating that mutations are likely to cause a neuronal defect rather than craniofacial. Variant is not identified in homozygosity in the affected sibling.

SLC6A19	NM_001003841.3:c.535G>C	p.(Glu179Gln)	Chr5:121235 6	0/31 (100%)	3.19E-5	19.6	Hartnup disorder (biallelic), Hyperglycinuria (monoallelic), Iminoglycinuria (biallelic)		There are eight heterozygotes in gnomAD. Variants in *SLC6A19* cause Hartnup disease. Null mutants in mice phenocopy human Hartnup disease (adipose, digestive, decreased body size, anomalies of homeostasis and renal/ urinary systems (Broer et al., 2011)).
SPTLC1	NM_001281303.2:c.888T>C	p.(=)	Chr9:92049960	0/35 (100%)	3.25E-5	7.3	Neuropathy (monoallelic)		There are 49 heterozygotes in gnomAD. This variant is annotated as likely benign in ClinVar. Variants in this gene are associated with neuropathy.
THSD7B	NM_001316349.2:c.3664A>G	p.(Met1222Val)	Chr2:137618490	0/31 (100%)	7.92E-4	4.8	-		3 homozygotes reported in gnomAD.
TOR1B	NM_001317893.2:c.549C>G	p.(His183Gln)	Chr9:129807271	0/32 (100%)	9.03E-4	16.6	-		4 homozygotes reported in gnomAD.
TSPEAR	NM_144991.3:c.265A>G	p.(Ile89Val)	Chr21:44567823	0/36 (100%)	-	19.2	Deafness (biallelic), Ectodermal dysplasia (biallelic), Tooth agenesis (biallelic)	Heterozygous in both parents, homozygous in both affected siblings	There are eight heterozygotes reported in gnomAD. Pathogenic variants affect the EPTP domain whereas this residue does not intersect a functional domain. Variants in this gene are associated with deafness/ ectodermal dysplasia which are not evident in either affected children.
TWF2	NM_007284.4:c.35A>G	p.(Glu12Gly)	Chr3:52235055	0/17 (100%)	8.14E-4	28.9	-		1 homozygote in gnomAD.
TWF2	NM_007284.4:c.77G>A	p.(Arg26Gln)	Chr3:52235097	1/22 (96%)	8.71E-4	24.6	-		1 homozygote in gnomAD.
WHAMM	NM_001080435:c.1961C>T	p.(Pro654Leu)	Chr15:82830918	0/14 (100%)	8.00E-4	9.2	-		1 homozygote in gnomAD.
ZNF782	NM_001001662:c.961C>T	p.(Arg321Cys)		0/49 (100%)	7.29E-4	4	-		1 homozygote in gnomAD.

7.3.5 Analysis of patient RNA and protein from lymphoblastoid and fibroblast cells

Analysis of the effect of the *PTCH1* variant (c.1301T>C) on RNA was attempted by qPCR analysis of patient lymphoblastoid and fibroblast cells. Primers were designed to amplify transcripts within the hedgehog signalling pathway (*PTCH1, SMO, GLI1, GLI3*, and *HHIP*) of a control lymphoblastoid and fibroblast sample. However, the expression levels of the targets were too low in both tissues to assay, and it therefore would have been difficult to draw conclusions from the proband sample or their parents. Following this, protein stability was assayed using a western blot which showed a stable product in all controls, parents and proband (figure 7.7); this result was confirmed by biological and technical replicates (figure A7.2).

7.3.6 A homozygous PTCH1 missense variant in a patient with intellectual disability

In an attempt to identify further cases to corroborate the initial finding, I screened the 100kGP rare disease cohort for recessive variants in *PTCH1*. This analysis identified one patient enrolled under the normalised disease grouping of 'Intellectual Disability' with a homozygous missense variant (NM_000264.5:c.1397A>T; p.(Gln466Leu)) (figure A7.3). The variant is reported in one heterozygous allele in gnomAD (v2.1.1) but is not described in the UK Biobank data considering ~735,000 alleles and is conserved in both human and mouse PTCH proteins, and *Drosophila* Ptc (figure 7.5). Structural analysis suggests that it positions near to TM3 and does not display a key functional role in binding the SHH ligand, nor essential cholesterol moieties.

7.3.7 Cohort screening of unsolved patients

Additionally, 617 patients without a genetic cause for their craniosynostosis were screened for variants in *PTCH1* (see chapter 4 for further detail). While two novel

heterozygous variants were identified in the resequencing cohort (table A4.4), these were later discounted from further analyses: one missense variant encoding p.(Gln145Arg) affected an alternatively spliced exon and neighbouring heterozygous variants were reported as benign in ClinVar; the remaining variant was predicted to encode an in-frame insertion but was shown to be a sequencing error by analysis of reads on IGV. No additional biallelic variants were identified in *PTCH1* in the resequencing cohort.

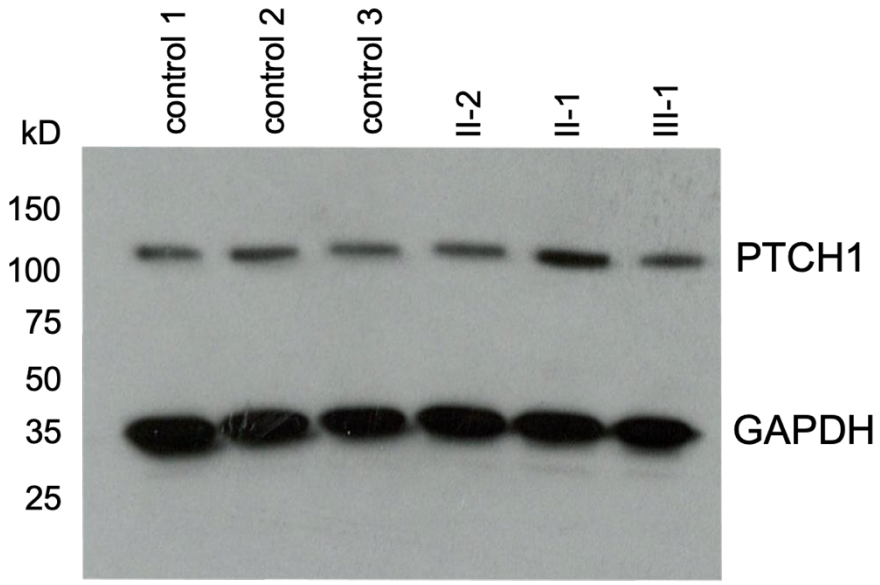


Figure 7.7 Protein analysis of lymphoblastoid patient cells. Western blot analysis of PTCH1 and GAPDH protein in three control lymphoblastoid cell lines, both parents (II-1, II-2) and the proband (III-1).

7.3.8 Assessment of the effect of p.(Phe434Ser) in a PTCH1 expression vector

Both identified homozygous variants in PTCH1 (c.1301T>C [encoding p.(Phe434Ser)] and c.1297A>T [encoding p.(Gln466Leu)]) were cloned into lentiviral vectors containing the full-length *PTCH1* (*PTCH1-FL*) or a vector with increased expression of *PTCH1* containing a deletion of aa 630–717 (*PTCH1ΔICL3*), encoding an inner circular loop. Both the *PTCH1-FL* and the *PTCH1ΔICL3* constructs were designed to contain a C-

terminal AVI or mVenus tag for binding and fluorescence assays, respectively. These constructs were given to the Siebold lab (University of Oxford) and the Rohatgi lab (Stanford University) for analysis of the ability of both p.(Phe434Ser) and p.(Gln466Leu) variant constructs to bind the SHH ligand or cholesterol moieties and determine expression of downstream targets. Early data by the Siebold lab indicated that both mutants expressed stable levels of protein and retained the ability to bind the SHH palmitoyl moiety, thus suggesting a defect in the ability of the mutant to bind cholesterol. However, further work to confirm this hypothesis is ongoing at Stanford.

7.3.9 Establishing a mouse model of p.(Phe434Ser)

Given the known consangunity, there is a 25% chance that both siblings share many of the homozygous variants identified in this screen and, as such, variant segreation alone is not strong enough evidence that the PTCH1 variant is pathogenic (table 7.1). Therefore, to ascertain if a Carpenter syndrome-like phenotype could be replicated in a mouse model with the PTCH1 variant, and thus provide further evidence of variant pathogenicity, mouse embryonic stem cells were targeted to introduce the variant substitution (mouse: c.1259T>C; p.(Phe420Ser)). A total of 82 mE14 cells were successfully targeted and subsequently DNA from these clones was amplified by PCR (figure A7.4). Following this, 21 clones, which showed unique products, were screened by restriction digest for evidence of the introduced SNP at the PAM sequence (encoding a BfaI restriction site). Digested clones were sent for dideoxy sequencing to confirm the presence of the heterozygous variant at c.1259T>C and a homozygous SNP at the PAM site. Subsequently, two clones (2B6 and 2D1) were taken forward for assessment of off-target effects surrounding the variant site, by increasing size of PCR amplicons (378 bp, 1123 bp, 3451 bp) (figure 7.8a), and BfaI restriction digest and dideoxy-sequencing of the ~3.5 kb product (figure A7.5). Karyotyping confirmed that both clones were 90% karyotypically normal and, following blastocyst injection of clone 2D1, 1 male chimera

was born. The male chimera was crossed with a wildtype C57BL/6 female to establish germline transmission. Next, the line will be backcrossed onto the C57BL/6 line before phenotypic assessment of the mouse sutures by computed tomography (CT) scans. Alizarin red and alcian blue staining will also be used for bone and cartilage assessment, respectively. However, owing to timing, this work is beyond the scope of this thesis.

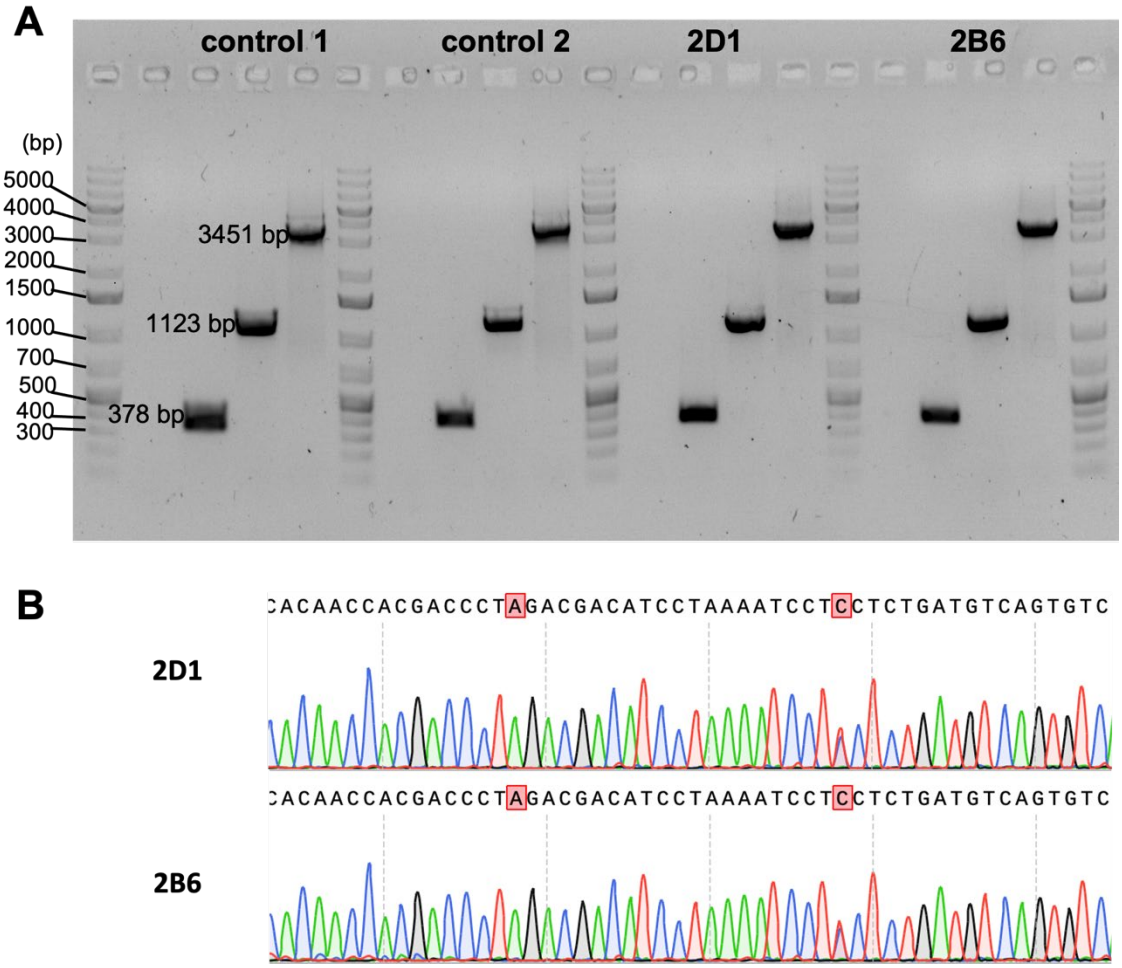


Figure 7.8 PCR analysis of two targeted mE14 clones: 2D1 and 2B6. (A) Increasing size of PCR amplicons were analysed to determine any local off-target effects of CRISPR targeting for two control samples and both targeted clones. In each sample, the first lane is a no template control. (B) Dideoxy-sequencing traces of both clones identifying a homozygous variant at the PAM site (G>A) and a heterozygous substitution at the variant site (T>C); the red box surrounding the nucleotide sequence denotes the variant positions.

7.4 Discussion

Here I present the first family, comprising two affected siblings, who both harbour a rare homozygous missense variant in *PTCH1* resulting in a Carpenter syndrome-like phenotype of multi-suture craniosynostosis, cleft palate, polysyndactyly of the hands and feet, and an epigastric hernia.

7.4.1 Proposed pathophysiological mechanism

The variant, encoding p.(Phe434Ser), affects a residue within the neck helix of the protein, which permits flexibility of the structure to change conformation based on the presence or absence of ligand binding (Gong et al., 2018). This non-conservative change of phenylalanine to serine would encode an amino acid with a smaller side chain and no hydrophobic aromatic ring, thus breaking van der Waals interactions with neighbouring residues (specifically Phe1145) (figure 7.9). Additionally, Phe434 sits in an interesting binding pocket for interaction with both the SHH ligand (Qian et al., 2019) and a key cholesterol molecule. Loss of either of these interactions would have conflicting outcomes for hedgehog signalling. An inability to bind its ligand would allow the normal inhibitory effect of PTCH1 on SMO, meaning that hedgehog signalling would be switched off. In contrast, loss of PTCH1 interaction with a key cholesterol molecule would render the protein unable to perform its normal inhibitory function, and therefore would increase hedgehog signalling.

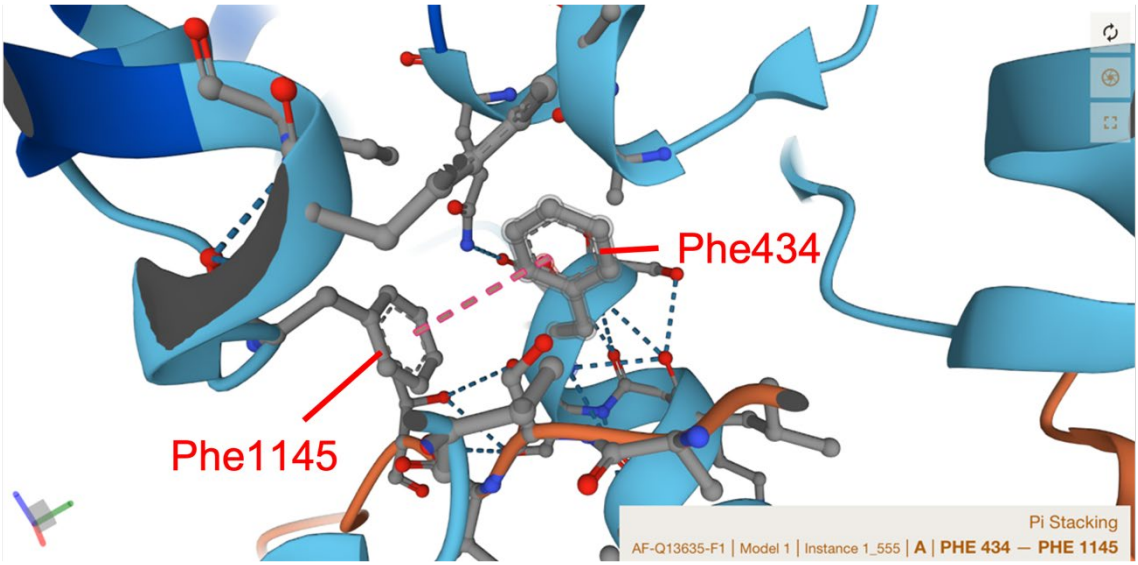


Figure 7.9 Interaction of Phe434 with neighbouring residues. A screenshot from https://alphafold.ebi.ac.uk/entry/Q13635 showing the interaction of Phe434 with neighbouring residues, specifically Phe1145, in a 3D structure.

Although the precise mechanism by which PTCH1 represses SMO is unclear, cumulative evidence suggests that PTCH1 inhibits SMO by diminishing a pool of accessible cholesterol and, in turn, represses hedgehog signalling. In support, overexpression of PTCH1 was shown to efflux a fluorescent analogue of cholesterol (Bidet et al., 2011) and sequence similarity of PTCH1 and NPC1 (a sterol transporter) suggest that both contain a tunnel to facilitate cholesterol transport. Given that Carpenter syndrome is typically characterised by recessive (partial or complete loss-of-function) variants in *RAB23* and *MEGF8* or may occasionally be associated with dominant heterozygous variants in *GLI3* (all negative regulators of active hedgehog signalling), an increase in cholesterol availability to SMO due to the loss of interaction of Phe434 with a key cholesterol molecule, resulting in an increase in hedgehog signalling, appears to be the most likely explanation. In further support, drugs targeting cholesterol trafficking, including astemizole (Hasanovic et al., 2020) and panicein A hydroquinone (Signetti et al., 2020), have been shown to interact with Phe434 and inhibit PTCH1 function.

Interestingly, a patient with metopic synostosis was previously described with a splicing variant which also affected the sterol-sensing domain of PTCH1 (see section 7.1) (Beltrami et al., 2020), suggesting a disturbance in cholesterol transport as a possible pathophysiological mechanism associated with variants in PTCH1 and craniosynostosis. The analysis of additional patients would be required to confirm this hypothesis.

7.4.2 Phenotype considerations

In this report, both parents were heterozygous for the variant allele but displayed no consistent phenotypic abnormalities. Nevertheless, monoallelic variants in *PTCH1* are associated with Gorlin syndrome which is known to present later in life (Martinez et al., 2019). Therefore, it would be important that both individuals are monitored for complications of Gorlin syndrome and take necessary precautions, including sunscreen and UV-protection. Both patients reported in this study display a remarkable similarity to the $Ptch^{DL}$ mouse, generated from a recessive ENU-mutagenesis screen (Feng et al., 2013) including craniosynostosis, polysyndactyly and an epigastric hernia. Additionally, both patients are documented with similarities to the $Ptch1^{mes}$ model (see section 7.1 for further information; table A7.1), the only other described hypomorphic mouse model of Ptch1, most notably the preaxial polydactyly and abnormal craniofacial development (Sweet et al., 1996, Makino et al., 2001).

However, to best of my knowledge, there are currently no animal models investigating the effect of homozygous missense variants in Ptch1. Other reported models include the complete loss of Ptch1 ($Ptch1^{LacZ/LacZ}$) (Goodrich et al., 1997), resulting in embryonic lethality from E8.5 owing to neural tube closure defects (table A7.1). Heterozygotes ($Ptch1^{LacZ/+}$) displayed no notable phenotype, although a small subset (1%) had polydactyly (Goodrich et al., 1997). Crossing the $Ptch1^{LacZ/+}$ mutant with a mouse heterozygous for the hypomorphic 32-bp deletion ($Ptch1^{+/mes}$) (Makino et al., 2001)

increased the severity of the homozygous mes phenotype (*Ptch1*$^{mes/mes}$). The compound heterozygotes (*Ptch1*$^{LacZ/mes}$) were observed at expected Mendelian ratios until E18.5; analysis of the phenotype at E18.5 showed increased body weight, abnormal anteroposterior axis formation of the limb buds, and polydactyly of all four feet (evident from E15.5), alongside ectopic Gli expression (an indicator of active hedgehog signalling) (Makino et al., 2001). Interestingly, there were no significant changes in Ptch1 expression in the double mutant (*Ptch1*$^{LacZ/mes}$) suggesting production of a stable protein but with reduced function. Additionally, intercrossing *Ptch1*^{LacZ} with a *Ptch1*^{DL} mouse resulted in a more severe cranial phenotype than the previously reported compound heterozygous mes model (*Ptch1*$^{LacZ/mes}$), including coronoid hyperplasia, abnormal sphenoid size, lambdoid suture synostosis, and absent presphenoid; the increase in severity is presumably due to the much larger deletion identified in the predominant *Ptch1*$^{DL-b}$ transcript compared to the mes mutant (~870 C-terminal aa deletion [*Ptch1*$^{DL-b}$] and ~150 aa [*Ptch1*^{mes}], see section 7.1) which is predicted to disrupt the sterol sensing domain alongside the remainder of the C-terminus of the protein. The less abundant *Ptch1*$^{DL-a}$ transcripts, which are predicted to encode a 12 aa in-frame insertion within the large intracellular loop of the protein, potentially disrupt *Ptch1*-mediated regulation of cell cycle progression by inhibiting its interaction with cyclin B1 (Barnes et al., 2005). In both allelic combinations (*Ptch1*$^{LacZ/mes}$ and *Ptch1*$^{LacZ/DL}$), in comparison to the homozygous null *Ptch1* mutant (*Ptch1*$^{LacZ/LacZ}$), mice survive until birth suggesting that the amount and function of the protein is sufficient to fulfil critical early developmental aspects of *Ptch1* function. Nevertheless, the phenotypic similarity of the hypomorphic Ptch1 models (*Ptch1*^{DL} and *Ptch1*^{mes}) to mutants displaying aberrant hedgehog signalling (including complete loss of the transcriptional repressor Gli3) (Rice et al., 2010), particularly the abnormal craniofacial development and polydactyly, suggests that there is insufficient protein to permit normal Ptch1 function (i.e., the inhibition of SMO).

There are two further reports of conditional mutants resulting in the deletion of exon 3 (Ellis et al., 2003) and exons 8–9 (*Ptch*^{flox}) (Uhmann et al., 2007). Ubiquitous deletion of exon 3 (using a Cre-recombinase under the transcriptional control of the human cytomegalovirus minimal promoter) resulted in mice with similar phenotypes to the first reported null allele, including neural tube defects and craniofacial abnormalities at E9.5; heterozygotes were indistinguishable from their wildtype littermates. Further information on the previously reported mouse models is provided in table A7.1. Given the absence of reported homozygous missense variants in animal models, I aimed to create a homozygous model of p.(Phe434Ser) in a C57BL/6 mouse, but experiments are ongoing to underpin the phenotype.

7.4.3 A second homozygous missense variant in PTCH1

Analysis of WGS data from the 100kGP identified a second homozygous missense variant in *PTCH1* (c.1397A>T), encoding p.(Gln466Leu), in a patient enrolled under the 'Intellectual Disability' cohort. Although this amino acid is shown to be highly conserved within the human and mouse PTCH family and *Drosophila* Ptc (figure 7.5), 3D structural analysis suggests that it does not interface a ligand or cholesterol binding site and therefore a specific mechanistic function is unknown. Importantly, the phenotype of this patient was not consistent with those described in mouse models with reduced Ptch1 function or with the two patients described in this study, particularly the presence of polydactyly. Despite the lack of evidence for a functional role of Gln466, the strong evolutionary conservation of this residue across the PTCH family and other species suggests that variants at this position could be disease-causing and it is therefore not possible to entirely exclude variant pathogenicity. Experimental work is ongoing in Oxford and Stanford to predict the likely effect of this variant. Nevertheless, the absence of features suggestive of Carpenter syndrome with the p.(Gln466Leu) allele indicates a very

specific mechanism associated with the p.(Phe434Ser) variant, likely due to a specific role in sterol transport.

It is perhaps surprising that additional patients have not already been identified with homozygous variants in *PTCH1* and a phenotype suggestive of Carpenter syndrome, given the high number of patients with heterozygous variants associated with NBCCS across the length of the protein. As such, this suggests some selective disadvantage to homozygosity in *PTCH1*. Excluding known benign polymorphisms there are 59 homozygous variants in ~75,000 individuals in gnomAD v3.1.2. Although this could indicate some tolerance to homozygosity, these variants localise to the alternatively spliced exon 1 or intersect non-functional regions of the protein, which may be more tolerant to variation than the sterol sensing domain (in which Phe434 intersects); it would be interesting to investigate this hypothesis in future studies.

7.4.4 Conclusion

To conclude, I present the first evidence of a homozygous missense variant in *PTCH1* in a sibling pair with a Carpenter syndrome-like phenotype. The patients displayed remarkable similarity to the phenotype reported in mouse models displaying reduced PTCH1 function, including craniosynostosis, polysyndactyly, and an umbilical hernia. The predicted pathophysiological mechanism is an increase in hedgehog signalling, likely due to the loss of interaction of Phe434 with a key cholesterol moiety, which is hypothesised to be essential for its negative regulation of SMO. The variant is unreported in exome and genome databases in gnomAD and the UK Biobank, suggesting that any variation at this position is not well tolerated and that homozygosity would be an extremely rare feature. Further cases would be required to confirm these findings, yet analysis of 617 patients without a genetic diagnosis for their craniosynostosis failed to identify any further individuals with homozygous variants in *PTCH1* (see chapter 4).

Hence the phenotype of the equivalent mutation in the mouse, which has been successfully targeted and backcrossed to establish germline transmission, will be determinative in this matter.

Chapter 8

Conclusions and Final Remarks

In this thesis I have analysed whole genome and exome sequencing data and undertaken targeted resequencing to identify likely pathogenic SNVs in patients with craniosynostosis. In addition, I have investigated the pathophysiological mechanisms associated with variants in three genes, *SPRY1*, *PRRX1* and *PTCH1*, delineating variants in these genes as novel causes of craniosynostosis.

8.1 Identification of diagnostic and candidate variants

8.1.1 Analysis of the 100kGP craniosynostosis cohort

The development of systematic bioinformatic pipelines to investigate the 100kGP data (chapter 3) was successful in identifying a likely pathogenic variant in eight families which were missed by the NHS triaged bioinformatic pipelines. Specifically, *de novo* variant analysis and the identification of failures in the calling of indels led to the greatest increase in diagnoses, accounting for 7/8 (87.5%) of the missing diagnoses found through my pipelines. This work, alongside the work of my colleagues, facilitated important changes to the current NHS England bioinformatic pipelines to augment diagnostic success rate (Hyder et al., 2021). Additionally, analysis of the 100kGP data highlighted seven candidate variants that were further scrutinised in a wider cohort of patients with craniosynostosis, without a genetic diagnosis (chapter 4).

