Target discovery for immunotherapy of cervical cancer and high-risk human papillomavirus infection



Xu Peng

Wolfson College
Nuffield Department of Medicine
University of Oxford
Trinity 2021

A thesis submitted for the degree of Doctor of Philosophy, University of Oxford

Abstract

Human papillomaviruses (HPVs) are small, double-stranded DNA viruses, and more than 170 types, including 15 types that are designated “high risk” (HR) or oncogenic, have been described. HR types are present in more than 99% of cervical cancers and each year, over 570,000 women are diagnosed with the disease causing 311,000 death per year worldwide. The emergence of cancer immunotherapy has provided new possibilities for the treatment of cancer. For this strategy, it is important to understand which proteins are presented on human leukocyte antigens (HLA) by the tumour and, therefore, could be recognised by T-cells and used as antigens in vaccination approaches or adoptive and recombinant T cell therapies.

High-performance mass-spectrometric techniques have provided the ability to globally investigate which peptides are presented by HLA in cells and tissues. In this study, these technological advances were applied for the identification of HLA class I peptidomes from cervical cancer cell lines and primary cervical cancer resections. We explored how to integrate whole exome sequencing data for the identification of *neo*antigens spanning tumour-specific mutation events, RNA sequencing data for definition of viral open reading frames and identification of viral HLA-presented peptides, and finally, the use of LC-MS *de novo* sequencing approaches that combine a *in silico* binding and peptide

retention time prediction for identification of non-canonical ligands (MAR-S linear model tool).

In our human cervical cancer line panel, 55,474 peptides (48,278 unique peptides） were identified from 15,129 (10,920 unique proteins) proteins. More than 84% of the peptides had the typical length of HLA-I peptides (8 to 14 amino acids), with the majority (60%) being nine amino acids long. In ten human primary cervical tumour resections, 78,618 peptides (52,382 unique peptides) from 9,765 proteins (9,215 unique proteins). More than 88% of these peptides had the typical length of HLA-I peptides (8 to 14 amino acids), and the majority (60%) were nine amino acids long. KDM5B and H1.2 were identified as shared cancer-testis, tumour-associated antigens which were presented by all 5 cell lines and 10 tumour resections. Using DNA exome sequencing-based protein databases, 6 peptides were identified to originate from 5 different *neo*antigens in C33, DoTc2 and SiHa cell lines and the presentation of E1, and cryptic peptide epitopes originating from viral transcripts in cervical cancer cells is reported for the first time. RNA sequencing and de novo assembly supported the discovery of 5 viral peptides in CaSki and HeLa cells. Immunogenicity of the viral peptides was demonstrated by ELIspot in HPV-positive women with and without cervical intraepithelial neoplasia (CIN). The importance of HPV-E1 as a tumour-specific antigen in cervical cancer is highlighted further by demonstration of consistently high expression levels in primary human tumour

resections relative to E6 and E7. We finally identified 12 lncRNA-related HLA peptide candidates in the primary cervical tumour samples, which warrant further validation.

Together, we conclude that E1 is an important target to consider for future immunotherapeutic approaches to cervical tumour treatment, and suggest further antigen candidates that are suitable for the consideration as molecular targets in next-generation immunotherapeutic development.

Acknowledgements

It would not have been possible to finish this thesis without the help of the people around me, to only some of whom it is possible to give a particular mention here.

First, I would like to express my sincere and deep gratitude to my supervisors Prof. Nicola Ternette and Prof. Lucy Dorrell for providing me the invaluable guidance throughout my DPhil. They are more than supervisor; they are my life mentors and best friends.

Special thanks to my first supervisor, Prof. Nicola Ternette, who gave me the opportunity to start my academic journey in this exciting field. Her wisdom, insight, academic rigor, and professionalism have deeply inspired me to become an independent researcher. I would also like to deeply thank my co-supervisor, Prof. Lucy Dorrell for her insightful discussion, support and encouragement in my DPhil study. It was a great privilege and honor to work and research under their guidance.

I would also like to thank the member of Ternette group, Dorrell Group and Kessler group for their advice, help and support. I wish to thank Dr. Annalisa Nicastri, Dr. Robert Parker, Dr. Hanqing Liao, Dr. Rebecca Doherty, Dr. Fergus

Poynton, Dr. Yoanna Ariosa-Morejon, Mr. Isaac Woodhouse, Dr. Thierry Schmidlin, Dr. Arun Tailor, Dr. Julius Muller, Dr. Gemma Hancock, Dr. Jakub Kopycinski and Dr. Thomas Partridge.

Finally, I would like to thank my parents and family for their love and support during my life. I would not have been able to study at Oxford without their support.

Table of Contents

Abstract ……i

Acknowledgements……iv

List of Figures……xii

List of Tables ……..……..……..xv

List of Abbreviations……..xvi

Author's contributions ……..……xix

Chapter 1: Introduction……1

1.1 Cervical cancer and Human Papillomavirus infection……..1

1.2 Diagnosis, treatment and prevention strategies for cervical cancer…….3

1.2.1 Diagnosis of infection and pre-invasive disease……3

1.2.2 Treatment of pre-invasive disease……..4

1.2.3 Treatment of cervical cancer……5

1.2.4 Prophylactic vaccines……6

1.3 Human papillomavirus (HPV) biology and lifecycle……7

1.3.1 HPV genome……..7

1.3.2 Entry, early infection stage and malignant cell transformation…….8

1.3.3 Late-stage infection, assembly and virion release……11

1.3.4 Mechanisms of oncogenesis in HPV infection……12

1.4 Integration of Human Papillomavirus into the Human genome……14

1.5 MHC Class I/II……18

1.6 Mass spectrometry……..……..……..23

1.6.1 Mass spectrometer components and function…….23

1.6.2 LC-MS for proteomics…….25

1.7 Immunotherapy……...26

1.7.1 Checkpoint inhibition……..26

1.7.1.1 CTLA4 blockade……..27

1.7.1.2 PD-1 and PD-L1……..27

1.7.2 Second generation immunotherapeutic approaches……29

1.7.2.1 Cancer vaccination……..30

1.7.2.2 Adoptive cell therapy (ACT)……31

1.8 Tumour Antigen Discovery ……33

1.8.1 Spanning tumour-specific mutations and INDELS……..35

1.8.2 Viral antigens……36

1.8.3 Intron-retention as a source of neoantigen……...37

1.8.4 Antigen sourced from alternative reading frames……37

1.8.5 Post-translational modification of antigens……..38

1.8.6 lncRNAs……39

1.9 Overall aim of the thesis……43

Chapter 2: Materials and Methods……..44

2.1 Materials ……..44

2.1.1 Plasticware and general reagents……44

2.1.2 Cell lines and tissue culture reagents ……...45

2.1.3 Peptides ……46

2.1.4 Chemical reagents ……50

2.1.5 Enzyme-linked immunospot assay reagents (ELISPOT)……51

2.2 Methods……53

2.2.1 Human tissue samples and cell culture……53

2.2.1.1 Human tissue samples……53

2.2.1.2 Cervical cancer cell culture……56

2.2.1.3 TC-1 cell culture……...56

2.2.2 Cell DNA purification methods……..56

2.2.3 RNA purification methods (cell)……....57

2.2.4 RNA purification methods (tissues)……..……58

2.2.5 Exome and RNA sequencing……59

2.2.6 Exome variant calling……..60

2.2.7 *De novo* assembly……...61

2.2.8 Flow cytometry……..……63

2.2.9 Preparation of antibody-conjugated immunoresin……63

2.2.10 MHC class I immunoprecipitation ……64

2.2.11 High performance liquid chromatography peptide fractionation……65

2.2.12 LC-tandem mass spectrometry (LC-MS/MS)……65

2.2.13 Mass spectrometry data analysis……67

2.2.14 Generation of STCLs and IFNγ ELISpot……..67

2.2.15 MAR-S ……69

2.2.16 Mouse tumor challenge……70

2.2.17 Statistics and data presentation……70

Chapter 3: Identification of tumor-specific antigens in five human cervical cancer cell lines……..71

3.1 Introduction……..71

3.2 Results……..72

3.2.1 Identification of tumour-specific antigens in five human cervical cancer cell lines……..72

3.2.1.1 Determination of HLA expression in five cervical cancer cell lines…..72

3.2.1.2 Preparation of personalized protein sequence databases……77

3.2.1.3 Sequencing of the immunopeptidomes of the cervical cancer cell line panel……80

3.2.1.4 Gibbs cluster and NetMHCpan successfully categorized the identified peptides to 15 alleles……88

3.2.2 Several important oncogenic driver genes and HPV master regulators were identified in the five human cervical cancer cell lines……92

3.2.3 Several important CT, Raven and protein atlas antigens were identified in the five human cervical cancer cell lines……..95

3.2.4 Identification of *neo*antigens in comparison to the reference proteome……..109

3.2.5 Viral peptides identified in five human cervical cancer cell lines……110

3.2.6 HPV18 E1 protein is expressed in HeLa cells……115

3.2.7 ELISpot response analysis of patients' Short-term cell lines (STCL) of confirmed peptides……..……116

3.3 Discussion……..118

Chapter 4: Identification of tumor-specific antigens in Ten human cervical cancer resections ……..……..……131

4.1 Introduction……131

4.2 Results……133

4.2.1 Identification of mouse tumour antigens in the TC-1 cell line and mouse tumour…….133

4.2.2 Identification of tumour-specific antigens in ten human cervical cancer tumours……140

4.2.2.1 Determination of HLA expression of ten human cervical cancer tumours……140

4.2.2.2 Preparation of personalized protein sequence databases……..141

4.2.2.3 Sequencing of the immunopeptidomes of the human primary tumour panel……142

4.2.2.4 Assignment of HLA-allele orgin for identified peptide ligands with Gibbs cluster and NetMHCpan……..…149

4.2.3 Viral peptides identified in ten human cervical cancer tumours……..166

4.2.4 Several important CT,Raven and protein atlas antigens were identified in

the ten human cervical cancer tumours……167

4.2.5 Long non-coding RNA in ten human cervical cancer tumours……..171

4.3 Discussion……180

Chapter 5: General discussion……..184

6.0 References……...186

7.0 Appendix……..205

8.0 List of publications…….340

LIST OF FIGURES

Figure 1. HPV screen coverage in UK (2011-2020)……4

Figure 2. HPV-16/18 genome ……...8

Figure 3. HPV integration……17

Figure 4. MHC class I/II antigen presentation pathways……22

Figure 5. A comparison of CTLA-4 and PD-1……33

Figure 6. Workflow of MHC-I Peptidomics……74

Figure 7. FACS experiment of five human cell lines……77

Figure 8. RNAseq data and viral peptide antigens identified in cervical cancer cell lines……79

Figure 9. Eluted peptides identified from five human cancer cell lines……82

Figure 10. Comparison of eluted peptides identified in two biological replicates from five human cervical cancer cell lines……..85

Figure 11. Comparison of eluted peptides identified in four technical replicates from five human cervical cancer cell lines……..86

Figure 12. Compare the overlap of sequences for cells expressing in the same/different HLA allele……87

Figure 13. Amino acid distribution of all eluted 9-mer peptides from experiment……91

Figure 14. List of several important oncogenic driver genes and HPV master regulators identified in our experiment……94

Figure 15. Cancer testis antigens, Raven antigens and Human Protein

Atlas……..……..96

Figure 16. List of eight share antigens defined by the CT database…….104

Figure 17. List of eight share antigens defined by the Raven database…….108

Figure 18. Viral peptide antigens identified in HeLa and CaSki cells……113

Figure 19. Identification of a cryptic peptide sequence sourced within the human genome originating from a transcript spanning the viral integration site…...114

Figure 20. Immunogenicity of identified viral and cryptic peptide sequences……..117

Figure 21. Properties of the HLA-I peptidomes data set obtained from seven cell lines…….……..119

Figure 22.MHC-peptides identified in TC-1 cell lines……137

Figure 23. Mouse tumour experiment……139

Figure 24. Viral transcript expression levels in ten human cervical cancer tumours……144

Figure 25. Properties of the eluted peptides identified from ten human cancer tumour samples……..……146

Figure 26. Comparison of eluted peptides identified in two technical replicates from ten human cancer tumour samples…….……149

Figure 27. Size distribution and NetMHCpan binding prediction of peptides detected for patient 1 and 2……..……151

Figure 28. Size distribution and NetMHCpan binding prediction of peptides detected for patient 3 and 4…….……152

Figure 29. Size distribution and NetMHCpan binding prediction of peptides detected for patient 5 and 6……153

Figure 30. Size distribution and NetMHCpan binding prediction of peptides detected for patient 7 and 8……154

Figure 31. Size distribution and NetMHCpan binding prediction of peptides detected for patients 9 and 10……155

Figure 32. A-J. Amino acid distribution of all eluted 9-mer peptides from primary tumour resections for each patient……166

Figure 33. Identification of CT and Raven antigens in ten cervical cancer resections……168

Figure 34. Spectral matching analysis for lncRNA peptide sequences…...179

LIST OF TABLES

Table 1. List of five human cervical cancer cell lines genotypic information…78

Table 2. Tumour-specific *neo*antigens identified in cervical cancer panel….110

Table 3. List of viral peptides identified in five human cancer cell lines…….112

Table 4. List of viral peptides identified in dataset PXD004452 (Hela)……..115

Table 5. List of ten human cervical tumours genotypic information……145

Table 6. List of viral peptides identified in ten human cervical tumours……167

Table 7. List of peptides shared CT antigens KDM5B identified in five human cancer cell lines and ten human cervical tumours……170

Table 8. LncRNA-derived peptide sequence candidates identified in cervical tumour tissue resections ……..……..173

Table 9. Potential protein spliced source of LI9……174

Table 10. List of peptides defined by the CT database across the ten cervical cancer tumours……..205

Table 11. List of peptides defined by the RAVEN database across the ten cervical cancer tumours……218

Table 12. List of selected lncRNA peptides in ten human cervical tumour samples……287

Table.13. List of peptides shared RAVAN antigens HIST1H1C identified in five human cancer cell lines and ten human cervical tumors……297

Table 14. List of peptides were synthesised by GenScript Biotech (Netherlands) B.V for lncRNA peptides MS matching……336

LIST OF ABBREVIATIONS

ACT	cell adoptive cellular therapy
ANNS	artificial neural networks
APC	antigen-presenting cell
ATAD2	ATPase family AAA domain-containing protein 2
CA	capsid protein
CAR	chimeric antigen receptors
cDNA	complementary DNA
Cep55	Centrosomal protein of 55 kDa
CI-M6PR	cation-independent mannose 6-phosphate receptor
CIN	cervical intraepithelial neoplasia
CSGs	cancer-specific genes
CT	Cancer-testis
CTAs	Cancer testis antigens
CTL	cytotoxic T lymphocyte
CTLA-4	cytotoxic T-lymphocyte–associated antigen 4
DC	dendritic cell
DIA	data-independent acquisition
DMEM	Dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DRiP	defective ribosomal product
DTT	dithiothreitol
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
ELISPOT	enzyme-linked immunospot assay
ER	endoplasmic reticulum
ERAD	ER-associated degradation
ESI	electrospray ionization
FAM21	Family with sequence similarity 21
FBS/FCS	foetal bovine serum/foetal calf serum
FDR	false discovery rate
FEC	Flu, EBV, CMV (peptides)
FT-ICR	Fourier transform–ion cyclotron resonance
GP	glycoprotein
HATs	acetyltransferases
HBV	hepatitis B virus
HCD	higher-energy collisional dissociation
HCV	hepatitis C virus
HDACs	deacetylases

HIST1H1C	Histone H1.2
HLA	human leukocyte antigen
HOTAIR	HOX Transcript Antisense RNA
HPVs	Human papillomaviruses
HSDL1	Human hydroxysteroid dehydrogenase like 1 gene
hVPS34	late endosomal class III phosphatidylinositol 3'-kinase
IFI16	Interferon gamma inducible protein 16
IFN	interferon
Ig	immunoglobulin
Ii	invariant chain
INDELs	insertion and deletions
IP	immunoprecipitation
IT	ion trap
kDa	kilodalton
KDM5B	Lysine-specific demethylase 5B
KIF2C	Kinesin-like protein
LC-MS/MS	liquid chromatography tandem-mass spectrometry
LIT	linear ion trap
lncRNAs	long non-coding RNAs
MAGE	melanoma-associated antigen
MALAT1	Metastasis-associated lung adenocarcinoma transcript 1
MALDI	matrix assisted laser desorption/ionization
MFI	median fluorescence intensity
MHC	major histocompatibility complex
miRNA	microRNA
MS	mass spectrometer
NAAT	nucleic acid amplification tests
NC	nucleocapsid protein
NHS	National Health Service
NHSL1	NHS-like protein 1
NK	natural killer cells
NPCs	nucleoprotein complexes
NUF2	Kinetochore protein Nuf2
ORF	open reading frame
OVA	ovalbumin
PAS/PGS	protein A Sepharose/protein G Sepharose
PBMC	peripheral blood mononuclear cells
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PD-1	programmed death-1
PD-L1	Programmed death-ligand 1
PR	protease
RAB7A	Ras-related protein Rab-7a

RAVEN	Rich Analysis of Variable gene Expressions in Numerous tissues
RNA	ribonucleic acid
RP	regulatory complex (proteasome)
RPA	replication A proteins
RPC	replication factor C
RP-HPLC	reversed phase high performance liquid chromatography
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SFU	spot-forming units
SRNS	steroid-resistant nephrotic syndrome
TAAs	tumor-associated antigens
TAP	transporter associated with antigen processing
TCA	trichloroacetic acid
TCRs	T cell receptors
TFA	trifluoroacetic acid
TFH	T follicular helper cell
TIL	tumour-infiltrating lymphocytes
TLR	toll-like receptor
TOF	time-of-flight
TOPK	Lymphokine-activated killer T-cell-originated protein kinase
TRPV4	transient receptor potential proteins 4
TSAs	tumor-specific antigens
UEVLD	Ubiquitin-conjugating enzyme E2 variant 3
UTR	untranslated region
VSVG	vesicular stomatitis virus G protein
WES	whole exome sequencing
β2m	beta-2-microglobulin

Author's contributions

This thesis is original and independent work performed and written by myself (Xu Peng) with the following exceptions:

- Dr. Annalisa Nicastri (University of Oxford, UK), Dr. Robert Parker (University of Oxford, UK), Dr Nicola Ternette (University of Oxford, UK) operated mass spectrometers for liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis of all peptide samples
- Mr. Isaac Woodhouse (University of Oxford, UK) analysed the RNA and DNA sequence data, created Trinity and mutated databases and provided OptiType results.
- Dr. Hanqing Liao (University of Oxford, UK) assisted with data analysis of NetMHCpan software and MAR-S pipeline for discovery of lncRNA-derived peptides
- Dr. Julius Muller (University of Oxford, UK) analysed the RNA and DNA sequence data and first generation of Trinity databases.
- Dr. Gemma Hancock (University of Oxford, UK) and Dr. Jakub Kopycinski (University of Oxford, UK) performed ELISpot analyses

Chapter 1: Introduction

1.1 Cervical cancer and Human Papillomavirus infection

In the early 1980s, Harald zur Hausen was first to show the relationship between genital Human papillomaviruses (HPVs) infections and cervical cancer. Since then, more and more evidence was collected to demonstrate the association between HPV and cervical tumour development [1].

.

HPVs are non-enveloped, double-stranded DNA viruses are grouped in five major phylogenetic genera: Alpha (α)-, Beta(β)-, Gamma(γ)-, Mu(μ)-, and Nu(ν)- papillomaviruses. Different genera have <60% similarity within the L1 gene [2]. Alpha papillomaviruses have been associated with a wide spectrum of cancers (cervix, vagina, vulva, anal canal, penis and oropharynx) [3]. In the United States, HPV is associated with 70% of oropharyngeal cancers, 60% of penile cancers, and at least 90% of anal and cervical cancers [4]. Most of the genitally transmitted HPV types are alpha HPVs, for example, HPV6 and 11 are present in around 1% of the sexually active population which makes them the major sexually transmitted pathogens in the world [5]. All of the high-risk oncogenic HPVs are alpha papillomaviruses and present in approximately 10-25% women worldwide [6], including HPV 16/18. In contrast, papillomaviruses of the other four genera infect the cutaneous epithelia, some

of which some colonize human skin or cause wart development, and a potential association with non-melanoma skin cancer has been suggested [7-9] .

HPVs are small DNA viruses (approximately 8 kb in size), of which there are more than 170 types, including 15 types that are designated “high risk” (HR) or oncogenic [10, 11]. Importantly, HPVs are present in more than 99% of cervical cancers, which is the fourth most common type of female cancer worldwide [12-14]. At least 12 types of HPV are associated with cervical cancers (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59), with HPV 16/18 accounting for 70% and the other types being associated with the remaining 30% [15]. Each year, over 570,000 women are diagnosed as cervical cancer, and 311,000 die from this illness worldwide [16].

Approximately 90% of HPV infection are cleared within 12-48 months and most HPV infections are asymptomatic [17]. It has been widely assumed that clearance is mediated by innate and adaptive immune responses, however, random partitioning of viral episomes upon basal epithelial cell division may play an important role. Persistent infection can lead to the accumulation of genetic mutations within epithelial cells. Failure to clear the HPV infections could get progression to cervical intraepithelial neoplasia 1 (CIN1). CIN-1 (low-grade) involves the lower 1/3 or less of the epithelium and approximately

60% of CIN 1 lesions regress without treatment. In addition, 5% of CIN 2 (mild dysplasia basal 2/3 of epithelium) and 12% of CIN 3 (whole epithelium) cases will progress to invasive cancer if untreated [18].

1.2 Diagnosis, treatment and prevention strategies for cervical cancer

1.2.1 Diagnosis of infection and pre-invasive disease

There are two types of diagnostic tests for cervical cancer screening: cytology (the Papanikolaou test) and HPV nucleic acid amplification tests (NAAT) [19]. The Papanikolaou test detects cells with abnormal morphology in a cervical sample (swab). NAAT tests detect HPV DNA or RNA in cervical or vaginal samples. Three molecular methods can be used for detection of HPV infection: nonamplified hybridization assays, signal amplified hybridization assays and target amplification assays [19]. Among these assays, target amplification assays, such as polymerase chain reaction (PCR) and in situ PCR, are the most sensitive and specific [20, 21].

Cervical screening recommendations vary across countries and regions. According to the guidelines of the UK National Health Service (NHS) Cervical Screening Programme, women between the ages of 25 and 64 are invited for regular cervical screening (every 3 years for women aged 25-49 and every 5

years for women aged 50-64) [22]. Overall coverage for women was over 70% in 2020 (Fig.1) [22]. HPV NAAT tests were rolled out as the primary screening test in the UK in 2019,with cytology being carried out in case of a positive result [23]. Cytologic screening has been very successful in reducing the incidence rate of cervical cancer (50~60%) [24].

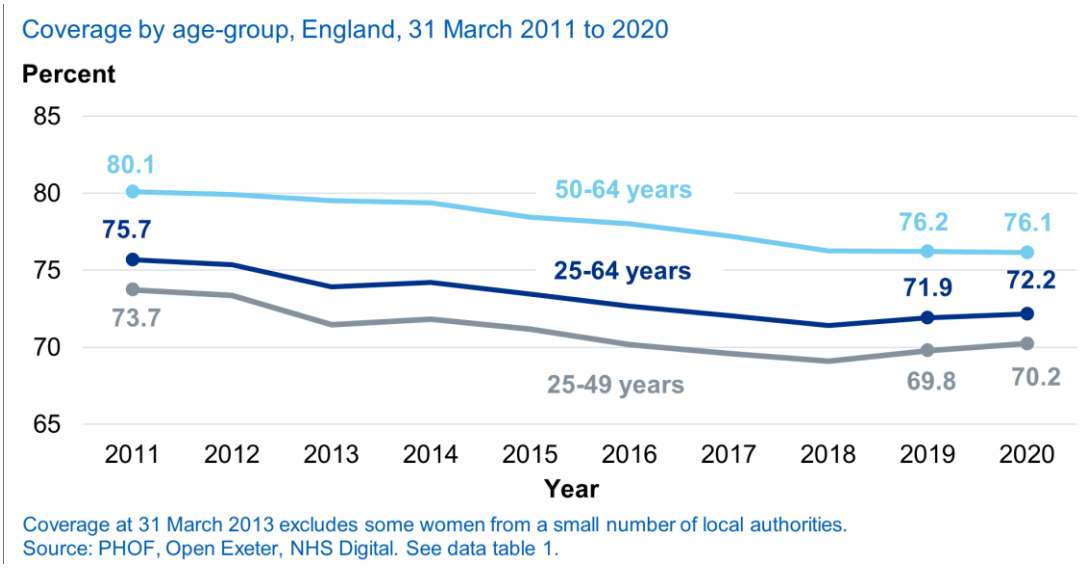


Figure 1. HPV screen coverage in UK (2011-2020) (This figure is a copy from ref 22) [22]

1.2.2 Treatment of pre-invasive disease

Regular screening tests are considered to be best methods to diagnose the cervical cancer early, which can greatly improve the chances of successful treatment. Early identification and intervention to prevent progression of CIN lessions are key factors in the control of cervical cancer morbidity and mortality.

If the cervical screening test results are abnormal, it will lead to further tests. Based on information along with your personal and family medical history, the doctor will arrange another test (such as screening test, colposcopy, endocervical scraping and cone biopsies) and provide the follow-up treatment for you.

1.2.3 Treatment of cervical cancer

There are several treatment choices for cervical cancer, such as surgery, radiation therapy and chemotherapy dependent on tumour stage, histologic features of the tumour and/or patient preference. In UK, the specialist care team will explain the treatments, benefits and side effects to the patients. They will discuss patients with the treatment may impact on their fertility and make a best treatment plan [25]. During and after any treatments, the patients will also have regular check-ups and tests [25].

Recently, cancer immunotherapy has provided remarkable new possibilities for the treatment of cancer. Cancer immunotherapy is a therapy that facilitates components of the immune system to attack tumour cells. Due to the immune system's capacity to target and destroy cells that present foreign or mutated antigens, immunotherapy has advantages over other untargeted approaches

like chemotherapy. Most current cancer immunotherapies aim to induce tumour-specific cytotoxic T cell responses in cancer patients in order to directly destroy tumour cells. Last year, the immune checkpoint inhibitor pembrolizumab (Keytruda) is already used to treat cervical cancer during or after treatment with chemotherapy [26]. In addition, the treatment using adoptive T cell therapy selected for E6 and E7 antigens or TCR-engineered T cells targeting E7 for patients with metastatic HPV-associated epithelial cancers have been shown to be able to induce complete tumour regression in selected treated patients [27-29].

1.2.4 Prophylactic vaccines

Prophylactic HPV vaccines have played a major role in reducing the incidence of oncogenic HPV infection and CIN lesions, which precede the development of cervical cancers. Three licensed vaccines are now available. They are defined as bivalent, quadrivalent or 9-valent, according to the number of HPV types they target. They prevent HPV infections by inducing humoral immunity through delivery of virus-like particles (VLPs) encoding late HPV viral capsid proteins L1 [30]. Cervarix (bivalent) was designed to prevent infection with HPV types 16 and 18, Gardasil protects against 4 HPV types, 16 and 18, together with two benign types (6 and 11) and Gardasil 9 protects against 9

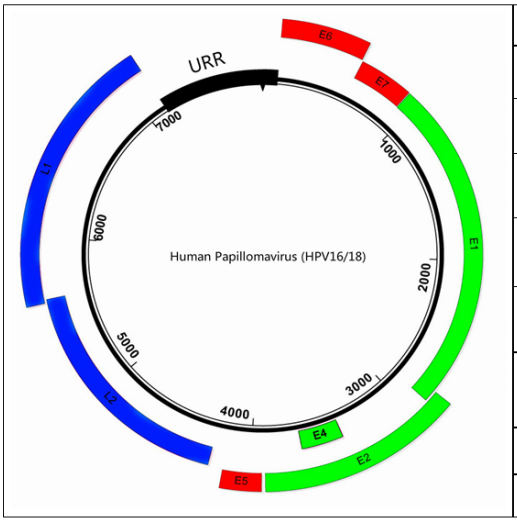
HPV types (16, 18, 6, 11, plus 5 other oncogenic types (31, 33, 45, 52 and 58) [30].

The combination of regular screening, development of effective therapeutic treatments, together with broad application of preventative vaccines across the globe has the potential to lead to the control of HPV disease and eventual elimination of cervical cancer worldwide.

1.3 Human papillomavirus (HPV) biology and lifecycle

1.3.1 HPV genome

Next to relying on the cellular translation machinery as typical for all viruses, HPVs must also depend on the host cell replication proteins to mediate viral DNA synthesis, because their genomes are small and do not encode polymerases or viral replication enzymes [12-14]. The HPV genome can be divided into an upstream regulatory region (URR), the 'exclusively early' genes (E6 and E7), 'early and late' genes (E1, E2, E4, and E5) and 'exclusively late' genes (L1 and L2) (Fig. 2) [12, 31-35]. Since HPV 16 and -18 are associated with 70% of all cervical cancers, most of the research has been focused on analysis of those high-risk HPV types[5].



Gene	Function
E1	Transcription factor, helicase activity
E2	Transcription factor, regulates viral copy number
E4	Facilitates virion release
E5	Downregulates surface MHC class I expression
E6	Downregulates cell cycle control through p53
E7	Keeps cells active through Rb inactivation
L1	Major capsid protein
L2	Minor capsid protein

Figure.2. HPV-16/18 genome. The left panel shows the ~8000-bp map of this oncogenic DNA virus and its genes. In the right panel, the key functions of the early and late genes are listed

1.3.2 Entry, early infection stage and malignant cell transformation

HPVs infect cells in the basal layer of keratinocytes in the stratified squamous epithelia [36]. Entry occurs through endocytosis, and heparan sulfate proteoglycans (HSPG) have been identified as the main initial contact molecules [37]. Basal layer cells are the only proliferating cells in normal epithelia, which are the differentiated cells in the supra-basal layers are no longer actively dividing. Following infection, HPV genomes are become extrachromosomal elements or episomes in low copy number [33, 38], and the virus continues to divide and spread together with the initially infected cell and its division cycle. Viral DNA may be cell cycle-independently amplified in

supra-basal cells, and thousands of viral genome copies may be present in these infected daughter cells.

The E1 protein is localised in the nucleus and is required for the initiation of viral DNA replication. E1 is the largest HPV protein and has the highest conservation throughout papillomaviruses, reflecting its vital function for the virus [39]. The main function of the E1 protein is the maintenance and amplification of viral genomes, which interact with several cellular proteins involved in the replication machinery, such as α-DNA polymerase, replication A proteins (RPA), replication factor C (RFC), nuclear proliferation cell antigen (PCNA), and topoisomerases I and II [40]. These work together to increase the copy number of the viral episome upon infection of epithelial layers. However, E1 is also associated with other cellular proteins that regulate epigenetic processes [41-43]. Leonardo et al. [44] reported that E1 proteins from HR and low-risk HPV types differentially regulate gene expression profiles. They showed that E1 from three viral subtypes down-regulated genes associated with the immune response, and all E1 proteins were able to inhibit the expression of IFNβ1 and IFNλ1, which are two key components in the antiviral immune response [44]. They promote the expression of interferon-stimulated (ISG) genes and assist in controlling the infectious processes through multiple mechanisms, such as protein translation inhibition, viral RNA degradation, and adaptive immune cells [45-47].

E1 and E2 are known to be the only two viral proteins necessary for HPV genome replication, making them an attractive target for drug development [48]. E2 plays an important role during productive infection and has multiple binding sites in the viral upstream regulatory region (URR). E2 binding to these non-coding regions of the viral genome [49] is essential for the recruitment of the viral E1 helicase to the viral origins of replication, which have specific E1 binding motifs [50]. E2 is also involved in genome segregation and is necessary for leading viral episomes to the correct mitotic chromosomes [51]. E2 acts as a transcription factor to regulate the viral early promoter and viral oncogenes E6 and E7 [50]. Previous research has indicated that both E1 and E2 are essential for viral genome maintenance in the basal layer [52, 53].

Besides E1 and E2, HPV E4 and E5 are also necessary for viral protein replication and genome amplification. E5 involves in cell cycle progression and activates late viral functions upon epithelial differentiation [54]. E5 also modulates cell signalling via the extracellular signal-regulated protein kinase (ERK1/2) and p38 MAP kinase, which increase during the epidermal growth factor (EGF)-mediated process [55]. The E4 open reading frame (ORF) is contained within the E2 ORF and modulates virus synthesis and genome amplification [56-58]. E4 proteins have also been shown to inhibit cell

proliferation during cell cycle growth phase 2 (G2) and be involved in virus assembly and release [59, 60].

In normal epithelia, basal cells migrating into the supra-basal cell layers before exit the cell cycle and undergo a process of terminal differentiation [38]. However, when oncoproteins E6 and E7 are expressed in the cells, they suppress the normal terminal differentiation and maintain the cells in the S phase, thereby contributing to the hyperproliferation of basal epithelial cells [61, 62].

1.3.3 Late-stage infection, assembly and virion release

Most viruses infect a host cell and produce progeny virus directly from the infected cell. However, in HPV infections, the new virions are generated when the infected cell has finished mitosis and one of the infected daughter cells has differentiated [36].

This process begins in mid or upper epithelial layers and requires all viral early proteins. A recent study suggested that E2 can recruit E1 for binding to the HPV upstream regulatory genome region, and that this recruitment is important for viral DNA replication [61]. However, the function of viral early

gene E4 and E5 in DNA replication and/or packaging are still not fully understood [54, 63]. Papillomaviruses have two structural proteins (L1 and L2) that are expressed in the upper layers of the infected tissue after viral genome amplification has been finalised [64]. The major capsid protein (L1) is expressed later than L2, and the L2 protein recruits L1 to nuclear structures known as promyelocytic leukaemia (PML) nuclear bodies [65]. Capsid proteins may accumulate at PML bodies and packaged there [66, 67].

After assembly, papillomaviruses need to be transported close to the epithelial surface for release from the infected cells [68]. Viral E4 protein can disturb keratin integrity [69, 70] and affect the assembly of the cornified envelope [71, 72]. Bryan and Brown have shown that viral shedding from the epithelial surface within a cornified squame may enhance viral survival [73].

1.3.4 Mechanisms of oncogenesis in HPV infection

The HPV oncoproteins E6 and E7 are crucial in the induction and maintenance of cellular transformation and are expressed in most HPV-associated cervical cancers [74].

E6 proteins in high-risk HPV types have evolved to target the tumour suppressor p53 [75, 76]. The E6 protein directly targets p53 for degradation

and leading to increased cellular proliferation, which contributes to the development of HPV-associated cancers [75, 76]

Retinoblastoma protein (pRb) is a negative regulator of the cell cycle, which associates with the E2F family of transcription factors to prevent the cell to enter the S-phase [62, 68]. The HPV protein E7 binds to pRb and disrupts interaction with E2F transcription factors that in turn remain active and stimulate cell cycle progression and DNA replication [68]. In all HPV types, E7 proteins bind Rb family members in terms of the high-risk E7 proteins have much higher affinity [20, 77-79]. Finally, the efficient of this binding was through a p53-dependent pathway to inhibit cell growth and apoptosis [77].

The high-risk E6 and E7 proteins work together to maintain host cell S phase competence in differentiating cells and interfere many cell cycles checkpoints. If cell persistently infected with HPV, which accumulated cellular mutations over time and subsequent progression to cancer [32, 75, 77, 80].

Deregulation of gene expression and epigenetic mechanisms also play an essential role in the oncogenic process, which is altering the organization of chromatin and the expression of genes [42]. Swinder et al.,[41] and Lee et al.,[42] reported that E1 interacts with the histone H1 and with the Ini1/hSNF5 complex, which plays an important role in the regulation of gene expression.

As a result, E1 could interact with chromatin modifiers such as acetyltransferases (HATs) and deacetylases (HDACs) to regulate gene expression [40].

The E5 protein cooperates with E6 and E7 function to promote hyperproliferation of infected cells and might facilitate malignant progression [32, 34, 76] through promoting epidermal growth factor receptor (EGFR)-induced signaling [81], downregulation of host HLA-A and HLA-B cell-surface expression [82, 83] and inhibition the cancer cell apoptosis [84, 85].

The HPV oncoproteins take advantage of these cellular signaling pathways to maintain genome replication in infected cells and facilitate viral replication and/or persistence. Those process could lead to the accumulation of cell mutations and increased genomic instability, which can finally induce the progress of cancer.

1.4 Integration of Human Papillomavirus into the Human genome

Human papillomaviruses genome can integrate into the host genome in persistent infection. The integration of the high-risk HPV16 and HPV18 types into the human genome is an essential step in malignant transformation, as it

interferes with host gene expression and also causes structural changes of the host genome. During the process of integration, deletions and/or rearrangements usually occur within the HPV genome and often result in high expression of HPV E6 and E7 viral oncogenes which promote cell division and proliferation and loss of function of p53 and pRB proteins [52, 53]. However, deletion and/or silencing of E1 or E2 may lead to a loss of transcriptional regulation and thus contributes to uncontrolled expression of E6 and E7 [52, 53].

HPV integration can be classified into two types [86] (Fig.3). Type 1 integration, which is a single viral genome integration into the host DNA, compared with Type 2 integration, which is integration and rearrangement of multiple viral genomes.

HPV genome integrate into cells will have a selective growth advantage, because the E6 and E7 open reading frames are virtually always retained, which were leading to expression of the E6 and E7 oncoproteins persistently [48]. Recent studies using genome-wide analysis of HPV integration in host genome provide further evidence that HPV integration occurs at common fragile sites (CFSs). The CFSs are chromosomal regions that are exhibit breakage and highly unstable in response to DNA replication stress. There are over 120 CFSs have been identified in human genome [87, 88]. The data

revealed that HPV integration not only occurs in more than 80% of HPV-positive cervical cancers [36], but also that 48% of the HPV16 and 63% of the HPV18 integrations occurred in CFS [89-91]. The identified junction sites between viral and human are contained repetitive and AT–rich sequences which were shown to have the potential to form stem–loop DNA secondary structures [92]. It was shown that HPV E1 and E2 proteins expression can cause genomic instability at the integration locus [93, 94] and HPV E2 associates with recombinational hot spots possibly facilitating integration at such sites [79, 95-97] .

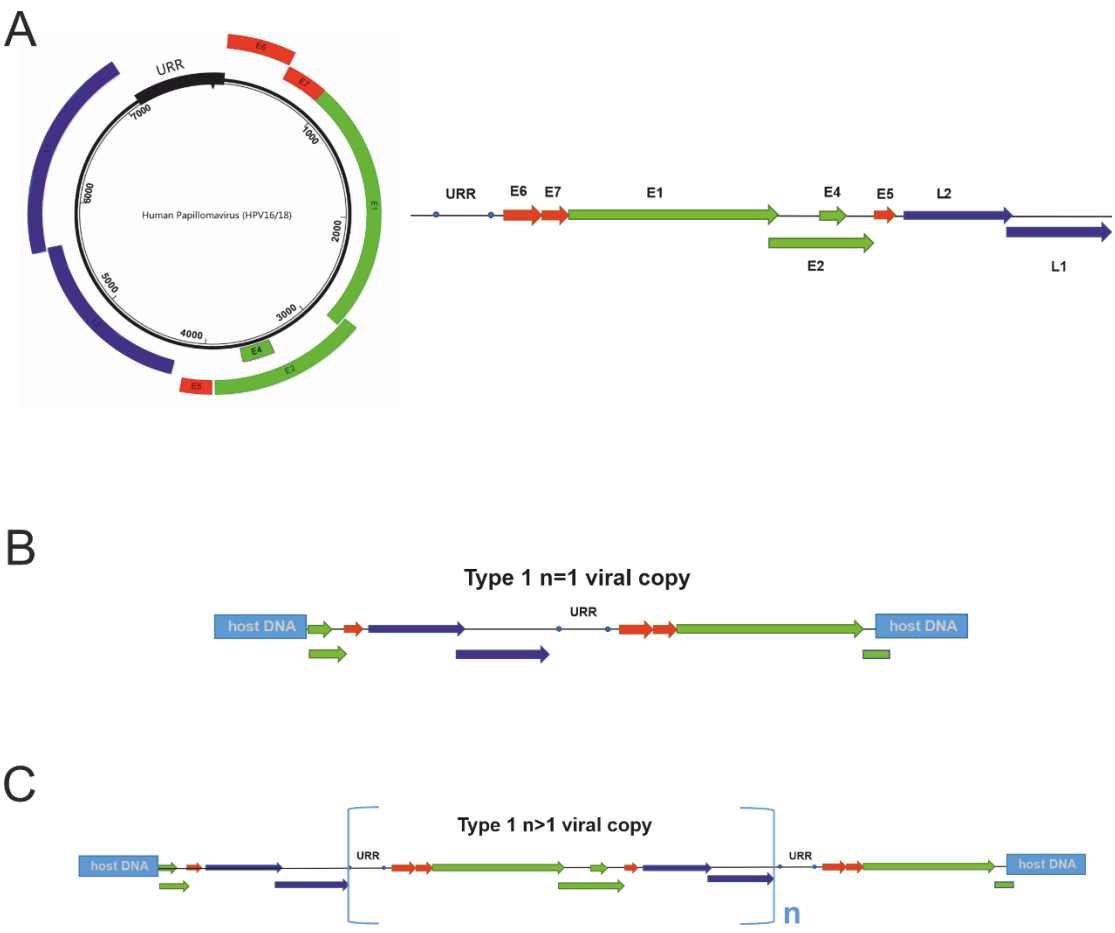


Figure.3. HPV integration. A. HPV genome. B. Linear HPV genome. Type 1 integration, a single viral genome is integrated into the host DNA. C. HPV Type 2 integration, multiple genomes are integrated in host genome [86].

1.5 MHC Class I/II

The Major Histocompatibility Complex (MHC) is a tightly linked cluster of genes that exists in all mammals. The molecular mechanism details of antigen processing and presentation by the MHC class I (MHC-I) and class II (MHC-II) molecules have been studied for many years. MHC class I and class II molecules are present on the surface of cells and display a large number of peptide epitopes for surveillance by CD8+ or CD4+ T cells, respectively. MHC class I molecules consist of one membrane-spanning α chain (heavy chain) and a β2-microglobulin (β_2 m) [98].

MHC-I molecules are present on the surface of all nucleated cells and display peptides to $CD8^+$ T cells, which are recognized by the T cells if foreign or abnormal. In humans, the classical MHC-I is referred to as the Human leukocyte antigen (HLA) complex which is encoded by three loci (HLA-A, B, and C) on Chromosome 6. In mice MHC is referred to as the H-2 (H-2D, H-2K and H-2L) complex and is located on Chromosome 17 [98, 99]. More than 10,000 HLA class I alleles have been identified so far, making HLA genes the most polymorphic genes in humans [100-103]. The presentation of antigenic peptides by MHC class I molecules is the result of a series of processes. First, protein turnover results in the degradation of proteins by the proteasome into peptides. These peptides are translocated to the endoplasmic reticulum (ER)

by the transporter associated with antigen presentation (TAP) protein complex. In the ER lumen, further processing can occur, and peptides are paired with MHC class I molecules by the MHC loading complex. The peptide–MHC class I complexes are released from the ER and transported via the Golgi apparatus to the plasma membrane to facilitate presentation to CD8+ T cells. If peptides and MHC class I molecules cannot associate in the ER, they are returned to the cytosol for degradation [104]. In addition to protein turnover, a portion of presented MHC peptides originate from DRiPs (Defective Ribosomal Products) which are misfolded or incorrectly translated proteins rapidly degraded into peptides to prevent protein aggregates [105]. This process has also been shown to directly couple viral transcript translation to the MHC class I presentation pathway (Fig.4)[106].

HLA-presented peptide lengths may vary between 8-15 amino acids, but the majority are nine amino acids long. The high degree of genetic variance of HLA-I alpha chains translates into allele-specific peptide-binding motifs defined by anchor positions [104]. Anchor positions are the amino acids which are responsible for binding to the MHC molecule which are usually the second and the C-terminal amino acids of a peptide [104, 107, 108]. The HLA-class I binding groove restricts the length of peptides, therefore longer sequences can be accommodated through the bulging of the central amino acid sequences.In humans, each cell has around 200,000 cell-surface-expressed HLA complexes,

which bind tens of thousands of unique peptide sequences [104]. Recent evidence has shown that MHC class I antigens are not always linear protein fragments but can also be formed by peptide ligation or proteasomal splicing [109-111]. Such spliced peptides fail to be predicted on the basis of the genomic sequence and are challenging to identify with conventional LC-MS analysis [112].

MHC class II molecules consist of two membrane-spanning chains, α and β, of similar size, encoded by three polymorphic genes (HLA-DR, HLA-DQ and HLA-DP in humans). Unlike MHC class I, MHC class II molecules, which are primarily expressed by professional APCs and present their peptide cargo to CD4+ T cells, follow an entirely different pathway to presentation. MHC class II α and β chains are first assembled in the endoplasmic reticulum (ER) assisted by the invariant chain (Ii) to stabilize the complex [113]. The Ii–MHC class II complex is then transported through the Golgi to the MHC class II compartment (MIIC), where Ii is digested by cathepsin proteases retaining only the class II-associated Ii peptide (CLIP). CLIP is then exchanged for an antigenic peptide with the help of the HLA-DM (known as H2-M in mice). HLA-DM has a highly similar structural like MHC class II proteins. However, due to closed binding groove, HLA-DM fail to bind any peptides. Finally, MHC class II molecules are transported to the plasma membrane to present peptides to CD4+ T cells (Fig.4) [114].

In three decades, great progress has been made on the underlying mechanisms of antigen presentation. However, there is still a need to understand how immunodominant peptides are generated and how exactly these peptides are selected for presentation. Modern technologies such as quantitative PCR, siRNA screens, flow cytometry and computational biology will be used to better understand the relevant molecular relationships and MHC class I/II antigen presentation pathways.

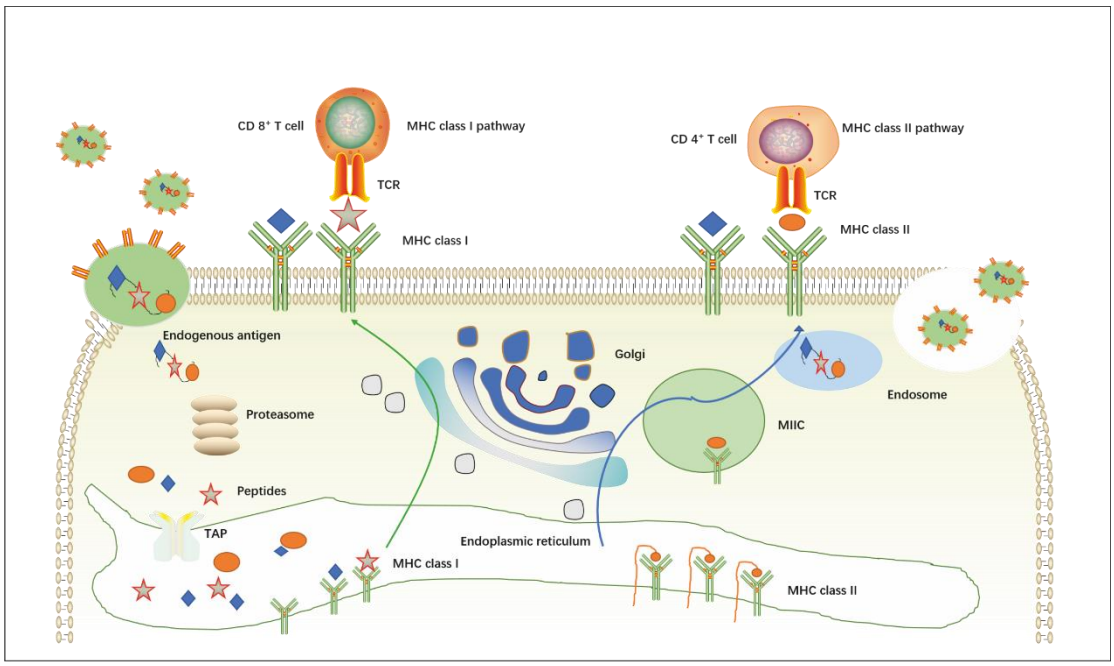


Figure.4 MHC class I/II antigen presentation pathways. The peptide MHC class I pathway (green arrow): First endogenous antigens are degraded by proteasome. Then, the resulting peptide are transported into the endoplasmic reticulum by the transporter associated with antigen processing (TAP) and loaded onto MHC class I molecules. The MHC class I complexes are transported to the plasma membrane via the Golgi to present the antigen to CD 8^{+}T cell. The peptide MHC class II pathway (blue arrow): exogenous antigens are taken up via the endosome pathway. MHC class II assemble and form a complex with the invariant chain (Ii) in endoplasmic reticulum (ER). Then, the complexes are transported to MHC class II compartment (MIIC) via Golgi and class II-associated Ii peptide (CLIP) fragment of Ii is exchanged with antigenic peptide. Finally, the MHC class II molecules are transported to the plasma membrane to present the antigen to CD 4^{+}T cell.

1.6 Mass spectrometry

Since the development of soft desorption ionization methods for mass spectrometric analyses of biological macromolecules by John B. Fenn and Koichi Tanaka who were recognized in the 2002 Nobel Prize in chemistry, mass spectrometry (MS) has become the useful method of choice for the analysis and identification of biomolecules. It is now used to identify, quantify, and characterize biomolecules like peptides and proteins, but also other classes of cellular metabolites like glycans, lipids and other small molecules. The use of high-performance mass-spectrometric techniques coupled with liquid chromatography (LC-MS) has become the methods of choice for peptide sequencing and has become central for antigen discovery research.

1.6.1 Mass spectrometer components and function

A mass spectrometer (MS) is made out of three basic parts: 1) ion source, 2) a mass analyzer, measuring the mass-to-charge ratio (m/z) of the ion. 3) a detector registers the number of ions at each m/z value. Over three decades, two techniques have expanded the range of potential applications of mass-spectrometric analysis, including electrospray ionization (ESI) [115] and matrix assisted laser desorption/ionization (MALDI) [116]. These techniques

have consequently improved the development of new mass analyzers and made MS instruments become the method of choice for the analysis of complex protein samples. There are presently many different types of mass analyzer used, such as an ion trap (IT), time-of-flight (TOF), quadrupole, Fourier transform ion cyclotron (FT-MS) and Orbitrap.

TOF analyzers record the flight time the of target ion when it flies through a specified length tube and this data is used to calculate the ion mass-to-charge ratio. TOF analyzers have greatly advanced in resolution and mass accuracy, which can follow the ESI and MALDI source to achieve high resolution and mass accuracy in MS or MS/MS. Unlike TOF analyzers, IT analyzers trap and accumulate ions enabling the selective isolation of parent ions which can be exploited for data-dependent acquisition. The disadvantages of IT include a lower resolution, lower ion trapping capacity and space-charging effects, all of which can be partially improved by linear ion trap (LIT) analyzers that have a higher dynamic range and achieve better overall sensitivity.

The development of Fourier transform–ion cyclotron resonance (FT-ICR) mass spectrometers which identifies rotational motion during excitation against a magnetic field to detect ionic mass-to-charge ratios and provided a remarkably high resolution and mass accuracy [117, 118]. A new mass analyzer called

orbitrap has now been introduced, achieving such high mass accuracy at high speed and accurate quantitative capacity.

1.6.2 LC-MS for proteomics

The coupling of advanced HPLC techniques alongside mass spectrometry has enabled the analysis of complex analyte mixtures. Here, the complex analyte mixture is separated on a preparative chromatographic column by applying solvent gradients, which are then eluted over a certain amount of time while fed live into the source of the mass spectrometer. This approach allows for the measurement of thousands of peptide molecules over 30 to 60 min gradient elution times in modern proteomics nanoflow applications.

Furthermore, in order to gather sequence information on individual peptide ions, an MS analyser is usually combined with an upstream mass filter device, such as a high performing quadrupole. The quadrupole allows isolation of a defined mass range from the incoming ion stream, allowing the isolation and further fragmentation analysis of individual ions species over time.

The combination of next generation hybrid orbitrap instruments and LIT devices combines the high accurate precursor mass measurement with the

possibility of acquisition of high sensitivity low-resolution tandem MS/MS spectra for improved dentification rates.

1.7 Immunotherapy

1.7.1 Checkpoint inhibition

The success of immune checkpoint blockade therapeutics has demonstrated the power of immune-modulatory therapies for the treatment of cancer [119]. The term ‘checkpoint blockade’ refers to the targeting of immune-inhibitory pathways that hamper the functionality of existing immune responses to cancer. These inhibitory molecules are the cytotoxic T-lymphocyte–associated antigen 4 (CTLA-4) and programmed death-1 (PD-1), which are expressed by T cells and can convey inhibitory signals [120-124]. Programmed death-ligand 1 (PD-L1) can be expressed on tumour cells. Its expression has been shown to be a mechanism through which tumour cells escape immune recognition.

Blockage of these molecules using monoclonal antibodies can re-activate and release pre-existing antitumour immune T cell responses from immunosuppression. This has been shown to be hugely effective in clinical trials for some tumour types [120-124].

1.7.1.1 CTLA4 blockade

T cells are activated when T cell receptors (TCRs) bind to antigens presented by the major histocompatibility complex (MHC) on antigen-presenting cells (APCs). B7-1 (CD80) or B7-2 (CD86) molecules on the APC bind to CD28 molecules on the T cell and provide a co-stimulatory signal that leads to T cell proliferation, resulting in increased T cell clone survival [125]. CTLA-4 is a CD28 homolog with a higher binding affinity for B7-1/2 [125], which prevents the normal signalling between CD28 and B7 from occurring [126, 127]. This, in turn, prohibits the T cell stimulatory signal from occurring.

Ipilimumab, an antibody against CTLA-4, was approved by the U.S. Food and Drug Administration (FDA) for the treatment of melanoma in 2011 [119]. This antibody blocks the action CTLA-4, thereby restoring the ability of the T cells to receive activation signals through CD28.

1.7.1.2 PD-1 and PD-L1

PD-1 pathways are thought to operate at different stages of the immune response. PD-1 is expressed more widely on activated T cells, B cells, and myeloid cells [120-122]. PD-1 binding to PD-L1 and programmed death-ligand 2 (PD-L2) inhibits T cell proliferation and reduces T cell survival [121]. The B7

ligands for CTLA-4 are expressed by APCs in the lymph nodes or spleen). However, PD-L1 and PD-L2 can be found on both leukocytes and nonlymphoid tissues [127, 128]. Tumour cells can express PD-L1 and PD-L2 to bind to PD-1, which inhibits TCR-mediated positive signalling and leads to the reduced proliferation and survival of T cells [121]. Two anti-PD-1 antibodies, pembrolizumab and nivolumab, were approved by the FDA in 2014 [119]. In clinical trials, PD-L1 and PD-1 inhibitors have demonstrated significant and long-lasting response rates [129].

The differences between the locations and effects of anti-CTLA-4 therapies and anti-PD-1 therapies suggest that combined treatment may enhance anti-tumour activity (Fig.5). Evidence shows that combination therapies lead to different patterns of immune activation [130]. Recent studies have trialled the combination of surgery, chemotherapy, and radiotherapy with either anti-CTLA-4 or anti-PD-1/PD-L1 antibodies [131, 132]. The release of dying tumour tissues induced by chemotherapy and radiotherapy could contribute to antigen release and presentation by dendritic cells (DCs), which, in turn, may increase anti-tumour immunity [133]. However, there are still many ongoing efforts to optimise these combination immunotherapies to increase therapeutic success for different types of cancer. Furthermore, the mechanisms are yet to be fully understood.

1.7.2 Second generation immunotherapeutic approaches

Traditional immunotherapies are untargeted and aim to remove inhibitory signals to activate new and re-activate existing immune responses to the tumour. However, due to the immune system's ability to specifically target and destroy cells that present foreign or mutated antigens, a targeted approach may be feasible and have significant advantages over traditional untargeted approaches. Recent studies have shown that the immune system can be used to recognise and kill tumour cells [134-136].

There are two major types of targeted approaches: cancer vaccinations and approaches that target a specific peptide-MHC complex using T cells or antibodies.

New tumour-specific immune responses can be generated by therapeutic cancer vaccination. Pre-existing tumour-specific immune responses can be further enhanced by using immune checkpoint blockades. In recent clinical trials, beneficial immunotherapy outcomes have been observed after the blockade of checkpoint inhibitory pathways, such as CTLA-4 and PD-L1 [120, 121, 123, 137-141]. The aim of such approaches is to enhance natural (pre-existing) T cell responses by lowering the threshold of T cell selection against self-antigens. Alternatively, adoptive cellular therapy (ACT), which is

based on passive adoptive treatments, can harvest and re-infuse autologous lymphocytes to anti-tumour cells [142, 143]. More recently, novel genetic engineering technologies have been used to improve ACT by designing T cells with specificity for tumour antigens [144, 145].

1.7.2.1 Cancer vaccination

Endogenous T cells can recognise peptide epitopes on the surface of malignant cells through the MHC I and II if these cells are sufficiently different from natural cells. Therapeutic cancer vaccines need to address the limits of T cell tolerance. In order to address T cell tolerance and immunogenicity, these vaccines should attempt to utilise antigens that are specific to the tumour cells (i.e., not presented by healthy tissues) and are ideally different from canonical self-proteins. There are several types of tumour-associated antigens that have the potential to elicit potent immune responses, such as viral or cryptic antigens that are uniquely expressed in tumour cells, antigens that result from a tumour-specific mutation, and antigens that are encoded by tumour-specific tissue enriched self-antigens (cancer germline genes) [146].

For virus-associated cancers, such as cervical cancer, epitopes derived from viral ORFs are attractive therapeutic targets due to their tumour-specificity and differences from self-antigens. Scientists have now successfully used

preventative vaccines to protect against cancer-causing infectious diseases, including the hepatitis B virus and HPV. However, these vaccines only work in a limited therapeutic capacity against tumours [147]. This highlights the fact that more work needs to be done to identify suitable tumour-specific antigens that are actively presented on tumour cells to form the basis for more efficient and effective vaccinations.

Recent studies have shown that the use of mass spectrometry, MHC presentation algorithms, and exome sequencing forms a valuable strategy to identify and predict *neo*antigens by MHC molecules [148, 149].

The ideal platform through which to deliver suitable antigens for tumour vaccination has yet to be determined. Unfortunately, most therapeutic cancer vaccines have been unsuccessful during initial randomised control trials. This is thought to be due to the poor understanding of the mechanisms of immunisation [150, 151].

1.7.2.2 Adoptive cell therapy (ACT)

As previously mentioned, the use of immune checkpoint blockades have been a breakthrough in cancer treatment and have currently been approved for several cancer indications. ACT comprises three different approaches:

adoptive transfer of tumour-infiltrating lymphocytes (TIL), gene-modified T cells expressing novel TCRs, and chimeric antigen receptors (CAR). In this approach, T cells are harvested from a patient's blood or tumour, then stimulated to grow and expand in an in vitro culture system. These are then reinfused into the patient to mediate an anti-tumour immune response. For TIL adoptive transfer treatment, it has been shown that CD137 and PD-1 removal by zinc finger nuclease engineering can increase the anti-tumour efficacy of modified T cells [152-154]. Furthermore, it has been demonstrated that tumour-infiltrating BRAFV600E-specific CD4+ T cells correlate with complete clinical responses in melanoma cases after adoptive TIL cell transfer [155]. In 2015, it was demonstrated that HPV antigens E6 and E7 can be used as targets for ACT for the treatment of HPV-associated cancers [27, 28]. Therapies directly targeting the *neo*antigens, such as viral antigens and tumour-specific mutations antigens, typically utilise recombinant TCR or CAR approaches. Unlike the adoptive transfer of TIL, TCR and CAR therapies directly facilitate a cytotoxic T cell response towards a specific antigen and the subsequent killing of the APC. CD19-targeted CARs are the most successful, as the B cell function is largely replaceable [156]. In these approaches, the target selection must balance antitumour efficacy with toxicity to normal tissues.

In conclusion, cancer therapy approaches have developed rapidly in recent years. However, extensive efforts must be made to optimise these treatments to enhance the anti-tumour effect and reduce associated toxicity.

Target	Name	Status*	Company
CTLA-4	Ipilimumab	Approved for the treatment of unresectable or metastatic melanoma Phase III: lung cancer, kidney cancer, and prostate cancer Phase II: cervical cancer, colorectal cancer, gastric cancer, pancreatic cancer, ovarian cancer, and urothelial cancer	Bristol-Myers Squibb
CTLA-4	Tremelimumab	Phase II studies in lung cancer	MedImmune/ AstraZeneca
PD-1	Pembrolizumab	Approved in the United States for treatment of unresectable or metastatic melanoma† Phase III: gastric/GEJ cancer, lung cancer, head and neck cancer, and urothelial cancer Phase II: colorectal cancer, glioblastoma, Merkel cell cancer, pancreatic cancer, and hematologic malignancies	Merck
PD-1	Nivolumab	Approved in the United States for second-line/third-line treatment of unresectable or metastatic melanoma† and for the treatment of metastatic non–small cell lung cancer‡ Phase III: gastric cancer, glioblastoma, head and neck cancer, kidney cancer, and lung cancer (nonsquamous) Phase II: cervical cancer, colorectal cancer, pancreatic cancer, and hematologic malignancies	Bristol-Myers Squibb
PD-1	Pidilizumab	Phase II: kidney cancer and hematologic malignancies	CureTech/ Medivation
PD-L1	Durvalumab	Phase III: head and neck cancer and lung cancer Phase II: colorectal cancer and glioblastoma	MedImmune/ AstraZeneca
PD-L1	Atezolimab	Phase III: bladder cancer and lung cancer Phase II: kidney cancer	Roche

*Only most advanced phase of development for any tumor type is listed; phase I or phase I/II indications are not listed. Includes both monotherapy and combination trials. Information from clinicaltrials.gov.

†With disease progression following ipilimumab and, if *BRAF* V600 mutation positive, a *BRAF* inhibitor. Or in combination with ipilimumab in BRAF WT patients.

‡With disease progression on or after platinum-based chemotherapy. Patients with EGFR or ALK genomic tumor aberrations should have disease progression on FDA-approved therapy for these aberrations prior to receiving pembrolizumab or nivolumab.

CTLA-4 indicates cytotoxic T-lymphocyte–associated antigen 4; GEJ, gastroesophageal junction; PD-1, programmed death 1; PD-L1, programmed death ligand 1.

Figure 5. A comparison of CTLA-4 and PD-1 (This figure is a copy from ref 157) [157]

1.8 Tumour Antigen Discovery

Immunotherapy is relying on T lymphocyte-mediated recognition of HLA-presented, tumour-associated antigens for recognition and elimination of the tumour tissue in the patient. Immunopeptidomics as a tool to investigate

and sequence peptides that are presented to the HLA on tumour tissue has become a major new direction for antigen discovery research.

Cancer cells arise from normal tissues due to somatic mutations that drive transformation of the tumour cells, with a resulting tumour-specific profile of protein expression and pathway activity. The challenge for antigen discovery is to find common antigens expressed highly by tumour cells across many patients in a specific indication, or to monitor in detail changes in individual tumours for tailored therapy design against so called “personalized” events in one patient. Tumour antigens have been classified into two main categories: Cancer testis antigens (CTAs) or tumour-associated antigens (TAAs), and tumour-specific antigens (TSAs) [158]. CTAs and TAAs are encoded by genes that can be normally expressed in healthy cells, but are highly expressed across cancers, and have been proven to be suitable immune cell targets. Examples are melanoma-associated antigen (MAGE) [153] and MART1 [154]. TSAs are antigens spanning specific alterations only present in the tumour, for example sequences spanning tumour-specific mutations, or open reading frames that are unique to the tumour arising from insertion and deletions (INDELs) or viral transcripts, which are absent in all normal tissue. These latter class of antigens are referred to as *neo*antigens, and are the most ideal targets for T cell immunotherapy because of their sole expression on tumour cells.

1.8.1 Spanning tumour-specific mutations

The rapid growth of next generation sequencing (NGS) has made the whole exome sequencing (WES) and RNA sequencing (RNA-seq) common approaches for the identification of tumour-specific, somatic non-synonymous mutations: single-nucleotide variants (SNVs) or INDELs mutations. HLA-associated sequences that span such a mutation, or are product of a novel open reading frame created by INDELs, are attractive therapy targets due to their tumour-specificity and absence on any other healthy tissue. For identification of such neoantigens, so far two approaches have been utilized: Firstly, prediction of HLA sequences spanning the tumour-specific mutation, and secondly, the integration of NGS with immunopeptidomics data. There are various algorithms that have been utilised to predict which peptides can be interacted with TCRs on T cells, such as NetMHCpan algorithm [159]. After input the mutated protein sequences, possible resulting HLA-presented peptides are ranked by their HLA-binding affinities, and then selected for immunogenicity screening to validate whether presentation actually occurs in the patient. However, those prediction screening methods still have some limitations, such as inaccurate sequence prediction, and the limited ability to predict post-translationally modified or spliced peptide sequences.

Immunopeptidomics is an unbiased approach to identify those candidates which significantly contribute to development of antigen-specific immunotherapies for cancer. In this pipeline, firstly the endogenous HLA molecule complexes from the patient's tumour surface are isolated, and associated peptides eluted. Peptides are then injected into liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS). Finally, the MS spectra are compared with database protein sequences which is built by NGS data from patients' tumours to identify the tumour specific *neo*antigens.

1.8.2 Viral antigens

Solid cancers can also raise from viral infection, such as cervical cancer, nasopharyngeal carcinomas and Merkel cell carcinomas [5, 160, 161]. Viral gene products that transform the cell are sources of HLA-presentation and T cell recognition. Importantly, those oncogenic viral proteins could form safe and attractive targets for cancer immunotherapy, because they are not present in any uninfected healthy tissue and are more different from self proteins than neoantigens spanning a single mutation which often just differ from self by a single amino acid.

1.8.3 Intron-retention as a source of *neo*antigen

There is evidence that unconventional antigens can arise specifically in tumours. Such antigens arise from either non-coding regions of the genome, such as mRNA splicing, RNA translation error and post-translational modifications of cellular proteins or peptides. The mRNA splicing in tumours may result in intron retention and novel exon-exon junctions (EEJs). Coulie et al shown that mutated intron sequence codes for an antigenic peptide that can recognized by cytolytic T lymphocytes on a human melanoma [162]. Robbins et al. also showed that the peptide which is translated from intron of the gp100 gene can be recognized by T cells [163]. Recent studies show that intron retention is a general cancer mechanism and can be a source of neoepitopes [164, 165], and cancer-specificity was further demonstrated by analysis of tumour WES and RNA sequencing data that showed that those novel EEJs were present only in cancer cells and predicted to generate more immunogenic epitopes than regular missense mutations [166, 167].

1.8.4 Antigen sourced from alternative reading frames

RNA translation errors can also lead to antigens originating from protein-coding genes in alternative open reading frames (aORFs) or by translation of presumably non-coding sequences, which can be either

tumour-associated or tumour-specific. There are several antigens have been found in human tumours [168-172].

1.8.5 Post-translational modification of antigens

Finally, tumour antigens could also arise from post-translational modifications of cellular proteins or peptides which may consist in specific modification of single amino acid residues, or peptides generated by proteasomal fusion. Those peptides may splice from two separate peptide fragments which original from the same protein (cis-splicing) or from different proteins (trans-splicing) [110, 111]. In 2004, Hanada et al. shown the human renal cancer antigen through post-translational protein splicing [173]. There are several immunopeptidomic studies found spliced peptides from other proteins [174-176]

In future, tumour immunotherapy will improve the treatment's efficacy and provide clinical benefit to treat cancer patients. However, there are still a lot of efforts needed to be made within the realm of cancer vaccines and making these therapies cheaper to the public.

1.8.6 lncRNAs

Due to cervical cancer is the second highest number of cancer deaths in female worldwide, it was a great need to search future targeted therapy and biomarkers for cervical cancer patients. Recent studies have demonstrated that long non-coding RNAs (lncRNAs) is a pivotal regulator in various biological processes which also importance in the diagnosis and treatment of cervical cancers. More than 70% of the mammalian genome can be transcribed into RNA, but only 1% of the genome is directly translated into the known canonical proteome [177]. NcRNAs have two subclasses which are small ncRNAs (20–200 nucleotides) and long ncRNAs (lncRNAs) (> 200 nucleotides to 100 kb) [178]. Date from Human GENCODE shows that Humans have over 16,000 LncRNA genes which could exceed to an estaimetd 100,000 human LncRNAs [179-181]. Most of the lncRNAs have 5′-end m^7G caps and 3′-end poly(A) tails and transcribed by RNA polymerase II and also by other RNA polymerases, which indicate they processed are similarly to mRNAs [181-184]. However, lncRNAs have weaker internal splicing signals and are spliced less efficiently than mRNAs [184]. The recent studies indicated that certain splicing regulators and shorter polypyrimidine tracts contribute to inefficient splicing [185, 186]. Unlike the mRNA, most lncRNAs are located in the nucleus and are less evolutionarily conserved [182-184]. Nonetheles, molecular mechanism and functions of lncRNAs still require rigorous

investigation. LncRNAs lack of protein coding capacity but regulate diverse biological processes including cell growth, survival, migration, invasion, and differentiation [178, 187], which also associated with the development of cancers. Thus, lncRNAs hold great potential as treatment targets and diagnostic biomarkers. LncRNAs regulate gene expression by cis- and trans-regulation mechanisms [188]. Most notably, lncRNAs mainly function at transcriptional, translational, and post-translational levels by interacting with proteins, RNA and DNA [178, 187]. lncRNAs regulate histone modifications and/or chromatin status, which also bind to transcriptional-related proteins preventing them from binding to their DNA targets [178, 187]. Recent studies have shown that upregulation of lncRNAs promotes the growth of cervical cancer cells [189]. Most notably, there are many studies have shown that several lncRNAs are involved in the progression of cervical cancer, such as HOX Transcript Antisense RNA (HOTAIR), H19, Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), RSU1P2, CDKN2B-AS1 and LINC01305 [189-202]. By definition, lncRNAs are not supposed to encode proteins, but more and more evidence shows that proteins may be generated through distinct mechanisms [203-205].

HOTAIR is well known oncogenic lncRNAs, which located in chromosome 12q13.3 [206]. HOTAIR and HPV16 E7 strongly interact with each other in cervical carcinogenesis [194] and overexpression upregulates the mammalian

target of rapamycin (mTOR) pathway in cervical carcinoma cells [207]. Another study demonstrated that HOTAIR also increases tumour growth and invasion in cervical cancer by targeting the Notch pathway [208]. In 2015, Kim et al. showed that HOTAIR may promote tumour aggressiveness through the upregulation of VEGF and MMP-9 and EMT-related genes [209]. HOTAIR modulated HLA-G expression by competitively binding miR-148a and overexpression promoted the proliferation, migration and invasion of cervical cancer cells [195].

H19 is upregulated in several kinds of cancer, such as: ovarian, lung cancer and hepatocellular carcinoma [198]. H19 could act as a molecular sponge of miR-138-5p which could suppress cervical cancer cell progression by targeting SIRT1 [198]. MALAT1, which is associated with the development of various tumours, has been shown to be downregulated in CaSki, HeLa and SiHa cells, which weakened cancer cell invasion and metastasis by inhibiting EMT and epigenetic silencing of miR-124 [189, 210].

In cervical cancer, lncRNAs are found to be key regulators of the pathogenesis and could form therapeutic targets [178]. LncRNAs can be detected in plasma, serum, and other vesicles in extracellular fluid, and could form a suitable class of biomarkers. There is evidence that cervical cancer patients with high expression of lncRNA had a poor overall survival [211]. Therefore, the study of

lncRNA biology has become to improve our understanding of cancer biology and offers a potential new source of neoantigens for cervical cancer immunotherapeutic targeting.

1.9 Overall aim of the thesis:

Target discovery for immunotherapeutic approaches to cervical cancer and high−risk human papillomavirus infection.

Aim1: Definition of HLA-bound tumour-specific, viral antigens (HPV-16 and HPV-18) and tissue enriched self-antigens in a panel of 5 cervical cancer cell lines by integration of LC-MS, mRNA sequencing and exome sequencing datasets. Putative expansion of the definition of tumour-specific antigens in primary cervical cancer tissue samples

Aim 2: Antigen Discovery with a specific focus on lncRNA derived peptide sequences using a *de novo* peptide sequencing approach

Aim 3: Confirmation of immunogenicity of identified viral sequences in human, HLA matched HPV-exposed or infected patient cohorts.