Since our publication on ways to improve the diagnostic success rate of craniosynostosis from the 100kGP data (Hyder et al., 2021), five additional cases would now be considered solved based on new functional evidence, recent publications, and improvements to diagnostic pipelines. This includes, (1) the functional analysis of the stop-gain variant in Spry-1 (chapter 5 of this thesis; (Tooze et al., 2022a)), (2) confirmation of the mosaic *BCL11B* variant (chapter 3, section 3.3.6) and comparison of the phenotype of this patient to other reported cases (unpublished work by the Clinical Genetics Group), (3) identification of additional patients with variants in *ARHGAP35* through GeneMatcher (Lippincott et al., 2022), (4) independent

bioinformatic analyses by Genomics England which identified a *de novo* variant in *HIST1H1E* that was annotated as a “badRead” (meaning that the variant was supported only by reads with low quality bases close to the variant position), and (5) the identification of a copy number variant in *ARX* through use of an alternative algorithm (SVrare; this work was completed by another DPhil student in our lab – Mr Yang Pei). In total this brings the diagnostic success rate to 39/114 (34%) – a 15% increase in under 2 years.

8.1.2 Resequencing analysis

Following the analysis of the 100kGP data, I screened 617 patients (chapter 4) for variants in 42 genes (including the 100kGP candidate genes, genes with more than one reported individual with a putative pathogenic variant, and genes with available supporting animal models), which was successful in identifying six likely pathogenic SNVs and two CNVs, bringing the overall diagnostic success rate to 1.3% (8/617). Although I was unable to corroborate findings from the 100kGP and subsequently delineate a new disease-associated gene, this screen provided a genetic diagnosis for eight families (*ALX4, MSX2*, three *PRRX1*, *SMAD3*, and two *SOX6*) and enhanced ongoing research surrounding variants identified in *PRRX1* (chapter 6) providing a significant uplift to this cohort (3/15 families). Candidate variants identified from this analysis were integrated into our in-house database, CranioDB, which may provide a source of variants to corroborate future findings by the CGG.

Results from the resequencing data in chapter 4, specifically the two *SOX6* and three *PRRX1* findings, contributed towards identifying a subset of genes which should be considered diagnostic-grade Green genes on PanelApp, thus flagging variants in these genes for routine analysis. Notably, the analysis of the diagnostic sensitivity achieved for craniosynostosis from the 100kGP showed that only 47% of variants had been identified through the panel-based approach in use at that time (Hyder et al., 2021). Although some of the missing diagnoses were attributable to failure to call variants included in the contemporaneous Green panel (33%

of missing diagnoses), an additional 22% of diagnoses were missed because the gene was considered an Amber or Red gene at the time of analysis, which did not take account of more recent discoveries indicating that the gene should have been prioritised more highly (Hyder et al., 2021). Following a review of the literature over the last 5 years on patients with craniosynostosis described in exome-, genome-, and panel-based gene analyses (table A8.1), alongside any additional case reports and results from my 100kGP (chapter 3) and resequencing analysis (chapter 4), I identified a further 16 genes (table 8.1) which should be promoted to Green status (Tooze et al., 2023), bringing the total number of Green genes to 81 (a 25% increase). Of the 16 newly proposed genes, four were classified as Amber (*MASP1, NFIA, PRRX1,* and *SOX6*), six genes were annotated as Red (*ADAMTSL4, AHDC1, FBN1, FGF9, KAT6B,* and *NFIX*) and six genes were not reported on PanelApp (*ARID1B, BCL11B, CDK13, FBXO11, IL6ST,* and *MAN2B1*) (Tooze et al., 2023). Further information on current Amber, Red and PanelApp 'null' genes are listed in tables A8.2–3. Since this publication, additional *de novo* variants have been identified in *CDK13*, *NFIX*, and *ARID1B* providing further evidence to support a role for these genes in craniosynostosis (Timberlake et al., 2023b). Importantly, the 16 genes identified in my analyses have since been submitted for review to be included in future updates of the craniosynostosis gene panel on PanelApp.

Table 8.1 Genes with sufficient evidence to be updated to Green PanelApp status

Gene	Current Panel	Mode of Inheritance	Broad Categories of Pathophysiology	Literature
MASP1	Amber	Biallelic	Bone osteogenesis, resorption, and homeostasis	Two reviews identify a prevalence of 27–31% of patients with craniosynostosis and 3MC syndrome and a variant in *MASP1* (Durmaz and Altiner, 2021, Atik et al., 2015).
NFIA	Amber	Monoallelic	Regulator of cell fate and differentiation	Four patients identified in independent screens (Tønne et al., 2022, Chen et al., 2022a, Tønne et al., 2021, Yoon et al., 2020).
PRRX1	Amber	Monoallelic	Regulator of cell fate and differentiation	There are 17 patients from 14 independent families with rare heterozygous variants in *PRRX1,* predicting loss-of-function variants or missense variants affecting the homeodomain (Tooze et al., 2022b).
SOX6	Amber	Monoallelic	Regulator of cell fate and differentiation	Seven independent families with loss-of-function variants in *SOX6* and craniosynostosis; five of these are published (Tolchin et al., 2020, Chen et al., 2022a, Tagariello et al., 2006) and two were identified in a screen of 617 patients without a genetic diagnosis to their craniosynostosis (chapter 4; table A8.2).
ADAMTSL4	Red	Biallelic	Regulator of the extracellular matrix	More than 12 cases of ectopia lentis and craniosynostosis are associated with recessive variants in *ADAMTSL4* (Gustafson et al., 2022).
AHDC1	Red	Monoallelic	Regulator of cell fate and differentiation	There are three individuals reported with bona fide craniosynostosis and variants in *AHDC1* (Gumus, 2020, Miller et al., 2017, Ritter et al., 2018), a further four individuals are described with variants in *AHDC1* and suspected craniosynostosis (Tønne et al., 2021, Danda et al., 2022).
FBN1	Red	Monoallelic	Regulator of the extracellular matrix	There are five likely pathogenic *de novo* variants reported in independent families and one deletion which includes *FBN1* (Adès et al., 2006, Miller et al., 2017, Takenouchi et al., 2013, Hiraki et al., 2008, Yoon et al., 2020).
FGF9	Red	Monoallelic	Regulator of cell fate and differentiation	Three likely pathogenic variants have been reported in independent families; one variant segregates in 12 individuals in the same family and another variant was inherited from an affected father (Sentchordi-Montané et al., 2021, Wu et al., 2009, Rodriguez-Zabala et al., 2017). A missense substitution, p.(Asn143Thr), in murine *Fgf9* results in a phenotype similar to multiple synostoses syndrome 3, with craniosynostosis (Murakami et al., 2002).

KAT6B	Red	Monoallelic	Chromatinopathy	Three loss-of-function variants have been identified in patients with craniosynostosis and a phenotype similar to Lin-Gettig syndrome (Tønne et al., 2021, Bashir et al., 2017).
NFIX	Red	Monoallelic	Regulator of cell fate and differentiation	Four families are reported with variants in *NFIX.* Only one of these variants does not affect a functional domain (p.(Met48Lys)), but it is reported in ClinVar as likely pathogenic for Malan overgrowth syndrome (Tønne et al., 2021, Timberlake et al., 2023a).
ARID1B	Absent	Monoallelic	Chromatinopathy	Four independent families have been described with loss-of-function variants in *ARID1B* and craniosynostosis with developmental delay (Hyder et al., 2021, Chen et al., 2022a, Suzuki et al., 2020, Mignot et al., 2016).
BCL11B	Absent	Monoallelic	Regulator of cell fate and differentiation	There are seven families with variants in *BCL11B* and confirmed craniosynostosis (Goos et al., 2019, Zhao et al., 2022, Eto et al., 2022).
CDK13	Absent	Monoallelic	Cell-cycle regulator/genome stability	Four independent cases identified within the literature in patients with craniosynostosis (Hyder et al., 2021, Bostwick et al., 2017, Tønne et al., 2021).
FBXO11	Absent	Monoallelic	Cell-cycle regulator/genome stability	Three independent cases confirmed in patients (Hyder et al., 2021, Gregor et al., 2018).
IL6ST	Absent	Biallelic	Bone osteogenesis, resorption, and homeostasis	Two cases of recessive variants in *IL6ST* and craniosynostosis with additional animal models which phenocopy the human presentation (Schwerd et al., 2020, Schwerd et al., 2017).
MAN2B1	Absent	Biallelic	Bone osteogenesis, resorption, and homeostasis	Three independent families with recessive variants in *MAN2B1* and craniosynostosis, although not all individuals with recessive variants in *MAN2B1* develop craniosynostosis (Tønne et al., 2021, Hyder et al., 2021, Lipiński et al., 2022).

8.1.3 Avenues for further research

Given the extensive research effort to identify causes of craniosynostosis across the last two decades, the identification of definitively pathogenic variants is increasingly difficult. However, possible avenues to explore further are: (1) mosaicism, by which some variants may be filtered out of bioinformatic pipelines owing to a low percentage of variant reads compared to the reference (as was the case for *BCL11B*; chapter 3, section 3.4.3.3), (2) non-penetrance, requiring a greater number of cases to be confident that variants in that gene are causally associated (for example *PRRX1*; chapter 6), (3) intronic and non-coding variants, for which many algorithms cannot confidentially score their pathogenicity, and (4) complex CNVs, which are routinely missed from pipelines owing to the limit of certain sequencing technologies (however, CNV analysis was not the focus on my thesis). Additionally, genetic diagnoses may be difficult to identify if the phenotype is very specific and the genetic target is small. For example, *MSX2* (267 aa in length), where a specific point mutation at proline 148 causes craniosynostosis; this gene was the first to be described in association with craniosynostosis 30 years ago, yet in this thesis I identified only the second patient in the UK to have a variant at this position. Another consideration is that a higher percentage of cases may be attributable to polygenic effects or environmental interactors for which the percentage contribution of these interactions to the overall phenotype would be difficult to measure.

8.2 Functional validation of variants in three genes

8.2.1 *SPRY1*

The analysis of recessive variants in the 100kGP cohort (chapter 3) facilitated the identification of a homozygous stop-gain in an individual with craniosynostosis, inner ear, and renal anomalies (chapter 5). As discussed in the introduction chapter, genes mutated in craniosynostosis tend to perturb fundamental developmental pathways

(chapter 1, section 1.8.1) and the identification of a variant in Spry-1, a negative regulator of FGF signalling, provides another clear example of this. Although there have been two previous reports of loss-of-function heterozygous variants in *SPRY1* in patients with craniosynostosis, the identification of a patient with a homozygous variant in Spry-1 (p.(Leu27*)), alongside functional evidence showing the complete loss of protein in the proband (chapter 5, section 5.3.2) and reduced levels in both heterozygous parents (who were phenotypically normal), is sufficient to dispute variant pathogenicity of the previously documented heterozygous loss-of-function variants. This chapter was accepted and published without revision in the Journal of Medical Genetics (Tooze et al., 2022a).

Resequencing 617 patients with craniosynostosis (chapter 4) failed to identify any further homozygous variants in Spry-1. One copy number gain (table 4.5) with an additional suspected mosaic frameshift variant (table A4.2) was identified in the same individual but this was considered likely coincidental for a number of reasons (see chapter 4, section 4.3.4). However, absence of additional cases is perhaps not surprising given the rarity of heterozygous loss-of-function alleles in *SPRY1* (12 in 250,000 alleles; gnomAD v2.1.1), making it unlikely that these would occur in homozygosity (unless there is known consanguinity). Additionally, the majority of patients included in the resequencing analysis had non-syndromic single suture synostosis. Given the syndromic presentation of the child with the p.(Leu27*) variant in *SPRY1*, it would not be anticipated that a *SPRY1* variant would be identified in patients with non-syndromic single suture presentations.

8.2.2 *PRRX1*

The pathophysiological mechanism associated with variants in *PRRX1* was investigated in chapter 6 and this work, together with a description of the patient cohort, has been

accepted for publication in Genetics in Medicine. In total, 17 individuals from 14 families with craniosynostosis were identified with likely pathogenic variants affecting *PRRX1*. Through immunofluorescence analyses I was able to ascertain the likely pathogenic mechanism to be the inability of the mutant protein to enter the nucleus. This provided support for all missense variants affecting the homeodomain, leaving the p.(Asp54Ala) variant (the only substitution outside of the homeodomain) as a variant of unknown significance. This analysis suggested that variants in *PRRX1* are a relatively frequent cause of craniosynostosis, but with reduced penetrance (estimated at 12.5%). The wealth of evidence provided for *PRRX1* is sufficient to promote *PRRX1* to a Green PanelApp gene, allowing newly recruited patients to the NHS Genomic Medicine Service to be screened routinely by the NHS diagnostic labs for variants in this gene.

While variants in *PRRX1* have previously been associated with a rare and lethal condition known as agnathia-otocephaly (including two heterozygous loss-of-function variants, one heterozygous missense substitution within the homeodomain, and a homozygous substitution of the OAR domain), my immunofluorescence analysis suggested that an inability of the mutant protein to enter the nucleus could not explain the more severe agnathia-otocephaly phenotype. Additionally, I was unable to replicate the luciferase assays reported in previous papers and identified several errors in the constructs used in these assays. All four reported variants associated with agnathia-otocephaly were annotated in the literature incorrectly and scrutiny of an RNA sequencing dataset from one of the authors showed incomplete data and failed to support variant identification. Given evidence in mouse models that loss of *PRRX2*, in addition to *PRRX1*, can exacerbate the phenotype, it seems plausible that a second hit in either *PRRX2* or a modifier gene could present in the more severe craniofacial phenotype, although further studies would be required to confirm this hypothesis.

8.2.3 *PTCH1*

To date, metopic craniosynostosis has been reported in five independent families with heterozygous loss-of-function variants in *PTCH1*, but currently no patients have been described with homozygous variants in this gene. In chapter 7, I identified a novel homozygous missense variant in *PTCH1*, encoding p.(Phe434Ser), which is suspected to cause Carpenter syndrome in two affected individuals within the same family. Consistent with the known consanguinity, the *PTCH1* missense variant was found to intersect one of the largest regions of homozygosity in the proband and segregated in their affected sibling. Structural analysis of the p.(Phe434Ser) variant suggested that interaction with a key cholesterol molecule would be disrupted, thus rendering the protein unable to suppress SMO and elevating hedgehog signalling. Western blot analysis confirmed that the protein was stable, but qPCR experiments were unable to ascertain the likely effect on downstream hedgehog signalling using patient lymphoblastoid and fibroblast cell lines. Further experiments aiming to underpin the interaction of p.(Phe434Ser) with SHH, or cholesterol molecules, are being undertaken by our collaborators (Siebold Lab, Wellcome Centre for Human Genetics; Rohatgi Lab, Stanford University).

Interestingly, an ENU-mutagenized mouse model (*Ptch1*^{DL}) and a prior screen of recessive causes of mesenchymal dysplasia (*Ptch*^{mes}) identified two hypomorphic Ptch1 mouse mutants which displayed abnormal craniofacial and snout morphology, polydactyly, and, in the *Ptch*^{DL} mutant, an omphalocele; these features are concordant with the human homozygous p.(Phe434Ser) presentation. There are currently no animal models which have investigated the effect of a homozygous substitution since all other reported models result in complete loss of Ptch1 and embryonic lethality. Given the unique position of Phe434, interfacing both the SHH ligand and an essential cholesterol molecule, I aimed to underpin whether homozygous substitution of p.(Phe434Ser) in a mouse model would result in the same phenotypic features seen in our patient and the

previously reported hypomorphic models (*Ptch1*^{DL} and *Ptch1*^{mes}), but the phenotypic characterisation of this mouse is still ongoing.

8.3 Final remarks

In summary, this thesis has provided an overview of the current knowledge of the aetiology of craniosynostosis and ways to identify likely pathogenic SNVs from sequencing data. Building on the availability of current bioinformatic algorithms and the genetic basis for many disorders, I have been able to identify recessive, *de novo*, and dominantly inherited variants in a number of craniosynostosis-associated genes, which have positively impacted the diagnostic success rate for craniosynostosis in the 100kGP. Additionally, candidate genes were highlighted for further scrutiny from the 100kGP and resequencing datasets, which may provide a focus for further exploratory work. Functional validation of variants in *SPRY1*, *PRRX1* and *PTCH1*, using a variety of assays, confirmed the pathophysiological mechanism associated with variants in these genes and has provided substantial evidence to promote the status of *PRRX1* to a Green PanelApp gene. While I have submitted a review for *SPRY1* to be included on PanelApp for craniosynostosis, as there is only one case report, it currently wouldn't meet the criteria for a Green gene. Overall, this work has achieved improved diagnostic pipelines for craniosynostosis, and rare disease more widely, and has also served to delineate several new genetic causes of craniosynostosis by providing a key evidence base to support variant pathogenicity.

References

ABARKAN, M., GAITAN, J., LEBRETON, F., PERRIER, R., JAFFREDO, M., MULLE, C., MAGNAN, C., RAOUX, M. & LANG, J. 2019. The glutamate receptor GluK2 contributes to the regulation of glucose homeostasis and its deterioration during aging. *Mol Metab,* 30**,** 152-160.

ADADEY, S. M., WONKAM-TINGANG, E., ALVES DE SOUZA RIOS, L., ABOAGYE, E. T., ESOH, K., MANYISA, N., DE KOCK, C., AWANDARE, G. A., MOWLA, S. & WONKAM, A. 2022. Cell-based analysis of CLIC5A and SLC12A2 variants associated with hearing impairment in two African families. *Front Genet,* 13**,** 924904.

ADÈS, L. C., SULLIVAN, K., BIGGIN, A., HAAN, E. A., BRETT, M., HOLMAN, K. J., DIXON, J., ROBERTSON, S., HOLMES, A. D., ROGERS, J. & BENNETTS, B. 2006. FBN1, TGFBR1, and the Marfan-craniosynostosis/mental retardation disorders revisited. *Am J Med Genet A,* 140**,** 1047-58.

ADZHUBEI, I. A., SCHMIDT, S., PESHKIN, L., RAMENSKY, V. E., GERASIMOVA, A., BORK, P., KONDRASHOV, A. S. & SUNYAEV, S. R. 2010. A method and server for predicting damaging missense mutations. *Nat Methods,* 7**,** 248-9.

ALBERS, C. A., NEWBURY-ECOB, R., OUWEHAND, W. H. & GHEVAERT, C. 2013. New insights into the genetic basis of TAR (thrombocytopenia-absent radii) syndrome. *Curr Opin Genet Dev,* 23**,** 316-23.

AMANO, K., DENSMORE, M., NISHIMURA, R. & LANSKE, B. 2014. Indian hedgehog signaling regulates transcription and expression of collagen type X via Runx2/Smads interactions. *J Biol Chem,* 289**,** 24898-910.

ANGEL, G., REYNDERS, J., NEGRON, C., STEINBRECHER, T. & MORNET, E. 2020. Large-scale in vitro functional testing and novel variant scoring via protein modeling provide insights into alkaline phosphatase activity in hypophosphatasia. *Human Mutation,* 41**,** 1250-1262.

ANTONOPOULOU, I., MAVROGIANNIS, L. A., WILKIE, A. O. M. & MORRISS-KAY, G. M. 2004. Alx4 and Msx2 play phenotypically similar and additive roles in skull vault differentiation. *Journal of Anatomy,* 204**,** 487-499.

ASHRAF, T., COLLINSON, M. N., FAIRHURST, J., WANG, R., WILSON, L. C. & FOULDS, N. 2015. Two further patients with the 1q24 deletion syndrome expand the phenotype: A possible role for the miR199-214 cluster in the skeletal features of the condition. *Am J Med Genet A,* 167A**,** 3153-60.

ATIK, T., KOPARIR, A., BADEMCI, G., FOSTER, J., 2ND, ALTUNOGLU, U., MUTLU, G. Y., BOWDIN, S., ELCIOGLU, N., TAYFUN, G. A., ATIK, S. S., OZEN, M., OZKINAY, F., ALANAY, Y., KAYSERILI, H., THIEL, S. & TEKIN, M. 2015. Novel MASP1 mutations are associated with an expanded phenotype in 3MC1 syndrome. *Orphanet J Rare Dis,* 10**,** 128.

AUTON, A., BROOKS, L. D., DURBIN, R. M., GARRISON, E. P., KANG, H. M., KORBEL, J. O., MARCHINI, J. L., MCCARTHY, S., MCVEAN, G. A. & ABECASIS, G. R. 2015. A global reference for human genetic variation. *Nature,* 526**,** 68-74.

BABBS, C., STEWART, H. S., WILLIAMS, L. J., CONNELL, L., GORIELY, A., TWIGG, S. R., SMITH, K., LESTER, T. & WILKIE, A. O. 2011. Duplication of the EFNB1 gene in

familial hypertelorism: imbalance in ephrin-B1 expression and abnormal phenotypes in humans and mice. *Hum Mutat,* 32**,** 930-8.

BABURAJENDRAN, N., PALASINGAM, P., NARASIMHAN, K., SUN, W., PRABHAKAR, S., JAUCH, R. & KOLATKAR, P. R. 2010. Structure of Smad1 MH1/DNA complex reveals distinctive rearrangements of BMP and TGF-beta effectors. *Nucleic Acids Res,* 38**,** 3477-88.

BARNES, E. A., HEIDTMAN, K. J. & DONOGHUE, D. J. 2005. Constitutive activation of the shh–ptc1 pathway by a patched1 mutation identified in BCC. *Oncogene,* 24**,** 902-915.

BASHIR, R. A., DIXIT, A., GOEDHART, C., PARBOOSINGH, J. S., INNES, A. M., FERREIRA, P., HASAN, S. U. & AU, P. B. 2017. Lin-Gettig syndrome: Craniosynostosis expands the spectrum of the KAT6B related disorders. *Am J Med Genet A,* 173**,** 2596-2604.

BASKIN, S. M., MORRIS, S. A., VARA, A., HECHT, J. T. & FARACH, L. S. 2020. The first reported case of Loeys-Dietz syndrome in a patient with biallelic SMAD3 variants. *Am J Med Genet A,* 182**,** 2755-2760.

BASSIR, S. H., GARAKANI, S., WILK, K., ALDAWOOD, Z. A., HOU, J., YEH, S. A., SFEIR, C., LIN, C. P. & INTINI, G. 2019. Prx1 Expressing Cells Are Required for Periodontal Regeneration of the Mouse Incisor. *Front Physiol,* 10**,** 591.

BASSON, M. A., AKBULUT, S., WATSON-JOHNSON, J., SIMON, R., CARROLL, T. J., SHAKYA, R., GROSS, I., MARTIN, G. R., LUFKIN, T., MCMAHON, A. P., WILSON, P. D., COSTANTINI, F. D., MASON, I. J. & LICHT, J. D. 2005. Sprouty1 is a critical regulator of GDNF/RET-mediated kidney induction. *Dev Cell,* 8**,** 229-39.

BECKER, J., SEMLER, O., GILISSEN, C., LI, Y., BOLZ, H. J., GIUNTA, C., BERGMANN, C., ROHRBACH, M., KOERBER, F., ZIMMERMANN, K., DE VRIES, P., WIRTH, B., SCHOENAU, E., WOLLNIK, B., VELTMAN, J. A., HOISCHEN, A. & NETZER, C. 2011. Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. *Am J Hum Genet,* 88**,** 362-71.

BELTRAMI, B., PRADA, E., TOLVA, G., SCUVERA, G., SILIPIGNI, R., GRAZIANI, D., BULFAMANTE, G., GERVASINI, C., MARCHISIO, P. & MILANI, D. 2020. Unexpected phenotype in a frameshift mutation of PTCH1. *Mol Genet Genomic Med,* 8**,** e987.

BEN-SALEM, S., SOBREIRA, N., AKAWI, N. A., AL-SHAMSI, A. M., JOHN, A., PRAMATHAN, T., VALLE, D., ALI, B. R. & AL-GAZALI, L. 2016. Gonadal mosaicism in ARID1B gene causes intellectual disability and dysmorphic features in three siblings. *Am J Med Genet A,* 170A**,** 156-61.

BERENGUER, M., DARNAUDERY, M., CLAVEROL, S., BONNEU, M., LACOMBE, D. & ROORYCK, C. 2018. Prenatal retinoic acid exposure reveals candidate genes for craniofacial disorders. *Sci Rep,* 8**,** 17492.

BERNHOFER, M., DALLAGO, C., KARL, T., SATAGOPAM, V., HEINZINGER, M., LITTMANN, M., OLENYI, T., QIU, J., SCHUTZE, K., YACHDAV, G., ASHKENAZY, H., BEN-TAL, N., BROMBERG, Y., GOLDBERG, T., KAJAN, L., O'DONOGHUE, S., SANDER, C., SCHAFFERHANS, A., SCHLESSINGER, A., VRIEND, G., MIRDITA, M., GAWRON, P., GU, W., JAROSZ, Y., TREFOIS, C., STEINEGGER, M., SCHNEIDER, R. & ROST, B. 2021. PredictProtein - Predicting Protein Structure and Function for 29 Years. *Nucleic Acids Res,* 49**,** W535-W540.

BERTOLA, D., YAMAMOTO, G., BUSCARILLI, M., JORGE, A., PASSOS-BUENO, M. R. & KIM, C. 2017. The recurrent PPP1CB mutation p.Pro49Arg in an additional Noonan-like syndrome individual: Broadening the clinical phenotype. *Am J Med Genet A,* 173**,** 824-828.

BIALEK, P., KERN, B., YANG, X., SCHROCK, M., SOSIC, D., HONG, N., WU, H., YU, K., ORNITZ, D. M., OLSON, E. N., JUSTICE, M. J. & KARSENTY, G. 2004. A Twist Code Determines the Onset of Osteoblast Differentiation. *Developmental Cell,* 6**,** 423–435.

BIDET, M., JOUBERT, O., LACOMBE, B., CIANTAR, M., NEHME, R., MOLLAT, P., BRETILLON, L., FAURE, H., BITTMAN, R., RUAT, M. & MUS-VETEAU, I. 2011. The hedgehog receptor patched is involved in cholesterol transport. *PLoS One,* 6**,** e23834.

BLAIR, E. M., WALSH, S., OLDRIDGE, M., WALL, S. A. & WILKIE, A. O. M. 2000. Newly Recognised Craniosynostosis Syndrome That Does Not Map to Known Disease Loci. *American Journal of Medical Genetics,* 95**,** 4-9.

BLUMEL, R., ZINK, M., KLOPOCKI, E. & LIEDTKE, D. 2019. On the traces of tcf12: Investigation of the gene expression pattern during development and cranial suture patterning in zebrafish (Danio rerio). *PLoS One,* 14**,** e0218286.

BOGERSHAUSEN, N. & WOLLNIK, B. 2018. Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. *Front Mol Neurosci,* 11**,** 252.

BONAVENTURE, J., ROUSSEAU, F., LEGEAI-MALLET, L., LE MERRER, M., MUNNICH, A. & MAROTEAUX, P. 1996. Common mutations in the fibroblast growth factor receptor 3 (FGFR3) gene account for achondroplasia, hypochondroplasia, and thanatophoric dwarfism. *American Journal of Medical Genetics,* 63**,** 148-154.

BOSTWICK, B. L., MCLEAN, S., POSEY, J. E., STREFF, H. E., GRIPP, K. W., BLESSON, A., POWELL-HAMILTON, N., TUSI, J., STEVENSON, D. A., FARRELLY, E., HUDGINS, L., YANG, Y., XIA, F., WANG, X., LIU, P., WALKIEWICZ, M., MCGUIRE, M., GRANGE, D. K., ANDREWS, M. V., HUMMEL, M., MADAN-KHETARPAL, S., INFANTE, E., COBAN-AKDEMIR, Z., MISZALSKI-JAMKA, K., JEFFERIES, J. L., ROSENFELD, J. A., EMRICK, L., NUGENT, K. M., LUPSKI, J. R., BELMONT, J. W., LEE, B., LALANI, S. R. & MEMBERS OF THE UNDIAGNOSED DISEASES, N. 2017. Phenotypic and molecular characterisation of CDK13-related congenital heart defects, dysmorphic facial features and intellectual developmental disorders. *Genome Medicine,* 9**,** 73.