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Plasticware and general reagents

2.1.1 Plasticware and general reagents

Item	Company, Country	Catalogue number
15, 50 ml falcon tubes	Corning Inc, US	430791/430829
175 cm2 tissue culture flasks	Corning Inc, US	431080
25 cm2 tissue culture flasks	TPP	90025
5, 10, 25 and 50 ml pipettes	Corning Inc, US	4487/4488/4489/4490
6-well plates; tissue culture-treated	Corning Inc, US	3516
75 cm2 tissue culture flasks	Corning Inc, US	430641
96-well flat-bottom plates; tissue culturetreated	Corning Inc, US	3596
Cell scrapers	BD Biosciences, US	353085
FACS tubes	BD Biosciences, US	352063
Protein Lo-Bind Eppendorf tubes; 1.5 ml/2 ml	Thermo Fisher Scientific, US	13-698-794/13-698-795

2.1.2 Cell lines and tissue culture reagents

2.1.2.1 Five human cervical cancer cell lines genotypic information

Cell line	HLA-A1	HLA-A2	HLA-B1	HLA-B2	HLA-Cw1	HLA-Cw2
C33 (uninfected)	02:01	02:01	07:02	44:02	05:01	07:02
CaSki (HPV-16)	02:01	03:01	07:02	37:01	07:02	07:02
Hela (HPV-18)	68:02	68:02	15:03	15:03	12:03	12:03
SiHa (HPV-16)	24:02	24:02	40:02	40:02	03:04	03:04
DoTc2 (HPV-16)	03:01	03:01	55:01	55:01	03:03	03:03

2.1.2.2 TC-1 cell line

The tumour cell line, TC-1, is a primary epithelial cells of C57BL/6 micecotransformed with HPV-16 E6 and E7 and c-Ha-ras oncogenes [212]

2.1.2.3 Cell culture reagents and buffers

Item	Company, Country	Catalogue number
Foetal bovine serum (FBS); heatinactivated	APS, US	S-001

HEPES ((4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)	Thermo Fisher Scientific, US	15630080
L-glutamine	ThermoFisher Scientific, US	25030081
MEM Non-Essential Amino Acids	Thermo Fisher Scientific, US	11140035
Penicillin/streptomycin	Thermo Fisher Scientific, US	15140122
Phosphate-buffered saline (PBS)	Sigma-Aldrich, US	D8537
RPMI 1640 medium	Thermo Fisher Scientific, US	11875093
Trypsin/EDTA	Lonza, CH	BE17-161E
Minimum Essential Medium	Sigma-Aldrich, US	56416C
Dulbecco's Modified Eagler's Medium	Sigma-Aldrich, US	D6429

2.1.3 Peptides

Peptides were synthesised by GenScript Biotech (Netherlands) B.V.Table 2.1.3 Synthesis peptide

Sequence	Description	AccuPep-Solubility Test	Company, Country
KQGAMLAVF	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
DTPEWIQRL	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
MKFRLTFY	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 8	SC1634	GenScript Biotech (Netherlands) B.V
RLFSTLITV	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
MLYQMTRTK	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%;	SC1634	GenScript Biotech (Netherlands)

	Length: 9		B.V
TDPSVLQPPTP	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 11	SC1634	GenScript Biotech (Netherlands) B.V
YMLDLQPET	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
YILALACHSR	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 10	SC1634	GenScript Biotech (Netherlands) B.V
ISLFMLHDTI	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 10	SC1634	GenScript Biotech (Netherlands) B.V
VTRAQQGLK	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
KSTISDPVSV	Chemical Peptide Synthesis:	SC1634	GenScript

	Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 10		Biotech (Netherlands) B.V
RVAWPLSAQ	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
LYTRSSNII	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
LHNRLVVFT	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
RVDLLVYNV	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 9	SC1634	GenScript Biotech (Netherlands) B.V
VALLSDNSLLK	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%;	SC1634	GenScript Biotech (Netherlands)

	Length: 11		B.V
APSRLPGV	Chemical Peptide Synthesis: Quantity:9 mg ; Gross Weight; Purity: >85%; Length: 8	SC1634	GenScript Biotech (Netherlands) B.V

2.1.4 Chemical reagents

All chemical reagents were resuspended and stored according to the manufacturer's instructions or as indicated in the section methods 2.2.

2.1.4Chemical reagents and drugs used

Chemical Name	Company, Country	Catalogue number
Acetic acid	Sigma-Aldrich, US	A6283
Acetone	Thermo Fisher Scientific, US	A/0560/PC17
Acetonitrile LC grade	Thermo Fisher Scientific, US	A/0627/17
Boric acid	Sigma-Aldrich, US	B0394
Chloroform	Sigma-Aldrich, US	650498
EDTA	Invitrogen, US	15575-020

Ethanol	Sigma-Aldrich, US	32221
IGEPAL CA630	Sigma-Aldrich, US	I8896
Glycine	Sigma-Aldrich, US	G7126
Methanol	Sigma-Aldrich, US	32213
Sodium dodecyl sulfate (SDS)	BDH, DE	444464T
Sodium hydroxide	Sigma-Aldrich, US	S5881
TRIzol Reagent	Invitrogen, US	15596026
Trifluoroacetic acid	Sigma-Aldrich, US	302031
Tris-HCl	Sigma-Aldrich, US	857645
Urea	Sigma-Aldrich, US	U6504

2.1.5 Enzyme-linked immunospot assay reagents (ELISPOT)

Reagents for ELISPOT were stored as recommended by the manufacturer's instructions.

2.1.5 ELISPOT assay reagents

Reagent	Company, Country	Catalogue number
3-amino-9-ethylcarbazole (AEC)	Sigma-Aldrich, US	A6926
Acetic acid	Sigma-Aldrich,	A8976

	US	
anti-human IFN-γ mAb 7-B6-1, biotinylated	Mabtech, SE	3420-6-250
anti-human IFNγ mAb 1-D1K, purified	Mabtech, SE	3420-3-250
Bovine serum albumin	Sigma-Aldrich, US	F9665
Hydrogen peroxide	Sigma-Aldrich, US	H1009
N-Ndimethylformamide for molecular biology>99%	Sigma-Aldrich, US	D4551
Sodium acetate	Sigma-Aldrich, US	S2404
Sterile D-PBS (1X)	Gibco US	14190-136
RPMI 1640	Gibco US	11875-093
200mM L-Glutamine	Gibco US	25030-024
IFN-ɣ human ELISpot kit	MABTECH Swedish	3420-2A
100X Penicillin/Streptomycin solution	Gibco US	15140-122
PHA	Sigma-Aldrich, US	L1668
Dimethyl sulphoxide (DMSO)	Sigma-Aldrich, US	D2650

TWEEN® 20	Sigma-Aldrich, US	P2287
Bovine Serum Albumin	Sigma-Aldrich, US	A4919

2.2 Methods

2.2.1 Human tissue samples and cell culture

2.2.1.1 Human tissue samples

PBMC

PBMC were obtained from women aged 25-55 years with HPV16 or HPV18 infection confirmed by genotyping of viral DNA in a self-taken vaginal sample, with or without colposcopic abnormalities associated with HPV infection. These women were participants in an observational study of natural immune responses to high-risk HPV, which was approved by the South West – Central Bristol Research Ethics Committee (Ref. 16/SW/0331) [213]. All women gave written informed consent. Tissue samples were obtained from Tissue Solutions Ltd.. Ethical approval was granted by the Central University Research Ethics Committee (CUREC) of the University of Oxford under reference R68126/RE001.

Human cervical cancer tissues

2.2.1.1 Detail of human cervical cancer tissues

Sample Nmber	Sex	Age	Ethnicity	Sample type	Matrix	Diagnosis	Histological diagnosis	Grade	TNM	Stage	Tumor size, cm	Metastases	Date of surgery	Tumor content %	Sample weight, g
1	F	39	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b2N0M0	IB2	4.5x4.5x1	0/27 lymph nodes	04/15/2019	90	0.7
2	F	24	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b1N0M0	IB1	3x2x1.2	0/17 lymph nodes	12/14/2018	100	0.8
3	F	59	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T2a1N1M0	IIIB	4x2.4x2.3	2/19 lymph nodes	12/03/2018	90	0.6
4	F	35	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b1N0M0	IB1	3x2	0/11 lymph nodes	03/04/2011	100	0.6
5	F	50	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b2N0M0	IB2	6x4x2.5	0/16 lymph nodes	10/20/2010	95	0.75
6	F	56	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b1N1M0	IIIB	3.5x3.5x2.5	3/3 lymph nodes	09/13/2010	100	0.9
7	F	53	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b1N0M0	IB1	2.5x2	no	07/08/2010	90	0.75
8	F	52	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b1N0M0	IB1	4x4x3	0/3 lymph nodes	06/01/2010	90	0.9
9	F	54	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G2	T1b1N0M0	IB1	2.3x1.5x1	no	05/24/2010	90	0.8
10	F	39	Caucasian	FF	cervix	cancer	squamous cell carcinoma	G3	T1b2N1M0	IIIB	4.5x1.3	1/3 lymph nodes	05/06/2010	100	0.9

2.2.1.2 Cervical cancer cell culture

C33 and SiHa were cultured in Minimum Essential Medium (Sigma) supplemented with 10% heat-inactivated fetal calf serum, 2 mM L-glutamine, 1 mM pyruvate, 0.1mM Non-Essential Amino Acids Solution and 100U penicillin/ml and was incubated at 37˚C in 5% CO2. HeLa and DoTC2 were cultured in Dulbecco's Modified Eagler's Medium (Sigma) supplemented with 10% heat-inactivated fetal calf serum, 2 mM L-glutamine and 100U penicillin/ml, and was incubated at 37˚C in 5% CO2. CaSki were cultured in RPMI 1640 (Sigma) supplemented with 10% heat-inactivated fetal calf serum, 2 mM L-glutamine, 1 mM pyruvate, 0.1mM HEPE Solution and 100U penicillin/ml, and was incubated at 37˚C in 5% CO2.

2.2.1.3 TC-1 cell culture

TC-1 were cultured in RPMI 1640 (Sigma) supplemented with 10% heat-inactivated fetal calf serum, 2 mM L-glutamine, 1 mM pyruvate, 0.1mM Non-Essential Amino Acids Solution and 100U penicillin/ml, and was incubated at 37˚C in 5% CO2.

2.2.2 Cell DNA purification methods

Genomic DNA was extracted from the five cervical cell lines using DNeasy ® Blood & Tissue Kits (Qiagen, Hilden, Germany). DNA extraction was performed following the manufacturer's instructions. Each 5 x 10^6 cervical cancer cells were centrifuged for 5 min at 300 x g and resuspend in 200 µl PBS. 20 µl proteinase K was added to the tube. And then, 200 µl Buffer AL was added to the sample and briefly vortexed. 200 µl ethanol was added and vortexed again. The mixture was pipetted into a DNeasy Mini spin column which placed in a 2 ml collection tube and centrifuged at 8000 rpm for 1 min. Spin column was placed in a new 2 ml collection tube. 500 µl Buffer AW1 was added and spin column was centrifuge for 1 min at 6000 x g. Spin column was placed in a new 2 ml collection tube. 500 µl Buffer AW2 was added and spin column was centrifuge for 3 min at 20,000 x g. Spin column was transferred to a new 2 ml microcentrifuge tube and added 200 µl Buffer AE. The sample was incubated in room temperature for 1 min and centrifuged for 5 min at 300 x g to elute the DNA. DNA concentration and purity were measured by Nanodrop (NanoDrop 2000 spectrophotometer, Thermo Fisher Scientific, US).

2.2.3 RNA purification methods (cell)

Genomic RNA was extracted from the five cervical cell lines using RNeasy® Mini Kits (Qiagen, Hilden, Germany). RNA extraction was performed following the manufacturer's instructions. Each 5 x 10^6 cervical cancer cells were

transferred to an RNase-free tube and centrifuged for 5 min at 300 x g after washed with PBS. The cells were resuspended in 175 µl of precooled (4ºC) buffer RLN and incubated on ice for 5 min. The sample was centrifuged at 4ºC for 2 min at 300 x g. The supernatant was transferred to a new tube and 600 µl Buffer RLT added. After vortexing, 430 µl ethanol was added to homogenize the sample. 700 µl of the sample were transferred to a RNeasy spin column which was placed in a 2 ml collection tube. After the sample was centrifuged for 15 s at 8000 x g, the flow-through was discarded. RNeasy spin columns were centrifuged for 15 s at 8000 x g after 700 µl buffer RW1 was added to wash the spin column membrane. The RNeasy spin columns were centrifuged for 15 s at 8000 x g again after adding 500 µl Buffer RPE to wash the spin column membrane. The RNeasy spin column was then centrifuged for 2 min at 8000 x g again after adding 500 µl Buffer RPE. Finally, the RNeasy spin column was placed in a new 1.5 ml collection tube. 30 µl RNase-free water was added and the spin column was centrifuged for 1 min at 8000 x g to elute the RNA. RNA concentration and purity were measured by Nanodrop (NanoDrop 2000 spectrophotometer, Thermo Fisher Scientific, US).

2.2.4 RNA purification methods (tissues)

Genomic RNA was extracted from the ten cervical tumour tissues using TRIzol Reagent (Invitrogen Thermo Fisher Scientific, US). RNA extraction was

performed following the manufacturer's instructions. Each 50 mg cervical tumour resection was transferred to an RNase-free tube and 1 mL of TRIzol Reagent was added. The tissues were homogenized using a homogenizer (Precellys evolution Homogenizer (Bertin Technologies SAS) using the following method: 6500:1X40-0015.). The sample was incubated on ice for 5 min and 0.2 mL chloroform was added. The sample was centrifuged at 4ºC for 15 min at 12,000 x g. The aqueous phase was transferred to a new tube. 0.5mL isopropanol was added to each tube and incubated for 10 minutes. The sample was centrifuged at 4ºC for 10 min at 12,000x g and the supernatant discarded. 1 mL ethanol was added to resuspend the sample. After vortexing, the sample was centrifuged at 4ºC for 5 min at 7500x g. The supernatant was discarded, the RNA pellet was air dried for 10 minutes. And then, RNA pellet was resuspended in 50µl RNase-free water, and RNA concentration and purity were measured by Nanodrop (NanoDrop 2000 spectrophotometer, Thermo Fisher Scientific, US).

2.2.5 Exome and RNA sequencing

Exome sequencing of all samples was performed in the Wellcome Centre Human Genetics, Oxford Genomics Centre on an Illumina HiSeq 4000 system using the SureSelect kit v6 (Agilent). RNA sequencing of all samples was performed in the Wellcome Centre Human Genetics, Oxford Genomics Centre

on an Illumina HiSeq 4000 system using the SmartSeq2 workflow.

2.2.6 Exome variant calling

Raw exome sequencing reads had adaptors removed and low quality bases trimmed using Trim_Galore v0.6.2 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore). Reads were then mapped to the UCSC hg38 genome sequence using Bowtie2 v2.3.5.1 [214]. Duplicates were marked using GATK v4.1.7.0 Mark Duplicates function.

Germline variant calling of these cell lines was necessary due to the absence of paired germline samples. Variants were called using Strelka v1.9.10, with indel candidates provided by Manta v1.6.0 [215, 216]. The variant calling was limited to the target capture regions described in the file "SureSelect_Human_All_Exon_V6_r2" for hg38 from Illumina. At this stage, output files were filtered to retain variants marked "PASS" by Strelka's filtering standards. Additional filtering set a threshold mutation read coverage of at least 10 reads (DP vcf field for SNVs, DPI for indels).

An in-house modified and updated version of neoantigen prediction tool MuPeXI [217], known as TUNAPASTA, was used to further process these variants. As part of the process, variant annotation was performed by

Ensembl-VEP v101.0 [218]. Mutations in transcripts identified as other than 'protein-coding' gene biotype were excluded, as were any transcripts that included the biotype "NMD_mediated_decay". Ensembl-VEP also provided global allele frequency information, used in the subsequent step of analysis. TUNAPASTA then generated output files including this collated information, along with short protein sequences (desired length 31 aa) fasta file outputs with paired normal and mutant sequences where mutations are located from the 16th aa (and onwards if frameshifted).

Downstream processing in R was used to further filter and prepare the fasta file. Normal and mutant peptide sequences are split into separate paired entries. In an attempt to emulate the nature of somatic variants being present in less than 1% of surveyed populations, mutations were retained at this threshold according to the "highest allele frequency observed in any population from 1000 genomes, ESP or gnomAD" projects. This final stage of filtering generated the provided fasta files of peptides representing the contextual protein sequences of likely-somatic mutations.

2.2.7 *De novo* assembly

Raw RNA sequencing reads had adaptors removed and low quality bases trimmed using Trim_Galore v0.6.2

(http://www.bioinformatics.babraham.ac.uk/projects/trim_galore) and quality control metrics assessed using FastQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). A custom STAR (v2.7.3a) [219] genome reference was created utilising UCSC hg38 genome sequence together with relevant HPV genome sequences from PaVE (https://pave.niaid.nih.gov/). Reads were then mapped to this custom genome reference using STAR with two passes, sorted and indexed by Samtools v1.10 [220]. Duplicates were marked using GATK v4.1.7.0 [221] MarkDuplicates function and viral-mapped reads were quantified using feature Counts from Subread v1.6.4 [222] to perform HPV-typing. This was then confirmed using the HPV genotyping tool HPVEM [223].

Unmapped reads from STAR were pooled with viral-mapped reads and used as inputs for Trinity v2.11.0 [224] for *de novo* transcript assembly. These resulting transcripts were filtered by aligning to each sample's respective HPV genome using NCBI BLASTn [225]. Figures displaying alignment of Trinity transcripts to the HPV genome were created using genoPlotR [226]. Finally, transcripts were translated into all 6 reading frames using the R package seqinr [227] and appended to the Human UniProt protein reference to create a sample-specific protein database for PEAKS analysis.

Genbank references HPV 16 (K02718.1), HPV 18 (X05015.1)

HPV-31(gi|333048) and HPV-45 (gi|397022) were used for: RNAseq pipeline including Trinity, BLAST alignment of Trinity transcripts, and the HPV genoplot alignment figures with the rainbow of annotated genes at the top. PAVE references were used for: HPV-typing with HPV-EM program, and the associated read counts (in the attached file) and coverage plots.

2.2.8 Flow cytometry

To determine HLA-I and HLA-DR surface expression, the cell lines were washed once in 200 ul PBS then stained with anti-HLA-I-PE (W6/32; BioLegend) or anti-HLA-DR-PE (L243, Biolegend) for 15 minutes at room temperature. Cells were then washed and fixed in 4% PFA. Staining data were acquired on a CyAn ADP Analyser with Summit software and analysis was performed on FlowJo software (FlowJo LLC).

2.2.9 Preparation of antibody-conjugated immunoresin

1ml Protein G-Sepharose beads (for HB-51) or Protein A-Sepharose beads (for W6/32) (GE Healthcare) were washed in 50 mM borate, 50 mM KCl (pH 8.0) solution and then incubated with 2 mg of antibody slowly rotating 1 hour in cold room. The beads were washed in a column format with 0.2 M

triethanolamine (pH 8.2), and the bound antibody was cross-linked by incubation with 40 mM dimethyl pimelimidate dihydrochloride (DMP) (Sigma) (pH 8.3) for 1 h at room temperature. Ice-cold 0.2 M Tris buffer (pH 8.0) was added to the mixture to stop the reaction. Unbound antibody was washed off the column by washing with 0.1 M citrate (pH 3.0), and the column was equilibrated in 50 mM Tris (pH 8.0) for further use [228].

2.2.10 MHC class I immunoprecipitation

All experiments were conducted at temperatures below 4 °C. For cell sample, $5x10^8$ cells pellets were lysed by using 10 ml lysis buffer ((0.5% IGEPAL 630, 150 mM NaCl, 50mM Tris [pH 8.0] plus protease inhibitor cocktail (Roche)) and gentle agitation for 30 minutes. For tissues sample, same quantity of beads as the volume of the tissue were add for each Eppendorf. 1 ml 2x lysis buffer (0.5% IGEPAL 630, 150 mM NaCl, 50mM Tris [pH 8.0] plus protease inhibitor cocktail (Roche)) was added to each tissue sample. Homogenized the tissue with Precellys evolution Homogenizer (Bertin Technologies SAS) using the following method: 6500:1X40-0015. The lysate was centrifuged at 300g for 10 min to remove nuclei and then at 15,000g for 60 min to pellet other insoluble material. 1ml HB51 or W6/32 monoclonal antibody cross-linked to protein G-Sepharose beads (GE) (1ml) to capture the MHC class I complexes. The column was washed by 50mM Tris buffer (pH 8.0) containing first 150 mM

NaCl, then 450 mM NaCl, and next no salt. MHC-peptide complexes were eluted by using 5 ml 10% acetic acid and dried. Containing first 150 mM NaCl, then 450 mM NaCl, and next no salt. HLA-peptide complexes were eluted by using 5 ml 10% acetic acid and dried [228].

2.2.11 High performance liquid chromatography peptide fractionation

After immunoprecipitation, peptide-MHC/HLA complexes were dried and resuspended in 120 µl of loading buffer (0.1% trifluoroacetic acid (TFA), 1% acetonitrile (ACN) in water). Samples were fractioned by RP-HPLC using an Ultimate 3000 high-performance liquid chromatography (HPLC) system (Thermo Scientific) and 4.6- by 50-mm ProSwift RP-1S column (Thermo Scientific). Samples were loaded into the column and eluted using a 10-min gradient from 3% to 30% ACN in 0.1% TFA at a flow rate of 500 nl/min and absorbance of protein at 280 nm. Fractions were collected from 2.5 to 17.5 minutes at a 1-minute interval. Fractions containing peptides were combined in two fractions (odd and even), dried, resuspended in 20 ul of loading buffer and analysed by LC-tandem mass spectrometry (LC-MS/MS) [228].

2.2.12 LC-tandem mass spectrometry (LC-MS/MS)

Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo Scientific):

Peptides were analyzed by nanoliquid chromatography – tandem mass spectrometry (LC-MS2) using an Ultimate 3000 RSLCnano System (PepMap C18 column, 2 μm particle size, 75 μm × 50 cm; Thermo Scientific) coupled with an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Scientific). A 60 min linear gradient from 3% to 25% ACN in 5% DMSO/0.1% formic acid at a flow rate of 250 nl/min was used for peptides elution. Mass spectrometry (MS) detection was performed with a resolution of 120,000 for full MS (300-1500m/z scan range), whereas precursors were selected using TopSpeed ion selection within a 2 second cycle time, an isolation width of 1.2 amu for fragmentation, resolution set at 30,000 and high-energy collisional dissociation (HCD) set at 28 for peptides with 2-4 charges and 32 for peptides with singly-charged.

Q Exactive™ HF Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Scientific): Peptides were analyzed using an Ultimate 3000 RSLCnano System coupled with an Q Exactive™ HF Hybrid Quadrupole- Orbitrap™ Mass Spectrometer (Thermo Scientific). Peptide were loading into a column (PepMap C18 column, 2 μm particle size, 75 μm × 50 cm; Thermo Scientific) and eluted using a 60-min linear gradient from 3% to 25% ACN in 1% DMSO/0.1% formic acid at a flow rate of 250 nl/min. Mass spectrometry (MS) detection was performed with a resolution of 120,000 for full MS (320-1600m/z scan range) and AGC target of 300,000. Precursors were isolated using an

isolation width of 1.6 amu and fragmented using high-energy collisional dissociation (HCD) set at 25 for peptides with 2-4 charges. Fragments resolution was set at 60,000 with AGC target of 50000.

2.2.13 Mass spectrometry data analysis

LC-MS/MS datasets were analyzed using Peaks v10. For in-house datasets the following parameters were set: precursor mass tolerance: 5 ppm; fragment mass tolerance: 0.03 Da; digestion: none; fixed and variable modifications: none. 6-frame translations of Trinity assemblies (Caski: 19, Dotc2: 24, SiHa: 16, Hela: 13 entries) mapping to the HPV genome were concatenated with the reviewed human entries in Uniprot (Swissprot version April 2018, containing 20,328 entries). For immunopeptidomics data, a threshold of -10lgP=15 was applied across all datasets, and the FDR was consistently reported as <5% (C33: 4.83%; Caski: 1.79%; Dotc2: 1.81%; Hela: 1.22%; Siha: 0.87%). For PXD004452 data analysis the following parameters were selected: precursor mass tolerance: 20 ppm; fragment mass tolerance: 0.05 Da; digestion: specific to the indicated enzyme used for digestion; fixed modifications: carbamidomethylation; variable modifications: none.

2.2.14 Generation of STCLs and IFNγ ELISpot

Standard operating procedures for the isolation and preservation of peripheral blood mononuclear cells for analysis of antigen-reactive T cell responses were applied. PBMCs ($1\text{-}2\times10^6$) from HPV-positive patient blood samples were cultured at 37°C and 5% CO2 for 10 days in R10 medium (RPMI with 10% fetal calf serum) plus IL-7 (25ng/ml) with HPV peptides identified from the cervical cancer panel (4μg/ml) or a pool of irrelevant FEC peptides (2μg/ml). Cells were supplemented with 1,800 U/ml IL-2 on days 3 and 7 and additional R10 on day 7. Cells were collected on day 10, washed four times with PBS and rested at 37°C and 5% CO2 for 29-34 hours in fresh R10. ELISpot plates (MSIPS 4510, Millipore) were coated with anti-human IFN-γ coating antibody (clone 1-D1K, MabTech) in the dark at 4°C overnight. Plates were blocked with R10 complete medium at 37°C for a minimum of 2 hours. Cells ($1x10^5$/well) were added and incubated with HPV peptides (2ug/ml) or FEC (4ug/ml). PHA at 5μg/ml and R10 (0.1% DMSO) were used as positive and negative control, respectively. Plates were incubated for 16-18 hours at 37 °C, after which a biotinylated anti-human IFN-γ detection antibody (clone 7-B6-1, MabTech) was added for 1 hour at room temperature. Plates were then washed 6 times and ABC complex added for 1 hour at room temperature followed by the addition of AMEC Complex Buffer. Spot-forming units were counted using an automated reader (AID). The magnitude of response was defined as the number of spot-forming units (SFU) per million PBMCs following background subtraction.

2.2.15 MAR-S

The MAR-S model is a linear model to improve Peaks de-novo MS2 peptide-spectrum matching (PSM) candidate selection. It incorporates two factors to rerank the Peaks (Bioinformatics Solutions) reported ALC candidate peptide lists. Firsty, HLA binding prediction, and secondly chromatographic peptide retention time prediction. For HLA binding prediction, the MAR-S model uses NetMHCpan 4.1, and a deep neural network based algorithm named DeepRTplus [229] was used to do RT prediction. Comprehensive evaluations using multiple datasets suggested that MAR-S model improves performance of *de novo* PSM candidate selection over the baseline ALC model as reported by Peaks.

After obtaining the peptide sequences, a blastp algorithm was applied against a comprehensive range of databases to find potential origins of these MAR-S best candidate peptide sequence identifications. Firstly, the SwissProt, Trembl and Ensembl human protein databases were searched including protein isoforms, and any peptides matching a database peptide with Levenshtein distance 1 or less were excluded from further investigation (these are accurate matches and peptides with one amino acid difference). The remaining peptides were further matched against lnc-RNA databases Lncipedia [230] and LNCRNAK [231]. If the original database comprised DNA/RNA sequences,

six-frame translation was applied to convert them into amino-acid sequences. We only select peptides with 100% identity and 100% coverage (Levenshtein distance 0) with a database peptide.

2.2.16 Mouse tumour challenge.

A trypsinized TC-1 cell pool was washed once with HBSS (Hank's balanced salt solution). Then, cells were diluted to the desired concentration (10x of injection dose) with HBSS. 100 µl of cells suspension ($1x10^6$ per animal) were subcutaneously injected in the inguinal region of a C57BL/6 mice (Jackson Laboratory). Tumour growth was checked twice a week by visual inspection or/and palpation. Total number of six mice divided into two group after all tumours reached 3mm or more in diameter and then immunized and boosted with ChAd- MVA (Group A) or PBS (Group C) for one week. Tumours were resected and frozen until further processing (Dr.Gemma Hancock developed this mouse model and new vaccine ChAd-MVA to test in the mouse model).

2.2.17 Statistics and data presentation

Statistical tests were performed, and graphs were created using GraphPad Prism 9 Software (GraphPad Software Inc, USA).

Chapter 3: Identification of tumour-specific antigens in five human cervical cancer cell lines

3.1 Introduction

Human papillomaviruses (HPVs) are small, double-stranded DNA viruses. There are more than 170 types, including 15 types that are designated “high risk” (HR) or oncogenic. HR types are present in more than 99% of cervical cancers. Each year, over 570,000 women are diagnosed with cervical cancer, and 311,000 die from this disease worldwide. The emergence of cancer immunotherapy has provided remarkable new possibilities for the treatment of cancer. In this strategy, it is important to understand which proteins are exclusively presented by the tumour and therefore can be used as antigens in vaccination approaches. Recently, high performance mass-spectrometric techniques have given us the ability to globally investigate how peptides are presented to the immune system. In this study, we took advantage of the new generation of improved mass spectrometers, specifically the Orbitrap Fusion Lumos (Thermo Scientific). Our workflow (Fig.6) made it possible to identify MHC-I peptidomes from cancer cell lines (~5×10^8). We also explored how to use *in silico* binding prediction and clustering tools on the set of identified HLA peptides from cancer-associated antigens or proteins, with a view to select vaccine candidates. In addition, to identify the mutated and viral peptides, we sequenced the exome and mRNA transcriptome of five human cervical cancer

lines. For immunopeptidomics data, a threshold of -10lgP=15 was applied across all datasets, and the FDR was consistently reported as <5% (C33: 4.83%; Caski: 1.79%; Dotc2: 1.81%; Hela: 1.22%; Siha: 0.87%). In these human cervical cancer lines, we identified 55,474 peptides (48,278 unique peptides) from 15,129 proteins (10,920 unique proteins). More than 84% of the peptides had the typical length of HLA-I peptides (8 to 14 amino acids) and most of the presented peptides are nine amino (60%) acids long. With the help of RNA-seq and the DNA exome sequencing-based protein databases, we identified 5 viral peptides from our sample and 6 peptides that come from 5 neoantigens. Such peptides are especially promising for developing immunotherapeutics and our protocol is expected to advance immunopeptidomics as an antigen discovery platform for personalized cancer immunotherapies.

3.2 RESULTS

3.2.1 Identification of tumour-specific antigens in five human cervical cancer cell lines

3.2.1.1 Determination of HLA expression in five cervical cancer cell lines

In order to investigate the nature of viral and tumour-associated antigens in cervical cancer, we utilized five human cervical cancer cell lines (CaSki, HeLa, SiHa, DoTc2, and C33) as our human cervical cancer model (Table 1). The

five human cervical cancer cell lines express a total of four different HLA-A alleles, six HLA-B alleles, and five HLA-C alleles (Table 1). HeLa is reported to be transformed with HPV-18, while CaSki, DoTc2, and SiHa are reported to be HPV-16 positive. C33 is an HPV negative cervical tumour cell line (Table 1). Accordingly, no viral transcript reads were detected in the RNA sequencing data (Fig. 8 A, B). We performed FACS analysis to confirm expression of HLA class I and II in all of the cell lines. DoTc2 expressed both HLA-I and HLA-II alleles, but the other four cell lines only had HLA-I receptors (Fig. 7).

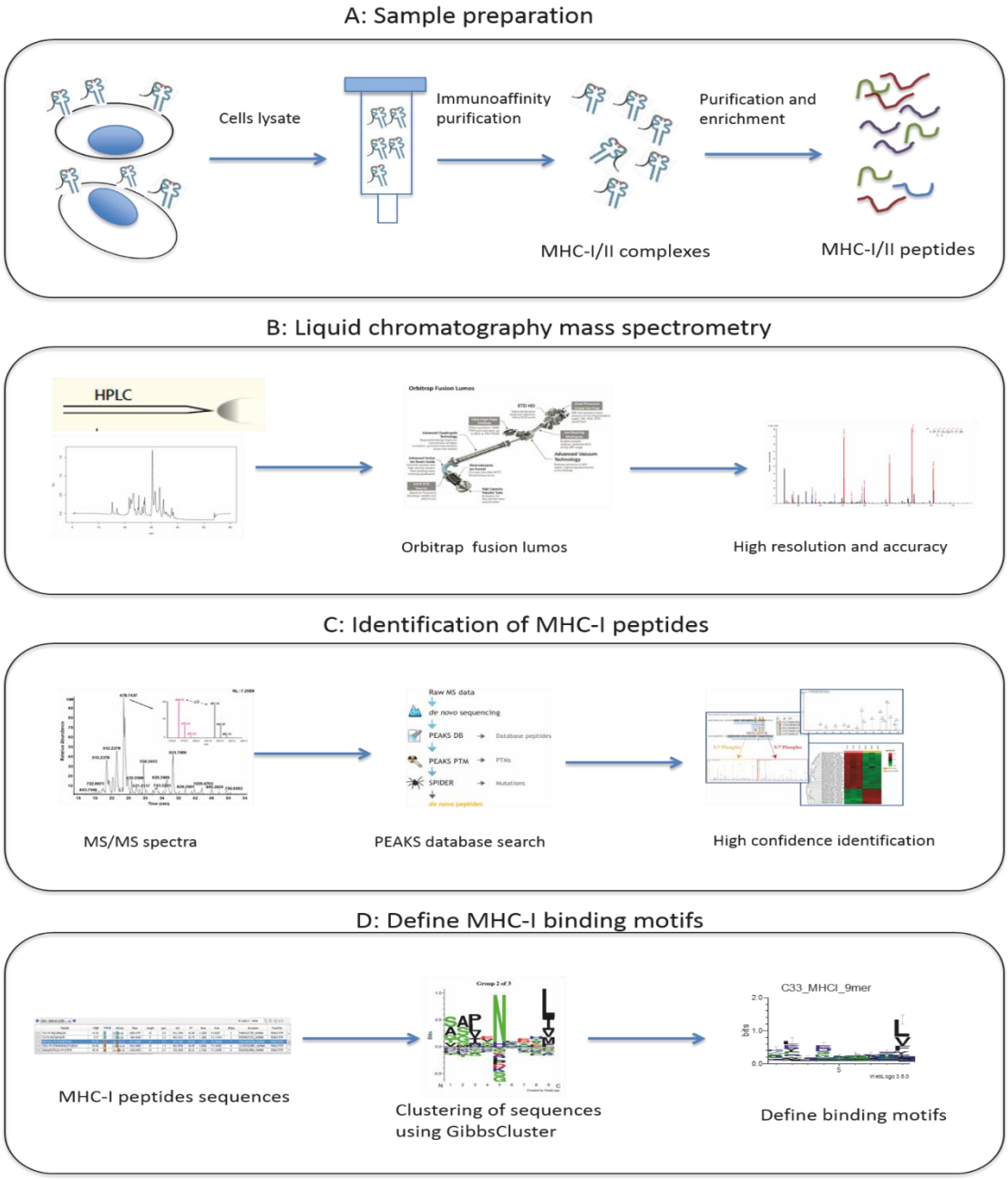
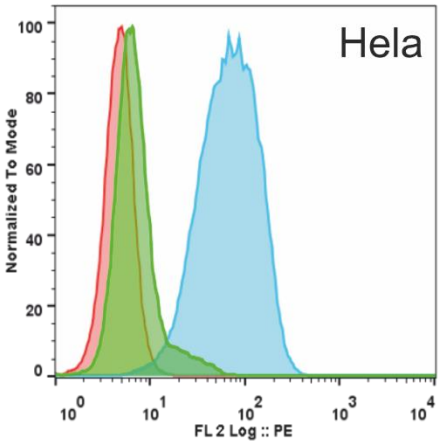
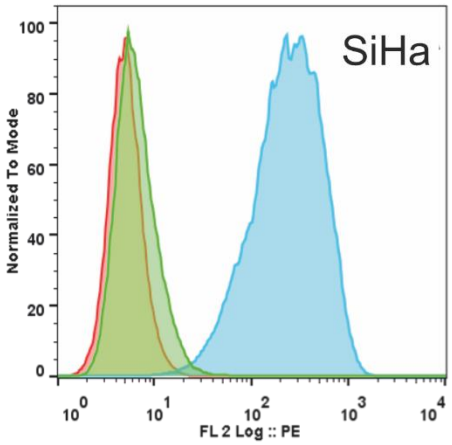
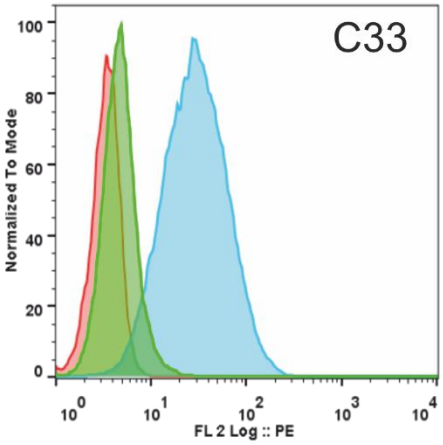


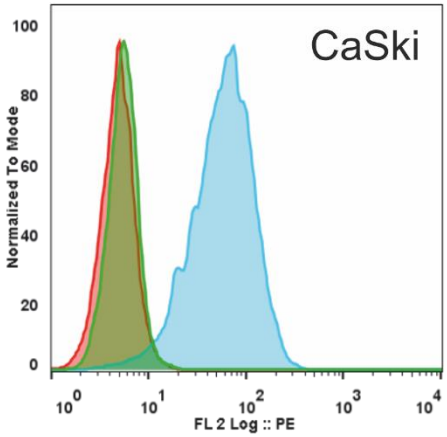
Figure 6. Workflow of MHC-I Peptidomics. Sample preparation. MHC-I complexes were immunoaffinity purified from cell lysates using anti-MHC-I (HB-51/W632) antibody cross-linked to Protein-A/G Sepharose beads. MHC-I peptides were purified from the beta-2- microglobulin and alpha chain by using

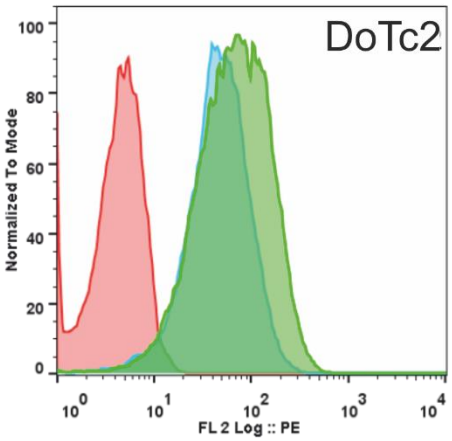
a preparative chromatography workflow. B. LC-MS/MS. The enriched mixtures of MHC-I peptides were measured using liquid chromatography mass spectrometry (LC-MS/MS), resulting in high resolution and high mass accuracy at the MS and MS/MS levels. C, Identification of MHC-I peptides. MHC-I peptides were analyzed with PEAKS software, allowing the identification of peptides with 1-4 positive charges. D, MHC-I stratification and consensus binding motifs. Using the NetMHCpan prediction engine and the Gibbs Cluster tool, the consensus binding motifs were defined from the identified peptides sequences.