BOTSTEIN, D. & RISCH, N. 2003. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Genet,* 33 Suppl**,** 228-37.

BOWDEN, T. A., ARICESCU, A. R., NETTLESHIP, J. E., SIEBOLD, C., RAHMAN-HUQ, N., OWENS, R. J., STUART, D. I. & JONES, E. Y. 2009. Structural plasticity of eph receptor A4 facilitates cross-class ephrin signaling. *Structure,* 17**,** 1386-97.

BOYADJIEV, S. A. 2007. Genetic analysis of non-syndromic craniosynostosis. *Orthod Craniofacial Res,* 10**,** 129-137.

BRAGDON, B. C., BENNIE, A., MOLINELLI, A., LIU, Y. & GERSTENFELD, L. C. 2022. Post natal expression of Prx1 labels appendicular restricted progenitor cell populations of multiple tissues. *J Cell Physiol,* 237**,** 2550-2560.

BRAY, S. J. & KAFATOS, F. C. 1991. Developmental function of Elf-1: an essential transcription factor during embryogenesis in Drosophila. *Genes & Development,* 5**,** 1672-1683.

BRICKMAN, J. M., CLEMENTS, M., TYRELL, R., MCNAY, D., WOODS, K., WARNER, J., STEWART, A., BEDDINGTON, R. S. & DATTANI, M. 2001. Molecular effects of novel mutations in Hesx1/HESX1 associated with human pituitary disorders. *Development,* 128**,** 5189-99.

BRISCOE, J. & THEROND, P. P. 2013. The mechanisms of Hedgehog signalling and its roles in development and disease. *Nat Rev Mol Cell Biol,* 14**,** 416-29.

BROER, A., JUELICH, T., VANSLAMBROUCK, J. M., TIETZE, N., SOLOMON, P. S., HOLST, J., BAILEY, C. G., RASKO, J. E. & BROER, S. 2011. Impaired nutrient signaling and body weight control in a Na+ neutral amino acid cotransporter (Slc6a19)-deficient mouse. *J Biol Chem,* 286**,** 26638-51.

BROUWER, A., TEN BERGE, D., WIEGERINCK, R. & MEIJLINK, F. 2003. The OAR/aristaless domain of the homeodomain protein Cart1 has an attenuating role in vivo. *Mechanisms of Development,* 120**,** 241-252.

BRUGGER, S. M., MERRILL, A. E., TORRES-VAZQUEZ, J., WU, N., TING, M. C., CHO, J. Y., DOBIAS, S. L., YI, S. E., LYONS, K., BELL, J. R., ARORA, K., WARRIOR, R. & MAXSON, R. 2004. A phylogenetically conserved cis-regulatory module in the Msx2 promoter is sufficient for BMP-dependent transcription in murine and Drosophila embryos. *Development,* 131**,** 5153-65.

BUKOWSKA-OLECH, E., GAWLINSKI, P., JAKUBIUK-TOMASZUK, A., JEDRZEJOWSKA, M., OBERSZTYN, E., PIECHOTA, M., BIELSKA, M. & JAMSHEER, A. 2021. Clinical and molecular characterization of craniofrontonasal syndrome: new symptoms and novel pathogenic variants in the EFNB1 gene. *Orphanet J Rare Dis,* 16**,** 286.

BURKARDT, D. D., ROSENFELD, J. A., HELGESON, M. L., ANGLE, B., BANKS, V., SMITH, W. E., GRIPP, K. W., MOLINE, J., MORAN, R. T., NIYAZOV, D. M., STEVENS, C. A., ZACKAI, E., LEBEL, R. R., ASHLEY, D. G., KRAMER, N., LACHMAN, R. S. & GRAHAM, J. M., JR. 2011. Distinctive phenotype in 9 patients with deletion of chromosome 1q24-q25. *Am J Med Genet A,* 155A**,** 1336-51.

BYRNE, E. F. X., SIRCAR, R., MILLER, P. S., HEDGER, G., LUCHETTI, G., NACHTERGAELE, S., TULLY, M. D., MYDOCK-MCGRANE, L., COVEY, D. F., RAMBO, R. P., SANSOM, M. S. P., NEWSTEAD, S., ROHATGI, R. & SIEBOLD, C. 2016. Structural basis of Smoothened regulation by its extracellular domains. *Nature,* 535**,** 517-522.

CAGLAYAN, A. O., PER, H., AKGUMUS, G., GUMUS, H., BARANOSKI, J., CANPOLAT, M., CALIK, M., YIKILMAZ, A., BILGUVAR, K., KUMANDAS, S. & GUNEL, M. 2013. Whole-exome sequencing identified a patient with TMCO1 defect syndrome and expands the phenotic spectrum. *Clin Genet,* 84**,** 394-5.

CALPENA, E., CUELLAR, A., BALA, K., SWAGEMAKERS, S. M. A., KOELLING, N., MCGOWAN, S. J., PHIPPS, J. M., BALASUBRAMANIAN, M., CUNNINGHAM, M. L., DOUZGOU, S., LATTANZI, W., MORTON, J. E. V., SHEARS, D., WEBER, A., WILSON, L. C., LORD, H., LESTER, T., JOHNSON, D., WALL, S. A., TWIGG, S. R. F., MATHIJSSEN, I. M. J., BOARDMAN-PRETTY, F., GENOMICS ENGLAND RESEARCH, C., BOYADJIEV, S. A. & WILKIE, A. O. M. 2020. SMAD6 variants in craniosynostosis: genotype and phenotype evaluation. *Genet Med,* 22**,** 1498-1506.

CALPENA, E., HERVIEU, A., KASERER, T., M.A, S., SWAGEMAKERS, A.C, J., GOOS, POPOOLA, O., ORTIZ-RUIZ, M. J., BARBARO-DIEBER, T., BOWNASS, L., BRILSTRA, E. H., BRIMBLE, E., FOULDS, N., GREBE, T. A., HARDER, A. V. E., LEES, M. M., MONAGHAN, K. G., NEWBURY-ECOB, R. A., ONG, K.-R., OSIO, D., SANTOS, F. J. R., RUZHNIKOV, M. R. Z., TELEGRAFI, A., BINSBERGEN, E. V., DOOREN, M. F. V., STUDY, T. D. D. D., SPEK, P. J. V. D., BLAGG, J., TWIGG, S. R. F., MATHIJSSEN, I. M. J., PAUL A. CLARKE & WILKIE, A. O. M. 2019. De Novo Missense Substitutions in the Gene Encoding CDK8, a Regulator of the Mediator Complex, Cause a Syndromic Developmental Disorder. *The American Journal of Human Genetics,* 104**,** 709-720.

CAMPS, M., NICHOLS, A., GILLIERON, C., ANTONSSON, B., MUDA, M., CHABERT, C., BOSCHERT, U. & ARKINSTALL, S. 1998. Catalytic Activation of the Phosphatase MKP-3by ERK2 Mitogen-Activated Protein Kinase. *Science,* 280**,** 1262-1265.

CANNAERTS, E., KEMPERS, M., MAUGERI, A., MARCELIS, C., GARDEITCHIK, T., RICHER, J., MICHA, D., BEAUCHESNE, L., TIMMERMANS, J., VERMEERSCH, P., MEYTEN, N., CHENIER, S., VAN DE BEEK, G., PEETERS, N., ALAERTS, M., SCHEPERS, D., VAN LAER, L., VERSTRAETEN, A. & LOEYS, B. 2019. Novel pathogenic SMAD2 variants in five families with arterial aneurysm and dissection: further delineation of the phenotype. *J Med Genet,* 56**,** 220-227.

CARBALLO, G. B., HONORATO, J. R., DE LOPES, G. P. F. & SPOHR, T. C. L. D. S. E. 2018. A highlight on Sonic hedgehog pathway. *Cell Communication and Signaling,* 16**,** 11.

CARMICHAEL, S. L., MA, C., RASMUSSEN, S. A., HONEIN, M. A., LAMMER, E. J., SHAW, G. M. & NATIONAL BIRTH DEFECTS PREVENTION, S. 2008. Craniosynostosis and maternal smoking. *Birth Defects Res A Clin Mol Teratol,* 82**,** 78-85.

CARPENTER, D., STONE, D. M., BRUSH, J., RYAN, A., ARMANINI, M., FRANTZ, G., ROSENTHAL, A. & SAUVAGE, F. J. D. 1998. Characterization of two patched receptors for the vertebrate hedgehog protein family. *Proc Natl Acad Sci U S A,* 99**,** 13630–13634.

CELIK, T., SIMSEK, P. O., SOZEN, T., OZYUNCU, O., UTINE, G. E., TALIM, B., YIGIT, S., BODUROGLU, K. & KAMNASARAN, D. 2012. PRRX1 is mutated in an otocephalic newborn infant conceived by consanguineous parents. *Clin Genet,* 81**,** 294-7.

CHASSAING, N., SORRENTINO, S., DAVIS, E. E., MARTIN-COIGNARD, D., IACOVELLI, A., PAZNEKAS, W., WEBB, B. D., FAYE-PETERSEN, O., ENCHA-RAZAVI, F., LEQUEUX, L., VIGOUROUX, A., YESILYURT, A., BOYADJIEV, S. A., KAYSERILI, H., LOGET, P., CARLES, D., SERGI, C., PUVABANDITSIN, S., CHEN, C. P., ETCHEVERS, H. C., KATSANIS, N., MERCER, C. L., CALVAS, P. & JABS, E. W. 2012. OTX2 mutations contribute to the otocephaly-dysgnathia complex. *J Med Genet,* 49**,** 373-9.

CHATRON, N., HADDAD, V., ANDRIEUX, J., DESIR, J., BOUTE, O., DIEUX, A., BAUMANN, C., DRUNAT, S., GERARD, M., BONNET, C., LEHEUP, B., TILL, M., ROSSI, M., FLORI, E., ALEMBIK, Y., STEWART, H., MCPARLAND, J., BERNARDINI, L., CASTELLUCCIO, P., ROOS, L., TUMER, Z., FAGAN, K., HACKETT, A., BAIN, N., VAN HAERINGEN, A., RUIVENKAMP, C., BENZACKEN, B., SANLAVILLE, D., EDERY, P., ABOURA, A. & SCHLUTH-BOLARD, C. 2015. Refinement of genotype-phenotype correlation in 18 patients carrying a 1q24q25 deletion. *Am J Med Genet A,* 167A**,** 1008-17.

CHEN, J., ZHANG, P., PENG, M., LIU, B., WANG, X., DU, S., LU, Y., MU, X., LU, Y., WANG, S. & WU, Y. 2022a. An additional whole-exome sequencing study in 102 panel-undiagnosed patients: A retrospective study in a Chinese craniosynostosis cohort. *Front Genet,* 13**,** 967688.

CHEN, P. C., YIN, J., YU, H. W., YUAN, T., FERNANDEZ, M., YUNG, C. K., TRINH, Q. M., PELTEKOVA, V. D., REID, J. G., TWOROG-DUBE, E., MORGAN, M. B., MUZNY, D. M., STEIN, L., MCPHERSON, J. D., ROBERTS, A. E., GIBBS, R. A., NEEL, B. G. & KUCHERLAPATI, R. 2014. Next-generation sequencing identifies rare variants associated with Noonan syndrome. *Proc Natl Acad Sci U S A,* 111**,** 11473-8.

CHEN, X., ZHU, Y., SONG, C., CHEN, Y., WANG, Y., LAI, M., ZHANG, C. & FANG, X. 2022b. MiR-424-5p targets HSP90AA1 to facilitate proliferation and restrain differentiation in skeletal muscle development. *Anim Biotechnol***,** 1-13.

CHESNEAU, B., EDOUARD, T., DULAC, Y., COLINEAUX, H., LANGEOIS, M., HANNA, N., BOILEAU, C., ARNAUD, P., CHASSAING, N., JULIA, S., JONDEAU, G., PLANCKE, A., KHAU VAN KIEN, P. & PLAISANCIÉ, J. 2020. Clinical and genetic data of 22 new patients with SMAD3 pathogenic variants and review of the literature. *Molecular Genetics & Genomic Medicine,* 8**,** e1132.

CHI, Y. 2005. Homeodomain Revisited: a Lesson from Disease-causing Mutations. *Hum Genet,* 116**,** 433-444.

CHING, S. T., CUNHA, G. R., BASKIN, L. S., BASSON, M. A. & KLEIN, O. D. 2014. Coordinated activity of Spry1 and Spry2 is required for normal development of the external genitalia. *Dev Biol,* 386**,** 1-11.

CHU, S. W., NOYES, M. B., CHRISTENSEN, R. G., PIERCE, B. G., ZHU, L. J., WENG, Z., STORMO, G. D. & WOLFE, S. A. 2012. Exploring the DNA-recognition potential of homeodomains. *Genome Res,* 22**,** 1889-98.

CINCA, K. P., DE PLANQUE, C. A., PETERS, N. C. J., VERSNEL, S. L. & MATHIJSSEN, I. M. J. 2022. Prenatal Ultrasound Parameters of Twins With Sagittal Suture Craniosynostosis Question Mechanical Constraint as the Leading Cause. *J Craniofac Surg,* 33**,** 2350-2353.

CINGOLANI, P., PLATTS, A., WANG LE, L., COON, M., NGUYEN, T., WANG, L., LAND, S. J., LU, X. & RUDEN, D. M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin),* 6**,** 80-92.

COHEN, M., KICHEVA, A., RIBEIRO, A., BLASSBERG, R., PAGE, K. M., BARNES, C. P. & BRISCOE, J. 2015. Ptch1 and Gli regulate Shh signalling dynamics via multiple mechanisms. *Nat Commun,* 6**,** 6709.

CONNERNEY, J., ANDREEVA, V., LESHEM, Y., MUENTENER, C., MERCADO, M. A. & SPICER, D. B. 2006. Twist1 dimer selection regulates cranial suture patterning and fusion. *Dev Dyn,* 235**,** 1345-57.

COPERTINO, D. W., JENKINSON, S., JONES, F. S. & EDELMAN, G. M. 1995. Structural and functional similarities between the promoters for mouse tenascin and chicken cytotactin. *Proc Natl Acad Sci U S A,* 92**,** 2131-2135.

CORNELISSEN, M., OTTELANDER, B., RIZOPOULOS, D., VAN DER HULST, R., MINK VAN DER MOLEN, A., VAN DER HORST, C., DELYE, H., VAN VEELEN, M. L., BONSEL, G. & MATHIJSSEN, I. 2016. Increase of prevalence of craniosynostosis. *J Craniomaxillofac Surg,* 44**,** 1273-9.

COUSIN, M. A., CREIGHTON, B. A., BREAU, K. A., SPILLMANN, R. C., TORTI, E., DONTU, S., TRIPATHI, S., AJIT, D., EDWARDS, R. J., AFRIYIE, S., BAY, J. C., HARPER, K. M., BELTRAN, A. A., MUNOZ, L. J., FALCON RODRIGUEZ, L., STANKEWICH, M. C., PERSON, R. E., SI, Y., NORMAND, E. A., BLEVINS, A., MAY, A. S., BIER, L., AGGARWAL, V., MANCINI, G. M. S., VAN SLEGTENHORST, M. A., CREMER, K., BECKER, J., ENGELS, H., ARETZ, S., MACKENZIE, J. J., BRILSTRA, E., VAN GASSEN, K. L. I., VAN JAARSVELD, R. H., OEGEMA, R., PARSONS, G. M., MARK, P., HELBIG, I., MCKEOWN, S. E., STRATTON, R., COGNE, B., ISIDOR, B., CACHEIRO, P., SMEDLEY, D., FIRTH, H. V., BIERHALS, T., KLOTH, K., WEISS, D., FAIRLEY, C., SHIEH, J. T., KRITZER, A., JAYAKAR, P., KURTZ-NELSON, E., BERNIER, R. A., WANG, T., EICHLER, E. E., VAN DE LAAR, I., MCCONKIE-ROSELL, A., MCDONALD, M. T., KEMPPAINEN, J., LANPHER, B. C., SCHULTZ-ROGERS, L. E., GUNDERSON, L. B., PICHURIN, P. N., YOON, G., ZECH, M., JECH, R., WINKELMANN, J., UNDIAGNOSED DISEASES, N., GENOMICS ENGLAND RESEARCH, C., BELTRAN, A. S., ZIMMERMANN, M. T., TEMPLE, B., MOY, S. S., KLEE, E. W., TAN, Q. K. & LORENZO, D. N. 2021. Pathogenic SPTBN1 variants cause an autosomal dominant neurodevelopmental syndrome. *Nat Genet,* 53**,** 1006-1021.

DANDA, S., DATAR, C., KHER, A., DESHPANDE, T., THOMAS, M. M. & OOMMEN, S. P. 2022. First reported cases with Xia-Gibbs syndrome from India harboring novel variants in AHDC1. *Am J Med Genet A,* 188**,** 2501-2504.

DAS NEVES, L., DUCHALA, C. S., TOLENTINO-SILVA, F., HAXHIU, M. A., COLMENARES, C., MACKLIN, W. B., CAMPBELL, C. E., BUTZ, K. G. & GRONOSTAJSKI, R. M. 1999. Disruption of the murine nuclear factor I-A gene (Nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum. *Proc Natl Acad Sci U S A,* 96**,** 11946-51.

DASOUKI, M., ANDREWS, B., PARIMI, P. & KAMNASARAN, D. 2013. Recurrent agnathia-otocephaly caused by DNA replication slippage in PRRX1. *Am J Med Genet A,* 161A**,** 803-8.

DAVIES, J. P. & IOANNOU, Y. A. 2000. Topological analysis of Niemann-Pick C1 protein reveals that the membrane orientation of the putative sterol-sensing domain is identical to those of 3-hydroxy-3-methylglutaryl-CoA reductase and sterol regulatory element binding protein cleavage-activating protein. *J Biol Chem,* 275**,** 24367-74.

DAWSON, S. J., PALMER, R. D., MORRIS, P. J. & LATCHMAN, D. S. 1998. Functional role of position 22 in the homeodomain of Brn-3 transcription factors. *Molecular Neuroscience,* 9**,** 2305-2309.

DE VRIES, M., CARPINELLI, M., RUTLAND, E., HATZIPANTELIS, A., PARTRIDGE, D., AUDEN, A., ANDERSON, P. J., DE GROEF, B., WU, H., OSTERWALDER, M., VISEL, A., JANE, S. M. & DWORKIN, S. 2020. Interrogating the Grainyhead-like 2 (Grhl2) genomic locus identifies an enhancer element that regulates palatogenesis in mouse. *Dev Biol,* 459**,** 194-203.

DEBNATH, S., YALLOWITZ, A. R., MCCORMICK, J., LALANI, S., ZHANG, T., XU, R., LI, N., LIU, Y., YANG, Y. S., EISEMAN, M., SHIM, J. H., HAMEED, M., HEALEY, J. H., BOSTROM, M. P., LANDAU, D. A. & GREENBLATT, M. B. 2018. Discovery of a periosteal stem cell mediating intramembranous bone formation. *Nature,* 562**,** 133-139.

DECIPHERING DEVELOPMENTAL DISORDERS STUDY 2017. Prevalence and architecture of de novo mutations in developmental disorders. *Nature,* 542**,** 433-438.
DECKELBAUM, R. A., HOLMES, G., ZHAO, Z., TONG, C., BASILICO, C. & LOOMIS, C. A. 2012. Regulation of cranial morphogenesis and cell fate at the neural crest-mesoderm boundary by engrailed 1. *Development,* 139**,** 1346-58.
DI ROCCO, F., BAUJAT, G., ARNAUD, E., RENIER, D., LAPLANCHE, J. L., DAIRE, V. C. & COLLET, C. 2014. Clinical spectrum and outcomes in families with coronal synostosis and TCF12 mutations. *Eur J Hum Genet,* 22**,** 1413-6.
DI ROCCO, F., ROSSI, M., VERLUT, I., SZATHMARI, A., BEURIAT, P. A., CHATRON, N., CHAUVEL-PICARD, J., MOTTOLESE, C., MONIN, P., VINCHON, M., GUERNOUCHE, S. & COLLET, C. 2023. Clinical interest of molecular study in cases of isolated midline craniosynostosis. *European Journal of Human Genetics*.
DJINOVIC-CARUGOA, K., GAUTELB, M., YLA, J. & YOUNGD, P. 2002. The spectrin repeat: a structural platform for cytoskeletal protein assemblies. *FEBS***,** 119-123.
DO, H. J., LIM, H. Y., KIM, J. H., SONG, H., CHUNG, H. M. & KIM, J. H. 2007. An intact homeobox domain is required for complete nuclear localization of human Nanog. *Biochem Biophys Res Commun,* 353**,** 770-5.
DOBSON, S. M., KISS, C., BORSCHNECK, D., HEATH, K. E., GROSS, A., GLUCKSMAN, M. J. & GUERIN, A. 2022. Novel FGF9 variant contributes to multiple synostoses syndrome 3. *Am J Med Genet A,* 188**,** 2162-2167.
DOHERTY, E. S., LACBAWAN, F., HADLEY, D. W., BREWER, C., ZALEWSKI, C., KIM, H. J., SOLOMON, B., ROSENBAUM, K., DOMINGO, D. L., HART, T. C., BROOKS, B. P., IMMKEN, L., LOWRY, R. B., KIMONIS, V., SHANSKE, A. L., JEHEE, F. S., BUENO, M. R., KNIGHTLY, C., MCDONALD-MCGINN, D., ZACKAI, E. H. & MUENKE, M. 2007. Muenke syndrome (FGFR3-related craniosynostosis): expansion of the phenotype and review of the literature. *Am J Med Genet A,* 143A**,** 3204-15.
DONNELLY, M., TODD, E., WHEELER, M., WINN, V. D. & KAMNASARAN, D. 2012. Prenatal diagnosis and identification of heterozygous frameshift mutation in PRRX1 in an infant with agnathia-otocephaly. *Prenat Diagn,* 32**,** 903-5.
DREYER, S. D., MORELLO, R., GERMAN, M. S., ZABEL, B., WINTERPACHT, A., LUNSTRUM, G. P., HORTON, W. A., OBERG, K. C. & LEE, B. 2000. LMX1B transactivation and expression in nail–patella syndrome. *Human Molecular Genetics,* 9**,** 1067-1074.
DURMAZ, C. D. & ALTINER, S. 2021. MASP1-related 3MC syndrome in a patient from Turkey. *Am J Med Genet A,* 185**,** 2267-2270.
ECKERT, M. A., LWIN, T. M., CHANG, A. T., KIM, J., DANIS, E., OHNO-MACHADO, L. & YANG, J. 2011. Twist1-induced invadopodia formation promotes tumor metastasis. *Cancer Cell,* 19**,** 372-86.
EL GHOUZZI, V., MERRER, M. L., LAJEUNIE, F. P.-S. E., BENIT, P., RENIER, D., BOURGEOIS, P., BOLCATO-BELLEMIN, A.-L., MUNNICH, A. & BONAVENTURE, J. 1997. Mutations of the TWIST gene in the Saethre-Chotzene syndrome. *Nature Genetics,* 15**,** 42-46.
ELEFTHERIADOU, O., BOGUSLAVSKYI, A., LONGMAN, M. R., COWAN, J., FRANCOIS, A., HEADS, R. J., WADZINSKI, B. E., RYAN, A., SHATTOCK, M. J. & SNABAITIS, A. K. 2017. Expression and regulation of type 2A protein phosphatases and alpha4 signalling in cardiac health and hypertrophy. *Basic Res Cardiol,* 112**,** 37.

ELLIS, T., SMYTH, I., RILEY, E., GRAHAM, S., ELLIOT, K., NARANG, M., KAY, G. F., WICKING, C. & WAINWRIGHT, B. 2003. Patched 1 conditional null allele in mice. *Genesis,* 36**,** 158-61.

ESCOBAR, L. M., ESCOBAR, J. D., BENDAHAN, Z. & CASTELLANOS, J. E. 2021. Retinoic and ascorbic acids induce osteoblast differentiation from human dental pulp mesenchymal stem cells. *J Oral Biol Craniofac Res,* 11**,** 143-148.

ETO, K., MACHIDA, O., YANAGISHITA, T., SHIMOJIMA YAMAMOTO, K., CHIBA, K., AIHARA, Y., HASEGAWA, Y., NAGATA, M., ISHIHARA, Y., MIYASHITA, Y., ASANO, Y., NAGATA, S. & YAMAMOTO, T. 2022. Novel BCL11B truncation variant in a patient with developmental delay, distinctive features, and early craniosynostosis. *Human Genome Variation,* 9**,** 43.

FAEDO, A., BORELLO, U. & RUBENSTEIN, J. L. 2010. Repression of Fgf signaling by sprouty1-2 regulates cortical patterning in two distinct regions and times. *J Neurosci,* 30**,** 4015-23.

FAN, L., DING, L., LAN, J., NIU, J., HE, Y. & SONG, L. 2019. Fibroblast Growth Factor-1 Improves Insulin Resistance via Repression of JNK-Mediated Inflammation. *Front Pharmacol,* 10**,** 1478.

FARMER, D. T., MLCOCHOVA, H., ZHOU, Y., KOELLING, N., WANG, G., ASHLEY, N., BUGACOV, H., CHEN, H. J., PARVEZ, R., TSENG, K. C., MERRILL, A. E., MAXSON, R. E., JR., WILKIE, A. O. M., CRUMP, J. G. & TWIGG, S. R. F. 2021. The developing mouse coronal suture at single-cell resolution. *Nat Commun,* 12**,** 4797.

FAYE-PETERSEN, O., DAVID, E., RANGWALA, N., SEAMAN, J. P., HUA, Z. & HELLER, D. S. 2006. Otocephaly: report of five new cases and a literature review. *Fetal Pediatr Pathol,* 25**,** 277-96.

FENG, W., CHOI, I., CLOUTHIER, D. E., NISWANDER, L. & WILLIAMS, T. 2013. The Ptch1(DL) mouse: a new model to study lambdoid craniosynostosis and basal cell nevus syndrome-associated skeletal defects. *Genesis,* 51**,** 677-89.

FENN, J. S., LORDE, N., WARD, J. M. & BOROVICKOVA, I. 2021. Hypophosphatasia. *J Clin Pathol,* 74**,** 635-640.

FITZGERALD, K. K., BHAT, A. M., CONARD, K., HYLAND, J. & PIZARRO, C. 2014. Novel SMAD3 Mutation in a Patient with Hypoplastic Left Heart Syndrome with Significant Aortic Aneurysm. *Case Reports in Genetics,* 2014**,** 591516.

FLEMING, N. I., JORISSEN, R. N., MOURADOV, D., CHRISTIE, M., SAKTHIANANDESWAREN, A., PALMIERI, M., DAY, F., LI, S., TSUI, C., LIPTON, L., DESAI, J., JONES, I. T., MCLAUGHLIN, S., WARD, R. L., HAWKINS, N. J., RUSZKIEWICZ, A. R., MOORE, J., ZHU, H. J., MARIADASON, J. M., BURGESS, A. W., BUSAM, D., ZHAO, Q., STRAUSBERG, R. L., GIBBS, P. & SIEBER, O. M. 2013. SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. *Cancer Res,* 73**,** 725-35.

FLORISSON, J. M., VERKERK, A. J., HUIGH, D., HOOGEBOOM, A. J., SWAGEMAKERS, S., KREMER, A., HEIJSMAN, D., LEQUIN, M. H., MATHIJSSEN, I. M. & VAN DER SPEK, P. J. 2013. Boston type craniosynostosis: report of a second mutation in MSX2. *Am J Med Genet A,* 161A**,** 2626-33.

FLÜCK, C. E., TAJIMA, T., PANDEY, A. V., ARLT, W., OKUHARA, K., VERGE, C. F., JABS, E. W., MENDONCA, B. B., FUJIEDA, K. & MILLER, W. L. 2004. Mutant P450 oxidoreductase causes disordered steroidogenesis with and without Antley-Bixler syndrome. *Nat Genet,* 36**,** 228-30.

FONTELES, C. S., FINNELL, R. H., LEI, Y., ZURITA-JIMENEZ, M. E., MONTEIRO, A. J., GEORGE, T. M. & HARSHBARGER, R. J. 2021. De novo ALX4 variant detected in child with non-syndromic craniosynostosis. *Braz J Med Biol Res,* 54**,** e11396.

FOOTZ, T., IDREES, F., ACHARYA, M., KOZLOWSKI, K. & WALTER, M. A. 2009. Analysis of mutations of the PITX2 transcription factor found in patients with Axenfeld-Rieger syndrome. *Invest Ophthalmol Vis Sci,* 50**,** 2599-606.

FRYBURG, J., HWANG, V. & LIN, K. 1995. Recurrent lambdoid synostosis within two families. *American Journal of Medical Genetics,* 58**,** 262–266.

FUKAMI, M., NAGAI, T., MOCHIZUKI, H., MUROYA, K., YAMADA, G., TAKITANI, K. & OGATA, T. 2010. Anorectal and urinary anomalies and aberrant retinoic acid metabolism in cytochrome P450 oxidoreductase deficiency. *Mol Genet Metab,* 100**,** 269-73.

FURTHAUER, M., LIN, W., ANG, S. L., THISSE, B. & THISSE, C. 2002. Sef is a feedback-induced antagonist of Ras/MAPK-mediated FGF signalling. *Nat Cell Biol,* 4**,** 170-4.

GALLO, E. M., LOCH, D. C., HABASHI, J. P., CALDERON, J. F., CHEN, Y., BEDJA, D., VAN ERP, C., GERBER, E. E., PARKER, S. J., SAULS, K., JUDGE, D. P., COOKE, S. K., LINDSAY, M. E., ROUF, R., MYERS, L., AP RHYS, C. M., KENT, K. C., NORRIS, R. A., HUSO, D. L. & DIETZ, H. C. 2014. Angiotensin II-dependent TGF-beta signaling contributes to Loeys-Dietz syndrome vascular pathogenesis. *J Clin Invest,* 124**,** 448-60.

GILISSEN, C., HEHIR-KWA, J. Y., THUNG, D. T., VAN DE VORST, M., VAN BON, B. W., WILLEMSEN, M. H., KWINT, M., JANSSEN, I. M., HOISCHEN, A., SCHENCK, A., LEACH, R., KLEIN, R., TEARLE, R., BO, T., PFUNDT, R., YNTEMA, H. G., DE VRIES, B. B., KLEEFSTRA, T., BRUNNER, H. G., VISSERS, L. E. & VELTMAN, J. A. 2014. Genome sequencing identifies major causes of severe intellectual disability. *Nature,* 511**,** 344-7.

GILISSEN, C., HOISCHEN, A., BRUNNER, H. G. & VELTMAN, J. A. 2012. Disease gene identification strategies for exome sequencing. *Eur J Hum Genet,* 20**,** 490-7.

GIMONA, M. & WINDER, S. J. 1998. Single calponin homology domains are not actin-binding domains. *Current Biology,* 8**,** R673-R675.

GLASS, G. E., O'HARA, J., CANHAM, N., CILLIERS, D., DUNAWAY, D., FENWICK, A. L., JEELANI, N. O., JOHNSON, D., LESTER, T., LORD, H., MORTON, J. E. V., NISHIKAWA, H., NOONS, P., SCHWIEBERT, K., SHIPSTER, C., TAYLOR-BEADLING, A., TWIGG, S. R. F., VASUDEVAN, P., WALL, S. A., WILKIE, A. O. M. & WILSON, L. C. 2019. ERF-related craniosynostosis: The phenotypic and developmental profile of a new craniosynostosis syndrome. *American Journal of Medical Genetics,* 179**,** 615-627.

GOLDIE, S. J., ARHATARI, B. D., ANDERSON, P., AUDEN, A., PARTRIDGE, D. D., JANE, S. M. & DWORKIN, S. 2016. Mice lacking the conserved transcription factor Grainyhead-like 3 (Grhl3) display increased apposition of the frontal and parietal bones during embryonic development. *BMC Dev Biol,* 16**,** 37.

GONG, X., QIAN, H., CAO, P., ZHAO, X., ZHOU, Q., LEI, J. & YAN, N. 2018. Structural basis for the recognition of Sonic Hedgehog by human Patched1. *Science,* 361.

GOODRICH, L. V., MILENKOVIC, L., HIGGINS, K. M. & SCOTT, M. P. 1997. Altered Neural Cell Fates and Medulloblastoma in Mouse patched Mutants. *Science,* 277**,** 1109-1113.

GOODWIN, S., MCPHERSON, J. D. & MCCOMBIE, W. R. 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet,* 17**,** 333-51.
GOOS, J. A. C. & MATHIJSSEN, I. M. J. 2019. Genetic Causes of Craniosynostosis: An Update. *Molecular Syndromology,* 10**,** 6-23.
GOOS, J. A. C., VOGEL, W. K., MLCOCHOVA, H., MILLARD, C. J., ESFANDIARI, E., SELMAN, W. H., CALPENA, E., KOELLING, N., CARPENTER, E. L., SWAGEMAKERS, S. M. A., VAN DER SPEK, P. J., FILTZ, T. M., SCHWABE, J. W. R., IWANIEC, U. T., MATHIJSSEN, I. M. J., LEID, M. & TWIGG, S. R. F. 2019. A de novo substitution in BCL11B leads to loss of interaction with transcriptional complexes and craniosynostosis. *Hum Mol Genet,* 28**,** 2501-2513.
GORLIN, R. J. 1995. Nevoid basal cell carcinoma syndrome. *Dermatol Clin,* 13**,** 113-25.
GRANADILLO, J. L., CHUNG, W. K., HECHT, L., CORSTEN-JANSSEN, N., WEGNER, D., NIJ BIJVANK, S. W. A., TOLER, T. L., PINEDA-ALVAREZ, D. E., DOUGLAS, G., MURPHY, J. J., SHIMONY, J. & SHINAWI, M. 2018. Variable cardiovascular phenotypes associated with SMAD2 pathogenic variants. *Hum Mutat,* 39**,** 1875-1884.
GRAND, K., SKRABAN, C. M., COHEN, J. L., DOWSETT, L., MAZZOLA, S., TARPINIAN, J., BEDOUKIAN, E., NESBITT, A., DENENBERG, B., LULIS, L., SANTANI, A., ZACKAI, E. H. & DEARDORFF, M. A. 2021. Nonlethal presentations of CYP26B1-related skeletal anomalies and multiple synostoses syndrome. *Am J Med Genet A,* 185**,** 2766-2775.
GREEN, A. C., KOCOVSKI, P., JOVIC, T., WALIA, M. K., CHANDRARATNA, R. A. S., MARTIN, T. J., BAKER, E. K. & PURTON, L. E. 2017. Retinoic acid receptor signalling directly regulates osteoblast and adipocyte differentiation from mesenchymal progenitor cells. *Exp Cell Res,* 350**,** 284-297.
GREGOR, A., SADLEIR, L. G., ASADOLLAHI, R., AZZARELLO-BURRI, S., BATTAGLIA, A., OUSAGER, L. B., BOONSAWAT, P., BRUEL, A. L., BUCHERT, R., CALPENA, E., COGNÉ, B., DALLAPICCOLA, B., DISTELMAIER, F., ELMSLIE, F., FAIVRE, L., HAACK, T. B., HARRISON, V., HENDERSON, A., HUNT, D., ISIDOR, B., JOSET, P., KUMADA, S., LACHMEIJER, A. M. A., LEES, M., LYNCH, S. A., MARTINEZ, F., MATSUMOTO, N., MCDOUGALL, C., MEFFORD, H. C., MIYAKE, N., MYERS, C. T., MOUTTON, S., NESBITT, A., NOVELLI, A., ORELLANA, C., RAUCH, A., ROSELLO, M., SAIDA, K., SANTANI, A. B., SARKAR, A., SCHEFFER, I. E., SHINAWI, M., STEINDL, K., SYMONDS, J. D., ZACKAI, E. H., REIS, A., STICHT, H. & ZWEIER, C. 2018. De Novo Variants in the F-Box Protein FBXO11 in 20 Individuals with a Variable Neurodevelopmental Disorder. *Am J Hum Genet,* 103**,** 305-316.
GROVA, M., LO, D. D., MONTORO, D., HYUN, J. S., CHUNG, M. T., WAN, D. C. & LONGAKER, M. T. 2012. Models of cranial suture biology. *J Craniofac Surg,* 23**,** 1954-8.
GUÉRIT, E., ARTS, F., DACHY, G., BOULOUADNINE, B. & DEMOULIN, J. B. 2021. PDGF receptor mutations in human diseases. *Cell Mol Life Sci,* 78**,** 3867-3881.
GUIMIER, A., GABRIEL, G. C., BAJOLLE, F., TSANG, M., LIU, H., NOLL, A., SCHWARTZ, M., EL MALTI, R., SMITH, L. D., KLENA, N. T., JIMENEZ, G., MILLER, N. A., OUFADEM, M., MOREAU DE BELLAING, A., YAGI, H., SAUNDERS, C. J., BAKER, C. N., DI FILIPPO, S., PETERSON, K. A., THIFFAULT, I., BOLE-FEYSOT, C., COOLEY, L. D., FARROW, E. G., MASSON, C., SCHOEN, P., DELEUZE, J. F., NITSCHKE, P., LYONNET, S., DE PONTUAL, L., MURRAY, S. A., BONNET, D., KINGSMORE, S. F., AMIEL, J., BOUVAGNET, P., LO, C. W. & GORDON, C. T. 2015. MMP21 is

mutated in human heterotaxy and is required for normal left-right asymmetry in vertebrates. *Nat Genet,* 47**,** 1260-3.
GUMUS, E. 2020. Extending the phenotype of Xia-Gibbs syndrome in a two-year-old patient with craniosynostosis with a novel de novo AHDC1 missense mutation. *Eur J Med Genet,* 63**,** 103637.
GUSTAFSON, J., BJORK, M., VAN RAVENSWAAIJ-ARTS, C. M. A. & CUNNINGHAM, M. L. 2022. Mechanism of Disease: Recessive ADAMTSL4 Mutations and Craniosynostosis with Ectopia Lentis. *Case Rep Genet,* 2022**,** 3239260.
GUY, G. R., JACKSON, R. A., YUSOFF, P. & CHOW, S. Y. 2009. Sprouty proteins: modified modulators, matchmakers or missing links? *J Endocrinol,* 203**,** 191-202.
HACOHEN, N., KRAMER, S., SUTHERLAND, D., HIROMI, Y. & KRASNOW, M. A. 1998. Sprouty Encodes a Novel Antagonist of FGF Signaling that Patterns Apical Branching of the Drosophila Airways. *Cell,* 92**,** 253-263.
HAHN, H., WICKING, C., ZAPHIROPOULOS, P. G., GAILANI, M. R., SHANLEY, S., CHIDAMBARAM, A., VORECHOVSKY, I., HOLMBERG, E., UNDEN, A. B., GILLIES, S., NEGUS, K., SMYTH, I., PRESSMAN, C., LEFFELL, D. J., GERRARD, B., GOLDSTEIN, A. M., DEAN, M., TOFTGARD, R., CHENEVIX-TRENCH, G., WAINWRIGHT, B. & BALE, A. E. 1996. Mutations of the Human Homolog of Drosophila patched in the Nevoid Basal Cell Carcinoma Syndrome. *Cell,* 85**,** 841-851.
HALLER, G., ALVARADO, D. M., WILLING, M. C., BRAVERMAN, A. C., BRIDWELL, K. H., KELLY, M., LENKE, L. G., LUHMANN, S. J., GURNETT, C. A. & DOBBS, M. B. 2015. Genetic Risk for Aortic Aneurysm in Adolescent Idiopathic Scoliosis. *J Bone Joint Surg Am,* 97**,** 1411-7.
HANAFUSA, H., TORII, S., YASUNAGA, T. & NISHIDA, E. 2002. Sprouty1 and Sprouty2 provide a control mechanism for the Ras/MAPK signalling pathway. *Nat Cell Biol,* 4**,** 850-8.
HARADA, M. & AKITA, K. 2020. Mouse fibroblast growth factor 9 N143T mutation leads to wide chondrogenic condensation of long bones. *Histochem Cell Biol,* 153**,** 215-223.
HARADA, M., MURAKAMI, H., OKAWA, A., OKIMOTO, N., HIRAOKA, S., NAKAHARA, T., AKASAKA, R., SHIRAISHI, Y., FUTATSUGI, N., MIZUTANI-KOSEKI, Y., KUROIWA, A., SHIROUZU, M., YOKOYAMA, S., TAIJI, M., ISEKI, S., ORNITZ, D. M. & KOSEKI, H. 2009. FGF9 monomer-dimer equilibrium regulates extracellular matrix affinity and tissue diffusion. *Nat Genet,* 41**,** 289-98.
HARITA, Y., KITANAKA, S., ISOJIMA, T., ASHIDA, A. & HATTORI, M. 2017. Spectrum of LMX1B mutations: from nail–patella syndrome to isolated nephropathy. *Pediatric Nephrology,* 32**,** 1845-1850.
HASAN, M. R., TAKATALO, M., MA, H., RICE, R., MUSTONEN, T. & RICE, D. P. 2020. RAB23 coordinates early osteogenesis by repressing FGF10-pERK1/2 and GLI1. *Elife,* 9.
HASANOVIC, A., SIMSIR, M., CHOVEAU, F. S., LALLI, E. & MUS-VETEAU, I. 2020. Astemizole Sensitizes Adrenocortical Carcinoma Cells to Doxorubicin by Inhibiting Patched Drug Efflux Activity. *Biomedicines,* 8.
HATA, A. & CHEN, Y. G. 2016. TGF-beta Signaling from Receptors to Smads. *Cold Spring Harb Perspect Biol,* 8.

HEMPEL, A., PAGNAMENTA, A. T., BLYTH, M., MANSOUR, S., MCCONNELL, V., KOU, I., IKEGAWA, S., TSURUSAKI, Y., MATSUMOTO, N., LO-CASTRO, A., PLESSIS, G., ALBRECHT, B., BATTAGLIA, A., TAYLOR, J. C., HOWARD, M. F., KEAYS, D., SOHAL, A. S., COLLABORATION, D. D. D., KUHL, S. J., KINI, U. & MCNEILL, A. 2016. Deletions and de novo mutations of SOX11 are associated with a neurodevelopmental disorder with features of Coffin-Siris syndrome. *J Med Genet,* 53**,** 152-62.

HIDE, T., HATAKEYAMA, J., KIMURA-YOSHIDA, C., TIAN, E., TAKEDA, N., USHIO, Y., SHIROISHI, T., AIZAWA, S. & MATSUO, I. 2002. Genetic modifiers of otocephalic phenotypes in Otx2 heterozygous mutant mice. *Development and Disease,* 129**,** 4347–4357.

HILDEBRANDT, F., HEERINGA, S. F., RUSCHENDORF, F., ATTANASIO, M., NURNBERG, G., BECKER, C., SEELOW, D., HUEBNER, N., CHERNIN, G., VLANGOS, C. N., ZHOU, W., O'TOOLE, J. F., HOSKINS, B. E., WOLF, M. T., HINKES, B. G., CHAIB, H., ASHRAF, S., SCHOEB, D. S., OVUNC, B., ALLEN, S. J., VEGA-WARNER, V., WISE, E., HARVILLE, H. M., LYONS, R. H., WASHBURN, J., MACDONALD, J., NURNBERG, P. & OTTO, E. A. 2009. A systematic approach to mapping recessive disease genes in individuals from outbred populations. *PLoS Genet,* 5**,** e1000353.

HIRAKI, Y., MORIUCHI, M., OKAMOTO, N., ISHIKAWA, N., SUGIMOTO, Y., EGUCHI, K., SAKAI, H., SAITSU, H., MIZUGUCHI, T., HARADA, N. & MATSUMOTO, N. 2008. Craniosynostosis in a patient with a de novo 15q15-q22 deletion. *Am J Med Genet A,* 146A**,** 1462-5.

HOLLAND, P. W., BOOTH, H. A. & BRUFORD, E. A. 2007. Classification and nomenclature of all human homeobox genes. *BMC Biol,* 5**,** 47.

HONG, Y., ZHANG, J., ZHANG, H., LI, X., QU, J., ZHAI, J., ZHANG, L., CHEN, F. & LI, T. 2016. Heterozygous PTCH1 Mutations Impact the Bone Metabolism in Patients With Nevoid Basal Cell Carcinoma Syndrome Likely by Regulating SPARC Expression. *J Bone Miner Res,* 31**,** 1413-28.

HOOPER, J. E. & SCOTT, M. I. 1989. The Drosophila patched Gene Encodes a Putative Membrane Protein Required for Segmental Patterning. *Cell,* 59**,** 751-765.

HOWARD, T. D., PAZNEKAS, W. A., GREEN, E. D., CHIANG, L. C., MA, N., LUNA, R. I. O. D., DELGADO, C. G., GONZALEZ-RAMOS, M., KLINE, A. D. & JABS, E. W. 1997. Mutations in TWIST, a basic helix–loop–helix transcription factor, in Saethre-Chotzen syndrome. *Nature Genetics,* 5**,** 36–41.

HU, Y. & BOULOUX, P. M. 2010. Novel insights in FGFR1 regulation: lessons from Kallmann syndrome. *Trends Endocrinol Metab,* 21**,** 385-93.

HUANG, P., NEDELCU, D., WATANABE, M., JAO, C., KIM, Y., LIU, J. & SALIC, A. 2016. Cellular Cholesterol Directly Activates Smoothened in Hedgehog Signaling. *Cell,* 166**,** 1176-1187 e14.

HUCKSTADT, V., CHINTON, J., GOMEZ, A., OBREGON, M. G. & GRAVINA, L. P. 2021. Noonan syndrome with loose anagen hair with variants in the PPP1CB gene: First familial case reported. *Am J Med Genet A,* 185**,** 1256-1260.

HUNG, I. H., SCHOENWOLF, G. C., LEWANDOSKI, M. & ORNITZ, D. M. 2016. A combined series of Fgf9 and Fgf18 mutant alleles identifies unique and redundant roles in skeletal development. *Dev Biol,* 411**,** 72-84.

HUNTER, A. G. W. & RUDD, N. L. 1976. Craniosynostosis. I. Sagittal synostosis; Its genetics and associated clinical findings in 214 patients who lacked involvement of the coronal suture. *Teratology,* 14**,** 185-193.

HYDER, Z., CALPENA, E., PEI, Y., TOOZE, R. S., BRITTAIN, H., TWIGG, S. R. F., CILLIERS, D., MORTON, J. E. V., MCCANN, E., WEBER, A., WILSON, L. C., DOUGLAS, A. G. L., MCGOWAN, R., NEED, A., BOND, A., TAVARES, A. L. T., THOMAS, E. R. A., GENOMICS ENGLAND RESEARCH, C., HILL, S. L., DEANS, Z. C., BOARDMAN-PRETTY, F., CAULFIELD, M., SCOTT, R. H. & WILKIE, A. O. M. 2021. Evaluating the performance of a clinical genome sequencing program for diagnosis of rare genetic disease, seen through the lens of craniosynostosis. *Genet Med,* 23**,** 2360-2368.

IBRAHIMI, O. A., ZHANG, F., ELISEENKOVA, A. V., LINHARDT, R. J. & MOHAMMADI, M. 2004. Proline to arginine mutations in FGF receptors 1 and 3 result in Pfeiffer and Muenke craniosynostosis syndromes through enhancement of FGF binding affinity. *Hum Mol Genet,* 13**,** 69-78.

IDRIZ, S., PATEL, J. H., RENANI, S. A., ALLAN, R. & VLAHOS, I. 2015. CT of Normal Developmental and Variant Anatomy of the Pediatric Skull: Distinguishing Trauma from Normality. *Radiographics,* 35**,** 1585-1601.

INGHAM, P. W. & MCMAHON, A. P. 2001. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev,* 15**,** 3059-87.

IP, C. K. M., NG, P. K. S., JEONG, K. J., SHAO, S. H., JU, Z., LEONARD, P. G., HUA, X., VELLANO, C. P., WOESSNER, R., SAHNI, N., SCOTT, K. L. & MILLS, G. B. 2018. Neomorphic PDGFRA extracellular domain driver mutations are resistant to PDGFRA targeted therapies. *Nat Commun,* 9**,** 4583.

IQBAL, R., ALOM, S., BINSAEID, J. & HARKY, A. 2021. Loeys-Dietz syndrome pathology and aspects of cardiovascular management: A systematic review. *Vascular,* 29**,** 3-14.

ISOJIMA, T., HARITA, Y., FURUYAMA, M., SUGAWARA, N., ISHIZUKA, K., HORITA, S., KAJIHO, Y., MIURA, K., IGARASHI, T., HATTORI, M. & KITANAKA, S. 2014. LMX1B mutation with residual transcriptional activity as a cause of isolated glomerulopathy. *Nephrol Dial Transplant,* 29**,** 81-8.

JABS, E. W., MÜLLER, U., LI, X., MA, L., LUO, W., HAWORTH, I. S., KLISAK, I., SPARKES, R., WARMAN, M. L., MULLIKEN, J. B., SNEAD, M. L. & MAXSON, R. 1993. A mutation in the homeodomain of the human MSX2 gene in a family affected with autosomal dominant craniosynostosis. *Cell,* 75**,** 443-450.

JAGANATHAN, K., KYRIAZOPOULOU PANAGIOTOPOULOU, S., MCRAE, J. F., DARBANDI, S. F., KNOWLES, D., LI, Y. I., KOSMICKI, J. A., ARBELAEZ, J., CUI, W., SCHWARTZ, G. B., CHOW, E. D., KANTERAKIS, E., GAO, H., KIA, A., BATZOGLOU, S., SANDERS, S. J. & FARH, K. K. 2019. Predicting Splicing from Primary Sequence with Deep Learning. *Cell,* 176**,** 535-548 e24.

JAMES, A. W., LEVI, B., XU, Y., CARRE, A. L. & LONGAKER, M. T. 2010. Retinoic acid enhances osteogenesis in cranial suture-derived mesenchymal cells: potential mechanisms of retinoid-induced craniosynostosis. *Plast Reconstr Surg,* 125**,** 1352-1361.

JANDL, N. M., SCHMIDT, T., ROLVIEN, T., STÜRZNICKEL, J., CHRYSOSTOMOU, K., VON VOPELIUS, E., VOLK, A. E., SCHINKE, T., KUBISCH, C., AMLING, M. & BARVENCIK,

F. 2021. Genotype–Phenotype Associations in 72 Adults with Suspected ALPL-Associated Hypophosphatasia. *Calcified Tissue International,* 108**,** 288-301.
JANSSEN, A., HOSEN, M. J., JEANNIN, P., COUCKE, P. J., DE PAEPE, A. & VANAKKER, O. M. 2013. Second family with the Boston-type craniosynostosis syndrome: novel mutation and expansion of the clinical spectrum. *Am J Med Genet A,* 161A**,** 2352-7.
JEHEE, F. S., JOHNSON, D., ALONSO, L. G., CAVALCANTI, D. P., DE SA MOREIRA, E., ALBERTO, F. L., KOK, F., KIM, C., WALL, S. A., JABS, E. W., BOYADJIEV, S. A., WILKIE, A. O. & PASSOS-BUENO, M. R. 2005. Molecular screening for microdeletions at 9p22-p24 and 11q23-q24 in a large cohort of patients with trigonocephaly. *Clin Genet,* 67**,** 503-10.
JENKINS, D., SEELOW, D., JEHEE, F. S., PERLYN, C. A., ALONSO, L. G., BUENO, D. F., DONNAI, D., JOSIFOVA, D., MATHIJSSEN, I. M., MORTON, J. E., ORSTAVIK, K. H., SWEENEY, E., WALL, S. A., MARSH, J. L., NURNBERG, P., PASSOS-BUENO, M. R. & WILKIE, A. O. 2007. RAB23 mutations in Carpenter syndrome imply an unexpected role for hedgehog signaling in cranial-suture development and obesity. *Am J Hum Genet,* 80**,** 1162-70.
JOHNSON, D., ISEKI, S., WILKIE, A. O. M. & MORRISS-KAY, G. M. 2000. Expression patterns of Twist and Fgfr1, -2 and -3 in the developing mouse coronal suture suggest a key role for Twist in suture initiation and biogenesis. *Mechanisms of Development,* 91**,** 341-345.
JOHNSON, D., KAN, S. H., OLDRIDGE, M., TREMBATH, R. C., ROCHE, P., ESNOUF, R. M., GIELE, H. & WILKIE, A. O. 2003. Missense mutations in the homeodomain of HOXD13 are associated with brachydactyly types D and E. *Am J Hum Genet,* 72**,** 984-97.
JOHNSON, R. L., ROTHMAN, A. L., XIE, J., GOODRICH, L. V., BARE, J. W., BONIFAS, J. M., QUINN, A. G., MYERS, R. M., COX, D. R., JR., E. H. E. & SCOTT, M. P. 1996. Human Homolog of Patched, a Candidate Gene for the Basal Cell Nevus Syndrome. *Science,* 272**,** 1668-1671.
JONES, F. S., MEECH, R., EDELMAN, D. B., OAKEY, R. J. & JONES, P. L. 2001. Prx1 Controls Vascular Smooth Muscle Cell Proliferation and Tenascin-C Expression and Is Upregulated With Prx2 in Pulmonary Vascular Disease. *Circulation Research,* 89**,** 131-138.
JONES, M., CHUNG, J., KIMONIS, V. & GOLD, J. A. 2017. A novel mutation of orthodenticle homeobox 2 contributing to a case of otocephaly initially diagnosed by prenatal ultrasound in the first trimester. *Clin Dysmorphol,* 26**,** 98-100.
JONSSON, H., SULEM, P., KEHR, B., KRISTMUNDSDOTTIR, S., ZINK, F., HJARTARSON, E., HARDARSON, M. T., HJORLEIFSSON, K. E., EGGERTSSON, H. P., GUDJONSSON, S. A., WARD, L. D., ARNADOTTIR, G. A., HELGASON, E. A., HELGASON, H., GYLFASON, A., JONASDOTTIR, A., JONASDOTTIR, A., RAFNAR, T., FRIGGE, M., STACEY, S. N., TH MAGNUSSON, O., THORSTEINSDOTTIR, U., MASSON, G., KONG, A., HALLDORSSON, B. V., HELGASON, A., GUDBJARTSSON, D. F. & STEFANSSON, K. 2017. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature,* 549**,** 519-522.
JUSTICE, C. M., CUELLAR, A., BALA, K., SABOURIN, J. A., CUNNINGHAM, M. L., CRAWFORD, K., PHIPPS, J. M., ZHOU, Y., CILLIERS, D., BYREN, J. C., JOHNSON,

D., WALL, S. A., MORTON, J. E. V., NOONS, P., SWEENEY, E., WEBER, A., REES, K. E. M., WILSON, L. C., SIMEONOV, E., KANEVA, R., YANEVA, N., GEORGIEV, K., BUSSARSKY, A., SENDERS, C., ZWIENENBERG, M., BOGGAN, J., ROSCIOLI, T., TAMBURRINI, G., BARBA, M., CONWAY, K., SHEFFIELD, V. C., BRODY, L., MILLS, J. L., KAY, D., SICKO, R. J., LANGLOIS, P. H., TITTLE, R. K., BOTTO, L. D., JENKINS, M. M., LASALLE, J. M., LATTANZI, W., WILKIE, A. O. M., WILSON, A. F., ROMITTI, P. A., BOYADJIEV, S. A. & NATIONAL BIRTH DEFECTS PREVENTION, S. 2020. A genome-wide association study implicates the BMP7 locus as a risk factor for nonsyndromic metopic craniosynostosis. *Hum Genet,* 139**,** 1077-1090.