Control

W632 HLA-I Ab

L243 HLA-II Ab

Figure 7. FACS experiment of five human cell lines. Unstained control is red, W632 Ab for HLA-I is blue and L243Ab for HLA-II is green. (n=1)

3.2.1.2 Preparation of personalized protein sequence databases for each cell lines

To identify the mutational load in the cell lines, we performed exome sequencing of all five cancer cell lines. Information was added to the databases for spectral interrogation and to allow for the identification of *neo*antigens and viral peptides in our samples. Furthermore, we performed mRNA sequencing in order to accurately define the nature of the viral transcript sequences. Viral transcript sequences were reconstructed by *de novo* assembly of mRNA transcript sequences that matched the HPV reference genomes (HPV-16 (NC_001526) and HPV-18 (NC_001357)) using the Trinity tool [224]. Both the Exome and mRNA sequencing data were analysed by Dr. Julius Muller and Dr. Issac Woodhouse (Fig. 8). We also downloaded the HPV-16 (NC_001526) and HPV-18 (NC_001357) genome and created a protein database by translating all open reading frames (ORF) over 8 aa long into all 6 frames as our 6Fr (6-frame translation database) reference.

Table 1. List of five human cervical cancer cell lines genotypic information

Cell line	HLA-A1	HLA-A2	HLA-B1	HLA-B2	HLA-Cw1	HLA-Cw2
C33 (uninfected)	02:01	02:01	07:02	44:02	05:01	07:02
CaSki (HPV-16)	02:01	03:01	07:02	37:01	07:02	07:02
Hela (HPV-18)	68:02	68:02	15:03	15:03	12:03	12:03
SiHa (HPV-16)	24:02	24:02	40:02	40:02	03:04	03:04
DoTc2 (HPV-16)	03:01	03:01	55:01	55:01	03:03	03:03

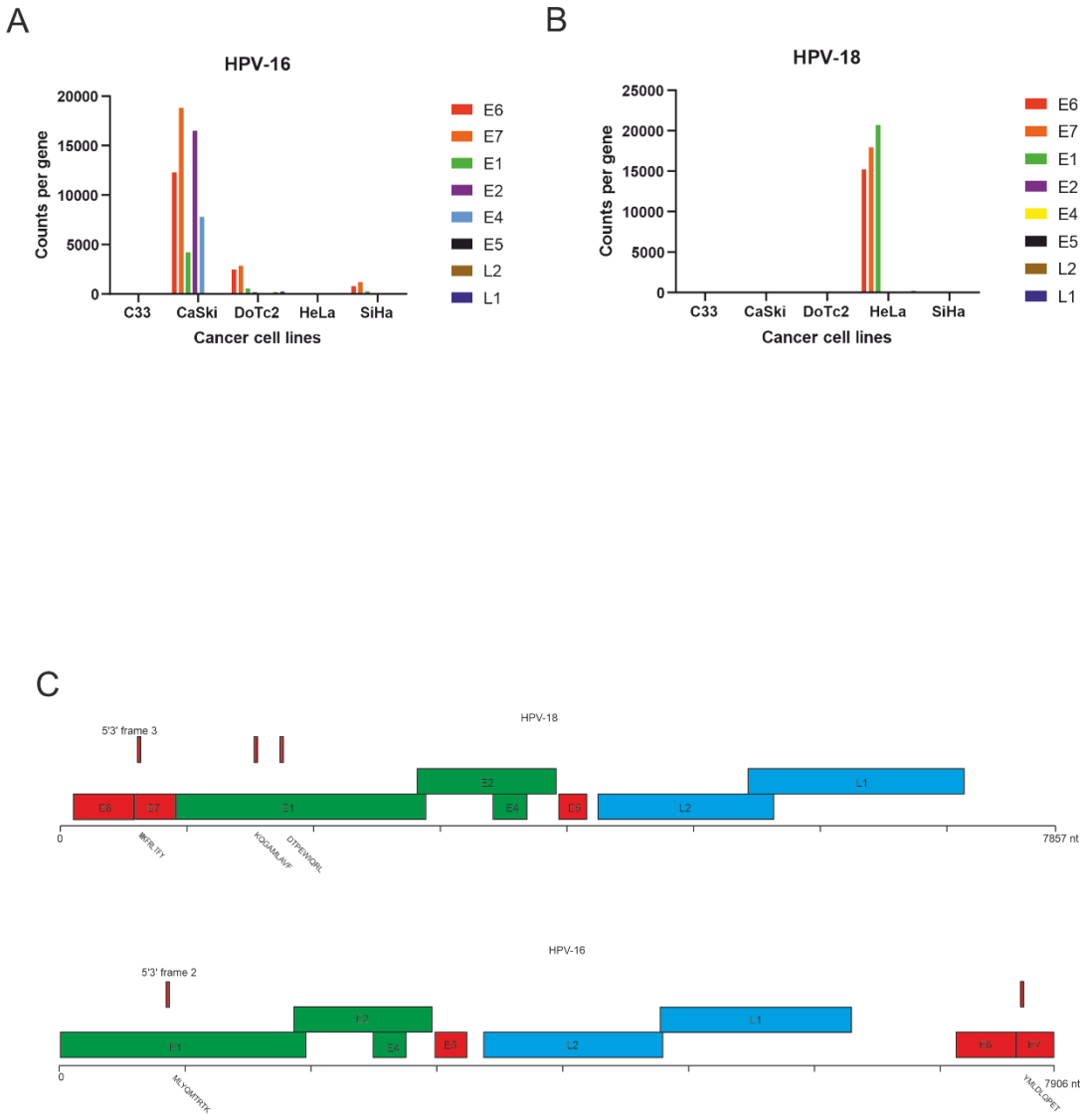


Figure 8. RNAseq data and viral peptide antigens identified in cervical cancer cell lines. A. Transcript abundance in counts per gene as measured by RNAseq for all five cell lines for HPV-16. B. Transcript abundance in counts per gene as measured by RNAseq for all five cell lines for HPV-18. C. Relative position of viral HLA-peptides identified mapping to HPV 16 (CaSki cells) and HPV 18 (in HeLa cells), respectively. The horizontal line indicates the nucleotide position in the regarding viral genome, horizontal boxes show the

open reading frames for the main viral proteins. Red vertical lines indicate the coding sequence for the identified HLA-associated peptide sequences.

3.2.1.3 Sequencing of the immunopeptidomes of the cervical cancer cell line panel

At a false discovery rate (FDR) threshold of 5%, we identified overall a total of 55,474 (48,278 unique peptides) peptides (Fig. 9 A) that originate from 15,129 (10,920 unique proteins) proteins (Fig. 9 B). More than 84% of the peptides have the typical length of HLA-I peptides (8 to 14 amino acids), and most of the presented peptides are nine amino (60%) acids long (Fig. 9 D). According technical and biological experiments, 5×10^8 cells per biological replicate obtained higher number of HLA peptides in a relatively short measurement time (C33 69%, CaSki 27%, DoTc2 27%, Hela 34%, SiHa 29%) (Fig. 10) (Fig. 11). Generally, unique repertoires of peptides are presented on the surface of cells that express different sets of HLA alleles according to the specificity of their binding motifs. Because the five human cervical cancer cell lines produce up to 15 unique cell surface receptors that bind and present the HLA class I peptidome (Table 1), more than 85% of the identified peptides are unique (Fig. 9 A). As expected, we observed that cells expressing similar alleles, such as DoTc2 and CaSki cells, shared dramatically more peptide sequences (3083) than the cells expressing different set of alleles, like SiHa and DoTc2 (1070)

(Fig. 9 A). In those 3083 peptides, more than 99% binding to the same allele (Fig. 12). This highlights the specificity of our protocol. HLA-I peptides are characterized by their typical binding motifs.

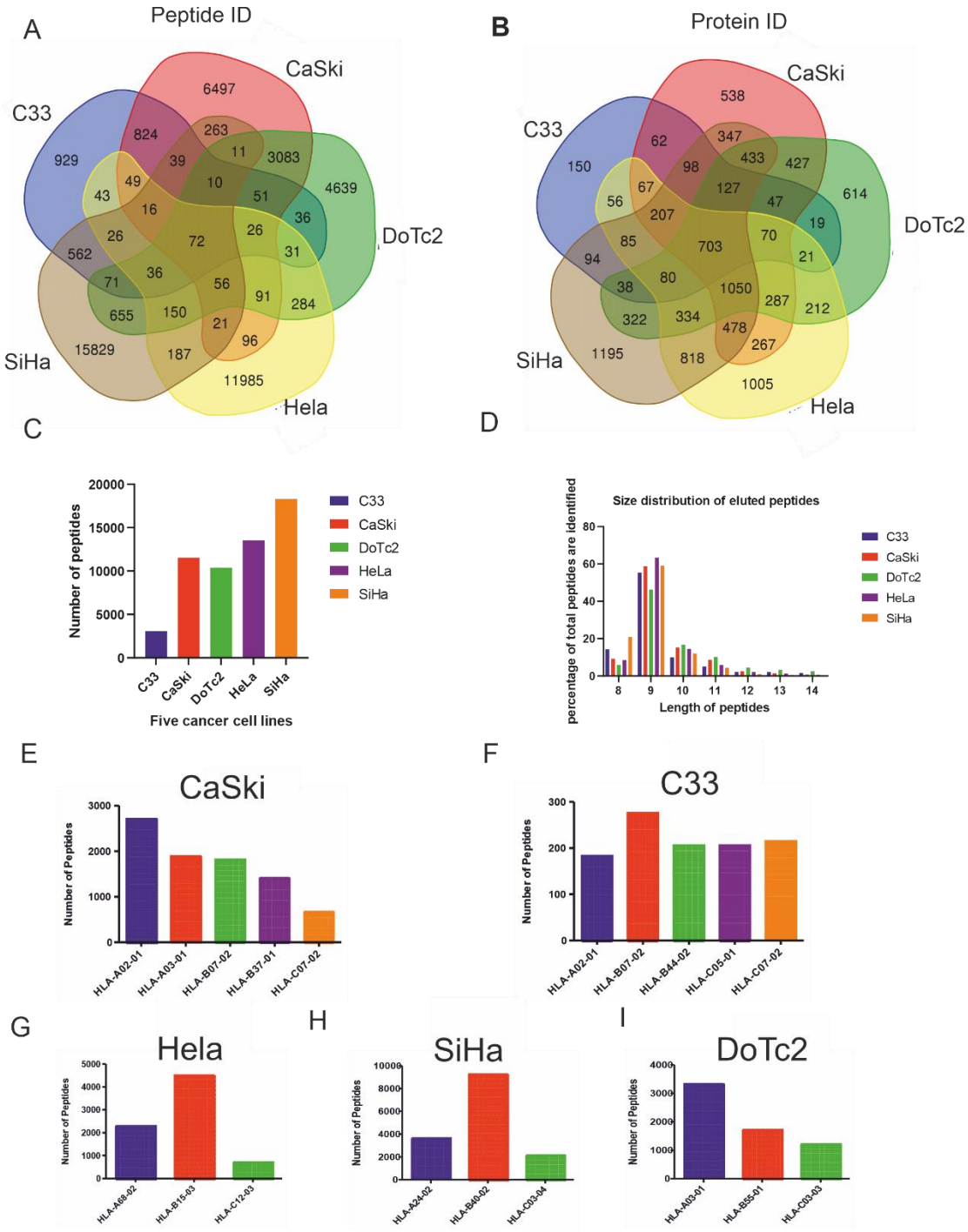
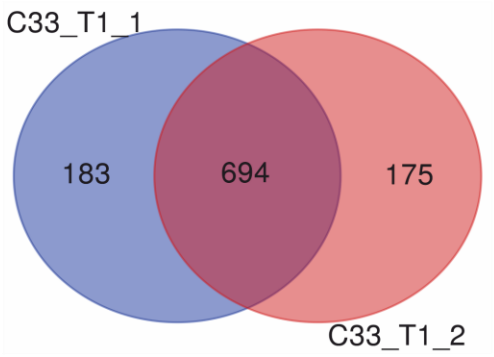
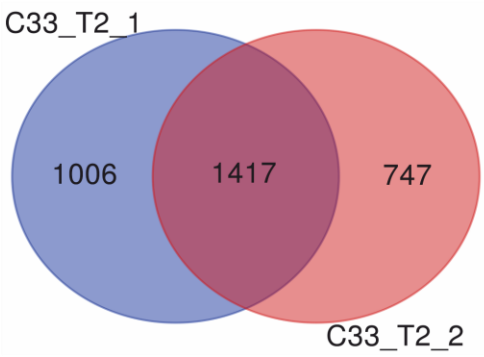


Figure 9. Eluted peptides identified from five human cancer cell lines. A. Overlap of peptides identified in five human cancer cell lines. B. Overlap of

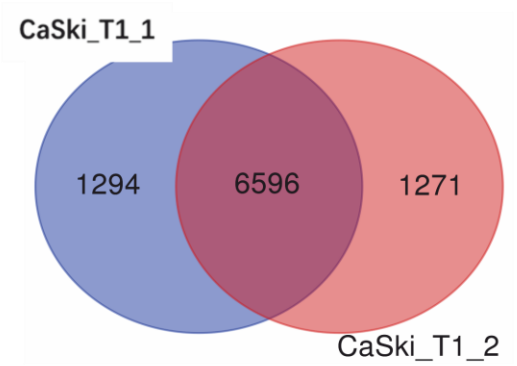
source proteins identified in five human cancer cell lines. C. Total number of peptides identified in five human cancer cell lines. D. Length distribution of 8-14 amino acids peptides in five human cancer cell lines. E-I. Number of eluted peptides predicted to bind to different HLA-I alleles present in each sample

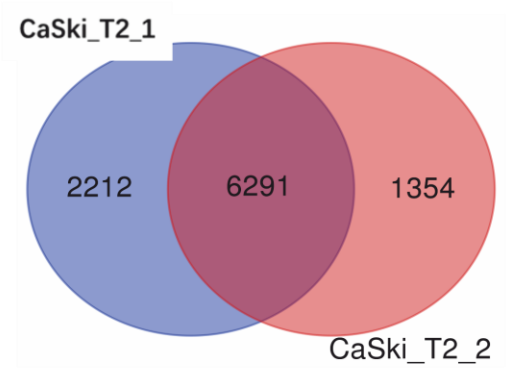
A


C33_T1_1
183
694
175
C33_T1_2

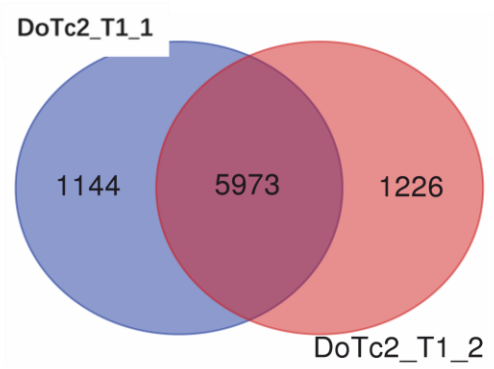

C33_T2_1
1006
1417
747
C33_T2_2

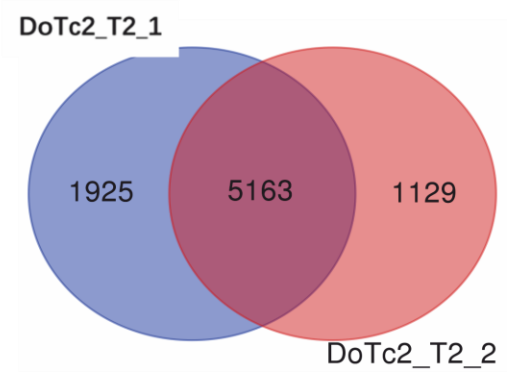
B


CaSki_T1_1
1294
6596
1271
CaSki_T1_2


CaSki_T2_1
2212
6291
1354
CaSki_T2_2

C


DoTc2_T1_1
1144
5973
1226
DoTc2_T1_2


DoTc2_T2_1
1925
5163
1129
DoTc2_T2_2

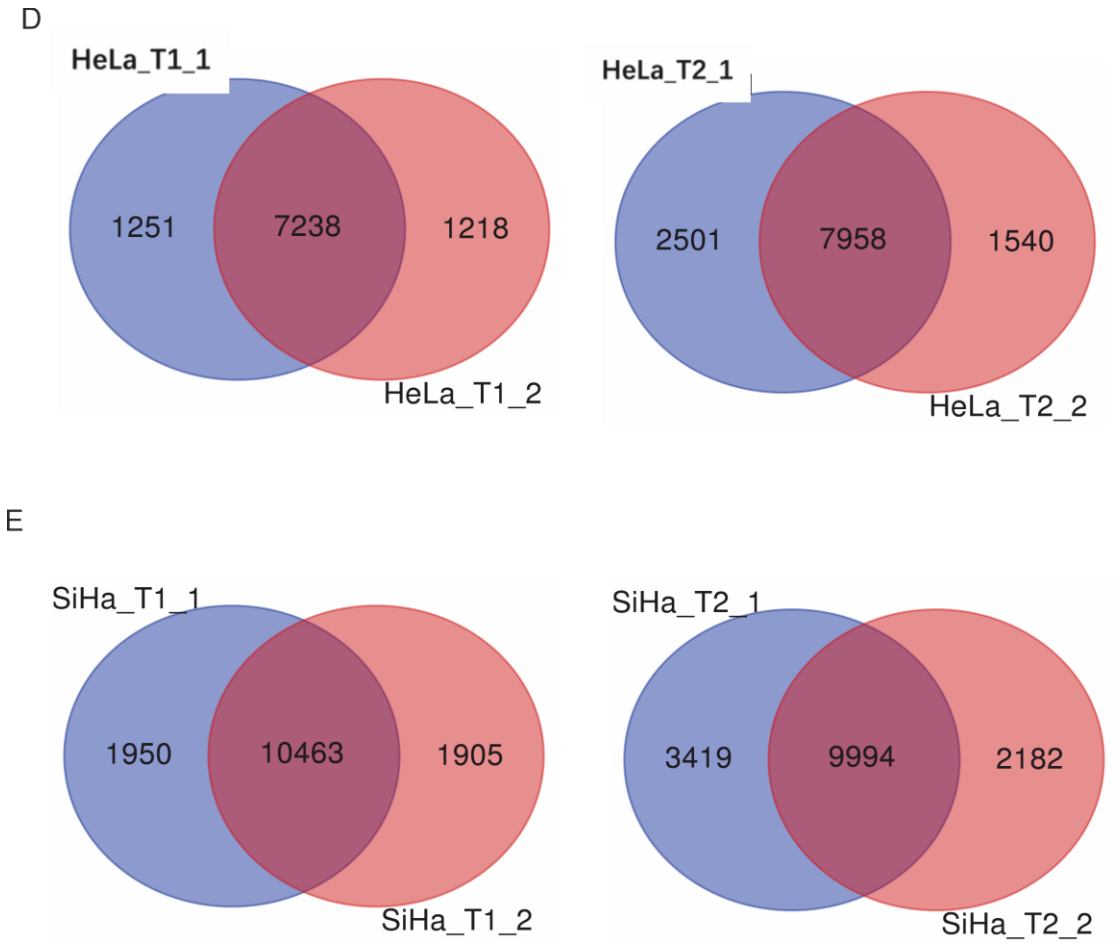


Figure 10. Comparison of eluted peptides identified in technical replicates from five human cervical cancer cell lines. Venn diagramms show the number of sequences identified in each individual replicate analysis.

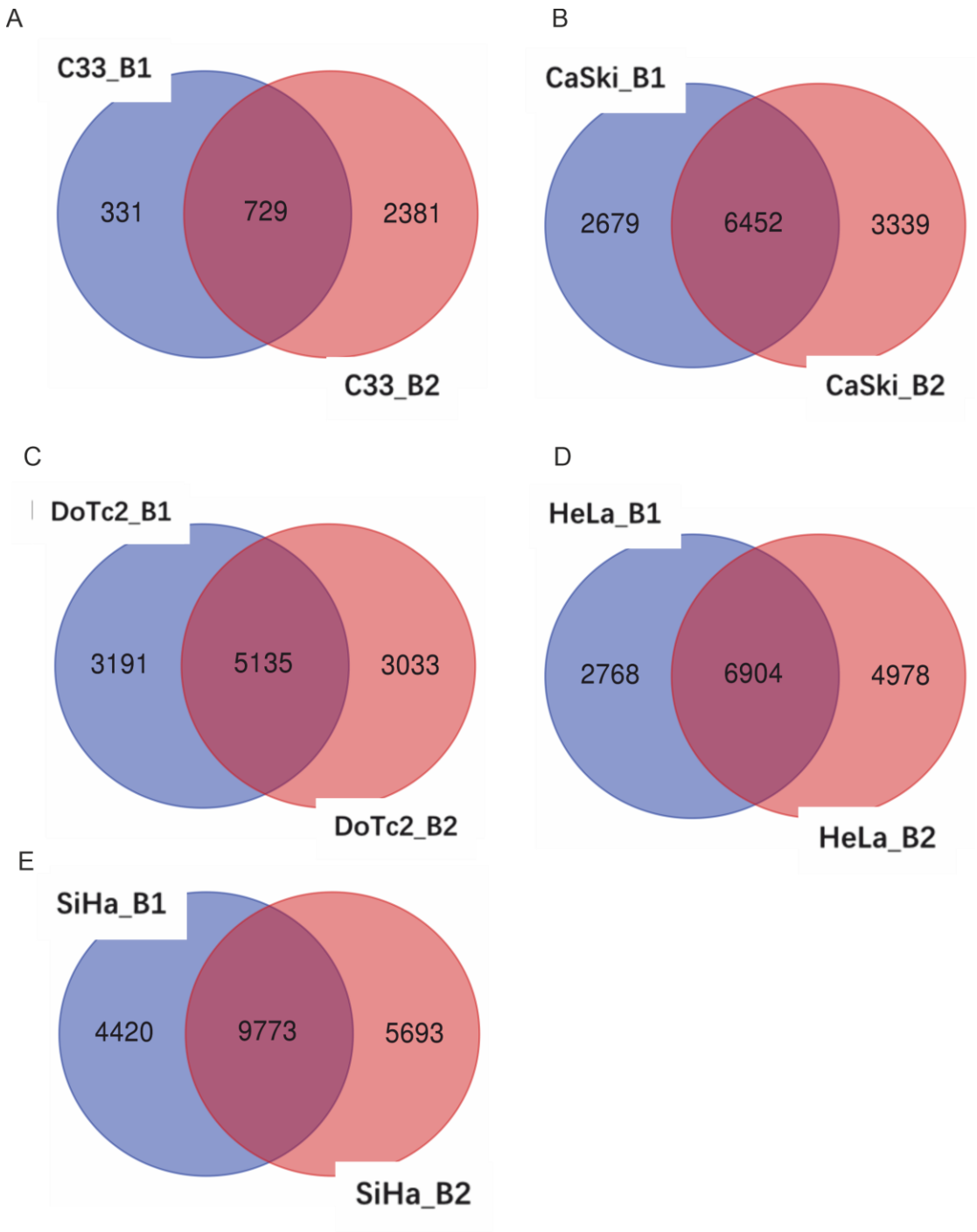
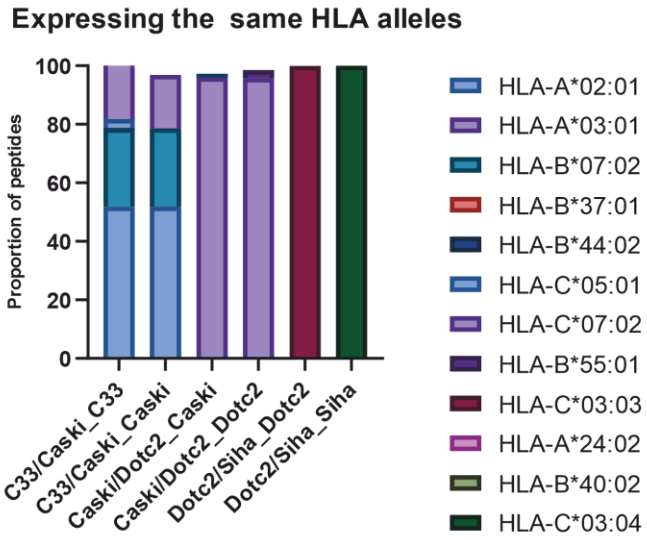


Figure 11. Comparison of eluted peptides identified in two biological replicates from five human cervical cancer cell lines. Venn diagramms show the number of sequences identified in each individual replicate analysis.

A



B

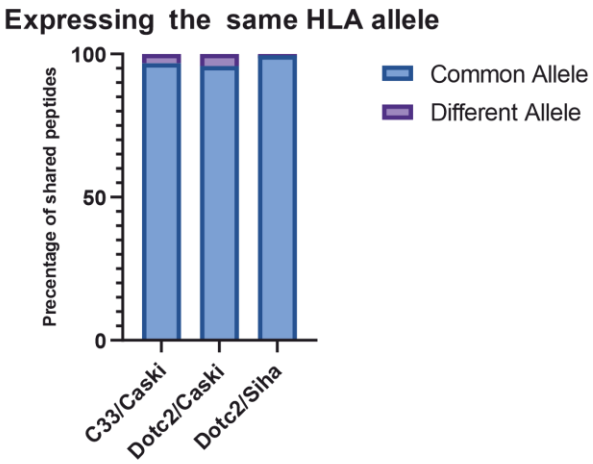


Figure 12. Compare the overlap of sequences for cells expressing the same/different HLA allele. A. Shown are bar graphs of the percentage of shared peptides in each indicated cell line pair according to their predicted allele of origin. Plotted are the allele annotations according the HLA types for the cell line indicated after the underscore B. Overall percentage of shared peptides and proportion originating from a common or a different HLA allele for indicated cell line pair.

3.2.1.4 Gibbs cluster and NetMHCpan successfully categorized the identified peptides to 15 alleles

We also employed advanced neural-network-based algorithms (Gibbs Cluster and NetMHCpan) [232, 233] to predict HLA-binding and to compare our data with known HLA ligand data from the IEDB (http://www.iedb.org) in order to evaluate the quality of our results.

NetMHCpan server predicts binding of peptides to any MHC molecule of known sequence using artificial neural networks (ANNs). The method is trained on a combination of more than 180,000 quantitative binding data (IEDB database) and MS derived MHC-eluted ligands. Gibbs clustering can define motifs directly from the identified peptides using unsupervised sequence clustering. The major difference between those two softwares is that Gibbs Cluster is a server for unsupervised alignment and clustering of peptide sequencee, while NetMHC prediction is based on prior knowledge of the HLA allele present in the sample.

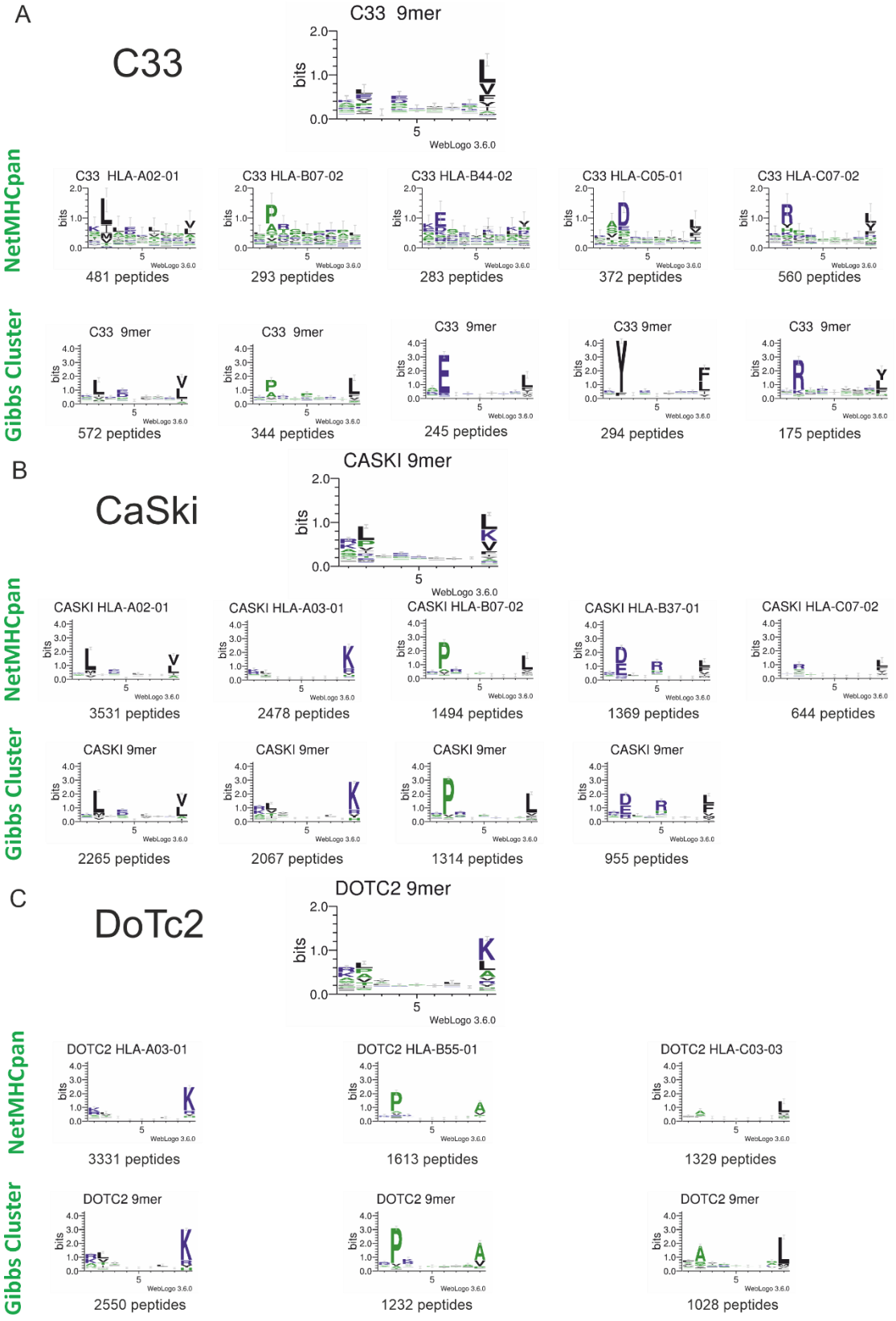
With the help of Gibbs Cluster (9-mer peptides only) and NetMHCpan, we successfully clustered the 32,008 and 47,108 peptides (8 to 14 amino acids) to

the 15 unique HLA-I receptors for each individual dataset and cell line (Fig 13 A-E), respectively.

Using the Gibbs clustering algorithm [232], and direct comparison with the known binding motifs for each HLA allele obtained from IEDB, we could successfully match the motifs obtained from our experiment with the regarding allele of origin. We thereby were able to define each allele's relative contribution to the overall immunopeptidome presentated by those cell lines. We could match peptides to all alleles for those cell lines that expressed three alleles only, i.e., SiHa, and DoTc2 cells (Fig. 13). However, clustering was limited in the cell lines C33, HeLa and CaSki, which express five distinct major HLA alleles. Here, the motifs of HLA-C alleles could not be resolved, probably because HLA-A and HLA-B exhibit higher expression levels than HLA-C, and too few peptide sequences were obtained. In our experiment, fewer than 14% of peptides binding to the HLA-C alleles were found among the total eluted peptides. Human HLA-C is poorly expressed for several reasons, including transcriptional and/or post-transcriptional control by microRNAs [98, 99, 101, 234].

It was notable that we obtained a HLA cluster for C33 cell line that did not match any of the HLA alleles present in these cells as determined by HLA

genotyping, which had characteristics of the HLA-A*24:02 supertype, and that no cluster for HLA-C*05:01 could be determined.



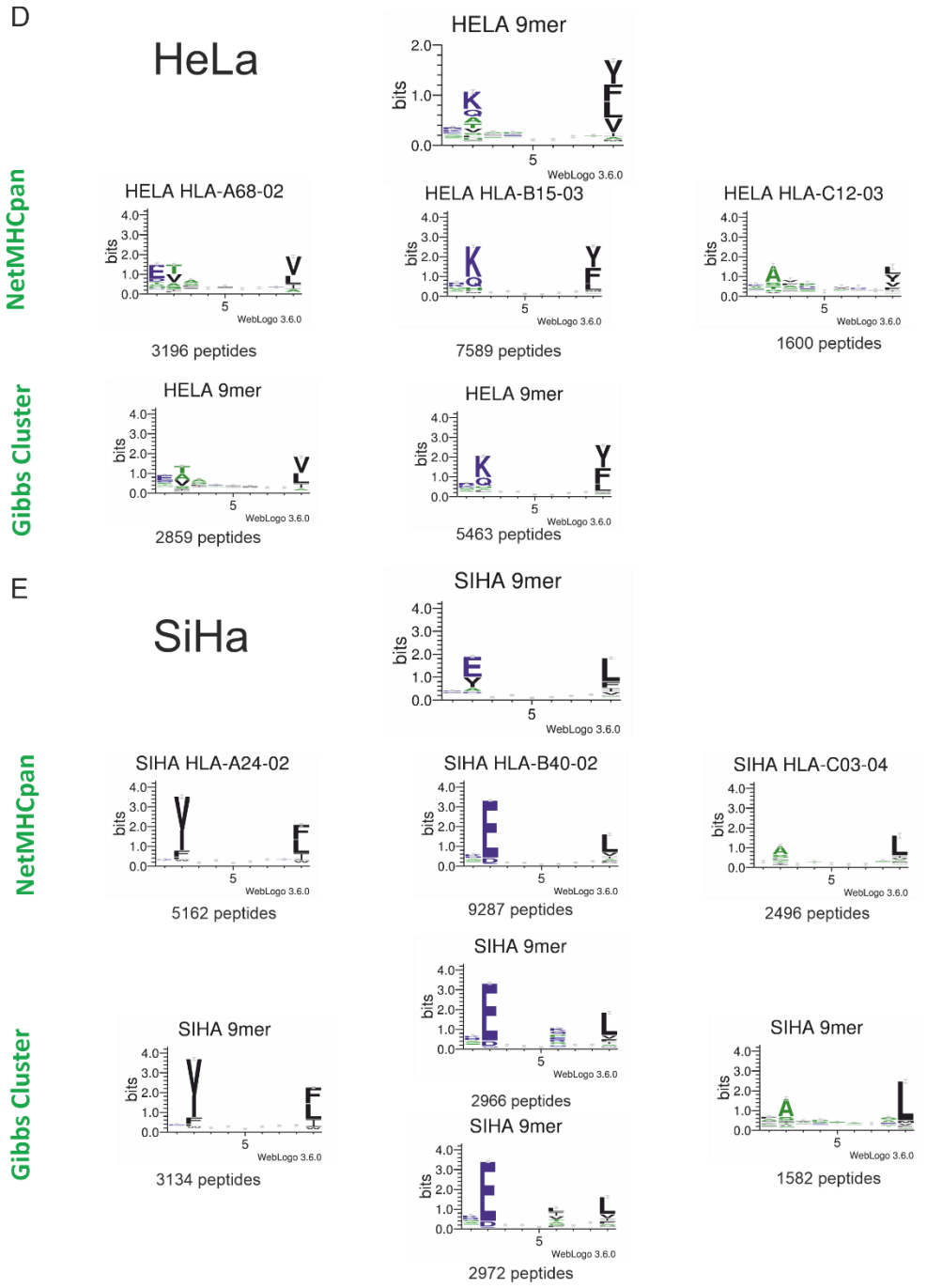


Figure 13. Amino acid distribution of all eluted 9-mer peptides from experiment. NetMHCpan prediction-obtained sequence motifs, GibbsCluster motifs, and IEDB database motifs for all five human cancer lines as indicated.

Gibbs cluster returned motifs were aligned with the respective likely allele of origin as predicted by NetMHCpan.

3.2.2 Several important oncogenic driver genes and HPV master regulators were identified in the five human cervical cancer cell lines

Interestingly, on the protein level, 703 HLA-I-associated proteins were identified to be present in all five cervical cancer cell lines. One protein, G2E3, which is a G2/M-phase specific E3 ubiquitin protein ligase, stood out, as it is defined as a tissue-enriched protein by the Human Proteome Atlas. More importantly, this protein has overall low expression in women in contrast to a rather high abundance in male testis tissue. G2E3 could therefore be a promising target to develop therapeutic approaches to cervical cancer in women.