JUSTICE, C. M., MUSOLF, A. M., CUELLAR, A., LATTANZI, W., SIMEONOV, E., KANEVA, R., PASCHALL, J., CUNNINGHAM, M., WILKIE, A. O. M., WILSON, A. F., ROMITTI, P. A. & BOYADJIEV, S. A. 2022. Targeted Sequencing of Candidate Regions Associated with Sagittal and Metopic Nonsyndromic Craniosynostosis. *Genes (Basel),* 13.

JUSTICE, C. M., YAGNIK, G., KIM, Y., PETER, I., JABS, E. W., ERAZO, M., YE, X., AINEHSAZAN, E., SHI, L., CUNNINGHAM, M. L., KIMONIS, V., ROSCIOLI, T., WALL, S. A., WILKIE, A. O., STOLER, J., RICHTSMEIER, J. T., HEUZE, Y., SANCHEZ-LARA, P. A., BUCKLEY, M. F., DRUSCHEL, C. M., MILLS, J. L., CAGGANA, M., ROMITTI, P. A., KAY, D. M., SENDERS, C., TAUB, P. J., KLEIN, O. D., BOGGAN, J., ZWIENENBERG-LEE, M., NAYDENOV, C., KIM, J., WILSON, A. F. & BOYADJIEV, S. A. 2012. A genome-wide association study identifies susceptibility loci for nonsyndromic sagittal craniosynostosis near BMP2 and within BBS9. *Nat Genet,* 44**,** 1360-4.

KAN, S. H., ELANKO, N., JOHNSON, D., CORNEJO-ROLDAN, L., COOK, J., REICH, E. W., TOMKINS, S., VERLOES, A., TWIGG, S. R., RANNAN-ELIYA, S., MCDONALD-MCGINN, D. M., ZACKAI, E. H., WALL, S. A., MUENKE, M. & WILKIE, A. O. 2002. Genomic screening of fibroblast growth-factor receptor 2 reveals a wide spectrum of mutations in patients with syndromic craniosynostosis. *Am J Hum Genet,* 70**,** 472-86.

KAPLANIS, J., SAMOCHA, K. E., WIEL, L., ZHANG, Z., ARVAI, K. J., EBERHARDT, R. Y., GALLONE, G., LELIEVELD, S. H., MARTIN, H. C., MCRAE, J. F., SHORT, P. J., TORENE, R. I., DE BOER, E., DANECEK, P., GARDNER, E. J., HUANG, N., LORD, J., MARTINCORENA, I., PFUNDT, R., REIJNDERS, M. R. F., YEUNG, A., YNTEMA, H. G., DECIPHERING DEVELOPMENTAL DISORDERS, S., VISSERS, L., JUUSOLA, J., WRIGHT, C. F., BRUNNER, H. G., FIRTH, H. V., FITZPATRICK, D. R., BARRETT, J. C., HURLES, M. E., GILISSEN, C. & RETTERER, K. 2020. Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature,* 586**,** 757-762.

KARCZEWSKI, K. J., FRANCIOLI, L. C., TIAO, G., CUMMINGS, B. B., ALFÖLDI, J., WANG, Q., COLLINS, R. L., LARICCHIA, K. M., GANNA, A., BIRNBAUM, D. P., GAUTHIER, L. D., BRAND, H., SOLOMONSON, M., WATTS, N. A., RHODES, D., SINGER-BERK, M., ENGLAND, E. M., SEABY, E. G., KOSMICKI, J. A., WALTERS, R. K., TASHMAN, K., FARJOUN, Y., BANKS, E., POTERBA, T., WANG, A., SEED, C., WHIFFIN, N., CHONG, J. X., SAMOCHA, K. E., PIERCE-HOFFMAN, E., ZAPPALA, Z., O'DONNELL-LURIA, A. H., MINIKEL, E. V., WEISBURD, B., LEK, M., WARE, J. S., VITTAL, C., ARMEAN, I. M., BERGELSON, L., CIBULSKIS, K., CONNOLLY, K. M., COVARRUBIAS, M., DONNELLY, S., FERRIERA, S., GABRIEL, S., GENTRY, J.,

GUPTA, N., JEANDET, T., KAPLAN, D., LLANWARNE, C., MUNSHI, R., NOVOD, S., PETRILLO, N., ROAZEN, D., RUANO-RUBIO, V., SALTZMAN, A., SCHLEICHER, M., SOTO, J., TIBBETTS, K., TOLONEN, C., WADE, G., TALKOWSKI, M. E., NEALE, B. M., DALY, M. J. & MACARTHUR, D. G. 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature,* 581**,** 434-443.

KARCZEWSKI, K. J., FRANCIOLI, L. C., TIAO, G., CUMMINGS, B. B., ALFÖLDI, J., WANG, Q., COLLINS, R. L., LARICCHIA, K. M., GANNA, A., BIRNBAUM, D. P., GAUTHIER, L. D., BRAND, H., SOLOMONSON, M., WATTS, N. A., RHODES, D., SINGER-BERK, M., SEABY, E. G., KOSMICKI, J. A., WALTERS, R. K., TASHMAN, K., FARJOUN, Y., BANKS, E., POTERBA, T., WANG, A., SEED, C., WHIFFIN, N., CHONG, J. X., SAMOCHA, K. E., PIERCE-HOFFMAN, E., ZAPPALA, Z., O'DONNELL-LURIA, A. H., MINIKEL, E. V., WEISBURD, B., LEK, M., WARE, J. S., VITTAL, C., ARMEAN, I. M., BERGELSON, L., CIBULSKIS, K., CONNOLLY, K. M., COVARRUBIAS, M., DONNELLY, S., FERRIERA, S., GABRIEL, S., GENTRY, J., GUPTA, N., JEANDET, T., KAPLAN, D., LLANWARNE, C., MUNSHI, R., NOVOD, S., PETRILLO, N., ROAZEN, D., RUANO-RUBIO, V., SALTZMAN, A., SCHLEICHER, M., SOTO, J., TIBBETTS, K., TOLONEN, C., WADE, G., TALKOWSKI, M. E., NEALE, B. M., DALY, M. J. & MACARTHUR, D. G. 2019. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-offunction intolerance across human protein-coding genes. *bioRxiv*.

KARCZEWSKI, K. J., SOLOMONSON, M., CHAO, K. R., GOODRICH, J. K., TIAO, G., LU, W., RILEY-GILLIS, B. M., TSAI, E. A., KIM, H. I., ZHENG, X., RAHIMOV, F., ESMAEELI, S., GRUNDSTAD, A. J., REPPELL, M., WARING, J., JACOB, H., SEXTON, D., BRONSON, P. G., CHEN, X., HU, X., GOLDSTEIN, J. I., KING, D., VITTAL, C., POTERBA, T., PALMER, D. S., CHURCHHOUSE, C., HOWRIGAN, D. P., ZHOU, W., WATTS, N. A., NGUYEN, K., NGUYEN, H., MASON, C., FARNHAM, C., TOLONEN, C., GAUTHIER, L. D., GUPTA, N., MACARTHUR, D. G., REHM, H. L., SEED, C., PHILIPPAKIS, A. A., DALY, M. J., DAVIS, J. W., RUNZ, H., MILLER, M. R. & NEALE, B. M. 2022. Systematic single-variant and gene-based association testing of thousands of phenotypes in 394,841 UK Biobank exomes. *Cell Genomics,* 2.

KESPOHL, B., SCHUMERTL, T., BERTRAND, J., LOKAU, J. & GARBERS, C. 2021. The cytokine interleukin-11 crucially links bone formation, remodeling and resorption. *Cytokine Growth Factor Rev,* 60**,** 18-27.

KESSLER, M. D., LOESCH, D. P., PERRY, J. A., HEARD-COSTA, N. L., TALIUN, D., CADE, B. E., WANG, H., DAYA, M., ZINITI, J., DATTA, S., CELEDON, J. C., SOTO-QUIROS, M. E., AVILA, L., WEISS, S. T., BARNES, K., REDLINE, S. S., VASAN, R. S., JOHNSON, A. D., MATHIAS, R. A., HERNANDEZ, R., WILSON, J. G., NICKERSON, D. A., ABECASIS, G., BROWNING, S. R., ZOLLNER, S., O'CONNELL, J. R., MITCHELL, B. D., NATIONAL HEART, L., BLOOD INSTITUTE TRANS-OMICS FOR PRECISION MEDICINE, C., GROUP, T. O. P. G. W. & O'CONNOR, T. D. 2020. De novo mutations across 1,465 diverse genomes reveal mutational insights and reductions in the Amish founder population. *Proc Natl Acad Sci U S A,* 117**,** 2560-2569.

KEUPP, K., LI, Y., VARGEL, I., HOISCHEN, A., RICHARDSON, R., NEVELING, K., ALANAY, Y., UZ, E., ELCIOGLU, N., RACHWALSKI, M., KAMACI, S., TUNCBILEK, G., AKIN, B., GROTZINGER, J., KONAS, E., MAVILI, E., MULLER-NEWEN, G., COLLMANN, H., ROSCIOLI, T., BUCKLEY, M. F., YIGIT, G., GILISSEN, C., KRESS, W., VELTMAN, J.,

HAMMERSCHMIDT, M., AKARSU, N. A. & WOLLNIK, B. 2013. Mutations in the interleukin receptor IL11RA cause autosomal recessive Crouzon-like craniosynostosis. *Mol Genet Genomic Med,* 1**,** 223-37.

KINNEBREW, M., LUCHETTI, G., SIRCAR, R., FRIGUI, S., VITI, L. V., NAITO, T., BECKERT, F., SAHEKI, Y., SIEBOLD, C., RADHAKRISHNAN, A. & ROHATGI, R. 2021. Patched 1 reduces the accessibility of cholesterol in the outer leaflet of membranes. *Elife,* 10.

KLOPOCKI, E., LOHAN, S., BRANCATI, F., KOLL, R., BREHM, A., SEEMANN, P., DATHE, K., STRICKER, S., HECHT, J., BOSSE, K., BETZ, R. C., GARACI, F. G., DALLAPICCOLA, B., JAIN, M., MUENKE, M., NG, V. C., CHAN, W., CHAN, D. & MUNDLOS, S. 2011. Copy-number variations involving the IHH locus are associated with syndactyly and craniosynostosis. *Am J Hum Genet,* 88**,** 70-5.

KOBOLDT, D. C. 2020. Best practices for variant calling in clinical sequencing. *Genome Med,* 12**,** 91.

KOELLING, N., BERNKOPF, M., CALPENA, E., MAHER, G. J., MILLER, K. A., RALPH, H. K., GORIELY, A. & WILKIE, A. O. M. 2019. Amplimap: a versatile tool to process and analyze targeted NGS data. *Bioinformatics,* 35**,** 5349-5350.

KOLEDOVA, Z., ZHANG, X., STREULI, C., CLARKE, R. B., KLEIN, O. D., WERB, Z. & LU, P. 2016. SPRY1 regulates mammary epithelial morphogenesis by modulating EGFR-dependent stromal paracrine signaling and ECM remodeling. *Proc Natl Acad Sci U S A,* 113**,** E5731-40.

KONG, A., FRIGGE, M. L., MASSON, G., BESENBACHER, S., SULEM, P., MAGNUSSON, G., GUDJONSSON, S. A., SIGURDSSON, A., JONASDOTTIR, A., JONASDOTTIR, A., WONG, W. S., SIGURDSSON, G., WALTERS, G. B., STEINBERG, S., HELGASON, H., THORLEIFSSON, G., GUDBJARTSSON, D. F., HELGASON, A., MAGNUSSON, O. T., THORSTEINSDOTTIR, U. & STEFANSSON, K. 2012. Rate of de novo mutations and the importance of father's age to disease risk. *Nature,* 488**,** 471-5.

KONG, J. H., YOUNG, C. B., PUSAPATI, G. V., PATEL, C. B., HO, S., KRISHNAN, A., LIN, J. I., DEVINE, W., MOREAU DE BELLAING, A., ATHNI, T. S., ARAVIND, L., GUNN, T. M., LO, C. W. & ROHATGI, R. 2020. A Membrane-Tethered Ubiquitination Pathway Regulates Hedgehog Signaling and Heart Development. *Dev Cell,* 55**,** 432-449 e12.

KORBERG, I., NOWINSKI, D., BONDESON, M. L., MELIN, M., KOLBY, L. & STATTIN, E. L. 2020. A progressive and complex clinical course in two family members with ERF-related craniosynostosis: a case report. *BMC Med Genet,* 21**,** 90.

KOWATSCH, C., WOOLLEY, R. E., KINNEBREW, M., ROHATGI, R. & SIEBOLD, C. 2019. Structures of vertebrate Patched and Smoothened reveal intimate links between cholesterol and Hedgehog signalling. *Curr Opin Struct Biol,* 57**,** 204-214.

KRAWITZ, P. M., SCHWEIGER, M. R., RODELSPERGER, C., MARCELIS, C., KOLSCH, U., MEISEL, C., STEPHANI, F., KINOSHITA, T., MURAKAMI, Y., BAUER, S., ISAU, M., FISCHER, A., DAHL, A., KERICK, M., HECHT, J., KOHLER, S., JAGER, M., GRUNHAGEN, J., DE CONDOR, B. J., DOELKEN, S., BRUNNER, H. G., MEINECKE, P., PASSARGE, E., THOMPSON, M. D., COLE, D. E., HORN, D., ROSCIOLI, T., MUNDLOS, S. & ROBINSON, P. N. 2010. Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. *Nat Genet,* 42**,** 827-9.

KRONENBERG, H. M. 2004. Twist Genes Regulate Runx2 and Bone Formation. *Developmental Cell,* 6**,** 317-318.

KROOK, M. A., REESER, J. W., ERNST, G., BARKER, H., WILBERDING, M., LI, G., CHEN, H. Z. & ROYCHOWDHURY, S. 2021. Fibroblast growth factor receptors in cancer: genetic alterations, diagnostics, therapeutic targets and mechanisms of resistance. *Br J Cancer,* 124**,** 880-892.

KWIEK, N. C., THACKER, D. F., DATTO, M. B., MEGOSH, H. B. & HAYSTEAD, T. A. 2006a. PITK, a PP1 targeting subunit that modulates the phosphorylation of the transcriptional regulator hnRNP K. *Cell Signal,* 18**,** 1769-78.

KWIEK, N. C., THACKER, D. F., DATTO, M. B., MEGOSH, H. B. & HAYSTEAD, T. A. J. 2006b. Protein Phosphatase 6 Regulatory Subunits Composed of Ankyrin Repeat Domains. *Cellular Signalling,* 18**,** 1769-1778.

KYRYLKOVA, K., IWANIEC, U. T., PHILBRICK, K. A. & LEID, M. 2016. BCL11B regulates sutural patency in the mouse craniofacial skeleton. *Dev Biol,* 415**,** 251-260.

LAFLAMME, C., FILION, C. & LABELLE, Y. 2004. Functional characterization of SIX3 homeodomain mutations in holoprosencephaly: interaction with the nuclear receptor NR4A3/NOR1. *Hum Mutat,* 24**,** 502-8.

LAJEUNIE, E., BARCIK, U., THORNE, J. A., GHOUZZI, V. E., BOURGEOIS, M. & RENIER, D. 2001. Craniosynostosis and fetal exposure to sodium valproate. *Journal of Neurosurgery,* 95**,** 778-782.

LAJEUNIE, E., CRIMMINS, D. W., ARNAUD, E. & RENIER, D. 2005. Genetic considerations in nonsyndromic midline craniosynostoses: a study of twins and their families. *Journal of Neurosurgery,* 103**,** 353-356.

LAJEUNIE, E., MERRER, M. L., BONAITI-PELLIE, C., MARCHAC, D. & RENIER, A. D. 1995. Genetic Study of Nonsyndromic Coronal Craniosynostosis. *American Journal of Medical Genetics,* 55**,** 500-504.

LAJEUNIE, E., MERRER, M. L., BONAYTI-PELLIE, C., MARCHAC, D. & RENIER, D. 1996. Genetic Study of Scaphocephaly. *American Journal of Medical Genetics,* 62**,** 282-285.

LAJEUNIE, E., MERRER, M. L., MARCHACA, D. & RENIER, D. 1998. Syndromal and Nonsyndromal Primary Trigonocephaly: Analysis of a Series of 237 Patients. *American Journal of Medical Genetics,* 75**,** 211-215.

LANA-ELOLA, E., RICE, R., GRIGORIADIS, A. E. & RICE, D. P. 2007. Cell fate specification during calvarial bone and suture development. *Dev Biol,* 311**,** 335-46.

LAUE, K., POGODA, H. M., DANIEL, P. B., VAN HAERINGEN, A., ALANAY, Y., VON AMELN, S., RACHWALSKI, M., MORGAN, T., GRAY, M. J., BREUNING, M. H., SAWYER, G. M., SUTHERLAND-SMITH, A. J., NIKKELS, P. G., KUBISCH, C., BLOCH, W., WOLLNIK, B., HAMMERSCHMIDT, M. & ROBERTSON, S. P. 2011. Craniosynostosis and multiple skeletal anomalies in humans and zebrafish result from a defect in the localized degradation of retinoic acid. *Am J Hum Genet,* 89**,** 595-606.

LEE, E., LE, T., ZHU, Y., ELAKIS, G., TURNER, A., LO, W., VENSELAAR, H., VERRENKAMP, C. A., SNOW, N., MOWAT, D., KIRK, E. P., SACHDEV, R., SMITH, J., BROWN, N. J., WALLIS, M., BARNETT, C., MCKENZIE, F., FRECKMANN, M. L., COLLINS, F., CHOPRA, M., GREGERSEN, N., HAYES, I., RAJAGOPALAN, S., TAN, T. Y., STARK, Z., SAVARIRAYAN, R., YEUNG, A., ADÈS, L., GATTAS, M., GIBSON, K., GABBETT, M., AMOR, D. J., LATTANZI, W., BOYD, S., HAAN, E., GIANOUTSOS, M., COX, T.

C., BUCKLEY, M. F. & ROSCIOLI, T. 2018. A craniosynostosis massively parallel sequencing panel study in 309 Australian and New Zealand patients: findings and recommendations. *Genet Med,* 20**,** 1061-1068.

LEE, K. K. L., STANIER, P. & PAUWS, E. 2019. Mouse Models of Syndromic Craniosynostosis. *Mol Syndromol,* 10**,** 58-73.

LENNON, M. J., JONES, S. P., LOVELACE, M. D., GUILLEMIN, G. J. & BREW, B. J. 2017. Bcl11b-A Critical Neurodevelopmental Transcription Factor-Roles in Health and Disease. *Front Cell Neurosci,* 11**,** 89.

LENTON, K., JAMES, A. W., MANU, A., BRUGMANN, S. A., BIRKER, D., NELSON, E. R., LEUCHT, P., HELMS, J. A. & LONGAKER, M. T. 2011. Indian hedgehog positively regulates calvarial ossification and modulates bone morphogenetic protein signaling. *Genesis,* 49**,** 784-96.

LEUSSINK, B., BROUWER, A., KHATTABI, M. E., POELMANN, R. E., GROOT, A. C. G.-D. & MEIJLIR, F. 1995. Expression patterns of the paired-related homeobox genes MHox/Prxl and S8/Prx2 suggest roles in development of the heart and the forebrain. *Mechanisms of Development,* 52**,** 51-64.

LEVI, B., WAN, D. C., WONG, V. W., NELSON, E., HYUN, J. & LONGAKER, M. T. 2012. Cranial suture biology: from pathways to patient care. *J Craniofac Surg,* 23**,** 13-9.

LI, C., SCOTT1, D. A., HATCH, E., TIAN, X. & MANSOUR, S. L. 2007. Dusp6 (Mkp3) is a negative feedback regulator of FGF stimulated ERK signaling during mouse development. *Development,* 134**,** 167-176.

LI, Y., KLENA, N. T., GABRIEL, G. C., LIU, X., KIM, A. J., LEMKE, K., CHEN, Y., CHATTERJEE, B., DEVINE, W., DAMERLA, R. R., CHANG, C., YAGI, H., SAN AGUSTIN, J. T., THAHIR, M., ANDERTON, S., LAWHEAD, C., VESCOVI, A., PRATT, H., MORGAN, J., HAYNES, L., SMITH, C. L., EPPIG, J. T., REINHOLDT, L., FRANCIS, R., LEATHERBURY, L., GANAPATHIRAJU, M. K., TOBITA, K., PAZOUR, G. J. & LO, C. W. 2015. Global genetic analysis in mice unveils central role for cilia in congenital heart disease. *Nature,* 521**,** 520-4.

LIEM, R. K. 2016. Cytoskeletal Integrators: The Spectrin Superfamily. *Cold Spring Harb Perspect Biol,* 8.

LINDEBOOM, R. G. H., VERMEULEN, M., LEHNER, B. & SUPEK, F. 2019. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat Genet,* 51**,** 1645-1651.

LIPIŃSKI, P., RÓŻDŻYŃSKA-ŚWIĄTKOWSKA, A., IWANICKA-PRONICKA, K., PERKOWSKA, B., POKORA, P. & TYLKI-SZYMAŃSKA, A. 2022. Long-term outcome of patients with alpha-mannosidosis - A single center study. *Mol Genet Metab Rep,* 30**,** 100826.

LIPPINCOTT, M. F., XU, W., SMITH, A. A., MIAO, X., LAFONT, A., SHENNIB, O., FARLEY, G. J., SABBAGH, R., DELANEY, A., STAMOU, M., PLUMMER, L., SALNIKOV, K., GEORGOPOULOS, N. A., MERICQ, V., QUINTON, R., MAU-THEM, F. T., NAMBOT, S., HAMAD, A., BRITTAIN, H., TOOZE, R. S., CALPENA, E., WILKIE, A. O. M., WILLEMS, M., CROWLEY, W. F., BALASUBRAMANIAN, R., LAMARCHE-VANE, N., DAVIS, E. E. & SEMINARA, S. B. 2022. The p190 RhoGAPs, ARHGAP35, and ARHGAP5 are implicated in GnRH neuronal development: Evidence from patients with idiopathic hypogonadotropic hypogonadism, zebrafish, and in vitro GAP activity assay. *Genet Med,* 24**,** 2501-2515.

LIU, B., CHEN, S., JOHNSON, C. & HELMS, J. A. 2014. A ciliopathy with hydrocephalus, isolated craniosynostosis, hypertelorism, and clefting caused by deletion of Kif3a. *Reprod Toxicol,* 48**,** 88-97.

LIU, W., SELEVER, J., MURALI, D., SUN, X., BRUGGER, S. M., MA, L., SCHWARTZ, R. J., MAXSON, R., FURUTA, Y. & MARTIN, J. F. 2005. Threshold-specific requirements for Bmp4 in mandibular development. *Dev Biol,* 283**,** 282-93.

LO MUZIO, L. 2008. Nevoid basal cell carcinoma syndrome (Gorlin syndrome). *Orphanet J Rare Dis,* 3**,** 32.

LOEYS, B. L., CHEN, J., NEPTUNE, E. R., JUDGE, D. P., PODOWSKI, M., HOLM, T., MEYERS, J., LEITCH, C. C., KATSANIS, N., SHARIFI, N., XU, F. L., MYERS, L. A., SPEVAK, P. J., CAMERON, D. E., DE BACKER, J., HELLEMANS, J., CHEN, Y., DAVIS, E. C., WEBB, C. L., KRESS, W., COUCKE, P., RIFKIN, D. B., DE PAEPE, A. M. & DIETZ, H. C. 2005. A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in TGFBR1 or TGFBR2. *Nat Genet,* 37**,** 275-81.

LORENZO, C. & MCCORMICK, F. 2020. SPRED proteins and their roles in signal transduction, development, and malignancy. *Genes Dev,* 34**,** 1410-1421.

LU, M.-F., CHENG, H.-T., LACY, A. R., KERN, M. J., ARGAO, E. A., POTTER, S. S., OLSON, E. N. & MARTIN, J. F. 1999. Paired-Related Homeobox Genes Cooperate in Handplate and Hindlimb Zeugopod Morphogenesis. *Developmental Biology,* 205**,** 145-157.

LU, W., QUINTERO-RIVERA, F., FAN, Y., ALKURAYA, F. S., DONOVAN, D. J., XI, Q., TURBE-DOAN, A., LI, Q. G., CAMPBELL, C. G., SHANSKE, A. L., SHERR, E. H., AHMAD, A., PETERS, R., RILLIET, B., PARVEX, P., BASSUK, A. G., HARRIS, D. J., FERGUSON, H., KELLY, C., WALSH, C. A., GRONOSTAJSKI, R. M., DEVRIENDT, K., HIGGINS, A., LIGON, A. H., QUADE, B. J., MORTON, C. C., GUSELLA, J. F. & MAAS, R. L. 2007. NFIA haploinsufficiency is associated with a CNS malformation syndrome and urinary tract defects. *PLoS Genet,* 3**,** e80.

LUCHETTI, G., SIRCAR, R., KONG, J. H., NACHTERGAELE, S., SAGNER, A., BYRNE, E. F., COVEY, D. F., SIEBOLD, C. & ROHATGI, R. 2016. Cholesterol activates the G-protein coupled receptor Smoothened to promote Hedgehog signaling. *Elife,* 5.

MA, L., GOLDEN, S., WU, L. & MAXSON, R. 1996. The molecular basis of Boston-type craniosynostosis: The Pro148→His mutation in the N-terminal arm of the MSX2 homeodomain stabilizes DNA binding without altering nucleotide sequence preferences. *Human Molecular Genetics,* 5**,** 1915-1920.

MAIK-RACHLINE, G., HACOHEN-LEV-RAN, A. & SEGER, R. 2019. Nuclear ERK: Mechanism of Translocation, Substrates, and Role in Cancer. *Int J Mol Sci,* 20.

MAKINO, S., MASUYA, H., ISHIJIMA, J., YADA, Y. & SHIROISHI, T. 2001. A spontaneous mouse mutation, mesenchymal dysplasia (mes), is caused by a deletion of the most C-terminal cytoplasmic domain of patched (ptc). *Dev Biol,* 239**,** 95-106.

MARTIN, A. R., WILLIAMS, E., FOULGER, R. E., LEIGH, S., DAUGHERTY, L. C., NIBLOCK, O., LEONG, I. U. S., SMITH, K. R., GERASIMENKO, O., HARALDSDOTTIR, E., THOMAS, E., SCOTT, R. H., BAPLE, E., TUCCI, A., BRITTAIN, H., DE BURCA, A., IBANEZ, K., KASPERAVICIUTE, D., SMEDLEY, D., CAULFIELD, M., RENDON, A. & MCDONAGH, E. M. 2019. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. *Nat Genet,* 51**,** 1560-1565.

MARTIN, H. C., JONES, W. D., MCINTYRE, R., SANCHEZ-ANDRADE, G., SANDERSON, M., STEPHENSON, J. D., JONES, C. P., HANDSAKER, J., GALLONE, G., BRUNTRAEGER, M., MCRAE, J. F., PRIGMORE, E., SHORT, P., NIEMI, M., KAPLANIS, J., RADFORD, E. J., AKAWI, N., BALASUBRAMANIAN, M., DEAN, J., HORTON, R., HULBERT, A., JOHNSON, D. S., JOHNSON, K., KUMAR, D., LYNCH, S. A., MEHTA, S. G., MORTON, J., PARKER, M. J., SPLITT, M., TURNPENNY, P. D., VASUDEVAN, P. C., WRIGHT, M., BASSETT, A., GERETY, S. S., WRIGHT, C. F., FITZPATRICK, D. R., FIRTH, H. V., HURLES, M. E. & BARRETT, J. C. 2018. Quantifying the contribution of recessive coding variation to developmental disorders. *Science,* 362**,** 1161–1164.

MARTIN, J. F., BRADLEY, A. & OLSON, E. N. 1995. The paired-like homeo box gene MHox is required for early events of skeletogenesis in multiple lineages. *Genes and Development,* 9**,** 1237-1249.

MARTINEZ, M. F., ROMANO, M. V., MARTINEZ, A. P., GONZALEZ, A., MUCHNIK, C., STENGEL, F. M., MAZZUOCCOLO, L. D. & AZURMENDI, P. J. 2019. Nevoid Basal Cell Carcinoma Syndrome: PTCH1 Mutation Profile and Expression of Genes Involved in the Hedgehog Pathway in Argentinian Patients. *Cells,* 8.

MARUYAMA, T., JEONG, J., SHEU, T. J. & HSU, W. 2016. Stem cells of the suture mesenchyme in craniofacial bone development, repair and regeneration. *Nat Commun,* 7**,** 10526.

MATSUO, I., KURATANI, S., KIMURA, C., TAKEDA, N. & AIZAWA, S. 1995. Mouse Otx2 functions in the formation and patterning of rostral head. *Genes & Development,* 9**,** 2646-2658.

MAVROGIANNIS, L. A., TAYLOR, I. B., DAVIES, S. J., RAMOS, F. J., OLIVARES, J. L. & WILKIE, A. O. 2006. Enlarged parietal foramina caused by mutations in the homeobox genes ALX4 and MSX2: from genotype to phenotype. *Eur J Hum Genet,* 14**,** 151-8.

MAXIMY, A. A. D., NAKATAKE, Y., MONCADA, S., ITOH, N., THIERY, J. P. & BELLUSCI, S. 1999. Cloning and expression pattern of a mouse homologue of Drosophila sprouty in the mouse embryo. *Mechanisms of Development,* 81**,** 213-216.

MCLAREN, W., GIL, L., HUNT, S. E., RIAT, H. S., RITCHIE, G. R., THORMANN, A., FLICEK, P. & CUNNINGHAM, F. 2016. The Ensembl Variant Effect Predictor. *Genome Biol,* 17**,** 122.

MCMULLAN, D., ELLARD, S., WILLIAMS, M., BAPLE, E., ELMSLIE, F., THOMAS, E., DEANS, S., FRATTER, C. & MEIN, R. 2021. Guidelines for rare disease whole genome sequencing & next generation sequencing panel interpretation & reporting. *London: NHS England and NHS Improvement*.

MCNEILL, A., IOVINO, E., MANSARD, L., VACHE, C., BAUX, D., BEDOUKIAN, E., COX, H., DEAN, J., GOUDIE, D., KUMAR, A., NEWBURY-ECOB, R., FALLERINI, C., RENIERI, A., LOPERGOLO, D., MARI, F., BLANCHET, C., WILLEMS, M., ROUX, A. F., PIPPUCCI, T. & DELPIRE, E. 2020. SLC12A2 variants cause a neurodevelopmental disorder or cochleovestibular defect. *Brain,* 143**,** 2380-2387.

MERRILL, A. E., BOCHUKOVA, E. G., BRUGGER, S. M., ISHII, M., PILZ, D. T., WALL, S. A., LYONS, K. M., WILKIE, A. O. & MAXSON, R. E., JR. 2006. Cell mixing at a neural crest-mesoderm boundary and deficient ephrin-Eph signaling in the pathogenesis of craniosynostosis. *Hum Mol Genet,* 15**,** 1319-28.

MICHA, D., GUO, D. C., HILHORST-HOFSTEE, Y., VAN KOOTEN, F., ATMAJA, D., OVERWATER, E., CAYAMI, F. K., REGALADO, E. S., VAN UFFELEN, R., VENSELAAR, H., FARADZ, S. M., VRIEND, G., WEISS, M. M., SISTERMANS, E. A., MAUGERI, A., MILEWICZ, D. M., PALS, G. & VAN DIJK, F. S. 2015. SMAD2 Mutations Are Associated with Arterial Aneurysms and Dissections. *Hum Mutat,* 36**,** 1145-9.

MIGNOT, C., MOUTARD, M. L., RASTETTER, A., BOUTAUD, L., HEIDE, S., BILLETTE, T., DOUMMAR, D., GAREL, C., AFENJAR, A., JACQUETTE, A., LACOMBE, D., VERLOES, A., BOLE-FEYSOT, C., NITSCHKE, P., MASSON, C., FAUDET, A., LESNE, F., BIENVENU, T., ALBY, C., ATTIE-BITACH, T., DEPIENNE, C., NAVA, C. & HERON, D. 2016. ARID1B mutations are the major genetic cause of corpus callosum anomalies in patients with intellectual disability. *Brain,* 139**,** e64.

MILILLO, A., LA CARPIA, F., COSTANZI, S., D'URBANO, V., MARTINI, M., LANUTI, P., VISCHINI, G., LAROCCA, L. M., MARCHISIO, M., MISCIA, S., AMOROSO, A., GURRIERI, F. & SANGIORGI, E. 2015. A SPRY2 mutation leading to MAPK/ERK pathway inhibition is associated with an autosomal dominant form of IgA nephropathy. *Eur J Hum Genet,* 23**,** 1673-8.

MILILLO, A., MOLINARIO, C., COSTANZI, S., VISCHINI, G., LA CARPIA, F., LA GRECA, F., RIGANTE, D., GAMBARO, G., GURRIERI, F. & SANGIORGI, E. 2018. Defective activation of the MAPK/ERK pathway, leading to PARP1 and DNMT1 dysregulation, is a common defect in IgA nephropathy and Henoch-Schonlein purpura. *J Nephrol,* 31**,** 731-741.

MILLER, K. A., TWIGG, S. R., MCGOWAN, S. J., PHIPPS, J. M., FENWICK, A. L., JOHNSON, D., WALL, S. A., NOONS, P., REES, K. E., TIDEY, E. A., CRAFT, J., TAYLOR, J., TAYLOR, J. C., GOOS, J. A., SWAGEMAKERS, S. M., MATHIJSSEN, I. M., VAN DER SPEK, P. J., LORD, H., LESTER, T., ABID, N., CILLIERS, D., HURST, J. A., MORTON, J. E., SWEENEY, E., WEBER, A., WILSON, L. C. & WILKIE, A. O. 2017. Diagnostic value of exome and whole genome sequencing in craniosynostosis. *J Med Genet,* 54**,** 260-268.

MING, Q., ROSKE, Y., SCHUETZ, A., WALENTIN, K., IBRAIMI, I., SCHMIDT-OTT, K. M. & HEINEMANN, U. 2018. Structural basis of gene regulation by the Grainyhead/CP2 transcription factor family. *Nucleic Acids Res,* 46**,** 2082-2095.

MIRAOUI, H., DWYER, A. A., SYKIOTIS, G. P., PLUMMER, L., CHUNG, W., FENG, B., BEENKEN, A., CLARKE, J., PERS, T. H., DWORZYNSKI, P., KEEFE, K., NIEDZIELA, M., RAIVIO, T., CROWLEY, W. F., JR., SEMINARA, S. B., QUINTON, R., HUGHES, V. A., KUMANOV, P., YOUNG, J., YIALAMAS, M. A., HALL, J. E., VAN VLIET, G., CHANOINE, J. P., RUBENSTEIN, J., MOHAMMADI, M., TSAI, P. S., SIDIS, Y., LAGE, K. & PITTELOUD, N. 2013. Mutations in FGF17, IL17RD, DUSP6, SPRY4, and FLRT3 are identified in individuals with congenital hypogonadotropic hypogonadism. *Am J Hum Genet,* 92**,** 725-43.

MOEDE, T., LEIBIGER, B., POUR, H. G., BERGGREN, P.-O. & LEIBIGER, I. B. 1999. Identification of a nuclear localization signal, RRMKWKK, in the homeodomain transcription factor PDX-1. *FEBS,* 461**,** 229-234.

MORNET, E., TAILLANDIER, A., DOMINGUES, C., DUFOUR, A., BENALOUN, E., LAVAUD, N., WALLON, F., ROUSSEAU, N., CHARLE, C., GUBERTO, M., MUTI, C. & SIMON-BOUY, B. 2020. Hypophosphatasia: a genetic-based nosology and new insights

in genotype-phenotype correlation. *European Journal of Human Genetics,* 29**,** 289-299.

MORRISS-KAY, G. M. & WILKIE, A. O. M. 2005. Growth of the normal skull vault and its alteration in craniosynostosis: insights from human genetics and experimental studies. *Journal of Anatomy***,** 637-653.

MORTON, J. E., FRENTZ, S., MORGAN, T., SUTHERLAND-SMITH, A. J. & ROBERTSON, S. P. 2016. Biallelic mutations in CYP26B1: A differential diagnosis for Pfeiffer and Antley-Bixler syndromes. *Am J Med Genet A,* 170**,** 2706-10.

MOTOYAMA, J., TAKABATAKE, T., TAKESHIMA, K. & HUI, C.-C. 1998. Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog. *Nature Genetics,* 18**,** 104-106.

MURAKAMI, H., OKAWA, A., YOSHIDA, H., NISHIKAWA, S., MORIYA, H. & KOSEKI, H. 2002. Elbow knee synostosis (Eks): a new mutation on mouse Chromosome 14. *Mamm Genome,* 13**,** 341-4.

MUTAI, H., WASANO, K., MOMOZAWA, Y., KAMATANI, Y., MIYA, F., MASUDA, S., MORIMOTO, N., NARA, K., TAKAHASHI, S., TSUNODA, T., HOMMA, K., KUBO, M. & MATSUNAGA, T. 2020. Variants encoding a restricted carboxy-terminal domain of SLC12A2 cause hereditary hearing loss in humans. *PLoS Genet,* 16**,** e1008643.

NAGAYAMA, T., OKUHARA, S., OTA, M. S., TACHIKAWA, N., KASUGAI, S. & ISEKI, S. 2013. FGF18 accelerates osteoblast differentiation by upregulating Bmp2 expression. *Congenit Anom (Kyoto),* 53**,** 83-8.

NAJMI, L. A., AUKRUST, I., FLANNICK, J., MOLNES, J., BURTT, N., MOLVEN, A., GROOP, L., ALTSHULER, D., JOHANSSON, S., BJØRKHAUG, L. & NJØLSTAD, P. R. 2017. Functional Investigations of HNF1A Identify Rare Variants as Risk Factors for Type 2 Diabetes in the General Population. *Diabetes,* 66**,** 335-346.

NAKANO, Y., GUERRERO, I., HIDALGO, A., TAYLOR, A., WHITTLET, J. R. S. & INGHAM, P. W. 1989. A protein with several possible membrane-spanning domains encoded by the Drosophila segment polarity gene patched. *Nature,* 341**,** 508-513.

NG, P. C. & HENIKOFF, S. 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res,* 31**,** 3812-4.

NG, S. B., BUCKINGHAM, K. J., LEE, C., BIGHAM, A. W., TABOR, H. K., DENT, K. M., HUFF, C. D., SHANNON, P. T., JABS, E. W., NICKERSON, D. A., SHENDURE, J. & BAMSHAD, M. J. 2010. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet,* 42**,** 30-5.

NIEMINEN, P., MORGAN, N. V., FENWICK, A. L., PARMANEN, S., VEISTINEN, L., MIKKOLA, M. L., VAN DER SPEK, P. J., GIRAUD, A., JUDD, L., ARTE, S., BRUETON, L. A., WALL, S. A., MATHIJSSEN, I. M., MAHER, E. R., WILKIE, A. O., KREIBORG, S. & THESLEFF, I. 2011. Inactivation of IL11 signaling causes craniosynostosis, delayed tooth eruption, and supernumerary teeth. *Am J Hum Genet,* 89**,** 67-81.

NIETHAMER, T. K. & BUSH, J. O. 2019. Getting direction(s): The Eph/ephrin signaling system in cell positioning. *Dev Biol,* 447**,** 42-57.

NIEUWENHUIS, E., MOTOYAMA, J., BARNFIELD, P. C., YOSHIKAWA, Y., ZHANG, X., MO, R., CRACKOWER, M. A. & HUI, C. C. 2006. Mice with a targeted mutation of patched2 are viable but develop alopecia and epidermal hyperplasia. *Mol Cell Biol,* 26**,** 6609-22.

NORRIS, R. & KERN, M. 2001a. Identification of Domains Mediating Transcription Activation, Repression, and Inhibition in the Paired-Related Homeobox Protein, Prx2. *DNA and Cell Biology,* 20**,** 89-99.

NORRIS, R. A. & KERN, M. J. 2001b. The Identification of Prx1 Transcription Regulatory Domains Provides a Mechanism for Unequal Compensation by the Prx1 and Prx2 Loci. *Journal of Biological Chemistry,* 276**,** 26829-26837.

NORRIS, R. A., SCOTT, K. K., MOORE, C. S., STETTEN, G., BROWN, C. R., JABS, E. W., WULFSBERG, E. A., YU, J. & KERN, M. J. 2000. Human PRRX1 and PRRX2 genes: cloning, expression, genomic localization, and exclusion as disease genes for Nager syndrome. *Mamm Genome,* 11**,** 1000-5.

NOYES, M. B., CHRISTENSEN, R. G., WAKABAYASHI, A., STORMO, G. D., BRODSKY, M. H. & WOLFE, S. A. 2008. Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell,* 133**,** 1277-89.

NÜSSLEIN-VOLHARD, C. & WIESCHAUS, E. 1980. Mutations affecting segment number and polarity in Drosophila. *Nature,* 287**,** 785-801.

OGURA, Y., UEHARA, T., UJIBE, K., YOSHIHASHI, H., YAMADA, M., SUZUKI, H., TAKENOUCHI, T., KOSAKI, K. & HIRATA, H. 2022. The p.Thr395Met missense variant of NFIA found in a patient with intellectual disability is a defective variant. *Am J Med Genet A,* 188**,** 1184-1192.

OHBAYASHI, N., SHIBAYAMA, M., KUROTAKI, Y., IMANISHI, M., FUJIMORI, T., ITOH, N. & TAKADA, S. 2002. FGF18 is required for normal cell proliferation and differentiation during osteogenesis and chondrogenesis. *Genes Dev,* 16**,** 870-9.

OPPERMAN, L. A. 2000. Cranial Sutures as Intramembranous Bone Growth Sites. *Developmental Dynamics,* 219**,** 472-485.

ORNITZ, D. M. & MARIE, P. J. 2015. Fibroblast growth factor signaling in skeletal development and disease. *Genes and Development,* 29**,** 1463-1486.

OUYANG, Z., CHEN, Z., ISHIKAWA, M., YUE, X., KAWANAMI, A., LEAHY, P., GREENFIELD, E. M. & MURAKAMI, S. 2014. Prx1 and 3.2kb Col1a1 promoters target distinct bone cell populations in transgenic mice. *Bone,* 58**,** 136-145.

PAN, A., CHANG, L., NGUYEN, A. & JAMES, A. W. 2013. A review of hedgehog signaling in cranial bone development. *Front Physiol,* 4**,** 61.

PATAT, O., VAN RAVENSWAAIJ-ARTS, C. M., TANTAU, J., CORSTEN-JANSSEN, N., VAN TINTELEN, J. P., DIJKHUIZEN, T., KAPLAN, J. & CHASSAING, N. 2013. Otocephaly-Dysgnathia Complex: Description of Four Cases and Confirmation of the Role of OTX2. *Mol Syndromol,* 4**,** 302-5.

PEDERSEN, B. S., BROWN, J. M., DASHNOW, H., WALLACE, A. D., VELINDER, M., TRISTANI-FIROUZI, M., SCHIFFMAN, J. D., TVRDIK, T., MAO, R., BEST, D. H., BAYRAK-TOYDEMIR, P. & QUINLAN, A. R. 2021. Effective variant filtering and expected candidate variant yield in studies of rare human disease. *NPJ Genom Med,* 6**,** 60.

PEDERSEN, B. S., LAYER, R. M. & QUINLAN, A. R. 2016. Vcfanno: fast, flexible annotation of genetic variants. *Genome Biol,* 17**,** 118.

PEHLIVAN, D., KARACA, E., AYDIN, H., BECK, C. R., GAMBIN, T., MUZNY, D. M., BILGE GECKINLI, B., KARAMAN, A., JHANGIANI, S. N., CENTERS FOR MENDELIAN, G., GIBBS, R. A. & LUPSKI, J. R. 2014. Whole-exome sequencing links TMCO1 defect syndrome with cerebro-facio-thoracic dysplasia. *Eur J Hum Genet,* 22**,** 1145-8.

PERCIVAL, C. J., MARANGONI, P., TAPALTSYAN, V., KLEIN, O. & HALLGRIMSSON, B. 2017. The Interaction of Genetic Background and Mutational Effects in Regulation of Mouse Craniofacial Shape. *G3 (Bethesda),* 7**,** 1439-1450.

PEYRARD-JANVID, M., LESLIE, E. J., KOUSA, Y. A., SMITH, T. L., DUNNWALD, M., MAGNUSSON, M., LENTZ, B. A., UNNEBERG, P., FRANSSON, I., KOILLINEN, H. K., RAUTIO, J., PEGELOW, M., KARSTEN, A., BASEL-VANAGAITE, L., GORDON, W., ANDERSEN, B., SVENSSON, T., MURRAY, J. C., CORNELL, R. A., KERE, J. & SCHUTTE, B. C. 2014. Dominant mutations in GRHL3 cause Van der Woude Syndrome and disrupt oral periderm development. *Am J Hum Genet,* 94**,** 23-32.

POLI, M. C., EBSTEIN, F., NICHOLAS, S. K., DE GUZMAN, M. M., FORBES, L. R., CHINN, I. K., MACE, E. M., VOGEL, T. P., CARISEY, A. F., BENAVIDES, F., COBAN-AKDEMIR, Z. H., GIBBS, R. A., JHANGIANI, S. N., MUZNY, D. M., CARVALHO, C. M. B., SCHADY, D. A., JAIN, M., ROSENFELD, J. A., EMRICK, L., LEWIS, R. A., LEE, B., UNDIAGNOSED DISEASES NETWORK, M., ZIEBA, B. A., KURY, S., KRUGER, E., LUPSKI, J. R., BOSTWICK, B. L. & ORANGE, J. S. 2018. Heterozygous Truncating Variants in POMP Escape Nonsense-Mediated Decay and Cause a Unique Immune Dysregulatory Syndrome. *Am J Hum Genet,* 102**,** 1126-1142.

QI, X., SCHMIEGE, P., COUTAVAS, E., WANG, J. & LI, X. 2018. Structures of human Patched and its complex with native palmitoylated sonic hedgehog. *Nature,* 560**,** 128-132.

QIAN, H., CAO, P., HU, M., GAO, S., YAN, N. & GONG, X. 2019. Inhibition of tetrameric Patched1 by Sonic Hedgehog through an asymmetric paradigm. *Nat Commun,* 10**,** 2320.

RADHAKRISHNAN, A., ROHATGI, R. & SIEBOLD, C. 2020. Cholesterol access in cellular membranes controls Hedgehog signaling. *Nat Chem Biol,* 16**,** 1303-1313.

RAHBARI, R., WUSTER, A., LINDSAY, S. J., HARDWICK, R. J., ALEXANDROV, L. B., TURKI, S. A., DOMINICZAK, A., MORRIS, A., PORTEOUS, D., SMITH, B., STRATTON, M. R., CONSORTIUM, U. K. & HURLES, M. E. 2016. Timing, rates and spectra of human germline mutation. *Nat Genet,* 48**,** 126-133.

RAN, F. A., HSU, P. D., WRIGHT, J., AGARWALA, V., SCOTT, D. A. & ZHANG, F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc,* 8**,** 2281-2308.

REARDON, W., WINTER, R. M., RUTLAND, P., PULLEYN, L. J., JONES, B. M. & MALCOLM, S. 1994. Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nature Genetics,* 8**,** 98–103.

REIFF, T., TSAROVINA, K., MAJDAZARI, A., SCHMIDT, M., DEL PINO, I. & ROHRER, H. 2010. Neuroblastoma phox2b variants stimulate proliferation and dedifferentiation of immature sympathetic neurons. *J Neurosci,* 30**,** 905-15.

REIJNDERS, M. R. F., MILLER, K. A., ALVI, M., GOOS, J. A. C., LEES, M. M., DE BURCA, A., HENDERSON, A., KRAUS, A., MIKAT, B., DE VRIES, B. B. A., ISIDOR, B., KERR, B., MARCELIS, C., SCHLUTH-BOLARD, C., DESHPANDE, C., RUIVENKAMP, C. A. L., WIECZOREK, D., DECIPHERING DEVELOPMENTAL DISORDERS, S., BARALLE, D., BLAIR, E. M., ENGELS, H., LUDECKE, H. J., EASON, J., SANTEN, G. W. E., CLAYTON-SMITH, J., CHANDLER, K., TATTON-BROWN, K., PAYNE, K., HELBIG, K., RADTKE, K., NUGENT, K. M., CREMER, K., STROM, T. M., BIRD, L. M., SINNEMA, M., BITNER-GLINDZICZ, M., VAN DOOREN, M. F., ALDERS, M., KOOPMANS, M., BRICK, L., KOZENKO, M., HARLINE, M. L., KLAASSENS, M., STEINRATHS, M., COOPER, N. S., EDERY, P., YAP, P., TERHAL, P. A., VAN DER SPEK, P. J.,

LAKEMAN, P., TAYLOR, R. L., LITTLEJOHN, R. O., PFUNDT, R., MERCIMEK-ANDREWS, S., STEGMANN, A. P. A., KANT, S. G., MCLEAN, S., JOSS, S., SWAGEMAKERS, S. M. A., DOUZGOU, S., WALL, S. A., KURY, S., CALPENA, E., KOELLING, N., MCGOWAN, S. J., TWIGG, S. R. F., MATHIJSSEN, I. M. J., NELLAKER, C., BRUNNER, H. G. & WILKIE, A. O. M. 2018. De Novo and Inherited Loss-of-Function Variants in TLK2: Clinical and Genotype-Phenotype Evaluation of a Distinct Neurodevelopmental Disorder. *Am J Hum Genet,* 102**,** 1195-1203.

REINDERS, M. G., VAN HOUT, A. F., COSGUN, B., PAULUSSEN, A. D., LETER, E. M., STEIJLEN, P. M., MOSTERD, K., VAN GEEL, M. & GILLE, J. J. 2018. New mutations and an updated database for the patched-1 (PTCH1) gene. *Mol Genet Genomic Med,* 6**,** 409-415.

RENIER, D., G CINALLI, LAJEUNIE, E., ARNAUD, E. & MARCHAC, D. 1997. Oxycephaly, a severe craniosynostosis. Apropos of a series of 129 cases. *Arch Pediatr,* 4.

RENTZSCH, P., WITTEN, D., COOPER, G. M., SHENDURE, J. & KIRCHER, M. 2019. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res,* 47**,** D886-D894.

RIBES, V., OTTO, D. M., DICKMANN, L., SCHMIDT, K., SCHUHBAUR, B., HENDERSON, C., BLOMHOFF, R., WOLF, C. R., TICKLE, C. & DOLLE, P. 2007. Rescue of cytochrome P450 oxidoreductase (Por) mouse mutants reveals functions in vasculogenesis, brain and limb patterning linked to retinoic acid homeostasis. *Dev Biol,* 303**,** 66-81.

RICE, D. P., CONNOR, E. C., VELTMAAT, J. M., LANA-ELOLA, E., VEISTINEN, L., TANIMOTO, Y., BELLUSCI, S. & RICE, R. 2010. Gli3Xt-J/Xt-J mice exhibit lambdoid suture craniosynostosis which results from altered osteoprogenitor proliferation and differentiation. *Hum Mol Genet,* 19**,** 3457-67.

RIEUBLAND, C., HOLMES, A. D., CARAMINS, M., ROSCIOLI, T. & AMOR, D. J. 2011. Lambdoid synostosis and craniofacial dysmorphism with normal intellect: A novel syndrome? *Am J Med Genet A,* 155A**,** 233-4.

RITTER, A. L., MCDOUGALL, C., SKRABAN, C., MEDNE, L., BEDOUKIAN, E. C., ASHER, S. B., BALCIUNIENE, J., CAMPBELL, C. D., BAKER, S. W., DENENBERG, E. H., MAZZOLA, S., FIORDALISO, S. K., KRANTZ, I. D., KAPLAN, P., IERARDI-CURTO, L., SANTANI, A. B., ZACKAI, E. H. & IZUMI, K. 2018. Variable Clinical Manifestations of Xia-Gibbs syndrome: Findings of Consecutively Identified Cases at a Single Children's Hospital. *Am J Med Genet A,* 176**,** 1890-1896.

RODRIGUEZ-ZABALA, M., AZA-CARMONA, M., RIVERA-PEDROZA, C. I., BELINCHON, A., GUERRERO-ZAPATA, I., BARRAZA-GARCIA, J., VALLESPIN, E., LU, M., DEL POZO, A., GLUCKSMAN, M. J., SANTOS-SIMARRO, F. & HEATH, K. E. 2017. FGF9 mutation causes craniosynostosis along with multiple synostoses. *Hum Mutat,* 38**,** 1471-1476.

ROSENFELD, J. A., XIAO, R., BEKHEIRNIA, M. R., KANANI, F., PARKER, M. J., KOENIG, M. K., VAN HAERINGEN, A., RUIVENKAMP, C., ROSMANINHO-SALGADO, J., ALMEIDA, P. M., SA, J., PINTO BASTO, J., PALEN, E., OETJENS, K. F., BURRAGE, L. C., XIA, F., LIU, P., ENG, C. M., UNDIAGNOSED DISEASES, N., YANG, Y., POSEY, J. E. & LEE, B. H. 2021. Heterozygous variants in SPTBN1 cause intellectual disability and autism. *Am J Med Genet A,* 185**,** 2037-2045.

ROTHBERG, J. M., HINZ, W., REARICK, T. M., SCHULTZ, J., MILESKI, W., DAVEY, M., LEAMON, J. H., JOHNSON, K., MILGREW, M. J., EDWARDS, M., HOON, J.,

SIMONS, J. F., MARRAN, D., MYERS, J. W., DAVIDSON, J. F., BRANTING, A., NOBILE, J. R., PUC, B. P., LIGHT, D., CLARK, T. A., HUBER, M., BRANCIFORTE, J. T., STONER, I. B., CAWLEY, S. E., LYONS, M., FU, Y., HOMER, N., SEDOVA, M., MIAO, X., REED, B., SABINA, J., FEIERSTEIN, E., SCHORN, M., ALANJARY, M., DIMALANTA, E., DRESSMAN, D., KASINSKAS, R., SOKOLSKY, T., FIDANZA, J. A., NAMSARAEV, E., MCKERNAN, K. J., WILLIAMS, A., ROTH, G. T. & BUSTILLO, J. 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature,* 475**,** 348-52.