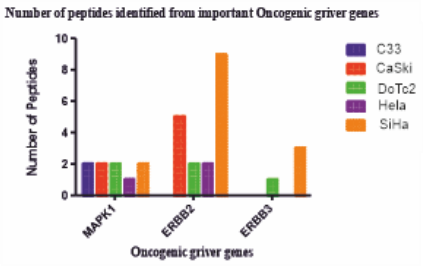
It is well known that HPV causes the majority of cervical tumours and HPV oncoproteins are being targeted by immunotherapies. Recently, clinical trials showed the importance of targeting non-viral antigens in HPV-driven cancer [235]. Researchers also found that mutated peptides from known oncogenic driver genes (KRAS, MAPK1, PIK3CA, ERBB2, and ERBB3) and HPV master regulators (EGR3, FOSB, NR4A2, PRDM1, SOX9, OVOL1, MNT, PA2G4, ENO1, TEAD4, FOXO4, IDO1, and ZNF365) are predicted to be potentially

effective targets in cervical cancer [236-238]. In our experiment, we identified several presentation antigens that came from the oncogenic driver genes MAPK1, ERBB2, and ERBB3 (Fig. 14). In addition, we also found HLA-associated peptides from HPV master regulators in our results, such as from EGR3, FNR4A2, PRDM1, SOX9, OVOL1, MNT, PA2G4, ENO1, and IDO1 (Fig. 14), which contribute to tumour progression. These regulators have been shown to regulate multiple downstream effects of HPV and played crucial roles in relation to mutations and *neo*antigen load [238].

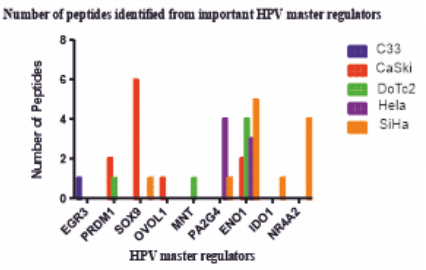
List of several important oncogenic driver genes

List of several important HPV master regulators

A



B



C

Gene	peptides	Antigen presentation
MAPK1	AAGAGPEM	SIHA HLA-C03-04
	KYIHSANVL	SIHA HLA-A24-02
	YRAPEIML	CASKI HLA-C07-02
	RPIFPGKHYL	CASKI HLA-B07-02
	ALPEHSPAK	DOTC2 HLA-A03-01
	EEKIQYLVY	DOTC2 HLA-B55-01
	YRAPEIML	C33 HLA-C07-02
	KYIHSANVL	C33 HLA-C07-02
	SAYDNVNKV	HELA HLA-C12-03
ERBB2	CPSGVKPDL	SIHA HLA-C03-04
	IEHGQAVSL	SIHA HLA-B40-02
	IYGPPQYNI	SIHA HLA-A24-02
	KTFQNTVEL	SIHA HLA-C03-04
	NKLITNDTF	SIHA HLA-B40-02
	RDLAARNVL	SIHA HLA-B40-02
	SLEQVPKEI	SIHA HLA-C03-04
	SYAKHSANMNF	SIHA HLA-A24-02
	TELRKVKVL	SIHA HLA-B40-02
	FLPANFGRLT	CASKI HLA-B07-02
	KIDENQLMY	CASKI HLA-A03-01
	KMLPKTMNR	CASKI HLA-A03-01
	REQLIDYLM	CASKI HLA-B37-01
	SPKANKEIL	CASKI HLA-B07-02
	SLPTHDPSPL	DOTC2 HLA-B55-01
	SPLTSIISA	DOTC2 HLA-B55-01
	SKVDEREKY	HELA HLA-B15-03
	AKHSANMNF	HELA HLA-B15-03
ERBB3	AMGACPASEQG	SIHA HLA-A24-02
	KLEEVELEP	SIHA HLA-B40-02
	VYIEKNDKL	SIHA HLA-A24-02
	HVRQHRGALGP	DOTC2 HLA-B55-01

D

Gene	peptides	Antigen presentation
EGR3	SAGILGVP	C33 HLA-B07-02
NR4A2	KEVVRTDSL	SIHA HLA-B40-02
	NYVATTHMI	SIHA HLA-A24-02
	KEVVRTDSL	SIHA HLA-B40-02
	NYVATTHMI	SIHA HLA-A24-02
PRDM1	EFEEKCTYI	CASKI HLA-C07-02
	KTLPYPLKK	CASKI HLA-A03-01
	KTLPYPLKK	DOTC2 HLA-A03-01
SOX9	HISPNAIFK	CASKI HLA-A03-01
	HPGVPATHGQ	CASKI HLA-C07-02
	KLADQYPHL	CASKI HLA-A02-01
	SPNAIFKAL	CASKI HLA-C07-02
	TPASAGHVW	CASKI HLA-C07-02
	YPHLHNAEL	CASKI HLA-C07-02
	KQQAPPPPP	SIHA HLA-B40-02
OVOL1	SLLQGSPHL	CASKI HLA-A02-01
MNT	VPAPAHLVA	DOTC2 HLA-B55-01
PA2G4	AQFKFTVL	HELA HLA-B15-03
	IAFPTSISV	HELA HLA-C12-03
	IAFPTSIS	HELA HLA-C12-03
	YKRDPSKQY	HELA HLA-B15-03
	KEGEFVAQF	SIHA HLA-B40-02
ENO1	EAMRIGAEV	HELA HLA-A68-02
	EIFDSRGNPTV	HELA HLA-A68-02
	LKIHAREIF	HELA HLA-B15-03
	EGGFAPNIL	SIHA HLA-C03-04
	GAVEKGVPL	SIHA HLA-C03-04
	RYISPDQLADL	SIHA HLA-A24-02
	GAVEKGVPL	DOTC2 HLA-C03-03
	KTRYMGKGVSK	DOTC2 HLA-A03-01
	RIAKAVNEK	DOTC2 HLA-A03-01
	TIAPALVSK	DOTC2 HLA-A03-01
	TIAPALVSK	CASKI HLA-A03-01
	RIAKAVNEK	CASKI HLA-A03-01
	TEQEKIDKL	SIHA HLA-B40-02
	VYHNLKNVI	SIHA HLA-A24-02
IDO1	RYMPPAHRNF	SIHA HLA-A24-02

Figure 14. List of several important oncogenic driver genes and HPV master regulators identified in our experiment. A. Number of peptides identified that belong to several oncogenic driver genes. B. Number of peptides identified that belong to HPV master regulators. C. List oncogenic driver genes. D. List of HPV master regulators.

3.2.3 Several important CT, Raven and protein atlas antigens were identified in the five human cervical cancer cell lines

In order to shortlist known testis antigens that are presented by HLA in the interrogated cell lines, we utilized three different open-source databases: the Cancer-testis antigens database (CT database[239], http://www.cta.lncc.br/) and RAVEN[240] (Rich Analysis of Variable gene Expressions in Numerous tissues; https://github.com/JSGerke/RAVENsoftware), and The Human Protein Atlas[241] (HPA, https://www.proteinatlas.org/)[240].

We identified a total of 227 peptides from 43 proteins from CT database across the cell lines (Fig. 15). Eight proteins [Lymphokine-activated killer T-cell-originated protein kinase (TOPK), Kinesin-like protein KIF2C (KIF2C), Lysine-specific demethylase 5B (KDM5B), Kinetochore scaffold 1 (KNL1), U3 small nucleolar ribonucleoprotein protein Imp3 (Imp3), Kinetochore protein Nuf2 (NUF2), ATPase family AAA domain-containing protein 2 (ATAD2), and Centrosomal protein of 55 kDa (Cep55)] were presented by all five cervical cancer lines (Fig.16). In addition, we identified 169 antigens mapping to the RAVEN database (Fig. 15) (Fig.17), and four antigens were presented by all five lines [Histone H1.2 (HIST1H1C), Ras-related protein Rab-7a (RAB7A), Nuclear pore complex protein Nup107, and Protein OS9], and 4 proteins that were classified as 'tissue enriched', and 'detected in a single' tissue were

shared by 2 or 3 cell lines as defined by the Human Protein Atlas.

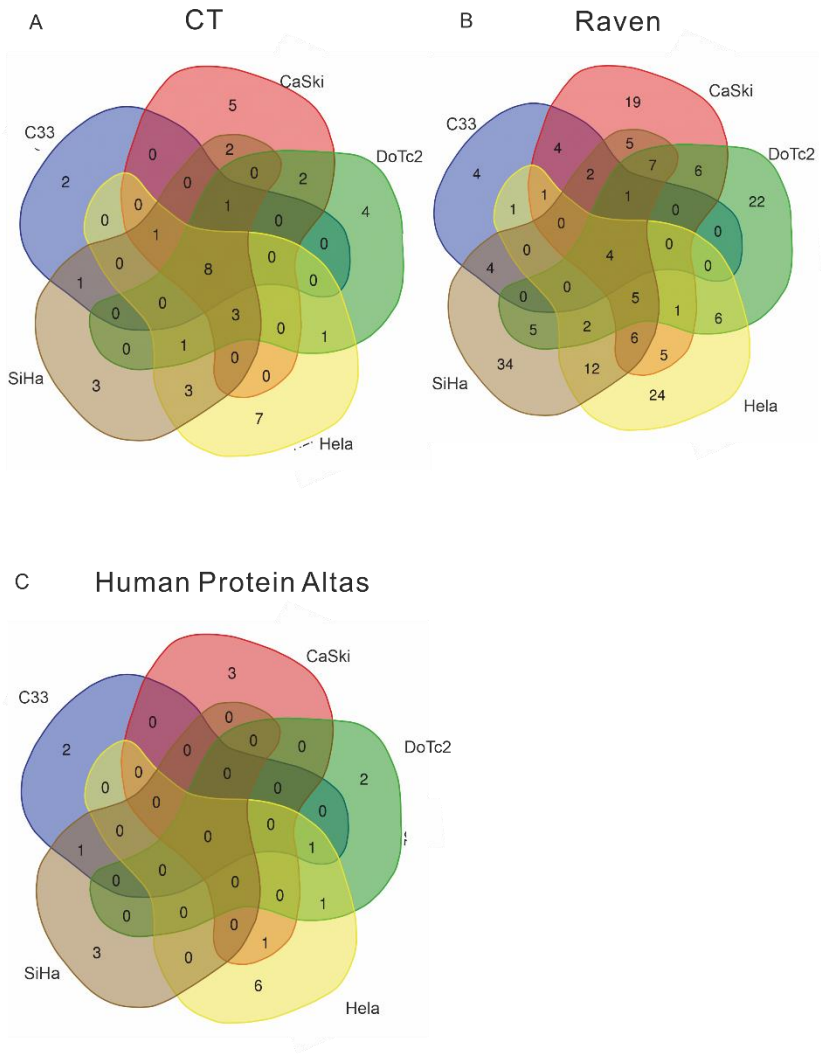
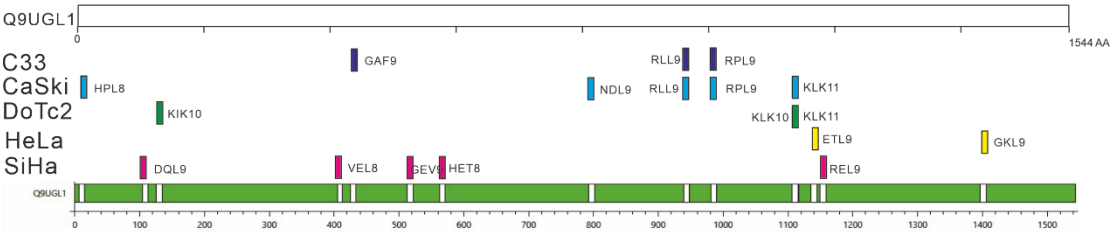


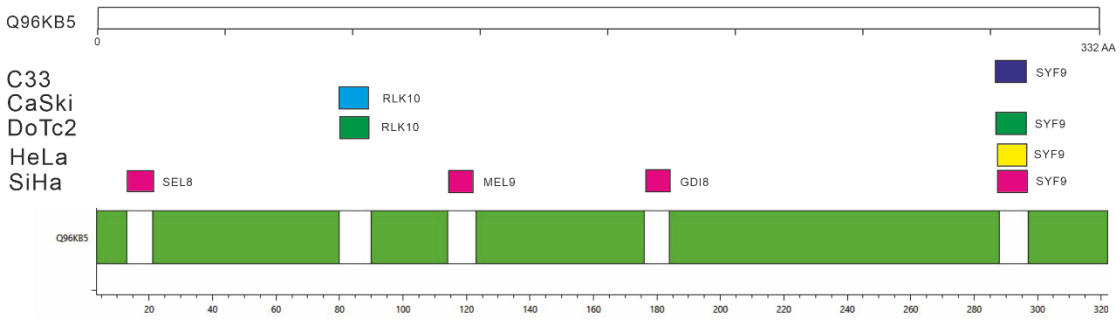
Figure 15. Cancer testis antigens, Raven antigens and Human Protein Atlas. A. Overlap between CT antigens (source proteins) identified in five human cancer cell lines. B. Overlap between RAVEN antigens identified in five human cancer cell lines. C. Overlap of Human Protein Atlas tissue-enriched antigens identified in five human cancer cell lines.

A



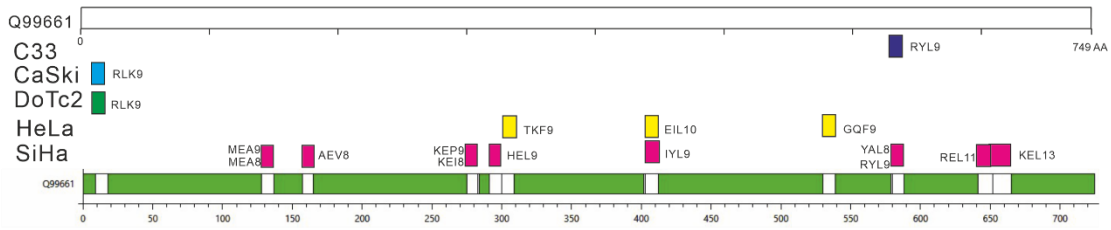
Name	Score	mass	Length	Antigens	Allele	Rank
KDM5B						
C33						
RPRHSLNSL	24.47	1078.6	9	Q9UGL1|KDM5B_HUMAN	HLA-B*07:02	0.0039
GADIASKEF	22.21	936.4552	9	Q9UGL1|KDM5B_HUMAN	HLA-C*05:01	0.2403
RLIDLGVGL	16.99	954.5862	9	Q9UGL1|KDM5B_HUMAN	HLA-A*02:01	0.183
CaSki						
NDLLRHLRL	33.13	1148.678	9	Q9UGL1|KDM5B_HUMAN	HLA-B*37:01	0.0896
KLKEPLPNGKK	29.43	1250.771	11	Q9UGL1|KDM5B_HUMAN	HLA-A*03:01	0.308
RLIDLGVGL	22.09	954.5862	9	Q9UGL1|KDM5B_HUMAN	HLA-A*02:01	0.183
RPRHSLNSL	20.04	1078.6	9	Q9UGL1|KDM5B_HUMAN	HLA-B*07:02	0.0039
HPGPRPAL	17.75	843.4715	8	Q9UGL1|KDM5B_HUMAN	HLA-B*07:02	0.1128
DoTc2						
KLKEPLPNGKK	43.48	1250.771	11	Q9UGL1|KDM5B_HUMAN	HLA-A*03:01	0.308
KILDLFQLNK	35.14	1230.734	10	Q9UGL1|KDM5B_HUMAN	HLA-A*03:01	0.0577
KLKEPLPNGK	29.7	1122.676	10	Q9UGL1|KDM5B_HUMAN	HLA-A*03:01	0.059
HeLa						
ETASAMATL	22.98	893.4164	9	Q9UGL1|KDM5B_HUMAN	HLA-A*68:02	0.0195
GKRDGINSL	23.58	958.5195	9	Q9UGL1|KDM5B_HUMAN	HLA-B*15:03	3.1334
SiHa						
GEPKTWYGV	35.72	1035.503	9	P41229|KDM5C_HUMAN:P29	HLA-B*40:02	0.9092
VEKEFWRL	19.73	1105.592	8	P41229|KDM5C_HUMAN:P29	HLA-B*40:02	1.471
DQIAKYWEL	32.68	1164.582	9	Q9UGL1|KDM5B_HUMAN	HLA-B*40:02	1.9612
HEVPVYRT	28.26	999.5137	8	Q9UGL1|KDM5B_HUMAN	HLA-B*40:02	4.379
REMEALQSL	22.13	1075.533	9	Q9UGL1|KDM5B_HUMAN	HLA-B*40:02	0.0063

B



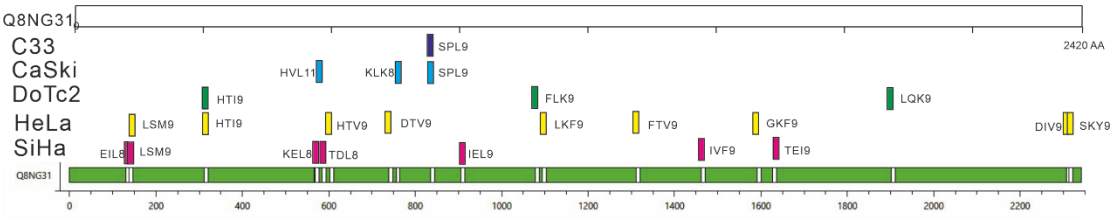
Name	Score	mass	Length	Antigens	Allele	Rank
Q96KB5						
C33						
SYQKVIELF	28.32	1125.607	9	Q96KB5|TOPK_HUMAN	HLA-C*07:02	0.1941
CaSki						
RLMDEAKILK	36.4	1215.701	10	Q96KB5|TOPK_HUMAN	HLA-A*03:01	0.0451
DoTc2						
RLMDEAKILK	32.05	1215.701	10	Q96KB5|TOPK_HUMAN	HLA-A*03:01	0.0451
SYQKVIELF	20.09	1125.607	9	Q96KB5|TOPK_HUMAN	HLA-C*03:03	16.8517
HeLa						
SYQKVIELF	21.34	1125.607	9	Q96KB5|TOPK_HUMAN	HLA-C*12:03	6.4656
SiHa						
SYQKVIELF	40.52	1125.607	9	Q96KB5|TOPK_HUMAN	HLA-A*24:02	0.0063
SEKKKSVL	39.71	917.5546	8	Q96KB5|TOPK_HUMAN	HLA-B*40:02	4.0335
MEYGGEKSL	36.26	1012.454	9	Q96KB5|TOPK_HUMAN	HLA-B*40:02	0.121
GDFETIKI	28.96	921.4807	8	Q96KB5|TOPK_HUMAN	HLA-B*40:02	3.8679

C



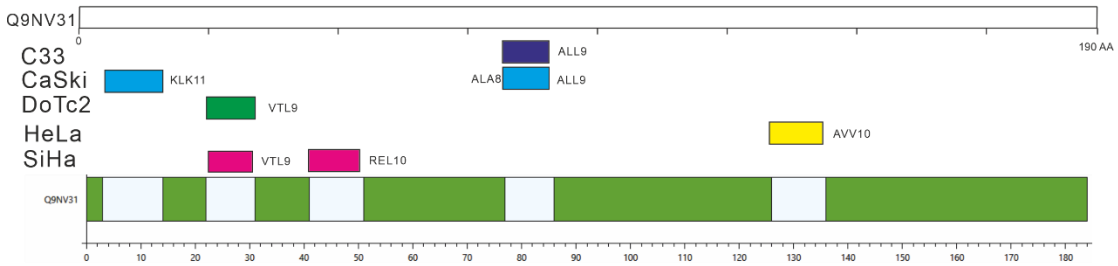
Name	Score	mass	Length	Antigens	Allele	Rank
Q99661						
C33						
RYADRVKEL	27.3	1148.63	9	Q5T7B8|KIF24_HUMAN:Q996(	HLA-C*07:02	0.0212
CaSki						
RLFPGLAIK	18.15	1013.639	9	Q99661|KIF2C_HUMAN	HLA-A*03:01	0.0099
DoTc2						
RLFPGLAIK	23.68	1013.639	9	Q99661|KIF2C_HUMAN	HLA-A*03:01	0.0099
HeLa						
EIYNGKLFDL	43.61	1210.623	10	Q99661|KIF2C_HUMAN	HLA-A*68:02	1.964
TKYLENQAF	41.04	1112.55	9	Q99661|KIF2C_HUMAN	HLA-B*15:03	0.0995
GQNKAHTPF	40.6	998.4933	9	Q99661|KIF2C_HUMAN	HLA-B*15:03	0.0092
SiHa						
IYNGKLFDL	43.08	1081.581	9	Q99661|KIF2C_HUMAN	HLA-A*24:02	0.2847
KEIIQQGPDWLEL	40.72	1567.825	13	Q99661|KIF2C_HUMAN	HLA-B*40:02	0.6905
RELEEKAMEEL	31.43	1375.665	11	Q99661|KIF2C_HUMAN	HLA-B*40:02	0.4511
AEIPLRMV	30.56	927.5211	8	Q99661|KIF2C_HUMAN	HLA-B*40:02	0.8209
HEPKLKVDL	26.68	1077.618	9	Q99661|KIF2C_HUMAN	HLA-B*40:02	1.3547
KEIDVISI	26.08	915.5276	8	Q99661|KIF2C_HUMAN	HLA-B*40:02	0.1521
KEIDVISIP	24.68	1012.58	9	Q99661|KIF2C_HUMAN	HLA-B*40:02	0.837
MEVELPAAA	22.77	929.4528	9	Q99661|KIF2C_HUMAN	HLA-B*40:02	0.3365
MEVELPAA	19.33	858.4156	8	Q99661|KIF2C_HUMAN	HLA-B*40:02	1.0554
YADRVKEL	37.6	992.5291	8	Q99661|KIF2C_HUMAN:Q5T7(	HLA-C*03:04	2.6601
RYADRVKEL	36.6	1148.63	9	Q99661|KIF2C_HUMAN:Q5T7(	HLA-A*24:02	0.5839

D



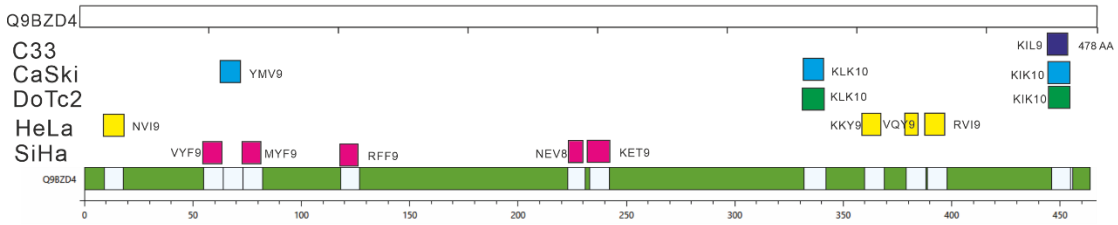
Name	Score	mass	Length	Antigens	Allele	Rank
Q8NG31						
C33						
SPIEKSGVL	20.56	928.5229	9	Q8NG31|KNL1_HUMAN	HLA-B*07:02	0.0367
CaSki						
SPIEKSGVL	31.33	928.5229	9	Q8NG31|KNL1_HUMAN	HLA-B*07:02	0.0367
KLAEPLRK	24.59	953.6022	8	Q8NG31|KNL1_HUMAN	HLA-A*03:01	0.7981
HVSKERIQQSL	19.84	1323.726	11	Q8NG31|KNL1_HUMAN	HLA-B*07:02	9.3136
DoTc2						
FLNEPLSSK	27.08	1033.544	9	Q8NG31|KNL1_HUMAN	HLA-A*03:01	0.2858
HTANIQTLI	18.06	1009.556	9	Q8NG31|KNL1_HUMAN	HLA-C*03:03	1.149
LQVHILIQK	16.34	1090.686	9	Q8NG31|KNL1_HUMAN	HLA-A*03:01	1.1005
HeLa						
LSAPIHTQM	41.55	996.5062	9	Q8NG31|KNL1_HUMAN	HLA-C*12:03	0.0505
HTANIQTLI	37.74	1009.556	9	Q8NG31|KNL1_HUMAN	HLA-A*68:02	0.0329
DIATILSKV	35.45	958.5699	9	Q8NG31|KNL1_HUMAN	HLA-A*68:02	0.088
HTSNLGSQV	32.84	941.4567	9	Q8NG31|KNL1_HUMAN	HLA-A*68:02	0.2427
SKVPLENNY	32.78	1062.535	9	Q8NG31|KNL1_HUMAN	HLA-B*15:03	1.6015
FTSSHATAV	29.13	919.4399	9	Q8NG31|KNL1_HUMAN	HLA-C*12:03	0.0293
LKNDKTIVF	24.07	1076.623	9	Q8NG31|KNL1_HUMAN	HLA-B*15:03	0.0639
DTAISSHTV	22.15	929.4454	9	Q8NG31|KNL1_HUMAN	HLA-A*68:02	0.0603
GKTGEFLAF	16.82	968.4967	9	Q8NG31|KNL1_HUMAN	HLA-B*15:03	0.2988
SiHa						
IYVIPQPHF	33.83	1112.602	9	Q8NG31|KNL1_HUMAN	HLA-A*24:02	0.1552
TEIHNINII	32.04	1065.582	9	Q8NG31|KNL1_HUMAN	HLA-B*40:02	0.2988
LSAPIHTQM	30.85	996.5062	9	Q8NG31|KNL1_HUMAN	HLA-C*03:04	0.0531
TDRKTELL	29.38	974.5396	8	Q8NG31|KNL1_HUMAN	HLA-B*40:02	14.8287
KERIQQSL	26.9	1000.567	8	Q8NG31|KNL1_HUMAN	HLA-B*40:02	1.0165
IEINHRPLL	24.36	1103.645	9	Q8NG31|KNL1_HUMAN	HLA-B*40:02	0.1356
EITGMNTL	18.68	877.4215	8	Q8NG31|KNL1_HUMAN	HLA-B*40:02	1.7354

E



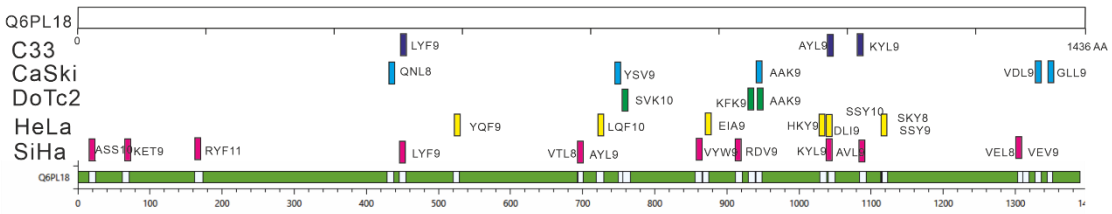
Name	Score	mass	Length	Antigens	Allele	Rank
Q9NV31						
C33						
ALLDKLYAL	28.57	1018.606	9	Q9NV31|IMP3_HUMAN	HLA-A*02:01	0.0086
CaSki						
KLKFHEQKLLK	35.34	1410.871	11	Q9NV31|IMP3_HUMAN	HLA-A*03:01	0.3908
ALLDKLYAL	35.15	1018.606	9	Q9NV31|IMP3_HUMAN	HLA-A*02:01	1.2795
ALLDKLYA	15.12	905.5222	8	Q9NV31|IMP3_HUMAN	HLA-A*02:01	2.0249
DoTc2						
VTDHNLHEL	41.4	1076.525	9	Q9NV31|IMP3_HUMAN	HLA-C*03:03	0.9872
HeLa						
AVAFVEQGHV	19.34	1055.54	10	Q9NV31|IMP3_HUMAN	HLA-A*68:02	0.8838
SiHa						
VTDHNLHEL	44.18	1076.525	9	Q9NV31|IMP3_HUMAN	HLA-C*03:04	0.9872
REDYTRYNQL	41.63	1356.642	10	Q9NV31|IMP3_HUMAN	HLA-B*40:02	0.2736

F



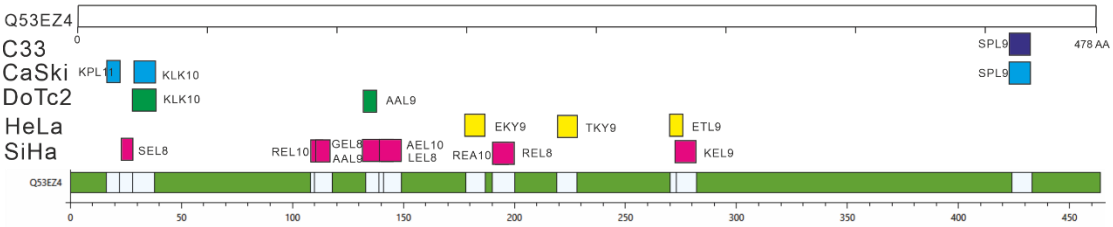
Name	Score	mass	Length	Antigens	Allele	Rank
Q9BZD4						
C33						
KIDEKTAEL	33.44	1045.566	9	Q9BZD4|NUF2_HUMAN	HLA-C*05:01	0.0172
CaSki						
KIDEKTAELK	38.92	1173.66	10	Q9BZD4|NUF2_HUMAN	HLA-A*03:01	0.611
YMMPVNSEV	36.76	1068.462	9	Q9BZD4|NUF2_HUMAN	HLA-A*02:01	0.0072
KLKTEENSFK	34.5	1222.656	10	Q9BZD4|NUF2_HUMAN	HLA-A*03:01	0.2032
DoTc2						
KIDEKTAELK	38.04	1173.66	10	Q9BZD4|NUF2_HUMAN	HLA-A*03:01	0.611
KLKTEENSFK	36.04	1222.656	10	Q9BZD4|NUF2_HUMAN	HLA-A*03:01	0.2032
HeLa						
KKHEDVKQY	47.28	1173.614	9	Q9BZD4|NUF2_HUMAN	HLA-B*15:03	0.7212
NVAEIVIHI	33.08	1006.581	9	Q9BZD4|NUF2_HUMAN	HLA-A*68:02	0.0191
VQEKRGAVY	21.19	1048.567	9	Q9BZD4|NUF2_HUMAN	HLA-B*15:03	0.1031
RVTTINQEI	18.18	1072.588	9	Q9BZD4|NUF2_HUMAN	HLA-C*12:03	1.7653
SiHa						
VYGIRLEHF	42.39	1132.603	9	Q9BZD4|NUF2_HUMAN	HLA-A*24:02	0.0395
MYPHLMEGF	39.27	1123.483	9	Q9BZD4|NUF2_HUMAN	HLA-A*24:02	0.0465
RFLSGIINF	31.1	1065.597	9	Q9BZD4|NUF2_HUMAN	HLA-A*24:02	0.0739
NELKLSVV	23.18	900.528	8	Q9BZD4|NUF2_HUMAN	HLA-B*40:02	1.3562
KEIQESLKT	22.37	1074.592	9	Q9BZD4|NUF2_HUMAN	HLA-B*40:02	1.3251

G



Name	Score	mass	Length	Antigens	Allele	Rank
Q6PL18						
C33						
AYAIIKEEL	32.15	1048.58	9	Q6PL18|ATAD2_HUMAN	HLA-C*07:02	0.5901
KYLTVKDYL	30.6	1141.638	9	Q6PL18|ATAD2_HUMAN	HLA-C*07:02	0.5948
LYPEVFEKF	29.64	1170.596	9	Q6PL18|ATAD2_HUMAN	HLA-C*07:02	0.2445
CaSki						
YSDDDVPSV	37.71	995.4083	9	Q6PL18|ATAD2_HUMAN	HLA-A*02:01	0.6238
AAKPPISKK	36.22	938.5912	9	Q6PL18|ATAD2_HUMAN	HLA-A*03:01	0.6912
GLSNHIAAL	29.67	894.4923	9	Q6PL18|ATAD2_HUMAN	HLA-A*02:01	0.4984
VDHERLKNL	26.7	1122.615	9	Q6PL18|ATAD2_HUMAN	HLA-B*37:01	0.5369
QNYNIFQL	17.2	1038.513	8	Q6PL18|ATAD2_HUMAN	HLA-B*37:01	2.9505
DoTc2						
AAKPPISKK	37.05	938.5912	9	Q6PL18|ATAD2_HUMAN	HLA-A*03:01	0.6912
SVYENGLSQK	21.06	1123.551	10	Q6PL18|ATAD2_HUMAN	HLA-A*03:01	0.0342
KFFEDLILK	15.67	1151.659	9	Q6PL18|ATAD2_HUMAN	HLA-A*03:01	1.3651
HeLa						
HKYLTVKDY	44.62	1165.613	9	Q6PL18|ATAD2_HUMAN	HLA-B*15:03	0.3834
SSSKYAPSYY	44.41	1151.513	10	Q6PL18|ATAD2_HUMAN	HLA-C*12:03	3.3861
LQRVFPHAEF	43.85	1242.651	10	Q6PL18|ATAD2_HUMAN	HLA-B*15:03	0.0347
SSKYAPSYY	38.36	1064.481	9	Q6PL18|ATAD2_HUMAN	HLA-C*12:03	0.6434
SKYAPSYY	34.12	977.4494	8	Q6PL18|ATAD2_HUMAN	HLA-B*15:03	0.9645
DLSSVISKI	32.12	960.5491	9	Q6PL18|ATAD2_HUMAN	HLA-A*68:02	1.9271
YQMRPSIIF	29.13	1153.595	9	Q6PL18|ATAD2_HUMAN	HLA-B*15:03	0.0055
EIVGPTLKA	25.29	926.5436	9	Q6PL18|ATAD2_HUMAN	HLA-A*68:02	0.7939
SiHa						
KYLTVKDYL	46.92	1141.638	9	Q6PL18|ATAD2_HUMAN	HLA-A*24:02	0.3894
VYVPHIHVW	44.98	1148.613	9	Q6PL18|ATAD2_HUMAN	HLA-A*24:02	0.0234
AVTSPGQAL	42.05	842.4498	9	Q6PL18|ATAD2_HUMAN	HLA-C*03:04	0.3999
LYPEVFEKF	39.22	1170.596	9	Q6PL18|ATAD2_HUMAN	HLA-A*24:02	0.0082
AYAIIKEEL	37.33	1048.58	9	Q6PL18|ATAD2_HUMAN	HLA-A*24:02	0.8492
RYSGVNQSMLF	37.17	1300.623	11	Q6PL18|ATAD2_HUMAN	HLA-A*24:02	0.0362
VTSPGQAL	31.95	771.4127	8	Q6PL18|ATAD2_HUMAN	HLA-C*03:04	2.7725
RDYGEIFNV	31.23	1111.53	9	Q6PL18|ATAD2_HUMAN	HLA-B*40:02	1.0929
VEKALAIL	23.19	855.5429	8	Q6PL18|ATAD2_HUMAN	HLA-B*40:02	0.8431
VEQQQLITV	22.31	1056.582	9	Q6PL18|ATAD2_HUMAN	HLA-B*40:02	0.2087
KEVETYHRT	18.26	1161.578	9	Q6PL18|ATAD2_HUMAN	HLA-B*40:02	1.2253
ASATGSLDLS	15.4	920.4451	10	Q6PL18|ATAD2_HUMAN	HLA-C*03:04	33.123

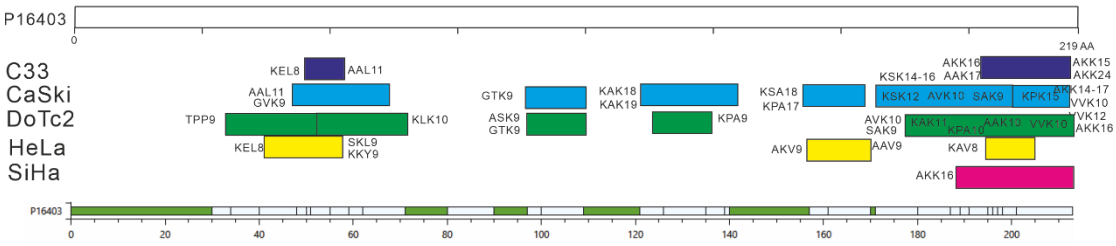
H



Name	Score	mass	Length	Antigens	Allele	Rank
Q53EZ4						
C33						
SPKSPTAAL	25.18	870.4811	9	Q53EZ4|CEP55_HUMAN	HLA-B*07:02	0.0064
CaSki						
SPKSPTAAL	43	870.4811	9	Q53EZ4|CEP55_HUMAN	HLA-B*07:02	0.0064
KPSNSKSETTL	37.7	1190.614	11	Q53EZ4|CEP55_HUMAN	HLA-B*07:02	0.33
KLKGEIAHLK	30.2	1135.708	10	Q53EZ4|CEP55_HUMAN	HLA-A*03:01	0.1092
DoTc2						
AATSRIAEL	36.91	930.5134	9	Q53EZ4|CEP55_HUMAN	HLA-C*03:03	0.162
KLKGEIAHLK	32.7	1135.708	10	Q53EZ4|CEP55_HUMAN	HLA-A*03:01	0.1092
HeLa						
ETQKEVHNL	42.37	1096.551	9	Q53EZ4|CEP55_HUMAN	HLA-A*68:02	1.3073
TKKPESEGY	36.45	1037.503	9	Q53EZ4|CEP55_HUMAN	HLA-B*15:03	4.4052
EKNQQWLVY	15.85	1206.603	9	Q53EZ4|CEP55_HUMAN	HLA-B*15:03	3.038
SiHa						
KEVHNLNQL	38.64	1093.588	9	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	0.112
AELESKTNTL	37.99	1104.566	10	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	0.1766
AATSRIAEL	35.95	930.5134	9	Q53EZ4|CEP55_HUMAN	HLA-C*03:04	0.162
REGERREQVL	32.41	1270.674	10	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	1.7572
SETTLEKL	24.78	919.4862	8	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	1.5848
REVYVKGL	23.99	962.5549	8	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	0.6029
GERREQVL	22.34	985.5305	8	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	2.9463
LESKTNTL	20.74	904.4866	8	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	1.3472
REVYVKGLLA	19.6	1146.676	10	Q53EZ4|CEP55_HUMAN	HLA-B*40:02	0.5207

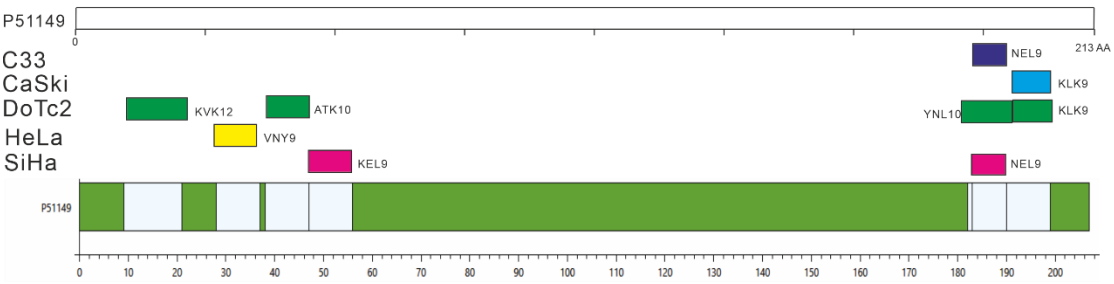
Figure 16. List of eight shared protein antigens defined by the CT database that were presented by all five cervical tumour cell lines. A-H**.** Relative position of peptides identified mapping to the parent protein coding sequence. The horizontal line and green horizontal box indicate the nucleotide position in the respective gene. White interruptions in the green box indicate the HLA-peptide coverage. Small boxes in different colors show the coding sequence for the CT-associated HLA-peptide sequence that were identified in the different cell lines.