SABHERWAL, N., SCHNEIDER, K. U., BLASCHKE, R. J., MARCHINI, A. & RAPPOLD, G. 2004. Impairment of SHOX nuclear localization as a cause for Leri-Weill syndrome. *J Cell Sci,* 117**,** 3041-8.

SACCHETTI, B., FUNARI, A., MICHIENZI, S., DI CESARE, S., PIERSANTI, S., SAGGIO, I., TAGLIAFICO, E., FERRARI, S., ROBEY, P. G., RIMINUCCI, M. & BIANCO, P. 2007. Self-renewing osteoprogenitors in bone marrow sinusoids can organize a hematopoietic microenvironment. *Cell,* 131**,** 324-36.

SAMOCHA, K. E., KOSMICKI, J. A., KARCZEWSKI, K. J., O'DONNELL-LURIA, A. H., PIERCE-HOFFMAN, E., MACARTHUR, D. G., NEALE, B. M. & DALY, M. J. 2017. Regional missense constraint improves variant deleteriousness prediction. *bioRxiv*.

SANCHEZ-LARA, P. A., CARMICHAEL, S. L., GRAHAM, J. M., JR., LAMMER, E. J., SHAW, G. M., MA, C., RASMUSSEN, S. A. & NATIONAL BIRTH DEFECTS PREVENTION STUDY 2010. Fetal constraint as a potential risk factor for craniosynostosis. *Am J Med Genet A,* 152A**,** 394-400.

SANGER, F., NICKLEN, S. & COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A,* 74**,** 5463-5467.

SCHEPERS, D., TORTORA, G., MORISAKI, H., MACCARRICK, G., LINDSAY, M., LIANG, D., MEHTA, S. G., HAGUE, J., VERHAGEN, J., VAN DE LAAR, I., WESSELS, M., DETISCH, Y., VAN HAELST, M., BAAS, A., LICHTENBELT, K., BRAUN, K., VAN DER LINDE, D., ROOS-HESSELINK, J., MCGILLIVRAY, G., MEESTER, J., MAYSTADT, I., COUCKE, P., EL-KHOURY, E., PARKASH, S., DINESS, B., RISOM, L., SCURR, I., HILHORST-HOFSTEE, Y., MORISAKI, T., RICHER, J., DESIR, J., KEMPERS, M., RIDEOUT, A. L., HORNE, G., BENNETT, C., RAHIKKALA, E., VANDEWEYER, G., ALAERTS, M., VERSTRAETEN, A., DIETZ, H., VAN LAER, L. & LOEYS, B. 2018. A mutation update on the LDS-associated genes TGFB2/3 and SMAD2/3. *Hum Mutat,* 39**,** 621-634.

SCHNEIDER, K. U., MARCHINI, A., SABHERWAL, N., ROTH, R., NIESLER, B., MARTTILA, T., BLASCHKE, R. J., LAWSON, M., DUMIC, M. & RAPPOLD, G. 2005. Alteration of DNA binding, dimerization, and nuclear translocation of SHOX homeodomain mutations identified in idiopathic short stature and Leri-Weill dyschondrosteosis. *Hum Mutat,* 26**,** 44-52.

SCHUBERT, M. S., THOMMANDRU, B., WOODLEY, J., TURK, R., YAN, S., KURGAN, G., MCNEILL, M. S. & RETTIG, G. R. 2021. Optimized design parameters for CRISPR Cas9 and Cas12a homology-directed repair. *Scientific Reports,* 11**,** 19482.

SCHWARZ, J. M., COOPER, D. N., SCHUELKE, M. & SEELOW, D. 2014. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods,* 11**,** 361-2.

SCHWARZE, K., BUCHANAN, J., FERMONT, J. M., DREAU, H., TILLEY, M. W., TAYLOR, J. M., ANTONIOU, P., KNIGHT, S. J. L., CAMPS, C., PENTONY, M. M., KVIKSTAD, E. M., HARRIS, S., POPITSCH, N., PAGNAMENTA, A. T., SCHUH, A., TAYLOR, J. C. &

WORDSWORTH, S. 2020. The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genet Med,* 22**,** 85-94.

SCHWARZE, K., BUCHANAN, J., TAYLOR, J. C. & WORDSWORTH, S. 2018. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med,* 20**,** 1122-1130.

SCHWERD, T., KRAUSE, F., TWIGG, S. R. F., ASCHENBRENNER, D., CHEN, Y.-H., BORGMEYER, U., MÜLLER, M., MANRIQUE, S., SCHUMACHER, N., WALL, S. A., JUNG, J., DAMM, T., GLÜER, C.-C., SCHELLER, J., ROSE-JOHN, S., JONES, E. Y., LAURENCE, A., WILKIE, A. O. M., SCHMIDT-ARRAS, D. & UHLIG, H. H. 2020. A variant in IL6ST with a selective IL-11 signaling defect in human and mouse. *Bone Research,* 8**,** 24.

SCHWERD, T., TWIGG, S. R. F., ASCHENBRENNER, D., MANRIQUE, S., MILLER, K. A., TAYLOR, I. B., CAPITANI, M., MCGOWAN, S. J., SWEENEY, E., WEBER, A., CHEN, L., BOWNESS, P., RIORDAN, A., CANT, A., FREEMAN, A. F., MILNER, J. D., HOLLAND, S. M., FREDE, N., MULLER, M., SCHMIDT-ARRAS, D., GRIMBACHER, B., WALL, S. A., JONES, E. Y., WILKIE, A. O. M. & UHLIG, H. H. 2017. A biallelic mutation in IL6ST encoding the GP130 co-receptor causes immunodeficiency and craniosynostosis. *J Exp Med,* 214**,** 2547-2562.

SENARATNE, T. N. & QUINTERO-RIVERA, F. 1993. NFIA-Related Disorder. *In:* ADAM, M. P., EVERMAN, D. B., MIRZAA, G. M., PAGON, R. A., WALLACE, S. E., BEAN, L. J. H., GRIPP, K. W. & AMEMIYA, A. (eds.) *GeneReviews(®).* Seattle (WA): University of Washington, Seattle.

SENTCHORDI-MONTANÉ, L., DIAZ-GONZALEZ, F., CÁTEDRA-VALLÉS, E. V. & HEATH, K. E. 2021. Identification of the third FGF9 variant in a girl with multiple synostosis-comparison of the genotype:phenotype of FGF9 variants in humans and mice. *Clin Genet,* 99**,** 309-312.

SEPPORTA, M.-V., PRAZ, V., BALMAS BOURLOUD, K., JOSEPH, J.-M., JAUQUIER, N., RIGGI, N., NARDOU-AUDERSET, K., PETIT, A., SCOAZEC, J.-Y., SARTELET, H., RENELLA, R. & MÜHLETHALER-MOTTET, A. 2022. TWIST1 expression is associated with high-risk neuroblastoma and promotes primary and metastatic tumor growth. *Communications Biology,* 5**,** 42.

SERGI, C. & KAMNASARAN, D. 2011. PRRX1 is mutated in a fetus with agnathia-otocephaly. *Clin Genet,* 79**,** 293-5.

SERGOUNIOTIS, P. I., URQUHART, J. E., WILLIAMS, S. G., BHASKAR, S. S., BLACK, G. C., LOVELL, S. C., WHITBY, D. J., NEWMAN, W. G. & CLAYTON-SMITH, J. 2015. Agnathia-otocephaly complex and asymmetric velopharyngeal insufficiency due to an in-frame duplication in OTX2. *J Hum Genet,* 60**,** 199-202.

SHARKIA, R., ZALAN, A., JABAREEN-MASRI, A., HENGEL, H., SCHOLS, L., KESSEL, A., AZEM, A. & MAHAJNAH, M. 2019. A novel biallelic loss-of-function mutation in TMCO1 gene confirming and expanding the phenotype spectrum of cerebro-facio-thoracic dysplasia. *Am J Med Genet A,* 179**,** 1338-1345.

SHARMA, V. P., FENWICK, A. L., BROCKOP, M. S., MCGOWAN, S. J., GOOS, J. A., HOOGEBOOM, A. J., BRADY, A. F., JEELANI, N. O., LYNCH, S. A., MULLIKEN, J. B., MURRAY, D. J., PHIPPS, J. M., SWEENEY, E., TOMKINS, S. E., WILSON, L. C., BENNETT, S., CORNALL, R. J., BROXHOLME, J., KANAPIN, A., WHOLE-GENOME SEQUENCES, C., JOHNSON, D., WALL, S. A., VAN DER SPEK, P. J., MATHIJSSEN, I.

M., MAXSON, R. E., TWIGG, S. R. & WILKIE, A. O. 2013. Mutations in TCF12, encoding a basic helix-loop-helix partner of TWIST1, are a frequent cause of coronal craniosynostosis. *Nat Genet,* 45**,** 304-7.

SHEA, K. L., XIANG, W., LAPORTA, V. S., LICHT, J. D., KELLER, C., BASSON, M. A. & BRACK, A. S. 2010. Sprouty1 regulates reversible quiescence of a self-renewing adult muscle stem cell pool during regeneration. *Cell Stem Cell,* 6**,** 117-29.

SHIMOJIMA, K., ONDO, Y., MATSUFUJI, M., SANO, N., TSURU, H., OYOSHI, T., HIGA, N., TOKIMURA, H., ARITA, K. & YAMAMOTO, T. 2016. Concurrent occurrence of an inherited 16p13.11 microduplication and a de novo 19p13.3 microdeletion involving MAP2K2 in a patient with developmental delay, distinctive facial features, and lambdoid synostosis. *Eur J Med Genet,* 59**,** 559-563.

SHIN, H. I., DIVIETI, P., SIMS, N. A., KOBAYASHI, T., MIAO, D., KARAPLIS, A. C., BARON, R., BRINGHURST, R. & KRONENBERG, H. M. 2004. Gp130-mediated signaling is necessary for normal osteoblastic function in vivo and in vitro. *Endocrinology,* 145**,** 1376-85.

SIGNETTI, L., ELIZAROV, N., SIMSIR, M., PAQUET, A., DOUGUET, D., LABBAL, F., DEBAYLE, D., DI GIORGIO, A., BIOU, V., GIRARD, C., DUCA, M., BRETILLON, L., BERTOLOTTO, C., VERRIER, B., AZOULAY, S. & MUS-VETEAU, I. 2020. Inhibition of Patched Drug Efflux Increases Vemurafenib Effectiveness against Resistant Braf(V600E) Melanoma. *Cancers (Basel),* 12.

SIM, J. C., WHITE, S. M. & LOCKHART, P. J. 2015. ARID1B-mediated disorders: Mutations and possible mechanisms. *Intractable Rare Dis Res,* 4**,** 17-23.

SIMS, N. A., JENKINS, B. J., NAKAMURA, A., QUINN, J. M., LI, R., GILLESPIE, M. T., ERNST, M., ROBB, L. & MARTIN, T. J. 2005. Interleukin-11 receptor signaling is required for normal bone remodeling. *J Bone Miner Res,* 20**,** 1093-102.

SINGH, R. P., DHARIWAL, D., BHUJEL, N., SHAIKH, Z., DAVIES, P., NISHIKAWA, H., SOLANKI, G. & DOVER, M. S. 2010. Role of parental risk factors in the aetiology of isolated non-syndromic metopic craniosynostosis. *Br J Oral Maxillofac Surg,* 48**,** 438-42.

SLATKO, B. E., GARDNER, A. F. & AUSUBEL, F. M. 2018. Overview of Next-Generation Sequencing Technologies. *Curr Protoc Mol Biol,* 122**,** e59.

SMEDLEY, D., JACOBSEN, J. O. B., JÄGER, M., KÖHLER, S., HOLTGREWE, M., SCHUBACH, M., SIRAGUSA, E., ZEMOJTEL, T., BUSKE, O. J., WASHINGTON, N. L., BONE, W. P., HAENDEL, M. A. & ROBINSON, P. N. 2015. Next-generation diagnostics and disease-gene discovery with the Exomiser. *Nature Protocols,* 10**,** 2004-2015.

SMEDLEY, D., SMITH, K. R., MARTIN, A., THOMAS, E. A., MCDONAGH, E. M., CIPRIANI, V., ELLINGFORD, J. M., ARNO, G., TUCCI, A., VANDROVCOVA, J., CHAN, G., WILLIAMS, H. J., RATNAIKE, T., WEI, W., STIRRUPS, K., IBANEZ, K., MOUTSIANAS, L., WIELSCHER, M., NEED, A., BARNES, M. R., VESTITO, L., BUCHANAN, J., WORDSWORTH, S., ASHFORD, S., REHMSTROM, K., LI, E., FULLER, G., TWISS, P., SPASIC-BOSKOVIC, O., HALSALL, S., FLOTO, R. A., POOLE, K., WAGNER, A., MEHTA, S. G., GURNELL, M., BURROWS, N., JAMES, R., PENKETT, C., DEWHURST, E., GRAF, S., MAPETA, R., KASANICKI, M., HAWORTH, A., SAVAGE, H., BABCOCK, M., REESE, M. G., BALE, M., BAPLE, E., BOUSTRED, C., BRITTAIN, H., DE BURCA, A., BLEDA, M., DEVEREAU, A., HALAI, D., HARALDSDOTTIR, E., HYDER, Z., KASPERAVICIUTE, D., PATCH, C., POLYCHRONOPOULOS, D., MATCHAN, A., SULTANA, R., RYTEN, M., TAVARES, A.

L. T., TREGIDGO, C., TURNBULL, C., WELLAND, M., WOOD, S., SNOW, C., WILLIAMS, E., LEIGH, S., FOULGER, R. E., DAUGHERTY, L. C., NIBLOCK, O., LEONG, I. U. S., WRIGHT, C. F., DAVIES, J., CRICHTON, C., WELCH, J., WOODS, K., ABULHOUL, L., AURORA, P., BOCKENHAUER, D., BROOMFIELD, A., CLEARY, M. A., LAM, T., DATTANI, M., FOOTITT, E., GANESAN, V., GRUNEWALD, S., COMPEYROT-LACASSAGNE, S., MUNTONI, F., PILKINGTON, C., QUINLIVAN, R., THAPAR, N., WALLIS, C., WEDDERBURN, L. R., WORTH, A., BUESER, T., COMPTON, C., DESHPANDE, C., et al. 2021. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. *N Engl J Med,* 385**,** 1868-1880.

SOROKIN, A., MOHAMMADI, M., HUANG, J. & SCHLESSINGER, J. 1994. Internalization of fibroblast growth factor receptor is inhibited by a point mutation at tyrosine 766. *Journal of Biological Chemistry,* 269**,** 17056-17061.

SPLINTER, K., ADAMS, D. R., BACINO, C. A., BELLEN, H. J., BERNSTEIN, J. A., CHEATLE-JARVELA, A. M., ENG, C. M., ESTEVES, C., GAHL, W. A., HAMID, R., JACOB, H. J., KIKANI, B., KOELLER, D. M., KOHANE, I. S., LEE, B. H., LOSCALZO, J., LUO, X., MCCRAY, A. T., METZ, T. O., MULVIHILL, J. J., NELSON, S. F., PALMER, C. G. S., PHILLIPS, J. A., 3RD, PICK, L., POSTLETHWAIT, J. H., REUTER, C., SHASHI, V., SWEETSER, D. A., TIFFT, C. J., WALLEY, N. M., WANGLER, M. F., WESTERFIELD, M., WHEELER, M. T., WISE, A. L., WORTHEY, E. A., YAMAMOTO, S., ASHLEY, E. A. & UNDIAGNOSED DISEASES, N. 2018. Effect of Genetic Diagnosis on Patients with Previously Undiagnosed Disease. *N Engl J Med,* 379**,** 2131-2139.

STAINS, J. P. & CIVITELLI, R. 2003. Genomic approaches to identifying transcriptional regulators of osteoblast differentiation. *Genome Biology,* 4**,** 222.

STEFANSSON, B., OHAMA, T., DAUGHERTY, A. E. & BRAUTIGAN, D. L. 2008. Protein phosphatase 6 regulatory subunits composed of ankyrin repeat domains. *Biochemistry,* 47**,** 1442-1451.

SUZUKI, T., SUZUKI, T., RAVEAU, M., MIYAKE, N., SUDO, G., TSURUSAKI, Y., WATANABE, T., SUGAYA, Y., TATSUKAWA, T., MAZAKI, E., SHIMOHATA, A., KUSHIMA, I., ALEKSIC, B., SHIINO, T., TOYOTA, T., IWAYAMA, Y., NAKAOKA, K., OHMORI, I., SASAKI, A., WATANABE, K., HIROSE, S., KANEKO, S., INOUE, Y., YOSHIKAWA, T., OZAKI, N., KANO, M., SHIMOJI, T., MATSUMOTO, N. & YAMAKAWA, K. 2020. A recurrent PJA1 variant in trigonocephaly and neurodevelopmental disorders. *Ann Clin Transl Neurol,* 7**,** 1117-1131.

SWEET, H. O., BRONSON, R. T., DONAHUE, L. R. & DAVISSON, M. T. 1996. Mesenchymal Dysplasia: A Recessive Mutation on Chromosome 13 of the Mouse. *Journal of Heredity,* 87**,** 87-95.

TAGARIELLO, A., HELLER, R., GREVEN, A., KALSCHEUER, V. M., MOLTER, T., RAUCH, A., KRESS, W. & WINTERPACHT, A. 2006. Balanced translocation in a patient with craniosynostosis disrupts the SOX6 gene and an evolutionarily conserved non-transcribed region. *J Med Genet,* 43**,** 534-40.

TAKABATAKE, T., OGAWA, M., TAKAHASHI, T., MIZUNO, M., OKAMOTO, M. & TAKESHIMA, K. 1997. Hedgehog and patched gene expression in adult ocular tissues. *FEBS Letters,* 410**,** 485-489.

TAKARADA, T., NAKAZATO, R., TSUCHIKANE, A., FUJIKAWA, K., IEZAKI, T., YONEDA, Y. & HINOI, E. 2016. Genetic analysis of Runx2 function during intramembranous ossification. *Development,* 143**,** 211-8.

TAKENOUCHI, T., HIDA, M., SAKAMOTO, Y., TORII, C., KOSAKI, R., TAKAHASHI, T. & KOSAKI, K. 2013. Severe congenital lipodystrophy and a progeroid appearance: Mutation in the penultimate exon of FBN1 causing a recognizable phenotype. *Am J Med Genet A,* 161A**,** 3057-62.

TAN, A., ABECASIS, G. R. & KANG, H. M. 2015. Unified representation of genetic variants. *Bioinformatics,* 31**,** 2202-4.

TANG, L., WU, X., ZHANG, H., LU, S., WU, M., SHEN, C., CHEN, X., WANG, Y., WANG, W., SHEN, Y., GU, M., DING, X., JIN, X., FEI, J. & WANG, Z. 2017. A point mutation in Fgf9 impedes joint interzone formation leading to multiple synostoses syndrome. *Hum Mol Genet,* 26**,** 1280-1293.

TANG, Y., KATURI, V., DILLNER, A., MISHRA, B., DENG, C.-X. & MISHRA, L. 2003. Disruption of Transforming Growth Factor– Signaling in ELF -Spectrin–Deficient Mice. *Science,* 299**,** 574-577.

TANIGUCHI, K., AYADA, T., ICHIYAMA, K., KOHNO, R., YONEMITSU, Y., MINAMI, Y., KIKUCHI, A., MAEHARA, Y. & YOSHIMURA, A. 2007. Sprouty2 and Sprouty4 are essential for embryonic morphogenesis and regulation of FGF signaling. *Biochem Biophys Res Commun,* 352**,** 896-902.

TAYLOR, J. C., MARTIN, H. C., LISE, S., BROXHOLME, J., CAZIER, J. B., RIMMER, A., KANAPIN, A., LUNTER, G., FIDDY, S., ALLAN, C., ARICESCU, A. R., ATTAR, M., BABBS, C., BECQ, J., BEESON, D., BENTO, C., BIGNELL, P., BLAIR, E., BUCKLE, V. J., BULL, K., CAIS, O., CARIO, H., CHAPEL, H., COPLEY, R. R., CORNALL, R., CRAFT, J., DAHAN, K., DAVENPORT, E. E., DENDROU, C., DEVUYST, O., FENWICK, A. L., FLINT, J., FUGGER, L., GILBERT, R. D., GORIELY, A., GREEN, A., GREGER, I. H., GROCOCK, R., GRUSZCZYK, A. V., HASTINGS, R., HATTON, E., HIGGS, D., HILL, A., HOLMES, C., HOWARD, M., HUGHES, L., HUMBURG, P., JOHNSON, D., KARPE, F., KINGSBURY, Z., KINI, U., KNIGHT, J. C., KROHN, J., LAMBLE, S., LANGMAN, C., LONIE, L., LUCK, J., MCCARTHY, D., MCGOWAN, S. J., MCMULLIN, M. F., MILLER, K. A., MURRAY, L., NEMETH, A. H., NESBIT, M. A., NUTT, D., ORMONDROYD, E., OTURAI, A. B., PAGNAMENTA, A., PATEL, S. Y., PERCY, M., PETOUSI, N., PIAZZA, P., PIRET, S. E., POLANCO-ECHEVERRY, G., POPITSCH, N., POWRIE, F., PUGH, C., QUEK, L., ROBBINS, P. A., ROBSON, K., RUSSO, A., SAHGAL, N., VAN SCHOUWENBURG, P. A., SCHUH, A., SILVERMAN, E., SIMMONS, A., SORENSEN, P. S., SWEENEY, E., TAYLOR, J., THAKKER, R. V., TOMLINSON, I., TREBES, A., TWIGG, S. R., UHLIG, H. H., VYAS, P., VYSE, T., WALL, S. A., WATKINS, H., WHYTE, M. P., WITTY, L., et al. 2015. Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. *Nat Genet,* 47**,** 717-726.

TEN BERGE, D., BROUWER, A., KORVING, J., MARTIN, J. F. & MEIJLINK, F. 1998. Prx1 and Prx2 in skeletogenesis: roles in the craniofacial region, inner ear and limbs. *Development,* 125**,** 3831-3842.

TENG, C. S., TING, M. C., FARMER, D. T., BROCKOP, M., MAXSON, R. E. & CRUMP, J. G. 2018. Altered bone growth dynamics prefigure craniosynostosis in a zebrafish model of Saethre-Chotzen syndrome. *Elife,* 7.

THORMANN, A., HALACHEV, M., MCLAREN, W., MOORE, D. J., SVINTI, V., CAMPBELL, A., KERR, S. M., TISCHKOWITZ, M., HUNT, S. E., DUNLOP, M. G., HURLES, M. E., WRIGHT, C. F., FIRTH, H. V., CUNNINGHAM, F. & FITZPATRICK, D. R. 2019.

Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. *Nat Commun,* 10**,** 2373.

TIMBERLAKE, A. T., CHOI, J., ZAIDI, S., LU, Q., NELSON-WILLIAMS, C., BROOKS, E. D., BILGUVAR, K., TIKHONOVA, I., MANE, S., YANG, J. F., SAWH-MARTINEZ, R., PERSING, S., ZELLNER, E. G., LORING, E., CHUANG, C., GALM, A., HASHIM, P. W., STEINBACHER, D. M., DILUNA, M. L., DUNCAN, C. C., PELPHREY, K. A., ZHAO, H., PERSING, J. A. & LIFTON, R. P. 2016. Two locus inheritance of non-syndromic midline craniosynostosis via rare SMAD6 and common BMP2 alleles. *Elife,* 5.

TIMBERLAKE, A. T., FUREY, C. G., CHOI, J., NELSON-WILLIAMS, C., YALE CENTER FOR GENOME, A., LORING, E., GALM, A., KAHLE, K. T., STEINBACHER, D. M., LARYSZ, D., PERSING, J. A. & LIFTON, R. P. 2017a. De novo mutations in inhibitors of Wnt, BMP, and Ras/ERK signaling pathways in non-syndromic midline craniosynostosis. *Proc Natl Acad Sci U S A,* 114**,** E7341-E7347.

TIMBERLAKE, A. T., FUREY, C. G., CHOI, J., NELSON-WILLIAMS, C., YALE CENTER FOR GENOME ANALYSIS, LORING, E., GALM, A., KAHLE, K. T., STEINBACHER, D. M., LARYSZ, D., PERSING, J. A. & LIFTON, R. P. 2017b. De novo mutations in inhibitors of Wnt, BMP, and Ras/ERK signaling pathways in non-syndromic midline craniosynostosis. *Proc Natl Acad Sci U S A,* 114**,** E7341-E7347.

TIMBERLAKE, A. T., KIZILTUG, E., JIN, S. C., NELSON-WILLIAMS, C., LORING, E., YALE CENTER FOR GENOME, A., ALLOCCO, A., MARLIER, A., BANKA, S., STUART, H., PASSOS-BUENOS, M. R., ROSA, R., ROGATTO, S. R., TONNE, E., STIEGLER, A. L., BOGGON, T. J., ALPEROVICH, M., STEINBACHER, D., STAFFENBERG, D. A., FLORES, R. L., PERSING, J. A., KAHLE, K. T. & LIFTON, R. P. 2023a. De novo mutations in the BMP signaling pathway in lambdoid craniosynostosis. *Hum Genet,* 142**,** 21-32.

TIMBERLAKE, A. T., MCGEE, S., ALLINGTON, G., KIZILTUG, E., WOLFE, E. M., STIEGLER, A. L., BOGGON, T. J., SANYOURA, M., MORROW, M., WENGER, T. L., FERNANDES, E. M., CALUSERIU, O., PERSING, J. A., JIN, S. C., LIFTON, R. P., KAHLE, K. T. & KRUSZKA, P. 2023b. De novo variants implicate chromatin modification, transcriptional regulation, and retinoic acid signaling in syndromic craniosynostosis. *The American Journal of Human Genetics*.

TIMBERLAKE, A. T. & PERSING, J. A. 2018. Genetics of Nonsyndromic Craniosynostosis. *Plast Reconstr Surg,* 141**,** 1508-1516.

TING, M. C., WU, N. L., ROYBAL, P. G., SUN, J., LIU, L., YEN, Y. & MAXSON, R. E., JR. 2009. EphA4 as an effector of Twist1 in the guidance of osteogenic precursor cells during calvarial bone growth and in craniosynostosis. *Development,* 136**,** 855-64.

TING, S. B., WILANOWSKI, T., AUDEN, A., HALL, M., VOSS, A. K., THOMAS, T., PAREKH, V., CUNNINGHAM, J. M. & JANE, S. M. 2003. Inositol- and folate-resistant neural tube defects in mice lacking the epithelial-specific factor Grhl-3. *Nat Med,* 9**,** 1513-9.

TOLCHIN, D., YEAGER, J. P., PRASAD, P., DORRANI, N., RUSSI, A. S., MARTINEZ-AGOSTO, J. A., HASEEB, A., ANGELOZZI, M., SANTEN, G. W. E., RUIVENKAMP, C., MERCIMEK-ANDREWS, S., DEPIENNE, C., KUECHLER, A., MIKAT, B., LUDECKE, H. J., BILAN, F., LE GUYADER, G., GILBERT-DUSSARDIER, B., KEREN, B., HEIDE, S., HAYE, D., VAN ESCH, H., KELDERMANS, L., ORTIZ, D., LANCASTER, E., KRANTZ, I. D., KROCK, B. L., PECHTER, K. B., ARKADER, A., MEDNE, L., DECHENE, E. T.,

CALPENA, E., MELISTACCIO, G., WILKIE, A. O. M., SURI, M., FOULDS, N., GENOMICS ENGLAND RESEARCH, C., BEGTRUP, A., HENDERSON, L. B., FORSTER, C., REED, P., MCDONALD, M. T., MCCONKIE-ROSELL, A., THEVENON, J., LE TANNO, P., COUTTON, C., TSAI, A. C. H., STEWART, S., MAVER, A., GORAZD, R., PICHON, O., NIZON, M., COGNE, B., ISIDOR, B., MARTIN-COIGNARD, D., STOEVA, R., LEFEBVRE, V. & LE CAIGNEC, C. 2020. De Novo SOX6 Variants Cause a Neurodevelopmental Syndrome Associated with ADHD, Craniosynostosis, and Osteochondromas. *Am J Hum Genet,* 106**,** 830-845.

TØNNE, E., DUE-TØNNESSEN, B. J., MERO, I. L., WIIG, U. S., KULSETH, M. A., VIGELAND, M. D., SHENG, Y., VON DER LIPPE, C., TVETEN, K., MELING, T. R., HELSETH, E. & HEIMDAL, K. R. 2021. Benefits of clinical criteria and high-throughput sequencing for diagnosing children with syndromic craniosynostosis. *Eur J Hum Genet,* 29**,** 920-929.

TØNNE, E., DUE-TØNNESSEN, B. J., VIGELAND, M. D., AMUNDSEN, S. S., RIBARSKA, T., ASTEN, P. M., SHENG, Y., HELSETH, E., GILFILLAN, G. D., MERO, I. L. & HEIMDAL, K. R. 2022. Whole-exome sequencing in syndromic craniosynostosis increases diagnostic yield and identifies candidate genes in osteogenic signaling pathways. *Am J Med Genet A,* 188**,** 1464-1475.

TØNNE, E., DUE-TØNNESSEN, B. J., WIIG, U., STADHEIM, B. F., MELING, T. R., HELSETH, E. & HEIMDAL, K. R. 2020. Epidemiology of craniosynostosis in Norway. *J Neurosurg Pediatr,* 26**,** 68-75.

TOOZE, R. S., CALPENA, E., TWIGG, S. R. F., D'ARCO, F., GENOMICS ENGLAND RESEARCH CONSORTIUM, WAKELING, E. L. & WILKIE, A. O. M. 2022a. Craniosynostosis, inner ear, and renal anomalies in a child with complete loss of SPRY1 (sprouty homolog 1) function. *J Med Genet*.

TOOZE, R. S., CALPENA, E., WEBER, A., WILSON, L. C., TWIGG, S. R. F. & WILKIE, A. O. M. 2023. Review of Recurrently Mutated Genes in Craniosynostosis Supports Expansion of Diagnostic Gene Panels. *Genes,* 14**,** 615.

TOOZE, R. S., MILLER, K., SWAGEMAKERS, S., MCGOWAN, S., BOUTE, O., COLLET, C., JOHNSON, D., LAFFARGUE, F., LEEUW, N. D., MORTON, J., NOONS, P., OCKELOEN, C., PHIPPS, J., TAN, T. Y., TIMBERLAKE, A., VANLERBERGHE, C., WALL, S., WEBER, A., WILSON, L., MATHIJSSEN, I., TWIGG, S. R. F. & WILKIE, A. O. M. 2022b. Pathogenic variants in the paired-related homeobox 1 (PRRX1) gene are associated with craniosynostosis. *European Society of Human Genetics.* Vienna.

TWIGG, S. R., KAN, R., BABBS, C., BOCHUKOVA, E. G., ROBERTSON, S. P., WALL, S. A., MORRISS-KAY, G. M. & WILKIE, A. O. 2004. Mutations of ephrin-B1 (EFNB1), a marker of tissue boundary formation, cause craniofrontonasal syndrome. *Proc Natl Acad Sci U S A,* 101**,** 8652-7.

TWIGG, S. R., VORGIA, E., MCGOWAN, S. J., PERAKI, I., FENWICK, A. L., SHARMA, V. P., ALLEGRA, M., ZARAGKOULIAS, A., SADIGHI AKHA, E., KNIGHT, S. J., LORD, H., LESTER, T., IZATT, L., LAMPE, A. K., MOHAMMED, S. N., STEWART, F. J., VERLOES, A., WILSON, L. C., HEALY, C., SHARPE, P. T., HAMMOND, P., HUGHES, J., TAYLOR, S., JOHNSON, D., WALL, S. A., MAVROTHALASSITIS, G. & WILKIE, A. O. 2013. Reduced dosage of ERF causes complex craniosynostosis in humans and mice and links ERK1/2 signaling to regulation of osteogenesis. *Nat Genet,* 45**,** 308-13.

TWIGG, S. R. F., HUFNAGEL, R. B., MILLER, K. A., ZHOU, Y., MCGOWAN, S. J., TAYLOR, J., CRAFT, J., TAYLOR, J. C., SANTORO, S. L., HUANG, T., HOPKIN, R. J., BRADY, A. F., CLAYTON-SMITH, J., CLERICUZIO, C. L., GRANGE, D. K., GROESSER, L., HAFNER, C., HORN, D., TEMPLE, I. K., DOBYNS, W. B., CURRY, C. J., JONES, M. C. & WILKIE, A. O. M. 2016. A Recurrent Mosaic Mutation in SMO, Encoding the Hedgehog Signal Transducer Smoothened, Is the Major Cause of Curry-Jones Syndrome. *Am J Hum Genet,* 98**,** 1256-1265.

TWIGG, STEPHEN R. F., LLOYD, D., JENKINS, D., ELÇIOGLU, NURSEL E., COOPER, CHRISTOPHER D. O., AL-SANNAA, N., ANNAGÜR, A., GILLESSEN-KAESBACH, G., HÜNING, I., KNIGHT, SAMANTHA J. L., GOODSHIP, JUDITH A., KEAVNEY, BERNARD D., BEALES, PHILIP L., GILEADI, O., MCGOWAN, SIMON J. & WILKIE, ANDREW O. M. 2012. Mutations in Multidomain Protein MEGF8 Identify a Carpenter Syndrome Subtype Associated with Defective Lateralization. *The American Journal of Human Genetics,* 91**,** 897-905.

TWIGG, S. R. F., MATSUMOTO, K., KIDD, A. M. J., GORIELY, A., TAYLOR, I. B., FISHER, R. B., JEANNETTE, A., HOOGEBOOM, M., MATHIJSSEN, I. M. J., LOURENC, M. T., MORTON, J. E. V., SWEENEY, E., WILSON, L. C., BRUNNER, H. G., MULLIKEN, J. B., WALL, S. A. & WILKIE, A. O. M. 2006. The Origin of EFNB1 Mutations in Craniofrontonasal Syndrome: Frequent Somatic Mosaicism and Explanation of the Paucity of Carrier Males. *The American Journal of Human Genetics,* 78**,** 999-1010.

TWIGG, S. R. F. & WILKIE, A. O. M. 2015. A Genetic-Pathophysiological Framework for Craniosynostosis. *The American Journal of Human Genetics,* 97**,** 359-377.

UEHARA, T., SANUKI, R., OGURA, Y., YOKOYAMA, A., YOSHIDA, T., FUTAGAWA, H., YOSHIHASHI, H., YAMADA, M., SUZUKI, H., TAKENOUCHI, T., MATSUBARA, K., HIRATA, H., KOSAKI, K. & TAKANO-SHIMIZU, T. 2021. Recurrent NFIA K125E substitution represents a loss-of-function allele: Sensitive in vitro and in vivo assays for nontruncating alleles. *Am J Med Genet A,* 185**,** 2084-2093.

UHMANN, A., DITTMANN, K., NITZKI, F., DRESSEL, R., KOLEVA, M., FROMMHOLD, A., ZIBAT, A., BINDER, C., ADHAM, I., NITSCHE, M., HELLER, T., ARMSTRONG, V., SCHULZ-SCHAEFFER, W., WIENANDS, J. & HAHN, H. 2007. The Hedgehog receptor Patched controls lymphoid lineage commitment. *Blood,* 110**,** 1814-23.

URS, S., VENKATESH, D., TANG, Y., HENDERSON, T., YANG, X., FRIESEL, R. E., ROSEN, C. J. & LIAW, L. 2010. Sprouty1 is a critical regulatory switch of mesenchymal stem cell lineage allocation. *FASEB J,* 24**,** 3264-73.

VAQUERO, M., CUESTA, S., ANERILLAS, C., ALTES, G., RIBERA, J., BASSON, M. A., LICHT, J. D., EGEA, J. & ENCINAS, M. 2019. Sprouty1 Controls Genitourinary Development via its N-Terminal Tyrosine. *J Am Soc Nephrol,* 30**,** 1398-1411.

VISSERS, L. E., DE LIGT, J., GILISSEN, C., JANSSEN, I., STEEHOUWER, M., DE VRIES, P., VAN LIER, B., ARTS, P., WIESKAMP, N., DEL ROSARIO, M., VAN BON, B. W., HOISCHEN, A., DE VRIES, B. B., BRUNNER, H. G. & VELTMAN, J. A. 2010. A de novo paradigm for mental retardation. *Nat Genet,* 42**,** 1109-12.

VU, H. L., PANCHAL, J., PARKER, E. E., LEVINE, N. S. & FRANCEL, P. 2001. The Timing of Physiologic Closure of the Metopic Suture: A Review of 159 Patients Using Reconstructed 3D CT Scans of the Craniofacial Region. *The Journal of Craniofacial Surgery,* 12**,** 527-532.

WAKAMIYA, M., MATSUURA, T., LIU, Y., SCHUSTER, G. C., GAO, R., XU, W., SARKAR, P. S., LIN, X. & ASHIZAWA, T. 2006. The role of ataxin 10 in the pathogenesis of spinocerebellar ataxia type 10. *Neurology,* 67**,** 607-613.

WALKER, K. A., SIMS-LUCAS, S. & BATES, C. M. 2016. Fibroblast growth factor receptor signaling in kidney and lower urinary tract development. *Pediatr Nephrol,* 31**,** 885-95.

WALTERS, M. E., LACASSIE, Y., AZAMIAN, M., FRANCISKOVICH, R., ZAPATA, G., HERNANDEZ, P. P., LIU, P., CAMPBELL, I. M., BOSTWICK, B. L. & LALANI, S. R. 2021. Vertical transmission of a large calvarial ossification defect due to heterozygous variants of ALX4 and TWIST1. *American Journal of Medical Genetics Part A,* 185**,** 916-922.

WANG, C.-C., LIN, S.-Y., HUANG, Y.-H., HSIEH, C.-H., CHANG, H.-H., CHEN, H.-Y., WENG, C.-W., CHANG, G.-C., YU, S.-L. & CHEN, J. J. 2021. Paired-like homeodomain 2B contributes to tumour progression and anti-autophagy in human lung cancer. *Am J Cancer Res,* 11**,** 4900-4918.

WHIFFIN, N., MINIKEL, E., WALSH, R., O'DONNELL-LURIA, A. H., KARCZEWSKI, K., ING, A. Y., BARTON, P. J. R., FUNKE, B., COOK, S. A., MACARTHUR, D. & WARE, J. S. 2017. Using high-resolution variant frequencies to empower clinical genome interpretation. *Genet Med,* 19**,** 1151-1158.

WIECZOREK, D., BÖGERSHAUSEN, N., BELEGGIA, F., STEINER-HALDENSTÄTT, S., POHL, E., LI, Y., MILZ, E., MARTIN, M., THIELE, H., ALTMÜLLER, J., ALANAY, Y., KAYSERILI, H., KLEIN-HITPASS, L., BÖHRINGER, S., WOLLSTEIN, A., ALBRECHT, B., BODUROGLU, K., CALIEBE, A., CHRZANOWSKA, K., COGULU, O., CRISTOFOLI, F., CZESCHIK, J. C., DEVRIENDT, K., DOTTI, M. T., ELCIOGLU, N., GENER, B., GOECKE, T. O., KRAJEWSKA-WALASEK, M., GUILLÉN-NAVARRO, E., HAYEK, J., HOUGE, G., KILIC, E., SIMSEK-KIPER, P. Ö., LÓPEZ-GONZÁLEZ, V., KUECHLER, A., LYONNET, S., MARI, F., MAROZZA, A., MATHIEU DRAMARD, M., MIKAT, B., MORIN, G., MORICE-PICARD, F., ÖZKINAY, F., RAUCH, A., RENIERI, A., TINSCHERT, S., UTINE, G. E., VILAIN, C., VIVARELLI, R., ZWEIER, C., NÜRNBERG, P., RAHMANN, S., VERMEESCH, J., LÜDECKE, H.-J., ZESCHNIGK, M. & WOLLNIK, B. 2013. A comprehensive molecular study on Coffin–Siris and Nicolaides–Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Human Molecular Genetics,* 22**,** 5121-5135.

WILK, K., YEH, S. A., MORTENSEN, L. J., GHAFFARIGARAKANI, S., LOMBARDO, C. M., BASSIR, S. H., ALDAWOOD, Z. A., LIN, C. P. & INTINI, G. 2017. Postnatal Calvarial Skeletal Stem Cells Expressing PRX1 Reside Exclusively in the Calvarial Sutures and Are Required for Bone Regeneration. *Stem Cell Reports,* 8**,** 933-946.

WILKIE, A., TANG, Z., ELANKO, N., WALSH, S., TWIGG, S. R. F., HURST, J. A., WALL, S. A., CHRZANOWSKA, K. H. & JR, R. E. M. 2000. Functional haploinsufficiency of the human homeobox gene MSX2 causes defects in skull ossification. *Nature Genetics,* 24**,** 387-390.

WILKIE, A. O., BYREN, J. C., HURST, J. A., JAYAMOHAN, J., JOHNSON, D., KNIGHT, S. J., LESTER, T., RICHARDS, P. G., TWIGG, S. R. & WALL, S. A. 2010. Prevalence and complications of single-gene and chromosomal disorders in craniosynostosis. *Pediatrics,* 126**,** e391-400.

WILKIE, A. O. M., JOHNSON, D. & WALL, S. A. 2017. Clinical genetics of craniosynostosis. *Current Opinion in Pediatrics,* 29**,** 622-628.

WILKIE, A. O. M., OLDRIDGE, S. F. S. M., POOLE, M. D., ASHWORTH, G. J., HOCKLEY, A. D., HAYWARD, R. D., DAVID, D. J., PULLEYN, L. J., RUTLAND, P., MALCOLM, S., WINTER, R. M. & REARDON, W. 1995. Apert syndrome results from localized mutations of FGFR2 and is allelic with Crouzon syndrome. *Nature,* 9**,** 165-172.

WILKINSON, D. G. 2014. Regulation of cell differentiation by Eph receptor and ephrin signaling. *Cell Adh Migr,* 8**,** 339-48.

WONG, W. S., SOLOMON, B. D., BODIAN, D. L., KOTHIYAL, P., ELEY, G., HUDDLESTON, K. C., BAKER, R., THACH, D. C., IYER, R. K., VOCKLEY, J. G. & NIEDERHUBER, J. E. 2016. New observations on maternal age effect on germline de novo mutations. *Nat Commun,* 7**,** 10486.

WRIGHT, K. D., MAHONEY ROGERS, A. A., ZHANG, J. & SHIM, K. 2015. Cooperative and independent functions of FGF and Wnt signaling during early inner ear development. *BMC Dev Biol,* 15**,** 15-33.

WU, A. C., MCMAHON, P. & LU, C. 2020. Ending the Diagnostic Odyssey-Is Whole-Genome Sequencing the Answer? *JAMA Pediatr,* 174**,** 821-822.

WU, T., CHEN, G., TIAN, F. & LIU, H. X. 2017. Contribution of cranial neural crest cells to mouse skull development. *Int J Dev Biol,* 61**,** 495-503.

WU, X. L., GU, M. M., HUANG, L., LIU, X. S., ZHANG, H. X., DING, X. Y., XU, J. Q., CUI, B., WANG, L., LU, S. Y., CHEN, X. Y., ZHANG, H. G., HUANG, W., YUAN, W. T., YANG, J. M., GU, Q., FEI, J., CHEN, Z., YUAN, Z. M. & WANG, Z. G. 2009. Multiple synostoses syndrome is due to a missense mutation in exon 2 of FGF9 gene. *Am J Hum Genet,* 85**,** 53-63.

WUYTS, W., REARDON, W., PREIS, S., HOMFRAY, T., RASORE-QUARTINO, A., CHRISTIANS, H., WILLEMS, P. J. & HUL, W. V. 2000. Identification of mutations in the MSX2 homeobox gene in families affected with foramina parietalia permagna. *Human Molecular Genetics,* 9**,** 1251–1255.

XIE, Y., SU, N., YANG, J., TAN, Q., HUANG, S., JIN, M., NI, Z., ZHANG, B., ZHANG, D., LUO, F., CHEN, H., SUN, X., FENG, J. Q., QI, H. & CHEN, L. 2020. FGF/FGFR signaling in health and disease. *Signal Transduct Target Ther,* 5**,** 181.

XIN, B., PUFFENBERGER, E. G., TURBEN, S., TAN, H., ZHOU, A. & WANG, H. 2010. Homozygous frameshift mutation in TMCO1 causes a syndrome with craniofacial dysmorphism, skeletal anomalies, and mental retardation. *Proc Natl Acad Sci U S A,* 107**,** 258-63.

XU, K., TZVETKOVA-ROBEV, D., XU, Y., GOLDGUR, Y., CHAN, Y. P., HIMANEN, J. P. & NIKOLOV, D. B. 2013. Insights into Eph receptor tyrosine kinase activation from crystal structures of the EphA4 ectodomain and its complex with ephrin-A5. *Proc Natl Acad Sci U S A,* 110**,** 14634-9.

YAGNIK, G., GHUMAN, A., KIM, S., STEVENS, C. G., KIMONIS, V., STOLER, J., SANCHEZ-LARA, P. A., BERNSTEIN, J. A., NAYDENOV, C., DRISSI, H., CUNNINGHAM, M. L., KIM, J. & BOYADJIEV, S. A. 2012. ALX4 gain-of-function mutations in nonsyndromic craniosynostosis. *Hum Mutat,* 33**,** 1626-9.

YAMADA, H., SHIMURA, M., TAKAHASHI, H., NARA, S., MORISHIMA, Y., GO, S., MIYASHITA, T., NUMABE, H. & KAWASHIMA, H. 2020. A familial case of overgrowth syndrome caused by a 9q22.3 microdeletion in a mother and daughter. *Eur J Med Genet,* 63**,** 103872.

YANG, S., NIE, T., SHE, H., TAO, K., LU, F., HU, Y., HUANG, L., ZHU, L., FENG, D., HE, D., QI, J., KUKAR, T., MA, L., MAO, Z. & YANG, Q. 2022. Regulation of TFEB nuclear localization by HSP90AA1 promotes autophagy and longevity. *Autophagy***,** 1-17.

YANG, X., PANDE, S., KOZA, R. A. & FRIESEL, R. 2021. Sprouty1 regulates gonadal white adipose tissue growth through a PDGFRalpha/beta-Akt pathway. *Adipocyte,* 10**,** 574-586.

YEERAVALLI, R., KAUSHIK, K. & DAS, A. 2021. TWIST1-mediated transcriptional activation of PDGFRβ in breast cancer stem cells promotes tumorigenesis and metastasis. *Biochim Biophys Acta Mol Basis Dis,* 1867**,** 166141.

YEN, H. Y., TING, M. C. & MAXSON, R. E. 2010. Jagged1 functions downstream of Twist1 in the specification of the coronal suture and the formation of a boundary between osteogenic and non-osteogenic cells. *Dev Biol,* 347**,** 258-70.

YEO, S. Y., LEE, K. W., SHIN, D., AN, S., CHO, K. H. & KIM, S. H. 2018. A positive feedback loop bi-stably activates fibroblasts. *Nat Commun,* 9**,** 3016.

YILMAZ, E., MIHCI, E., NUR, B., ALPER, O. M. & TACOY, S. 2019. Recent Advances in Craniosynostosis. *Pediatr Neurol,* 99**,** 7-15.

YOON, J. G., HAHN, H. M., CHOI, S., KIM, S. J., AUM, S., YU, J. W., PARK, E. K., SHIM, K. W., LEE, M. G. & KIM, Y. O. 2020. Molecular Diagnosis of Craniosynostosis Using Targeted Next-Generation Sequencing. *Neurosurgery,* 87**,** 294-302.

YOSHIDA, K., TAGA, T., SAITO, M., SUEMATSU, S., KUMANOGOH, A., TANAKA, T., FUJIWARA, H., HIRATA, M., YAMAGAMI, T., NAKAHATA, T., HIRABAYASHI, T., YONEDA, Y., TANAKA, K., WANG, W. Z., MORI, C., SHIOTA, K., YOSHIDA, N. & KISHIMOTO, T. 1996. Targeted disruption of gpl30, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. *Proc Natl Acad Sci U S A,* 93**,** 407–411.

YOSHIDA, T., VIVATBUTSIRI, P., MORRISS-KAY, G., SAGA, Y. & ISEKI, S. 2008. Cell lineage in mammalian craniofacial mesenchyme. *Mech Dev,* 125**,** 797-808.

YU, H. M., JERCHOW, B., SHEU, T. J., LIU, B., COSTANTINI, F., PUZAS, J. E., BIRCHMEIER, W. & HSU, W. 2005. The role of Axin2 in calvarial morphogenesis and craniosynostosis. *Development,* 132**,** 1995-2005.

YU, Y., YAO, R., WANG, L., FAN, Y., HUANG, X., HIRSCHHORN, J., DAUBER, A. & SHEN, Y. 2015. De novo mutations in ARID1B associated with both syndromic and non-syndromic short stature. *BMC Genomics,* 16.

YUSOFF, P., LAO, D. H., ONG, S. H., WONG, E. S., LIM, J., LO, T. L., LEONG, H. F., FONG, C. W. & GUY, G. R. 2002. Sprouty2 inhibits the Ras/MAP kinase pathway by inhibiting the activation of Raf. *J Biol Chem,* 277**,** 3195-201.

ZAIDI, S., CHOI, M., WAKIMOTO, H., MA, L., JIANG, J., OVERTON, J. D., ROMANO-ADESMAN, A., BJORNSON, R. D., BREITBART, R. E., BROWN, K. K., CARRIERO, N. J., CHEUNG, Y. H., DEANFIELD, J., DEPALMA, S., FAKHRO, K. A., GLESSNER, J., HAKONARSON, H., ITALIA, M. J., KALTMAN, J. R., KASKI, J., KIM, R., KLINE, J. K., LEE, T., LEIPZIG, J., LOPEZ, A., MANE, S. M., MITCHELL, L. E., NEWBURGER, J. W., PARFENOV, M., PE'ER, I., PORTER, G., ROBERTS, A. E., SACHIDANANDAM, R., SANDERS, S. J., SEIDEN, H. S., STATE, M. W., SUBRAMANIAN, S., TIKHONOVA, I. R., WANG, W., WARBURTON, D., WHITE, P. S., WILLIAMS, I. A., ZHAO, H., SEIDMAN, J. G., BRUECKNER, M., CHUNG, W. K., GELB, B. D., GOLDMUNTZ, E., SEIDMAN, C. E. & LIFTON, R. P. 2013. De novo mutations in histone-modifying genes in congenital heart disease. *Nature,* 498**,** 220-3.

ZARELLA, C. S., ALBINO, F. P., OH, A. K., WOOD, B. C., OLUIGBO, C. O., MYSEROS, J. S., MAGGE, S. N., KEATING, R. F. & ROGERS, G. F. 2016. Craniosynostosis Following Fetal Methotrexate Exposure. *J Craniofac Surg,* 27**,** 450-2.

ZHANG, J., WRIGHT, K. D., MAHONEY ROGERS, A. A., BARRETT, M. M. & SHIM, K. 2014. Compensatory regulation of the size of the inner ear in response to excess induction of otic progenitors by fibroblast growth factor signaling. *Dev Dyn,* 243**,** 1317-27.

ZHANG, W., ZENG, Q., XU, Y., YING, H., ZHOU, W., CAO, Q. & ZHOU, W. 2017. Exome sequencing identified a novel SMAD2 mutation in a Chinese family with early onset aortic aneurysms. *Clin Chim Acta,* 468**,** 211-214.

ZHANG, X., WAKELING, M., WARE, J. & WHIFFIN, N. 2020. Annotating high-impact 5'untranslated region variants with the UTRannotator. *Bioinformatics,* btaa783.

ZHANG, Y., BULKLEY, D. P., XIN, Y., ROBERTS, K. J., ASARNOW, D. E., SHARMA, A., MYERS, B. R., CHO, W., CHENG, Y. & BEACHY, P. A. 2018. Structural Basis for Cholesterol Transport-like Activity of the Hedgehog Receptor Patched. *Cell,* 175**,** 1352-1364 e14.

ZHANG, Z., ALPERT, D., FRANCIS, R., CHATTERJEE, B., YU, Q., TANSEY, T., SABOL, S. L., CUI, C., BAI, Y., KORIABINE, M., YOSHINAGA, Y., CHENG, J. F., CHEN, F., MARTIN, J., SCHACKWITZ, W., GUNN, T. M., KRAMER, K. L., DE JONG, P. J., PENNACCHIO, L. A. & LO, C. W. 2009. Massively parallel sequencing identifies the gene Megf8 with ENU-induced mutation causing heterotaxy. *Proc Natl Acad Sci U S A,* 106**,** 3219-24.

ZHAO, H., FENG, J., HO, T. V., GRIMES, W., URATA, M. & CHAI, Y. 2015. The suture provides a niche for mesenchymal stem cells of craniofacial bones. *Nat Cell Biol,* 17**,** 386-96.

ZHAO, H., FENG, J., SEIDEL, K., SHI, S., KLEIN, O., SHARPE, P. & CHAI, Y. 2014. Secretion of shh by a neurovascular bundle niche supports mesenchymal stem cell homeostasis in the adult mouse incisor. *Cell Stem Cell,* 14**,** 160-73.

ZHAO, X., WU, B., CHEN, H., ZHANG, P., QIAN, Y., PENG, X., DONG, X., WANG, Y., LI, G., DONG, C. & WANG, H. 2022. Case report: A novel truncating variant of BCL11B associated with rare feature of craniosynostosis and global developmental delay. *Front Pediatr,* 10**,** 982361.

ZHOU, B., LIU, C., XU, Z. & ZHU, G. 2012. Structural basis for homeodomain recognition by the cell-cycle regulator Geminin. *Proc Natl Acad Sci U S A,* 109**,** 8931-6.

ZHOU, Y., KOELLING, N., FENWICK, A. L., MCGOWAN, S. J., CALPENA, E., WALL, S. A., SMITHSON, S. F., WILKIE, A. O. M. & TWIGG, S. R. F. 2018. Disruption of TWIST1 translation by 5' UTR variants in Saethre-Chotzen syndrome. *Hum Mutat,* 39**,** 1360-1365.

ZHU, G., ZHANG, T., CHEN, M., YAO, K., HUANG, X., ZHANG, B., LI, Y., LIU, J., WANG, Y. & ZHAO, Z. 2021. Bone physiological microenvironment and healing mechanism: Basis for future bone-tissue engineering scaffolds. *Bioact Mater,* 6**,** 4110-4140.