A



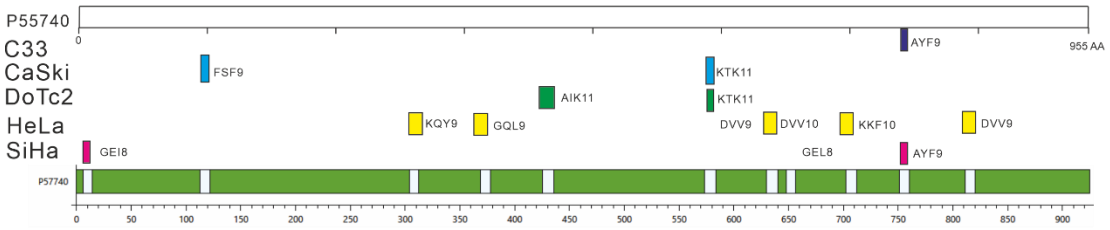
Name	Score	Mass	Length		Allele	Rank
P16403|H12_HUMAN:HIST1H1C						
C33						
AKPKVVKPKKAAPKKK	48.02	1745.177	16	P16403|H12_HUMAN	HLA-C*07	89.4812
AAKPKVVKPKKAAPKKK	36.48	1816.214	17	P16403|H12_HUMAN	HLA-C*07	77.2364
AKPKVVKPKKAAPKK	27.07	1617.082	15	P16403|H12_HUMAN	HLA-B*07	49.7117
AKAVKPKAAKPKVVKPKKAAPKKK	19.86	2538.694	24	P16403|H12_HUMAN		
AASKERSGVSL	46.3	1103.594	11	P16403|H12_HUMAN:P16402|H13_H	HLA-B*07	6.0227
KERSGVSL	29.03	874.4872	8	P16403|H12_HUMAN:P16402|H13_H	HLA-B*44	5.9517
CaSki						
AKPKVVKPKKAAPKKK	54.52	1745.177	16	P16403|H12_HUMAN		
KAGGTKPKKPVGAAKKPK	54.36	1790.125	18	P16403|H12_HUMAN		
KPKKAAKSAAKAVKPKA	50.94	1721.104	17	P16403|H12_HUMAN		
KSPKKAKVAKPKKAAK	47.71	1707.125	16	P16403|H12_HUMAN		
AAKPKVVKPKKAAPKKK	47.36	1816.214	17	P16403|H12_HUMAN		
KAGGTKPKKPVGAAKKPKK	46.07	1918.22	19	P16403|H12_HUMAN		
KSPKKAKVAKPKKAAKSA	45.93	1865.194	18	P16403|H12_HUMAN		
AKPKVVKPKKAAPK	44.82	1488.987	14	P16403|H12_HUMAN	HLA-A*03	4.9568
AKPKVVKPKKAAPKK	44.17	1617.082	15	P16403|H12_HUMAN		
KSPKKAKVAKPKKAA	41.98	1579.03	15	P16403|H12_HUMAN		
SAAKAVKPK	40.89	898.5599	9	P16403|H12_HUMAN	HLA-A*03	1.5436
VVKPKKAAPK	39.34	1064.707	10	P16403|H12_HUMAN	HLA-A*03	0.8116
KSPKKAKVAKPKKA	33.96	1507.993	14	P16403|H12_HUMAN	HLA-B*07	9.2818
KSPKKAKVAKPK	32.27	1308.861	12	P16403|H12_HUMAN	HLA-A*03	3.1304
VVKPKKAAPKKK	28.04	1320.897	12	P16403|H12_HUMAN	HLA-A*03	6.0763
KPKVVKPKKAAPKKK	23.98	1674.14	15	P16403|H12_HUMAN		
AVKPKAAKPK	43.62	1036.676	10	P16403|H12_HUMAN:P10412|H14_H	HLA-A*03	0.6597
AASKERSGVSL	48.5	1103.594	11	P16403|H12_HUMAN:P10412|H14_H	HLA-B*07	6.0227
GTGASGSFK	35.48	810.3871	9	P16403|H12_HUMAN:P10412|H14_H	HLA-A*03	0.2853
GVSLAALKK	22.41	885.5647	9	P16403|H12_HUMAN:P10412|H14_H	HLA-A*03	0.5109
DoTc2						
AVKPKAAKPK	43.54	1036.676	10	P16401|H15_HUMAN:P10412|H14_H	HLA-A*03	0.6597
KLGLKSLVSK	37.36	1071.702	10	P16401|H15_HUMAN:P10412|H14_H	HLA-A*03	0.0598
ASGSFKLNK	38.86	950.5185	9	P16401|H15_HUMAN:P10412|H14_H	HLA-A*03	0.4289
GTGASGSFK	38.6	810.3871	9	P16401|H15_HUMAN:P10412|H14_H	HLA-A*03	0.2853
SAAKAVKPK	46.72	898.5599	9	P16403|H12_HUMAN	HLA-A*03	1.5436
AKPKVVKPKKAAPKKK	46.42	1745.177	16	P16403|H12_HUMAN	HLA-A*03	28.2652
KPKKPVGAA	44.35	894.5651	9	P16403|H12_HUMAN	HLA-B*55	0.0216
VVKPKKAAPK	40.67	1064.707	10	P16403|H12_HUMAN	HLA-A*03	0.8116
AAKPKVVKPK	40.56	1064.707	10	P16403|H12_HUMAN	HLA-A*03	1.3972
KAAKPKVVKPK	26.24	1192.802	11	P16403|H12_HUMAN	HLA-A*03	2.2056
TPRKASGPP	18.11	909.5032	9	P16403|H12_HUMAN	HLA-B*55	0.9094
KPKVVKPKKA	17.22	1121.765	10	P16403|H12_HUMAN	HLA-B*55	3.2435
HeLa						
AAATVTKKV	38.78	887.544	9	P16403|H12_HUMAN	HLA-C*12	1.6447
KAAKPKVV	35.4	839.5592	8	P16403|H12_HUMAN	HLA-C*12	16.4383
AKKPAAATV	28.54	855.5178	9	P16403|H12_HUMAN	HLA-B*15	4.125
SKERSGVSL	32.76	961.5192	9	P16403|H12_HUMAN:P10412|H14_H	HLA-B*15	4.9342
KERSGVSL	19	874.4872	8	P16403|H12_HUMAN:P10412|H14_H	HLA-B*15	6.6686
KKALAAAGY	41.55	891.5178	9	P16403|H12_HUMAN:P10412|H14_H	HLA-B*15	0.2038
SiHa						
AKPKVVKPKKAAPKKK	50.05	1745.177	16	P16403|H12_HUMAN	HLA-A*24	90.9268
AKPKVVKPKKAAPK	47.15	1488.987	14	P16403|H12_HUMAN	HLA-A*24	78.9759
AAKPKVVKPKKAAPKKK	45.63	1816.214	17	P16403|H12_HUMAN	HLA-A*24	87.4009
AKPKVVKPKKAAPKK	42.36	1617.082	15	P16403|H12_HUMAN	HLA-A*24	93.8079
KPKVVKPKKAAPKKK	19.96	1674.14	15	P16403|H12_HUMAN	HLA-A*24	92.9651
ASGPPVSEL	35.63	855.4338	9	P16403|H12_HUMAN:P16402|H13_H	HLA-C*03	0.8887
KERSGVSLAAL	30.89	1129.646	11	P16403|H12_HUMAN:P16402|H13_H	HLA-B*40	0.5254
KERSGVSL	28.63	874.4872	8	P16403|H12_HUMAN:P16402|H13_H	HLA-B*40	0.781
KERSGVSLA	28.51	945.5243	9	P16403|H12_HUMAN:P16402|H13_H	HLA-B*40	0.8099
KERSGVSLAA	24.16	1016.561	10	P16403|H12_HUMAN:P16402|H13_H	HLA-B*40	0.6765
AALKKALAA	34.21	855.5541	9	P16403|H12_HUMAN:P16402|H13_H	HLA-C*03	2.5593
SELITKAV	32.44	859.5015	8	P16403|H12_HUMAN:P16402|H13_H	HLA-B*40	0.7546
SELITKAVA	23.58	930.5386	9	P16403|H12_HUMAN:P16402|H13_H	HLA-B*40	0.6151

B



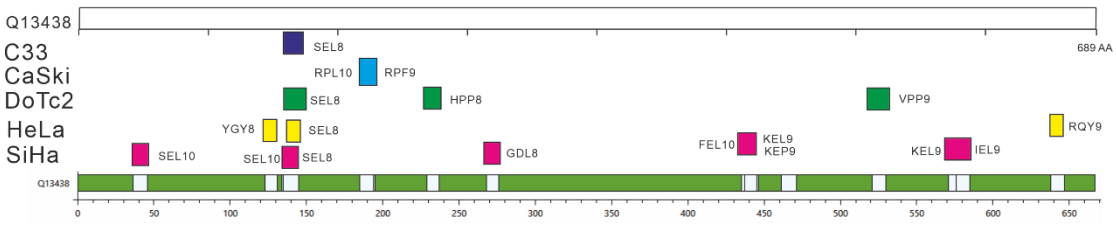
Name	Score	Mass	Length		Allele	Rank
P51149|RAB7A_HUMAN:RAB7A						
C33						
NEFPEPIKL	21.31	1085.576	9	P51149|RAB7A_HUMAN	HLA-B*44	1.071
CaSki						
KLDKNDRAK	29.72	1086.615	9	P51149|RAB7A_HUMAN	HLA-A*03	2.1069
DoTc2						
KLDKNDRAK	38.79	1086.615	9	P51149|RAB7A_HUMAN	HLA-A*03	2.1069
KVIILGDSGVGK	38.17	1184.713	12	P51149|RAB7A_HUMAN	HLA-A*03	0.8941
ATIGADFLTK	36.45	1035.56	10	P51149|RAB7A_HUMAN	HLA-A*03	0.1826
YNEFPEPIKL	26.58	1248.639	10	P51149|RAB7A_HUMAN	HLA-C*03	8.2645
HeLa						
VNKKFSNQY	45.21	1126.577	9	P51149|RAB7A_HUMAN	HLA-B*15	3.0857
SiHa						
NEFPEPIKL	34.75	1085.576	9	P51149|RAB7A_HUMAN	HLA-B*40	0.4049
KEVMVDDRL	30.25	1103.565	9	P51149|RAB7A_HUMAN	HLA-B*40	0.3287

C



Name	Score	Mass	Length		Allele	Rank
P57740|NU107_HUMAN:NUP107						
C33						
AYLEAHETF	18.87	1079.492	9	P57740|NU107_HUMAN	HLA-C*07	0.2226
CaSki						
KTYIQLLIREK	49.12	1403.85	11	P57740|NU107_HUMAN	HLA-A*03	0.29
FSSQRSGLF	30.49	1027.509	9	P57740|NU107_HUMAN	HLA-C*07	0.5274
DoTc2						
KTYIQLLIREK	49.59	1403.85	11	P57740|NU107_HUMAN	HLA-A*03	0.29
AIYAALSGNLK	38.49	1119.629	11	P57740|NU107_HUMAN	HLA-A*03	0.0578
HeLa						
DVATITKTVV	44.7	1045.602	10	P57740|NU107_HUMAN	HLA-A*68	0.3592
DVATITKTV	42.01	946.5335	9	P57740|NU107_HUMAN	HLA-A*68	0.142
GQAWRAATL	37.64	972.5141	9	P57740|NU107_HUMAN	HLA-B*15	0.6305
KQRQLTSY	29.62	1022.551	8	P57740|NU107_HUMAN	HLA-B*15	0.2669
KKHEAAKEVF	21.9	1185.651	10	P57740|NU107_HUMAN	HLA-B*15	0.8395
DVKEKMYNV	16.69	1124.554	9	P57740|NU107_HUMAN	HLA-A*68	0.7557
SiHa						
AYLEAHETF	46.63	1079.492	9	P57740|NU107_HUMAN	HLA-A*24	0.0099
GEFSHHDL	34.89	940.4039	8	P57740|NU107_HUMAN	HLA-B*40	0.5319
GEISSPVI	18.69	800.428	8	P57740|NU107_HUMAN	HLA-B*40	0.4655

D



Name	Score	Mass	Length		Allele	Rank
Q13438|OS9_HUMAN:OS9						
C33						
SEIKGEVL	28.55	873.4807	8	Q13438|OS9_HUMAN	HLA-B*44	2.7291
CaSki						
RPREAEVRFL	50.08	1271.71	10	Q13438|OS9_HUMAN	HLA-B*07	0.031
RPREAEVRF	44.62	1158.626	9	Q13438|OS9_HUMAN	HLA-B*07	0.0753
DoTc2						
VPKKPPPSP	41.59	945.5647	9	Q13438|OS9_HUMAN	HLA-B*55	3.8949
HPLLRPPP	32.36	925.5497	8	Q13438|OS9_HUMAN	HLA-B*55	4.3021
SEIKGEVL	24.89	873.4807	8	Q13438|OS9_HUMAN	HLA-C*03	38.4696
HeLa						
RQKELESNY	36.47	1165.573	9	Q13438|OS9_HUMAN	HLA-B*15	0.0884
YGRHIQQY	33.29	1063.52	8	Q13438|OS9_HUMAN	HLA-C*12	5.7095
SEIKGEVL	27.57	873.4807	8	Q13438|OS9_HUMAN	HLA-B*15	18.0721
SiHa						
DSEIKGEVL	35.51	988.5076	9	Q13438|OS9_HUMAN	HLA-C*03	13.5104
SEIKGEVL	32.96	873.4807	8	Q13438|OS9_HUMAN	HLA-B*40	0.4532
SEIKGEVLYL	29.1	1149.628	10	Q13438|OS9_HUMAN	HLA-B*40	0.207
RELENIIQET	29.07	1243.641	10	Q13438|OS9_HUMAN	HLA-B*40	0.6925
FEKELEGILL	27.95	1189.659	10	Q13438|OS9_HUMAN	HLA-B*40	0.3886
SEMRYGIEIL	27.02	1209.606	10	Q13438|OS9_HUMAN	HLA-B*40	0.0281
IEGLVKELL	26.91	1012.617	9	Q13438|OS9_HUMAN	HLA-B*40	0.6993
KELLEREGL	25.49	1085.608	9	Q13438|OS9_HUMAN	HLA-B*40	0.208
KELEGILL	22.95	913.5484	8	Q13438|OS9_HUMAN	HLA-B*40	0.1201
KELEGILLP	21.74	1010.601	9	Q13438|OS9_HUMAN	HLA-B*40	0.617
GDKIIEEL	18.82	915.4913	8	Q13438|OS9_HUMAN	HLA-B*40	5.0621

Figure 17. List of four shared antigens defined by the Raven database hat were presented by all five cervical tumour cell lines. A-D**.** Relative position of peptides identified mapping to the parent protein gene. The horizontal line and green horizontal box indicate the nucleotide position in the respective gene. White interruptions in the green box indicate the HLA-peptide coverage. Small boxes in difference colors show the coding sequence for the

Raven-associated HLA-peptide sequence that were identified in the different cell lines.

3.2.4 Identification of *neo*antigens in comparison to the reference proteome

In order to identify the tumour-specific *neo*peptides in cell datasets, we performed exome sequencing data and which in comparison to the human reference genome. Using protein sequence databases representing the effects of these mutations, we identified a total of 6 HLA-presented *neo*peptides spanning 5 distinct mutations between 3 of the cell lines (Table 2). These were all missense-derived except for a single frameshift-derived peptide that was originating from an alternative open reading frame that was generated by the deletion of a single nucleotide in the gene encoding Ubiquitin-conjugating enzyme E2 variant 3 (UEVLD). All peptides were predicted to be strong binders to one of the HLA alleles present in the respective tumour line (NetMHCpan rank score)

Table 2. Tumour-specific *neo*antigens identified in cervical cancer panel.

Altered amino acids are highlighted in red.

Peptide	Rank	HLA allele	-10lgP	Uniprot Acc	Mutation site	Gene Name	Description
C33							
KLWDIINVNTAA	0.622	A*02:01	26.78	Q3SXM5	178I>T	HSDL1	Inactive hydroxysteroid dehydrogenase-like protein 1
KLWDIINVNT	0.481	A*02:01	21.42	Q3SXM5	178I>T	HSDL1	Inactive hydroxysteroid dehydrogenase-like protein 1
DoTc2							
KVKGPTLSRK	0.006	A*03:01	26.3	Q16666	93A>T	IFI16	Gamma-interferon-inducible protein 16
RTDSLLRIPK	0.399	A*03:01	40.41	Q5SYE7	686R>L	NHSL1	NHS-like protein 1
SiHa							
KAVHPQSTV	0.289	C*03:04	33.59	Q8IX04	447K>X	UEVLD	Ubiquitin-conjugating enzyme E2 variant 3
IAEPSGRAL	0.002	C*03:04	21.96	Q5XKK7	3T>I	FAM219B	Protein FAM219B

3.2.5 Viral peptides identified in five human cervical cancer cell lines

In order to refine HPV-derived HLA peptides, we reconstructed viral transcript sequences by *de novo* assembly of RNA reads using Trinity [224]. With the help of this approach, we identified 5 peptides from the assembled viral transcripts (Fig.8) (Fig.18) (Table.3). In HeLa cells, we identified three peptide antigens originating from HPV 18. KQGAMLAVF (KF9) and DTPEWQRL (DL8) were derived from E1, while a peptide (MKFRLTFY, MY8) was coming from alternative reading frame in E7. In CaSki cells, we identified three HPV 16 associated peptides. YMLDLQPET (YT9) which is predicted to bind to HLA-A*02:01 and also been reported a T cell antigen [242]. The second one, MLYQMTRTK (MK9) maps to a short alternative reading frame within the E1 gene region. For validation of the correct annotation of peptide sequence from the LC-MS2 spectra, we performed spectral matching analysis using spectra from synthetic peptide standards. All peptides showed signal

distribution in agreement with their synthetic standard (Fig.19). The last peptide, RLFSTLITV (RV9), originating from a non-canonical reading frame in the human genome which spanned regions in HPV 16 E6, E7 and E2.

Since we could not map this amino acid sequence to any open reading frame encoded within the HPV genome, we interrogated the assembled transcript TRINITY_DN0_c0_g1_i6, which encoded the relevant open reading frame, in more detail. We found that the HPV-E2 sequence was partially truncated within this transcript, and that the transcript extended into a human gene region just upstream of the RUNX2 gene on chromosome 6, location chr6p12.3. This mapping was in line with previous observations detailing the proviral integration site in chr6p12.3[243, 244] (NCBI accession HE984566.1, https://pubmed.ncbi.nlm.nih.gov/23824673/). The RV9 peptide was derived from this additionally transcribed region of the transcript and is encoded in a reverse-oriented reading frame relative to the downstream RUNX2 coding region. We also confirmed the sequence using spectral matching. This result clearly demonstrates the validity of the chosen approach for discovery of antigens from non-conventional sources (Fig.19).

Table 3. List of viral peptides identified in five human cancer cell lines. *Trinity represents a novel method for the efficient and robust *de novo* reconstruction of transcriptomes from RNA-seq data. 6Fr Is 6 six-frame translation for HPV DNA sequences

Name	peptide	length	Trinity database accession ID	6Fr database	HPV protein	score	Type	Binder	Rank
Hela	KQGAMLAVF	9	Trinity (2 E1,3 E1 ,54 E1,77 E1)	89 [908 - 2884]	hpv18 [gene=E1]	20.48	Cell	HLA-B15:03	0.0657
	DTPEWIQRL	9	Trinity (2 E1,3 E1 ,54 E1,77 E1)	89 [908 - 2884]	hpv18 [gene=E1]	21.96	Cell	HLA-A68:02	0.3777
	MKFRLTFY	8	Trinity 53	19 [636 - 680]	ALT Frame E7	27.8	Cell	HLA-B15:03	1.976
Caski	RLFSTLITV	9	Trinity 28		E2s	20.17	Cell	HLA-A02:01	0.0445
	MLYQMTRTK	9		3 [74 - 145]	ALT Frame E1	34.24	Cell	HLA-A03:01	0.0348
	YMLDLQPET	9	Trinity 55		hpv16 [gene=E7]	5.11	Cell	HLA-A02:01	0.127

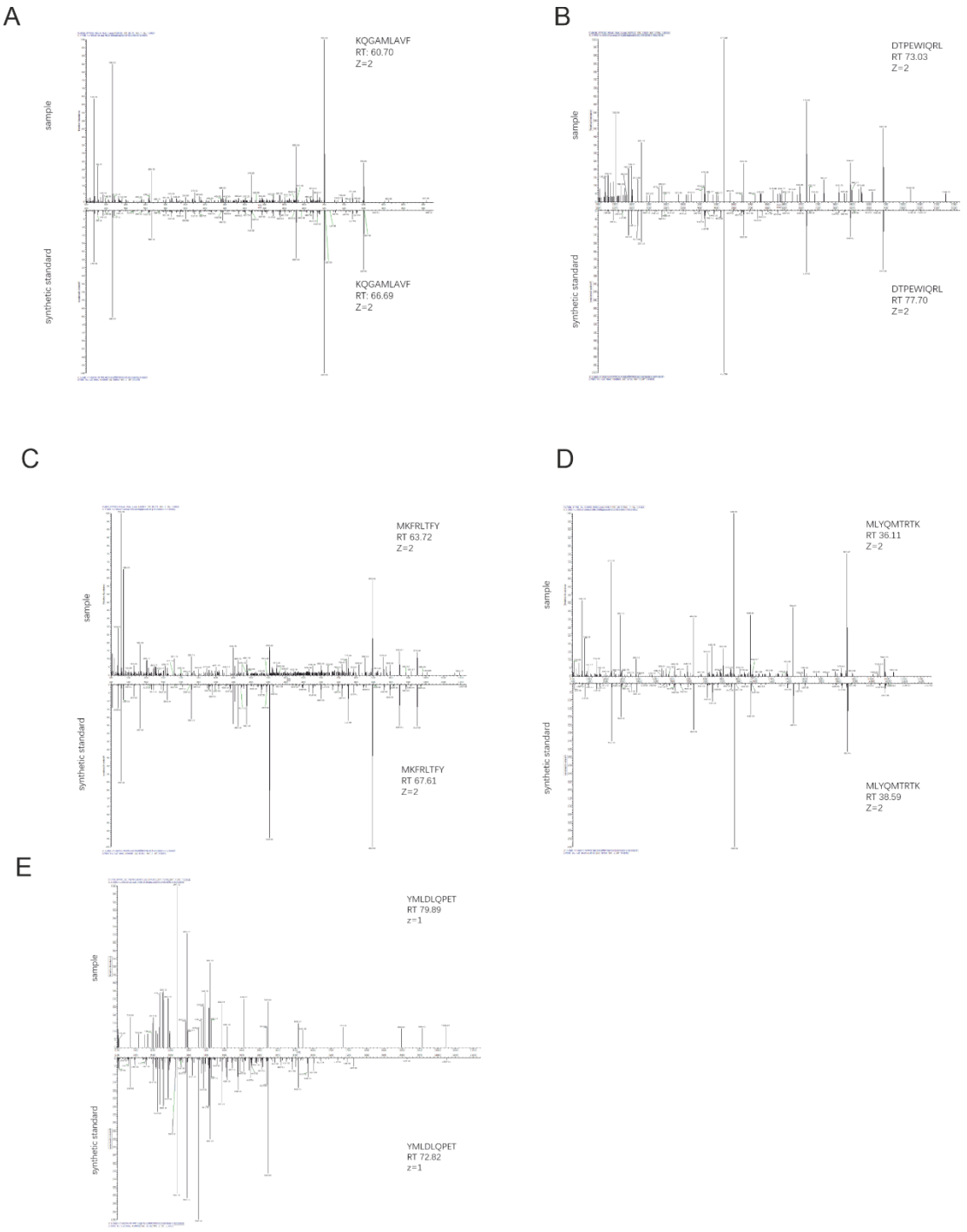


Figure 18. Viral peptide antigens identified in HeLa and CaSki cells. A-E. Spectral matching analysis for indicated peptide sequences. LC-MS fragment spectra are shown as measured in the experiment (sample) and measured from the equivalent synthetic peptide sequence counterpart (synthetic standard).

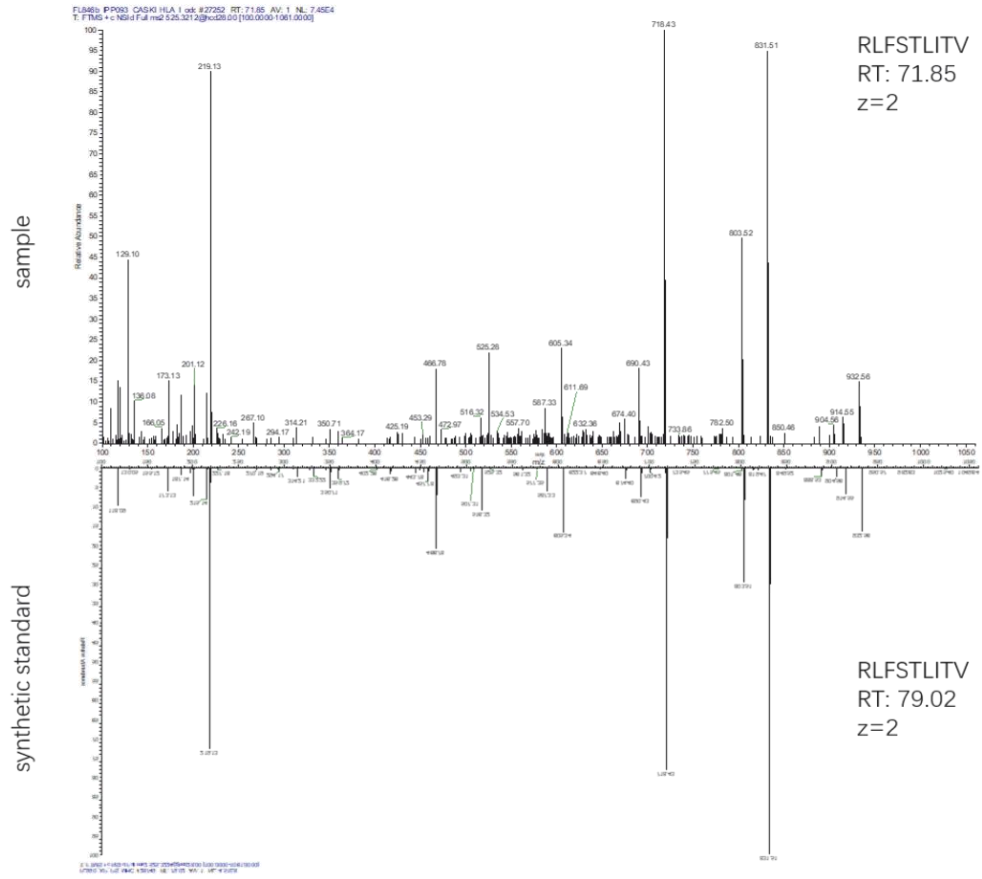


Figure 19. Identification of a cryptic peptide sequence sourced within the human genome originating from a transcript spanning the viral integration site. Spectral matching analysis for indicated peptide sequences. LC-MS fragment spectra are shown as measured in the experiment (sample) and measured from the equivalent synthetic peptide sequence counterpart (synthetic standard).

3.2.6 HPV18 E1 protein is expressed in HeLa cells

To confirm that E1 protein is expressed in HeLa cells, we downloaded HeLa deep proteomics data from dataset identifier PXD 004452[245]. In this manuscript, HeLa cell lysates were digested with several enzymes and resulting peptides were fractionated and interrogated with LC-MS/MS to obtain deep proteome coverage for this cell line. We found evidence for HPV-18 E1, E6, and E7 proteins in this dataset (Table 4). This finding confirms active production of HPV-E1 in HeLa cells, and further supports the findings of our work.

Table 4. List of viral peptides identified in dataset PXD004452 (Hela). *Trinity represents a novel method for the efficient and robust *de novo* reconstruction of transcriptomes from RNA-seq data. 6Fr Is 6 six-frame translation for HPV DNA sequences.

Name	peptide	length	Accession ID	Score	HPV Match Similarity	Human Match Similarity
Hela deep proteomics	HYSDSVYGDTLEK	13	Trinity 4/6Fr 15	80.81	E6 100%	69%
	FHNIAGHYR	9	Trinity 4/6Fr 15	49.45	E6 100%	71%
	QMNMSQWIR	9	Trinity 2/ 6Fr 89	63.21	E1 100%	66%
	QGAMLAVFK	9	Trinity 2/ 6Fr 89	25.51	E1 100%	66%
	HTMLCMCCK	9	6Fr 25	56.77	E7 100%	78%

3.2.7 ELISpot response analysis of patients' Short-term cell lines (STCL) of confirmed peptides

In order to validate the obtained viral sequences for their potential to elicit immune responses in infected individuals, we collected PBMCs from 7 women with cervical intraepithelial neoplasia (CIN, grades 1-3) or colposcopic evidence of HR-HPV changes without CIN (5 HPV 16 positive, and two HPV 18-positive on vaginal sampling for HPV DNA) and performed cultured ELISpot assays with the peptide sequences identified for the respective HPV types. We observed evidence for T cell responses for all identified peptide sequences in at least one out of the patients, confirming that the MS-identified peptides were immunogenic in the context of natural infection. Interestingly, one patient also showed a response against the non-canonical peptide derived from the human genome we identified in Caski cells (RV9) and confirmed its immunogenicity in vivo (Fig. 20).

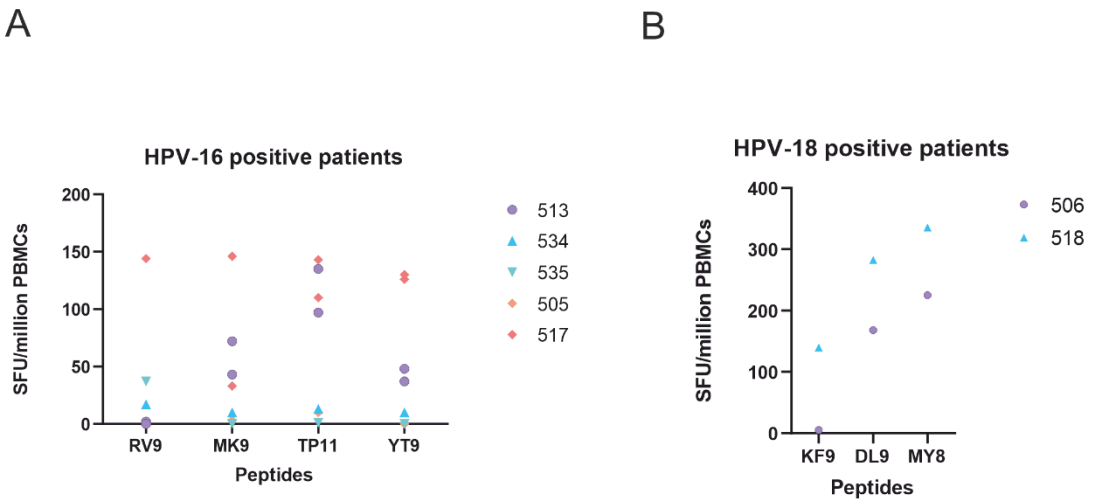


Figure 20: Immunogenicity of identified viral and cryptic peptide sequences. A. T cell response to peptides identified in HPV 16 positive patients as measured by ELISpot. (n=3-5) B. T cell responses to peptides identified in HPV 18 positive patients as measured by ELISpot (n=3-5)

3.3 Discussion

In these human cervical cancer lines, we identified 55,474（48,278 unique peptides）peptides from 15,129 (10,920 unique proteins) proteins. More than 84% of the peptides had the typical length of HLA-I peptides (8 to 14 amino acids) and most of the presented peptides are nine amino (60%) acids long. With the help of RNA-seq and the DNA exome sequencing-based protein databases, we identified 5 viral peptides from our sample and 6 peptides that come from 5 neoantigens.

Recently, high performance mass-spectrometric techniques have given us the ability to globally investigate how variety of antigens are presented to the immune system. This method has been demonstrated to successfully identifiy thousands of MHC-associated peptides. In 2014, Granados *et al.* identified 4468 MHC-I related peptides of PBMC from two non-twin HLA-identical female siblings [246]. Berlin *et al.* used mass spectrometry to define acute myeloid leukemia (AML)-associated peptide and successful identified 25,000 different naturally presented HLA ligands from 15 AML patients and 35 healthy controls [247]. Bassani-Sternberg *et al.* identified 2500 to 5200 peptides from seven human primary and cancer cell lines (Fig. 21) [248]. Therefore, our results are in line with the published results from in peer-reviewed literature.

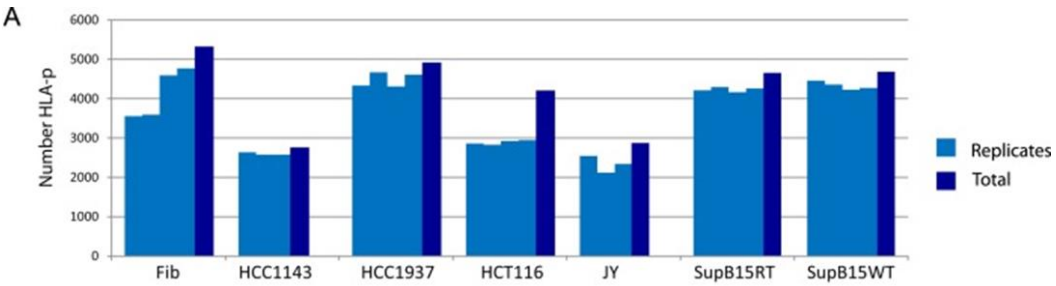


Figure 21. Properties of the HLA-I peptidomes data set obtained from seven cell lines. *A*, Number of HLA-I peptides identified in each peptidome sample. (This figure is a copy from ref 248) [248]

Human papillomaviruses (HPVs) are small DNA tumour viruses of the papillomaviruses family which present in more than 99% of cervical cancers [38]. Aside from the structural viral proteins L1 and 2, the HPV genome contains E1-7 genes which are necessary for viral DNA synthesis and transcriptional regulation. Most studies are focus on the HPV oncogenic proteins, E6 and E7, which are crucial in the induction and maintenance of cellular transformation. Both proteins are known to be presented by HLA on cervical tumour cell lines, and, despite their efficacy as vaccine antigens could not been demonstrated to date, these viral proteins have been shown to be potent therapeutic targets for TIL-based therapies.

The first HLA class I-restricted HPV T cell epitopes published in the literature

were identified through T cell immunogenicity screening in HLA-A*02:01-positive healthy donors of predicted peptide sequences from E6 and E7 ($E6_{29-38}$ (TIHDIILECV); $E7_{11-20}$ (YMLDLQPETT); $E7_{82-90}$ (LLMGTLGIV); $E7_{86-93}$(TLGIVCPI)) [249]. Overlapping peptide library screening led to the identification of two additional peptides $E6_{80-88}$ (ISEYRHYCY) and $E7_{44–52}$ (QAEPDRAHY), which were both presented by B*18:01 [250]. In 2011, the list was expanded and several HLA class I-bound E7 peptides were mapped in CaSki cells using a targeted LC-MS3 approach by screening a library of HPV 16 E7-derived peptides that were predicted to bind to HLA-A2. Such targeted MS approaches are highly sensitive, but are limited to detection of sequences included in the synthetic peptide library (17 sequences in this study). Authors reported $E7_{11–19}$ (YMLDLQPET, including C-terminally extended versions and an internal fragment), which was also detected in our study, $E7_{7–15}$ (TLHEYMLDL); $E7_{77–87}$ (RTLEDLLMGTL), $E7_{82–90}$ (LLMGTLGIV); $E7_{80–90}$ (EDLLMGTLGIV); $E7_{78–86}$ (TLEDLLMGT); $E7_{77–86}$ (RTLEDLLMGT); $E7_{81–90}$ (DLLMGTLGIV) [251].

Adoptive T cell therapy selected for E6 and E7 antigens has been shown to be able to induce complete tumour regression in selected treated patients [27], and adoptive T cell therapy in a patient showed a dominant T-cell clonotype targeting HLA-A*02:01–restricted targeting of $E6_{29–36}$ (TIHDIILECV) in a patient with metastatic anal cancer, who had been disease free 22 months

after successful treatment [28]. Furthermore, treatment using TCR-engineered T cells targeting E7 for patients with metastatic HPV-associated epithelial cancers showed a number of partial responders and robust tumour regression in 6/12 patients including 4 of 8 patients with anti-PD-1 refractory disease [29].

Here, we first report conclusive evidence that E1 peptides are presented as antigen in HPV-positive cervical tumour cells, and further showed that E1 is constitutively and consistently expressed in primary tumour tissue. E1 is the largest and highly conserved throughout papillomaviruses, reflecting that its function is vital to the virus. E1 not only regulate the HPV genome, but also the host genome to associate with the antiviral immune response[252]. Recently, a study showed that during the manifestation of HPV-associated cervical cancer, viral integration into the host genome often leads to the deletion of E2, E4, E5, L1, and L2 viral proteins and the constitutive upregulation of the E6 and E7 oncogenes [252]. However, the status of E1 in cervical cancer has not yet been reported.

E1 may be an additional potent antigen for the development of specific anti-tumour vaccines. E1 protein is localized in the nucleus and is required for the initiation of viral DNA replication. The main function of E1 protein known to date is its maintenance and amplification of viral genomes, and it's interaction

with several cellular proteins of the replication machinery such as α-DNA polymerase, replication A proteins (RPA), topoisomerases I and II, nuclear proliferation cell antigen (PCNA) and the replication factor C (RFC) [40] to increase the copy number of the viral episome upon infection of epithelial layers. It is known that E1 and E2 are the only two viral proteins necessary for HPV genome replication [236]. Therefore, therapeutical intervention with molecules specifically targeting and inhibiting the E1 function may effectively reduce HPV replication and cell transformation. We are here, however, highlighting the importance of HPV-E1 as a tumour-specific, HLA-presented antigen in cervical cancer opens opportunity to expand current vaccination and T-cell-based therapies to this antigen.

Recently studies have shown that integration of high-risk HPV DNA into the host genome can be associated with progression and invasiveness in cervical tumours. In addition, this HPV integration has been reported to occur at random chromosomal sites. In 2013, Xu, B., have reported the three novel CaSki DNA junctions in their study and they suggest that those DNA junctions probably originate from non-transcribed concatemeric HPV16 integrates. However, based on our results and analyses, we identified the peptide “RLFSTLITV” from our Trinity database which was translated from nucleotide sequence “CGCCTATTTTCAACTCTAATCACAGTA” (Fig.19). This sequence was found to be close to the CaSki viral-cellular junction HE984566 and

encoded by a transcript spanning this integration site. We found that the transcript was spanning HPV-16 E6, E7 and E2 reading frames upstream, and a 5'UTR region of the cellular gene RUNX2 located downstream of the integration site. Our results first indicate that cryptic sequences from viral transcripts spanning the integration site can be translated and presented on the HLA I molecules. However, due to this sequence being specific to the CaSki cell integration event, it is likely not to be present in other HPV-driven tumours. We did observe a significate response in our ELISpot experiment confirming that cryptic peptides from viral transcripts spanning the integration site can be immunogenic (Fig. 20), but therefore concluded that the observed reactivity is likely the result of a naive T cell response.

We found 5 mutated antigens in our experiment (Table 2), such as Interferon gamma inducible protein 16 (IFI16), Human hydroxysteroid dehydrogenase like 1 gene (HSDL1), NHS-like protein 1 (NHSL1), Ubiquitin-conjugating enzyme E2 variant 3 (UEVLD) and Family with sequence similarity 21 (FAM21). IFI16 is an a PYHIN protein which is involved in transcriptional regulation and innate immune sensor for intracellular DNA [253]. HSDL1 is encoding a 309 amino acid and located on human chromosome 16q24.2 [254]. It family and was found to be conserved in many other vertebrate species [255]. The study shown that HSDL1 is highly expressed in testis and ovary, which also predominantly increased in the prostate cancer tissue compared with the

normal prostate tissue [254]. NHSL1 is widely expressed in human body, but the function is still unknown [256]. UEVLD encoding 379 amino acids and maybe the negative regulator of polyubiquitination [257]. FAM21 could interact with various sorting nexin proteins and participates in NF-kB-dependent gene regulation [258, 259].

The emergence of cancer immunotherapy has provided remarkable new possibilities for the treatment of cancer. For this strategy, it is important to understand which proteins are exclusively presented by the tumour and not healthy tissues, and therefore can be used as antigens in T-cell-based therapies or vaccination approaches. Cancer-testis (CT) antigens are immunogenic proteins which preferentially expressed in tumours with limited expression in the rest of tissues which makes them preferred therapeutic targets. Utilizing databases that list known tumour-testis antigens, i.e. the CT database, we found four shared CT antigens presented by all cell lines that were associated with chromosome segregation and maintenance of genomic stability: KNL1, NUF2, Cep55, KIF2C. Cep55 localizes to the centrosome of interphase cells and to the midbody during cytokinesis [260]. Cep55 directly binds to microtubules in vitro which also a key regulator of cytokinesis essential for the midbody structure and for vesicle trafficking and fusion [261]. In addition, Cep55 participates in membrane abscission to form two daughter cells and plays a role in the development of the brain and kidney [262].

Furthermore, Cdk1, Erk2, and Plk1 cooperate in the mitotic phosphorylation of Cep55 to maintain genomic stability [260]. The spindle checkpoint is key regulator of mitotic progression. The assembly of a large proteinaceous structure termed the kinetochore is essential in chromosome segregation during mitosis. Defects in kinetochore proteins often lead to aneuploidy and cancer. Studies shown that the KNL1 protein counterpart is necessary for chromosome segregation and is required to localize a subset of outer kinetochore proteins [263]. The data also suggested that this kinetochore protein has a role not only in the association with checkpoint proteins Bub1 and BubR1, but also in the attachment to the spindle microtubules [264]. The kinetochore is a supramolecular complex which provides a chromosomal attachment point for the mitotic spindle and monitors the movements of chromosomes during mitosis. NUF2 is a kinetochore protein that forms a stable complex with HEC1, HsSPC24, and HsSPC25 (HsNDC80 complex) [265]. HsNDC80 complex is composed of four proteins: HEC1, HsNUF2, HsSPC24, and HsSPC25. HsNDC80 is also an essential kinetochore core component highly conserved across species, with a crucial role in proper chromosome congregation and segregation during mitosis [266-268]. Both the Nuf2-Ndc80 complex and its human homolog, NUF2-HEC1, have been shown to be essential for stable microtubule-kinetochore attachment, chromosome alignment, and also spindle checkpoint activation in mitosis [266-270]. Recent studies have shown that Homo sapiens (Hs) NUF2 is required for stable

kinetochore localization of centromere-associated protein E (CENP-E) in HeLa cells [266]. These results indicate that HsNUF2 and CENP-E are required for organization of stable microtubule-kinetochore attachment which is essential for chromosome segregation in mitosis [265]. Recent studies also have shown that that the main function of the HsNUF2-HEC1 complex is to stabilize the microtubule-kinetochore association, as cells lacking the HsNUF2-HEC1 complex often carry unstable microtubules [271]. Most solid tumours are aneuploid and show frequent mis-segregation of chromosomes which are due to maloriented attachment of chromosomes to spindle microtubules (MTs). Stable and dynamic MTs attachment at kinetochores are key to correct chromosome segregation which can normally adjust the position of malorientations. MTs dynamics are tightly control the spindle assembly which is associated with many MT binding proteins and their upstream regulators. Recent studies indicate that aurora kinases regulate MT plus-end stability in mitosis through control of KIF18B-KIF2C complex formation [272] and inhibition of aurora A and B led to a dramatic decrease in spindle MT stability. The studies also shown that two microtubule-depolymerizing kinesins, KIF18B and KIF2C, play an important role in kinetochore-microtubule dynamics during mitosis to correct malorientations and restore chromosome stability to chromosomally unstable tumour cell lines. [273].

The other identified, shared CT antigens were associated with transcriptional

regulation and cell signaling. ATAD2 plays an important role in normal growth and function of the female reproductive system and in the pathogenesis of the majority of human breast cancers [274, 275]. Imp3 is reported that mediate the association of the heterotrimeric complex with the U3 snoRNA [276], which is required for the early cleavage steps in pre-rRNA processing and essential for the formation of the small ribosomal subunit RNA (18S rRNA) in both fungi and vertebrates [277-279]. PDZ-binding kinase TOPK is a novel mitotic kinase and belongs to the protein kinase superfamily which is upregulated in a variety of proliferating cells and tissues. Recently studies have indicated TOPK was a possible new therapeutic target in many types of cancer. TOPK is association with cdk1/cyclin B and knockdown expression led to cytokinetic dysfunction [280, 281]. Recent studies have suggested that TOPK can also forms a complex with p53 after been phosphorylated and attenuates G2/M checkpoint during doxorubicin-induced DNA damage [282].

Epigenetics play a vital part in the regulation of gene transcription and genomic integrity [283]. Epigenetic changes, including DNA methylation and histone modifications, can alter gene expression without changing the DNA sequence [284, 285]. The histone lysine demethylase KDM5B that was first characterized in 2006, regulates gene expression and associated with tumour growth [286]. KDM5B is responsible for catalyzing the removal of methyl groups from histone H3 (H3K4me3/2/1), which shows a restricted expression

pattern in adult tissues and is primarily present in the testis and brain [287-292]. However, increased KDM5B levels were found in a variety of human cancers, such as breast [293], ovarian cancer [294], gastric cancer[295], bladder cancer[296], lung cancer[296], hepatocellular carcinoma [297], colorectal cancer [298], gliomas [299] and malignant melanoma [300]. Furthermore, specific KDM5B inhibitors have been recognized for their potential as therapeutic drugs for many types of cancer, and in recent years, many inhibitors have been developed. However, inhibitors of KDM5B have failed clinical trial due to secondary toxicities, and those compounds need to be further optimized and de-risked for cancer therapy.

Immunogenic peptides encoded by cancer-specific genes (CSGs) can be used as targets in immunotherapy. The four shared antigens identified matching to the Raven [301] database are related to ER stress and endosomal trafficing. The ER is the major site for protein folding, assembly, post-translationally modification and secretory pathways in eukaryotic cells. Nascent polypeptides in the ER are monitored by a quality control process which ensures only correctly folded proteins are allowed to exit the ER and to proceed to their site of function [302]. Misfolded or unassembled polypeptides are eliminated by a process known as ER-associated degradation (ERAD) pathway [303].

A recent study showed that OS-9 are upregulated in response to ER stress and associated both with components of the ERAD machinery and with ERAD substrates [304]. They also reported that OS-9 may bind terminally misfolded non-glycosylated proteins to transfer them to the ubiquitination machinery for degradation [304]. OS-9 also interacts with the cytosolic N-terminal tail of transient receptor potential proteins 4 (TRPV4) and as an auxiliary protein for TRPV4 maturation [305].

Rab7 is a transporter of vesicular stomatitis virus G protein (VSVG) and SV5 hemagglutinin–neuraminidase, from early to late endosomes [306]. It also reported that Rab7 is responsible for the transport of cathepsin D and cation–independent mannose 6–phosphate receptor (CI–M6PR) from the Golgi to late endocytic compartments [307]. In addition, Rab7 as an important regulator of late endosomal class III phosphatidylinositol 3′–kinase (hVPS34) function and regulates hVPS34 cycling between early and late endosomes [308-310]. Recent studies shown that Rab7 is also involved in regulation of late endocytic transport and membrane trafficking [311].

Histone H12 protein binds to linker DNA between nucleosomes forming the macromolecular structure known as the chromatin fiber [312]. Histones H1 are necessary for the condensation of nucleosome chains into higher-order

structured fibers. They also act as regulators of individual gene transcription through chromatin remodeling, nucleosome spacing and DNA methylation [312, 313].

Nuclear pore complexes (NPCs), which contains several distinct subcomplexes are the mediators of nucleocytoplasmic transport [314-317]. Mammalian NPCs are consistent of 47 Nups or Nup-associated proteins [318]. Nup107 is the core element part for the assembly of a subset of Nups into the NPC [319]. In addition, it is reported that mutations in the genes *NUP107* have associated with steroid-resistant nephrotic syndrome (SRNS) [320].

Immunotherapy is relying on T lymphocyte-mediated recognition of tumour antigens to eliminate tumour cells in patients. Next-generation sequencing and immunopeptidomics have greatly aided the discovery of novel antigens that are targetable in such next generation therapies, and we have here shown the discovery of so far unreported tumour-associated antigens, *neo*antigens spanning tumour-specific mutations and alternative open reading frames induced by INDELS, and finally, the identification of viral canonical and non-canonical tumour-specific antigens in a cervical tumour cell line panel.

Chapter 4: Identification of tumour-specific antigens in ten human cervical cancer resections

4.1 Introduction

The recent development in cancer immunotherapy treatment with immune checkpoint blockade has not only demonstrated the potency of utilizing T−cell reactivity for cancer therapy but, as therapies are only effective in a subgroup of patients, has also highlighted the need to understand the mechanism of action in more detail, and to expand the existing approaches to more specific therapies involving more suitable immune-targets. Immunotherapy relies on T cell recognition of antigens presented by human leukocyte antigens (HLA) on cancer cells.

For identification of clinically relevant antigens that are presented by HLA on tumour cells, it is important to directly interrogate primary tumour tissue. Such validation of the presentation of potential tumour-specific antigens identified through different approaches or in tumour cell lines directly in patient tissue is most relevant for the prioritization of antigens for therapeutic development.

Our workflow makes it possible to identify HLA-I peptidomes not only from cancer cell lines but also primary cancer tumours. We optimised the isolation of HLA complexes from tissue with primary tumours obtained from

proof-of-principle studies for HPV vaccination efficacy in the TC1 mouse model.

With the obtained protocol, we proceeded to purifiy and analyse HLA-associated peptides from 10 primary cervical tumour resections. We further integrated information from public databases to define the tumour-specificity of the identified antigens, with a view to shortlist the most commonly expressed, tumour-associated antigens across the analysed cohort. In addition, to identify the viral peptides, we also sequenced the mRNA transcriptome of ten human cervical cancer tumours in order to obtain the full viral transcriptome in each tumour and the associated viral open reading frame sequences for interrogation of or LC-MS data.

In these human cervical cancer tumours, we identified 78,618 peptides (52,382 unique peptides) from 9,765 proteins (9,215 unique proteins). More than 88% of the peptides had the typical length of HLA-I peptides (8 to 14 amino acids) and most of the presented peptides are nine amino (60%) acids long. Such peptides are especially promising for developing immunotherapeutics and our protocol is expected to advance immunopeptidomics as an antigen discovery platform for clinical personalized cancer immunotherapies.

4.2 RESULTS

4.2.1 Optimisation of tissue-processing protocol and using TC-1 cell line-derivedmouse tumour tissue.

First, we interrogated TC-1 mouse cell line and tumours to set up the protocol for the analysis of the human tumour samples.

The HPV oncogenic proteins, E6 and E7, are important in the induction and maintenance of cellular transformation and co-expressed in most HPV-containing cervical cancers [5, 77, 249, 321, 322]. Vaccines or immunological therapeutics targeting E6 and/or E7 proteins may provide an opportunity to treat HPV-associated cervical malignancy [20, 50, 77, 322]. In this study, we used the TC-1 mouse cell line that expresses HPV16-E6 and HPV16-E7 proteins, as a model of cervical cancer. Aim of the study was to identify T cell antigens suitable for immunotherapy with the focus on viral MHC class I-bound sequences. We purified MHC-I complexes and eluted bound peptides that were then analysed by LC-MS/MS in two technical replicates. The result showed that our immunopurification protocol and LC-MS/MS analysis workflow could isolate and identify a large number of peptides associated with MHC class I in mice. Moreover, those results expand the current knowledge of peptide presentation by MHC class I in cervical cancer and provide novel targets for T-cell vaccine development. A mouse experiment

was set up independent to our project by the Dorrell group in order to evaluate the efficacy of an HPV vaccine encoding conserved regions of HPV E proteins to control tumour growth in a therapeutic vaccination setting.

We performed immunopeptidomics on these samples with the aim to (i) set up a workflow to process tissues for HLA_peptide enrichment, and (ii) to identify peptide sequence suitable in their use for therapeutic vaccine design for proof-of principle studies to assess their protective capacity.

A total of 6 mice were divided into two groups and $1x10^6$ cells were implanted. (Fig.23 A). When tumours reached 3mm or more in diameter, animals were immunized and boosted with ChAd-MVA encoding a vaccine consistent of conserved regions of the HPV-E gene region (Group A) or PBS (Group C) for one week. (Experiments were performed by Dr.Gemma Hancock.)

To identify the cell line specific mutations and to sequence the accurate sequence of viral transcripts, we performed mRNA sequencing of the TC-1 cell line. Sequence interpretations of MS/MS spectra were also analysed by using PEAKS 8.5 (Bioinformatics Solutions Inc.) [323], which utilized a protein database included translations of sequencing data from TC-1 cell. In addition, viral transcript sequences were reconstructed by *de novo* assembly of mRNA

transcript sequences that matched the HPV genome using the Trinity tool [224] (Performed by Dr. Julius Muller). Information was added to the databases for spectral interrogation and to allow for the identification of *neo*antigens and viral peptides in these samples.

For comparison and validation, we first performed immunopeptidomics on the TC-1 mouse cell line, and we purified MHC-I complexes and eluted bound peptides that were then analysed by LC-MS/MS in two technical replicates. We identified a total of 2851 peptides associated with TC-1 cells MHC class I complexes (Fig.22 A). 2162 (79.58%) unique peptides and 1965 (71.47%) peptides were identified in each replicate analysis (Fig.22 A). Peptide-loaded MHC class I complexes were immunopurified from cell lysates by using HB-51 antibody. Monoclonal antibody HB-51, a mouse hybridoma cell-line, is directed against H-2 Kb and H-2 Db. MHC class I complexes were treated with acid in order to elute the peptides and peptides were separated from beta-2-microglobulin and alpha chain by reverse-phase HPLC. Finally, peptide fractions that did not contain beta-2-microglobuin were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Although the MHC class I -eluted peptides ranged between 5 and 29 amino acid in length, 65% of them were 8 or 9 amino acids long and 9-mers were overall the most abundant species (Fig. 22 B). Frequencies of individual amino acids in every epitope position were analysed for all 7- and 14-mer peptides, and the MHC

anchor amino acid preferences concurred with the MHC subtype of TC-1 cells (H-2 Kb and H-2 Db) (Fig. 22). To visualize and compare the dominant and subdominant motifs within our data to known binders in the IEDB, we downloaded the mouse MHC class I epitopes from IEDB and compared the 8 to 14 amino acids long motif (Fig.22 C). The 8- and 9-mer peptides showed the same pattern as IEDB database and the analysis shows that most 8-mer peptides originated from H-2 Kb and most 9-mer peptides from H-2 Db (Fig.22 C).

We then interrogated the mouse tumour resections, that were sourced from a failed therapeutic vaccine evaluation study utilising the TC-1 cell line for tissue protocol validation. We used a bead beater to homogenise the frozen tisses in lysis buffer. We performed all other steps as we did with the cell line. We identified total 4246 peptides associated with tumour MHC class I complexes (Fig.23 A), and 3758 (88.51%) unique peptides and 2571 (60.55%) peptides were identified in group A and group C, respectively (Fig.23 B). More than 70% of the peptides have the typical length of HLA-I peptides (8 to 14 amino acids) but most of the presented peptides are eight and nine (81.2%) amino acids long (Fig. 23 D). Most importantly, HPV-16 E7 viral peptide RAHYNIVTF was identified in mouse tumour (A31) and this peptide was a strong binder to H-2 Db. And this peptide was reported protected from tumour growth when given as a therapeutic vaccine in the TC-1 mouse model [324].

Hereby, we confirmed that we could sucessfully purify MHC_derived peptides from tumour tissue.

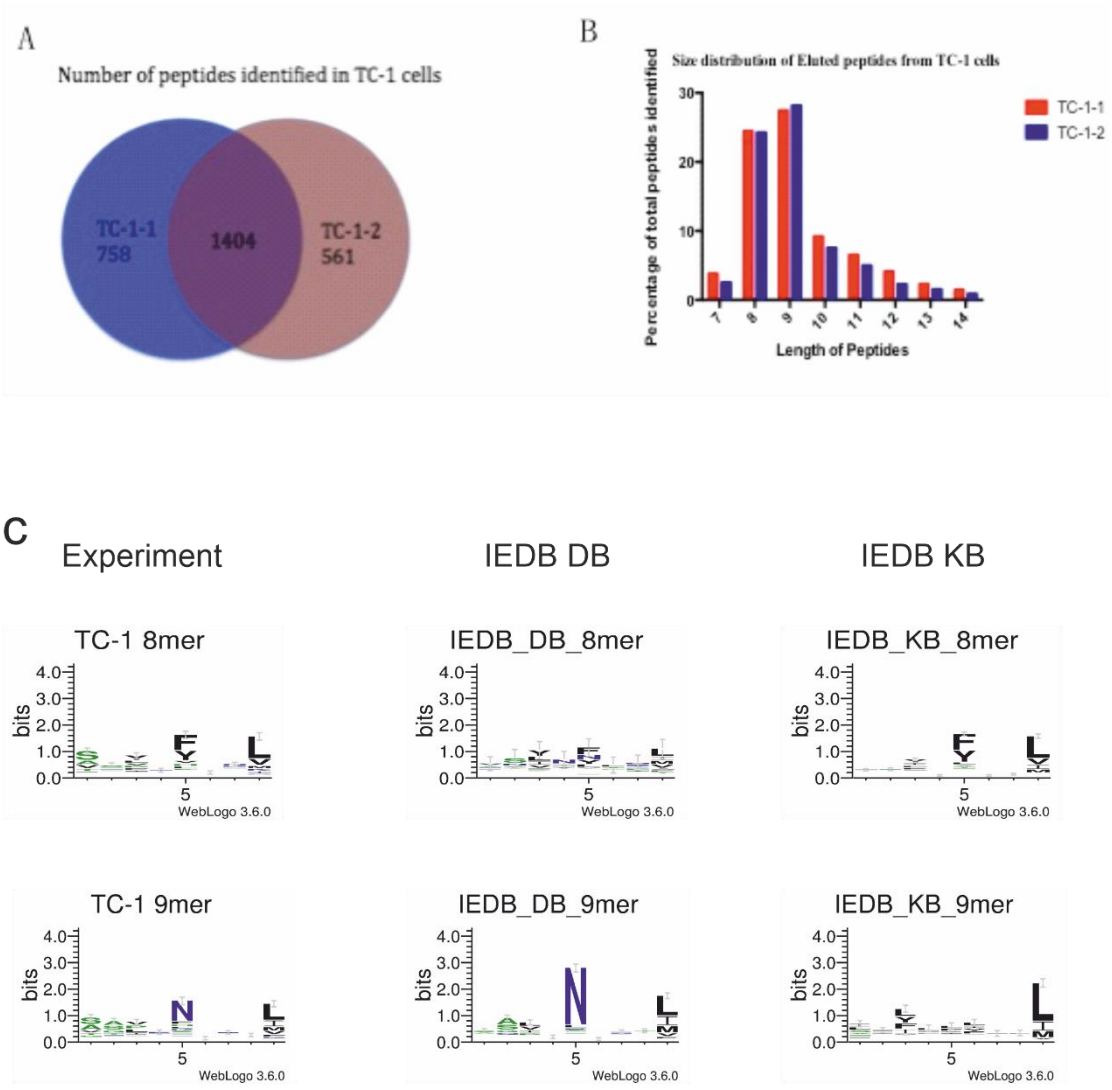


Figure 22. MHC-peptides identified in TC-1 cell lines. A. Number of peptides identified in TC-1 cell line. Blue is the first experiment and red is the

tech-replicate analysis. B. Size distribution of Eluted peptides from TC-1 cells. C. Amino acid distribution of all eluted 8- to 9mer peptides from the experiment in comparison with sequences in the IEDB database. The data is re-analysis by WebLogo [325]. X-axis :Add a label to the x-axis, or hide axis altogether. Y-axis:The vertical axis indicates the information content of a sequence position.

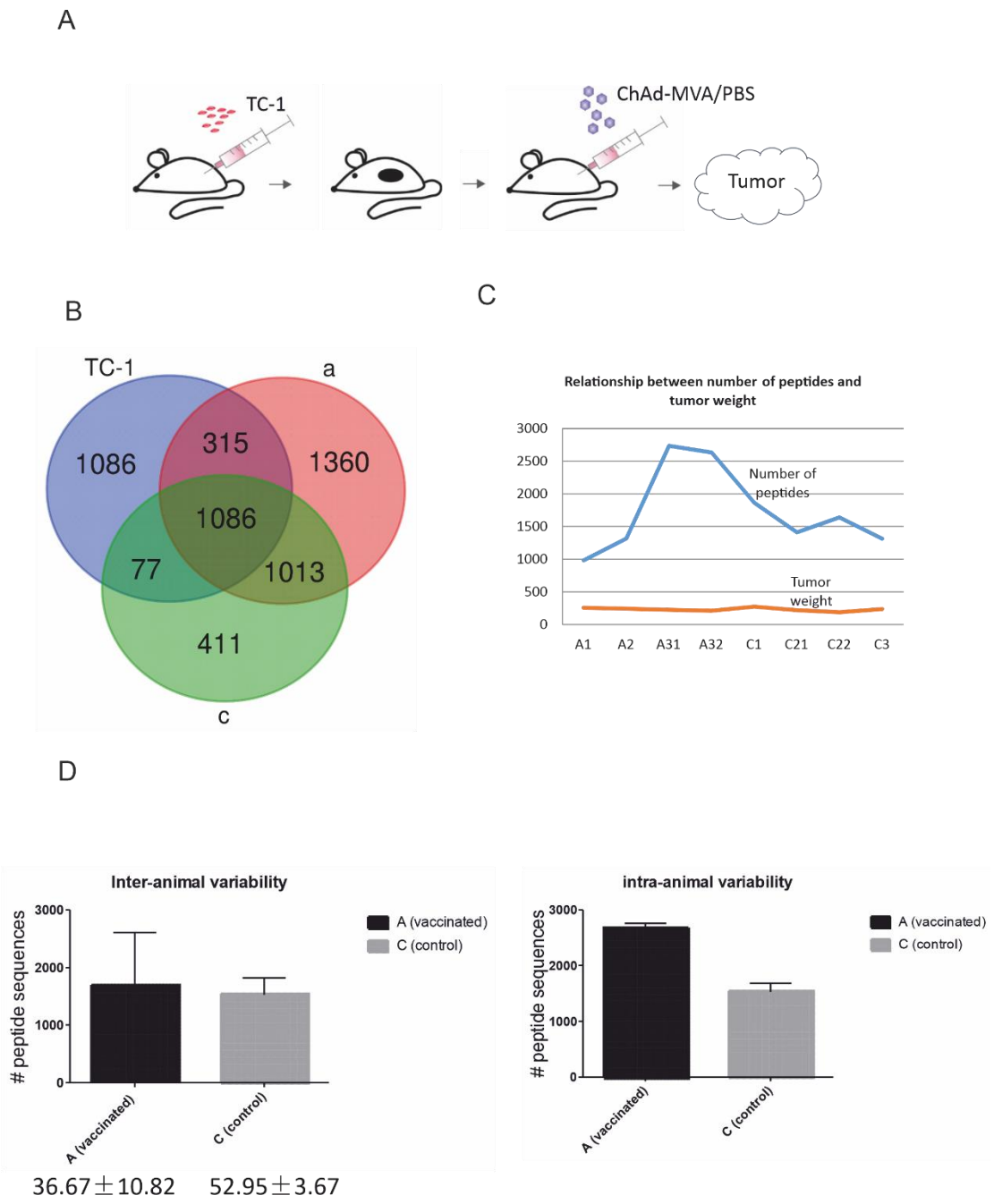


Figure 23. Mouse tumour experiment. A. TC-1 cells were injected into the animals subcutaneously in the inguinal region. Six mice were then divided into two groups after all tumours reached 3mm or more in diameter, and mice were immunized once with ChAd-MVA (Group A) or PBS (Group C). Tumours were harvested one week after vaccination. B. Overlap of peptides identified in our

experiments. Blue is the TC-1 cell line experiment, red is ChAd-MVA (Group A) and green is PBS (Group C). C. Relationship between the number of peptides identified and the tumour weight. D. Peptides variability of inter-animal and intra-animal (n=1)

4.2.2 Identification of tumour-specific antigens in ten human cervical cancer tumours

4.2.2.1 Determination of HLA expression in ten human cervical cancer tumours

We obtained 10 primary human tumour samples for investigation of the clinical relevance and nature of viral and tumour-associated antigen presentation directly in clinical samples. We started by sequencing all ten cervical tumours with Illumina RNA seq. We used OptiType [326] to genotype the patient's HLA background. OptiType is a novel HLA genotyping algorithm which is capable of producing accurate predictions (overall accuracy of 97%) from RNAseq data [326]. We validated the functionality of this approach using the cervical cell line panel, for which both RNAseq and conventional genotyping data was available and confirmed that the software can accurately match the genotype from the RNAseq datasets (see Chapter 3).

We found that the ten human cervical cancer cell tumours express a total of seven different HLA-A alleles, fourteen HLA-B alleles, and ten HLA-C alleles (Table 5). Interestingly, the patient 5 contained only one HLA-B allele and patient 6 and 10 had one HLA-C allele, only, indicating that these patients are likely homozygous in the respective gene locus.

We then proceeded to map the HPV types in each patient using the RNAseq data. There are seven patients who had evidence for HPV-16 sequences in their tumour transcriptome. Furthermore, we found one patient each with evidence for the presence of subtypes HPV-18, HPV-31, and HPV-45 in their transcriptome (Table 5) (Fig. 24).

4.2.2.2 Preparation of personalized protein sequence databases

In order to accurately define the sequence of the viral transcript sequences in the 10 human clinical cervical tumour resections, we applied our *de novo* assembly pipeline as described in Chapter 3. Viral transcript sequences were reconstructed by *de novo* assembly of mRNA transcript sequences that matched the HPV reference genomes (HPV-16 (NC_001526), HPV-18 (NC_001357), HPV-31(gi|333048) and HPV-45 (gi|397022)) using the Trinity tool [224]. The mRNA sequencing data were analysed by Mr. Isaac

Woodhouse (Fig. 24). The sequence information was added to the databases in form of all 6 frame translations of assembled viral transcripts for spectral interrogation. We also downloaded the HPV-16 (NC_001526), HPV-18 (NC_001357), HPV-31(gi|333048) and HPV-45 (gi|397022) genome and created a protein database by translating all open reading frames (ORF) over 8 aa long in all 6 frames as our 6Fr (6-frame translation) database reference.

4.2.2.3 Sequencing of the immunopeptidomes of the human primary tumour panel

At a false discovery rate (FDR) threshold of 5%, we identified overall a total of 78,618 peptides (52,382 unique peptides) from 9,765 proteins (9,215 unique proteins). More than 88% of the peptides have the typical length of HLA-I peptides (8 to 14 amino acids), and most of the presented peptides are nine amino (60%) acids long (Fig. 25 B). Technical replicate analysis also yielded a similar number of HLA peptides in a relatively short measurement time, leading to an increase in identifications for each patient (patient 1 18%, patient 2 34%, patient 3 19%, patient 4 31%, patient 5 23%, patient 6 29%, patient 7 43%, patient 8 16%, patient 9 20%, patient 10 26%) (Fig .26). Generally, unique repertoires of peptides are presented on the surface of cells that express different sets of HLA alleles according to the specificity of their binding motifs.

The ten human cervical cancer tumours contained 31 unique HLA alleles (Table 5), more than 88% of the identified peptides were found to predict to bind to the respective alleles present in each patient (Fig. 25 E). Therefore, the overlap between sequences identified in samples was low, and shared sequences were seen for patients with shared HLA alleles or supertypes as we had previously also observed for our cell line panel (Figure 25 C, D).

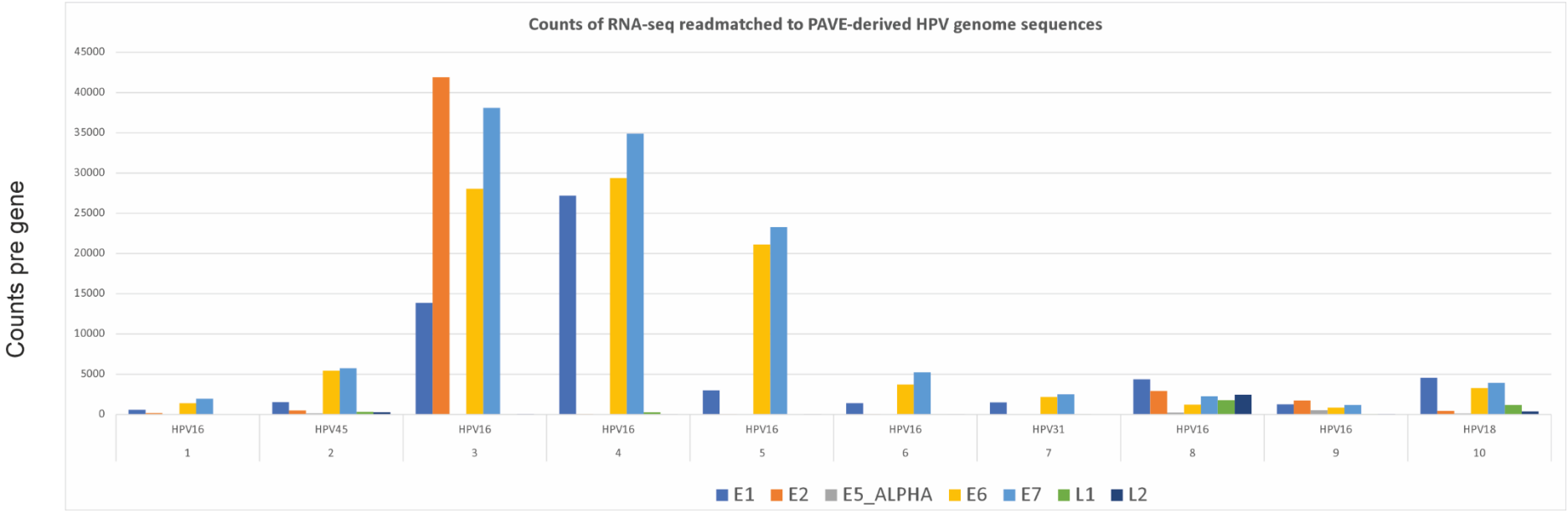


Figure 24. Viral transcript expression levels in ten human cervical cancer tumours. Shown are the transcript counts for the main HPV proteins in the ten human cervical cancer tumours.

Table 5. List of ten human cervical tumours genotypic information

Patient	A1	A2	B1	B2	C1	C2
1 (HPV16)	A*02:01	A*11:01	B*13:02	B*27:05	C*06:02	C*02:02
2 (HPV45)	A*25:01	A*01:01	B*40:01	B*18:01	C*03:04	C*12:03
3 (HPV16)	A*24:02	A*01:01	B*49:01	B*37:01	C*06:02	C*07:01
4 (HPV16)	A*03:01	A*02:01	B*40:02	B*07:02	C*07:02	C*02:02
5 (HPV16)	A*02:01	A*02:01	B*56:01	B*15:01	C*01:02	C*03:04
6 (HPV16)	A*26:01	A*01:01	B*37:01	B*13:02	C*06:02	C*06:02
7 (HPV31)	A*02:01	A*03:01	B*51:01	B*18:01	C*14:02	C*07:01
8 (HPV16)	A*11:01	A*03:01	B*07:02	B*35:01	C*04:01	C*07:02
9 (HPV16)	A*03:01	A*02:01	B*57:01	B*44:02	C*05:01	C*06:02
10 (HPV18)	A*24:02	A*01:01	B*13:02	B*50:01	C*06:02	C*06:02

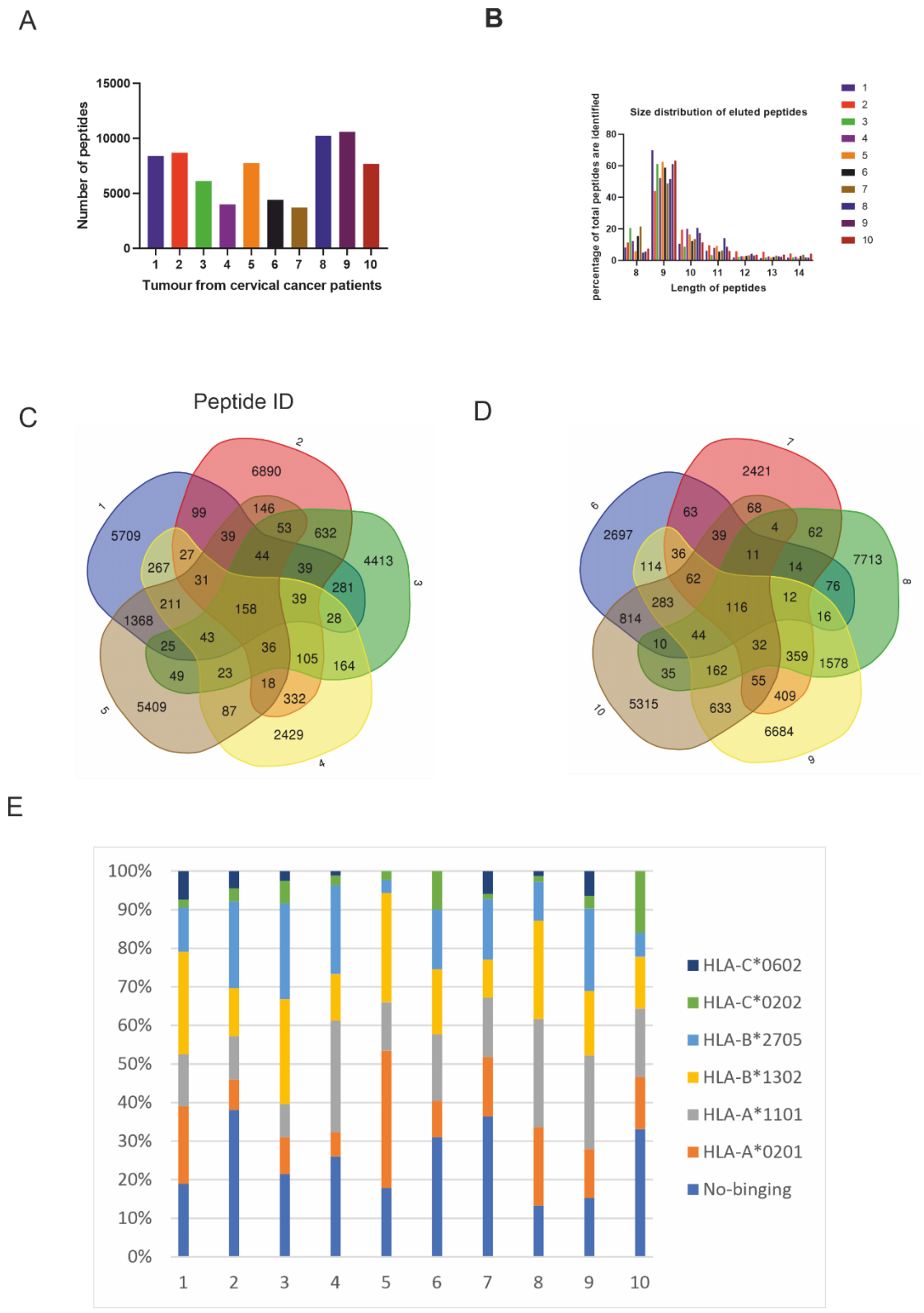
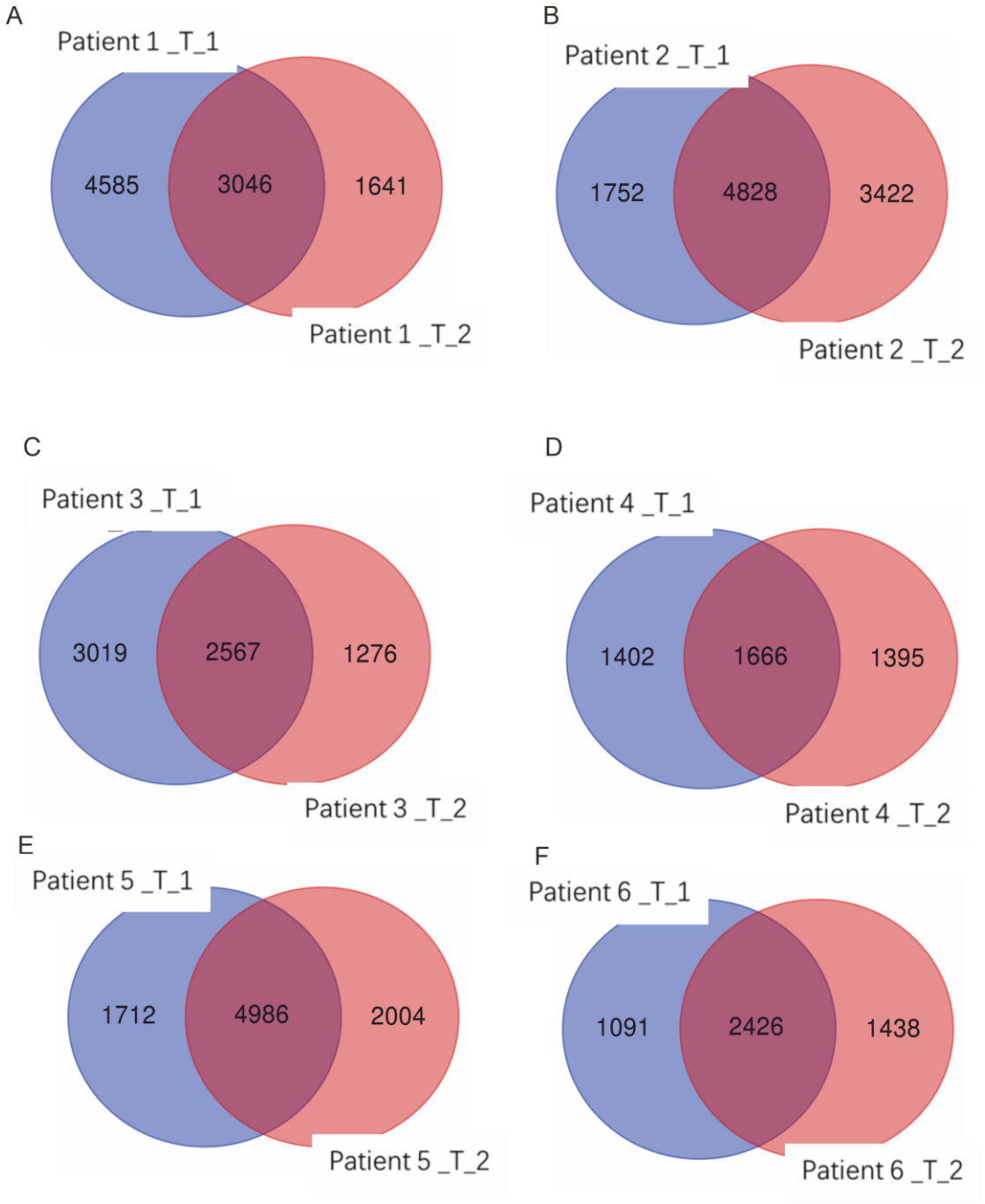


Figure 25. Properties of the eluted peptides identified from ten human cancer tumour samples. A.Total number of peptides identified in each

sample. B. Size distribution of 8-14 amino acids peptides for all samples. C.D. Overlap of peptides identified across samples in two technical replicate analyses. E. Proportion of eluted peptides predicted to bind to the HLA alleles in each patient as indicated.


A
Patient 1 _T_1
4585
3046
1641
Patient 1 _T_2
B
Patient 2 _T_1
1752
4828
3422
Patient 2 _T_2
C
Patient 3 _T_1
3019
2567
1276
Patient 3 _T_2
D
Patient 4 _T_1
1402
1666
1395
Patient 4 _T_2
E
Patient 5 _T_1
1712
4986
2004
Patient 5 _T_2
F
Patient 6 _T_1
1091
2426
1438
Patient 6 _T_2

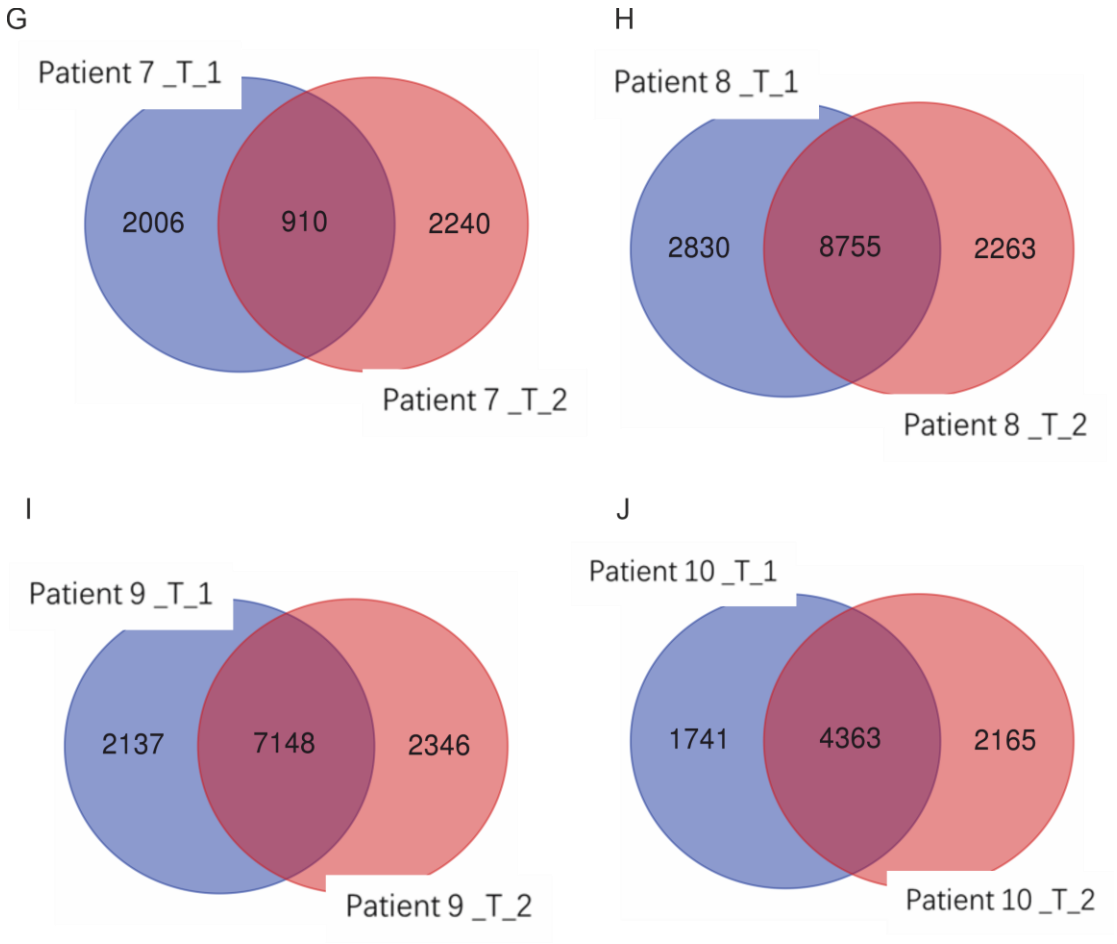


Figure 26. Comparison of eluted peptides identified in two technical replicates from ten human cancer tumour samples. Venn diagramms sho the number of sequences identified in each individual replicate analysis.

4.2.2.4 Assignment of HLA-allele orgin for identified peptide ligands with Gibbs cluster and NetMHCpan

We also employed advanced neural-network-based algorithms (Gibbs Cluster and NetMHCpan) [232, 233] to predict HLA-binding and to compare our data with known HLA ligand data from the IEDB (http://www.iedb.org) in order to

evaluate the quality of our results. With the help of NetMHCpan, we successfully clustered the 69,182 peptides (8 to 14 amino acids) to the 31 unique HLA-I receptors for each individual dataset and cell line (Fig 27-31). Using the Gibbs clustering algorithm [232], we could successfully resolve the motifs and their relative contribution to the total presentation from those tumour samples. However, clustering was limited in some of our data, and 9-mer peptides were clustering to less than five alleles in some patients. In addition, the majority of the lengths of the presented peptides were nine amino acids long. We compared the nine lengths of presented antigens data between the ten cervical cancer tumours and with the known HLA-binding alleles in the IEDB database. For example, similar motifs were obtained for HLA-A02:01 in patient 1 and patient 4 and for HLA-A01:01 in patient 2, patient 6 and patient 10 (Fig. 32). Moreover, we could assign the HLA-C07:02 allele motif for patient 8 based on our genotyping information (Table 5) (Fig.32 A-J).

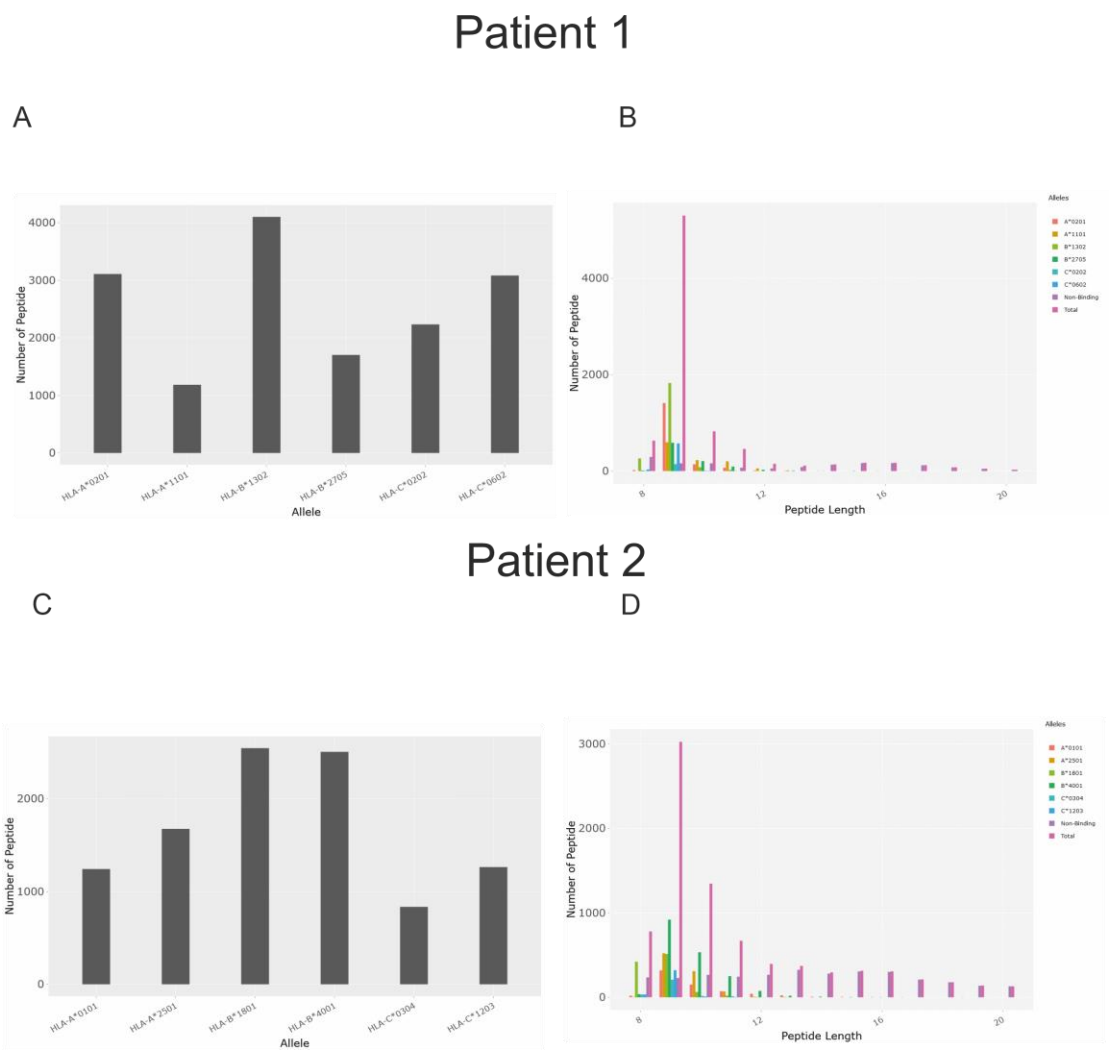


Figure 27. Size distribution and NetMHCpan binding prediction of peptides detected for patient 1 and 2. A. and C. NetMHCpan binding prediction for indicated patients. B. and D. Size distribution of peptides stratified to their allele of origin by the NetMHC pan binding prediction for the indicated patient.

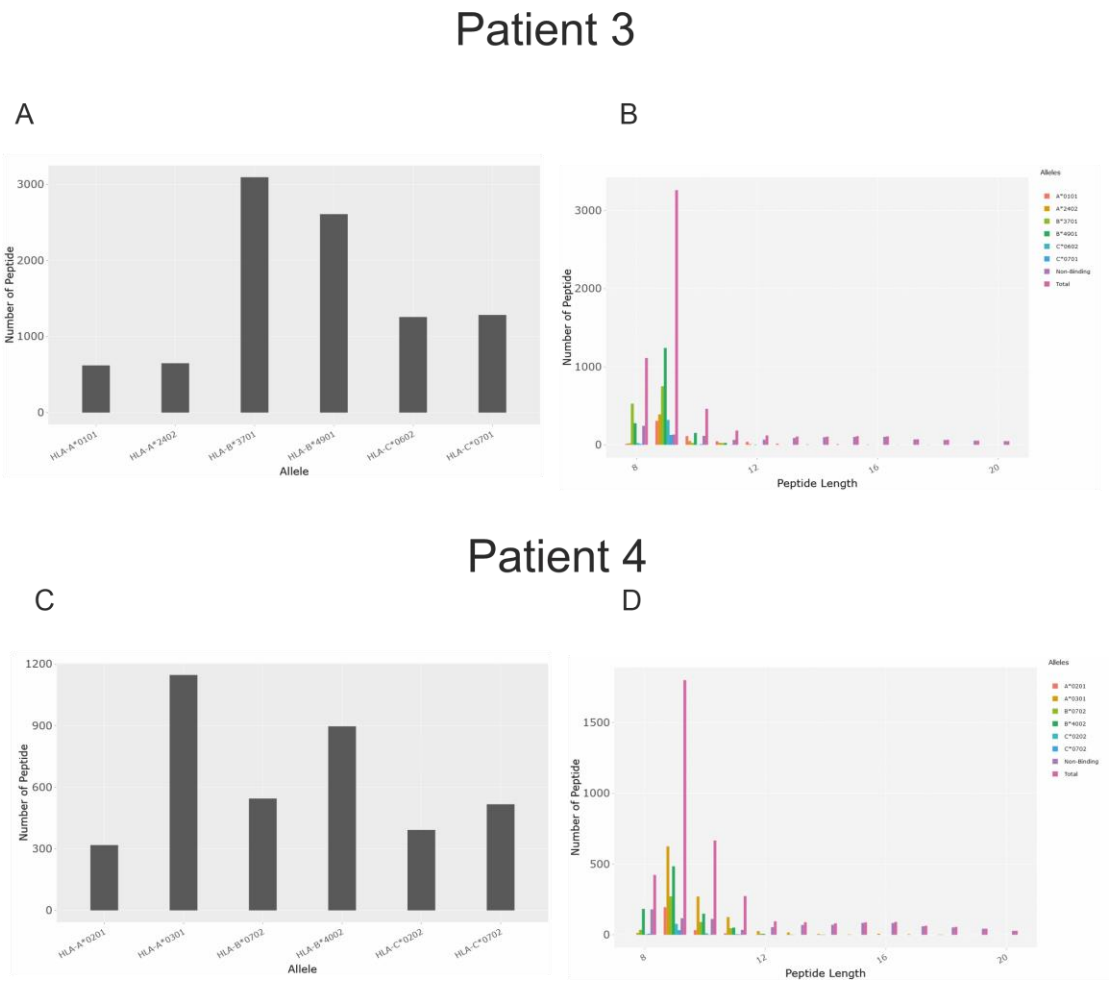


Figure 28. Size distribution and NetMHCpan binding prediction of peptides detected for patient 3 and 4. A. and C. NetMHCpan binding prediction for indicated patients. B. and D. Size distribution of peptides stratified to their allele of origin by the NetMHC pan binding prediction for the indicated patient.

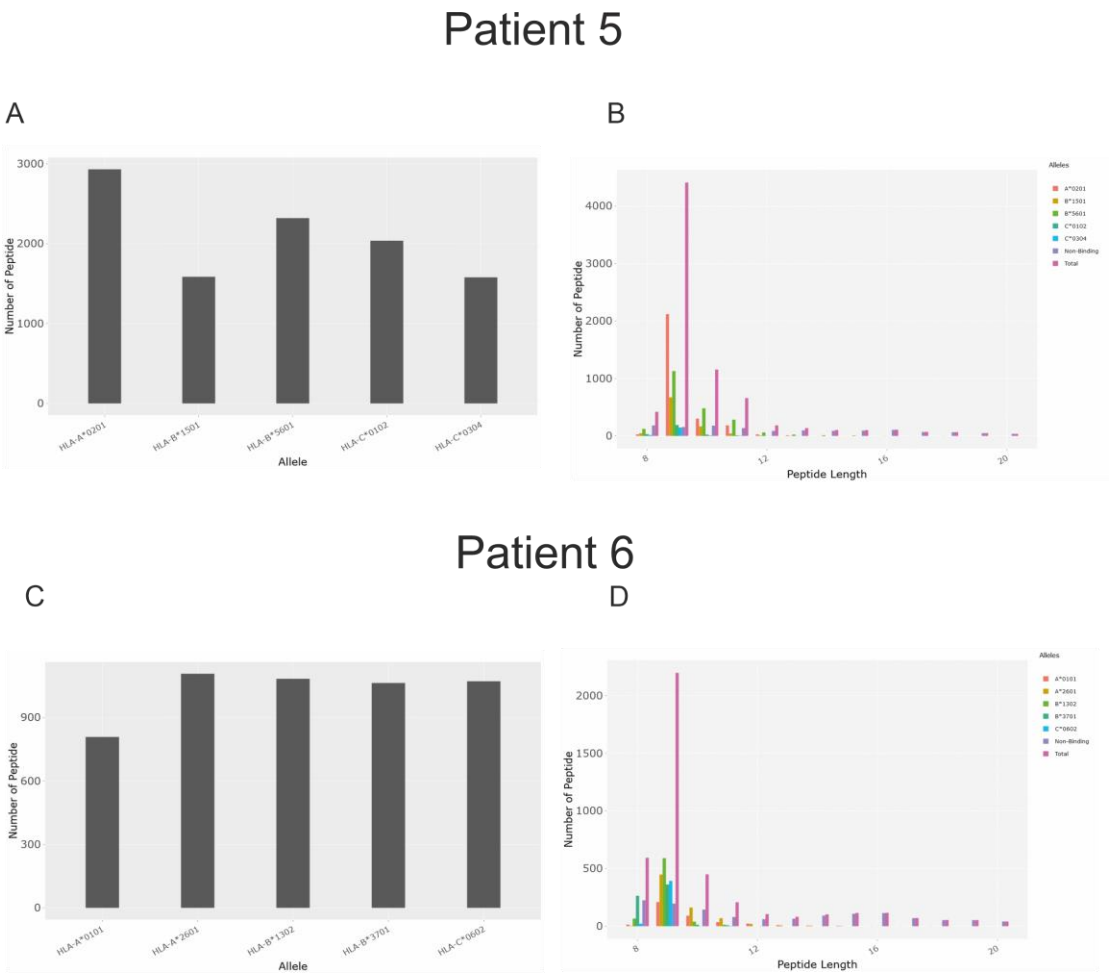


Figure 29. Size distribution and NetMHCpan binding prediction of peptides detected for patient 5 and 6. A. and C. NetMHCpan binding prediction for indicated patients. B. and D. Size distribution of peptides stratified to their allele of origin by the NetMHC pan binding prediction for the indicated patient.

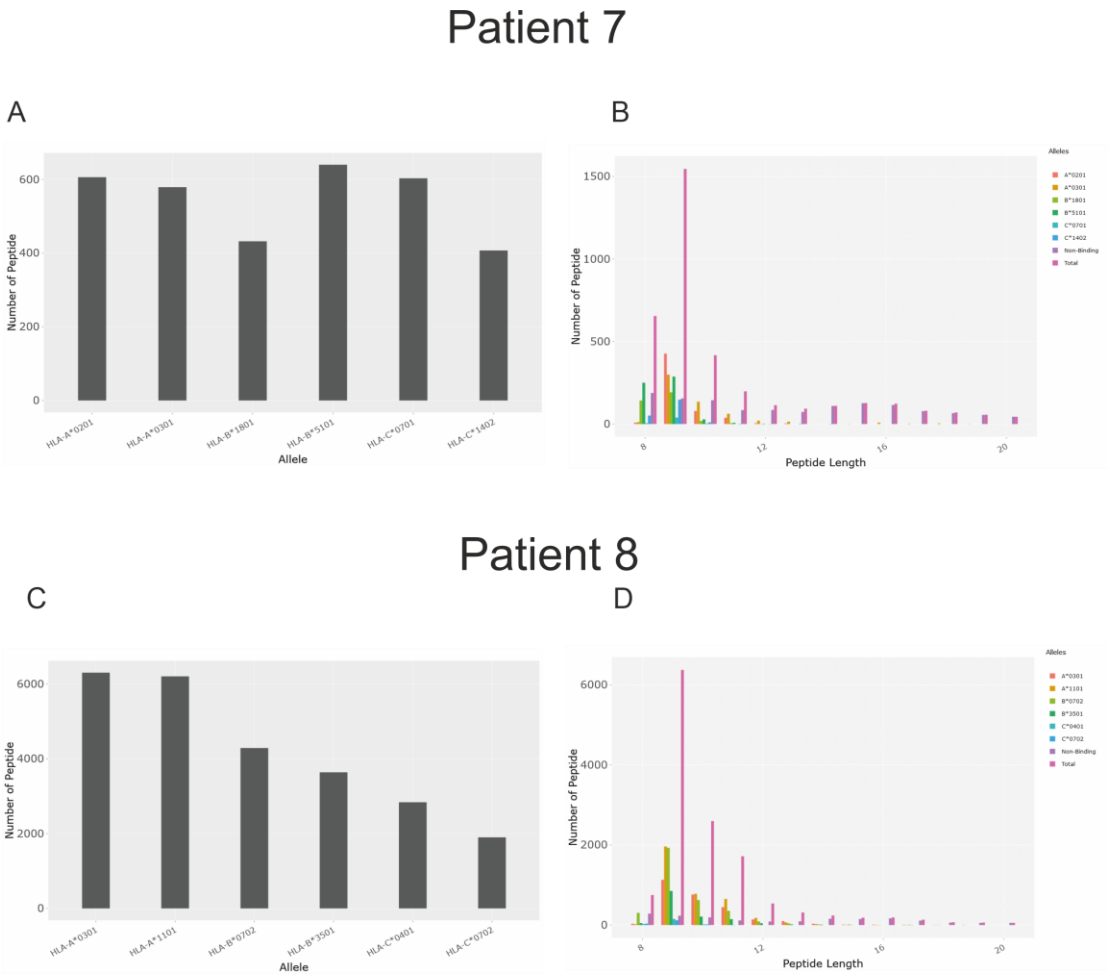


Figure 30. Size distribution and NetMHCpan binding prediction of peptides detected for patient 7 and 8. A. and C. NetMHCpan binding prediction for indicated patients. B. and D. Size distribution of peptides stratified to their allele of origin by the NetMHC pan binding prediction for the indicated patient.

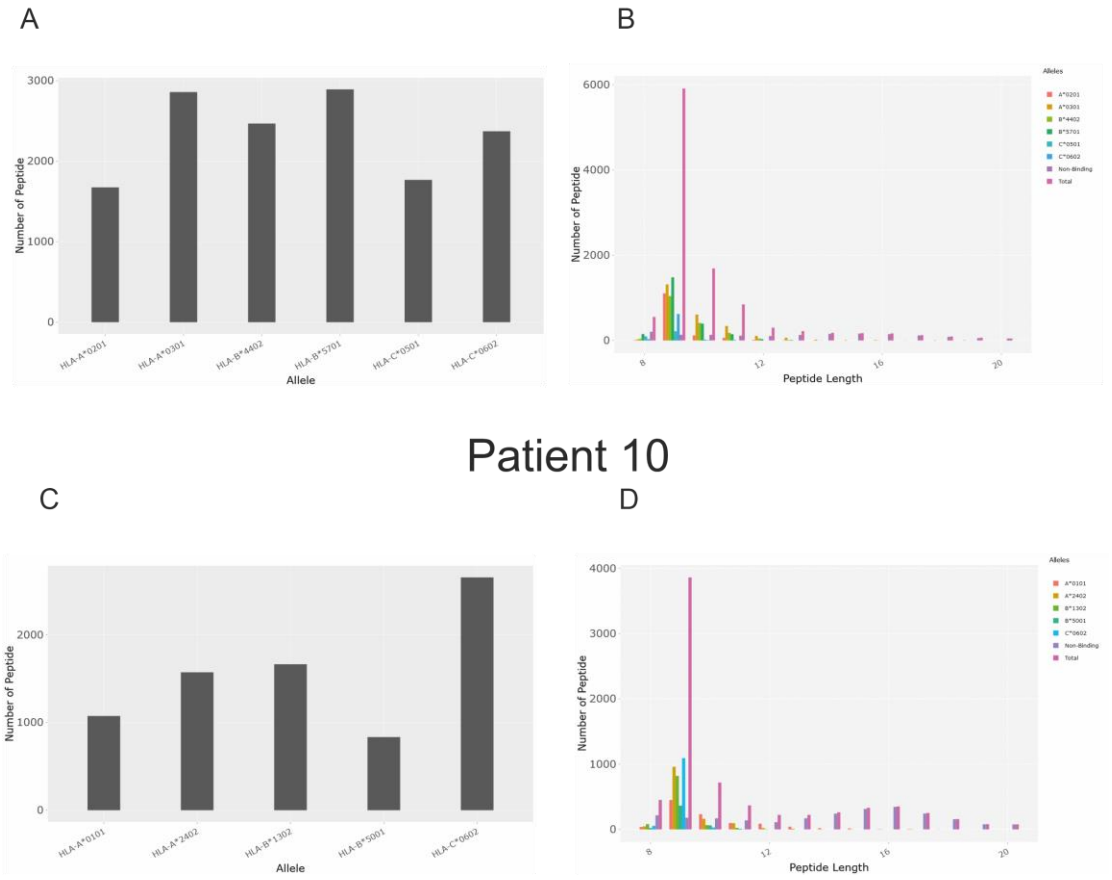
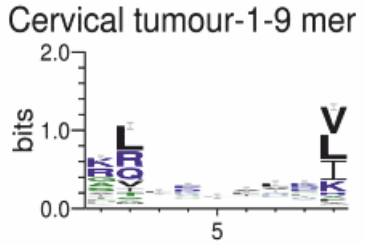


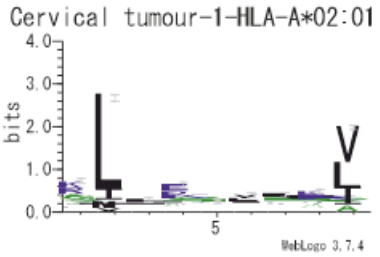
Figure 31. Size distribution and NetMHCpan binding prediction of peptides detected for patients 9 and 10. A. and C. NetMHCpan binding prediction for indicated patients. B. and D. Size distribution of peptides stratified to their allele of origin by the NetMHC pan binding prediction for the indicated patient.

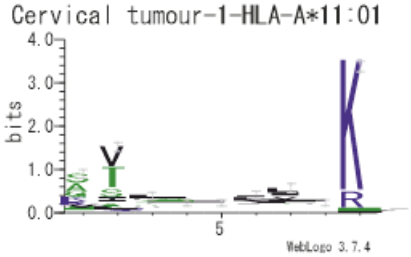
A

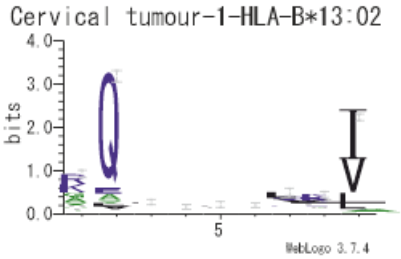
Patient 1

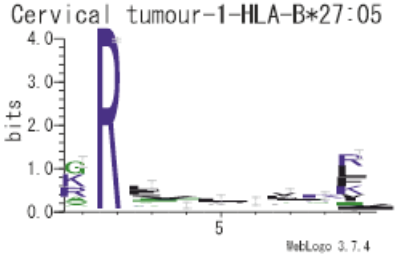


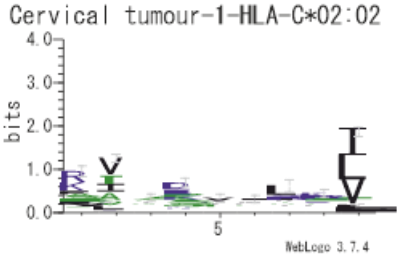
NetMHCpan

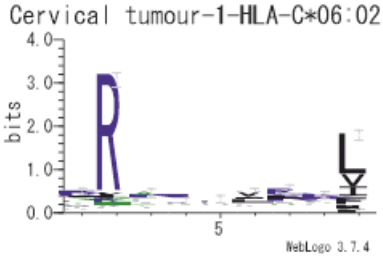




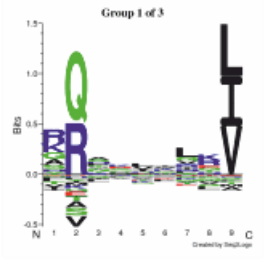


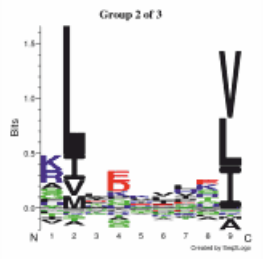


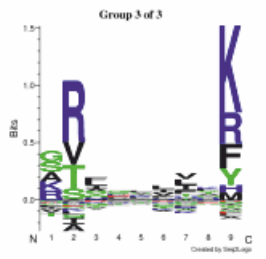




Gibbs Cluster

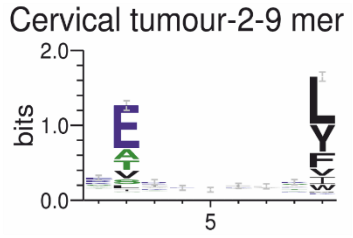




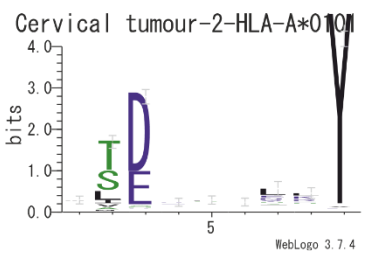


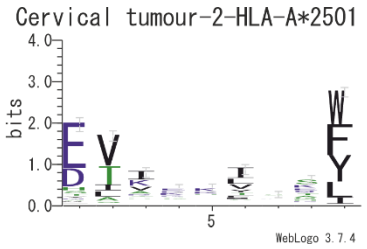
B

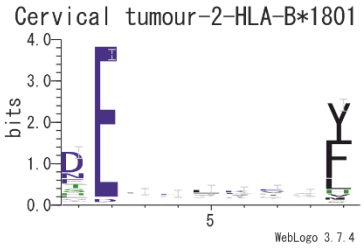
Patient 2

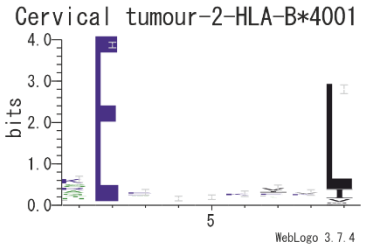


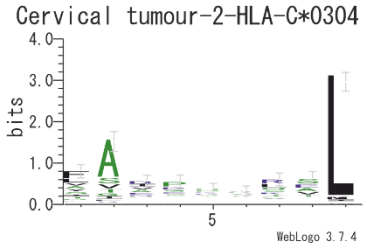
NetMHCpan

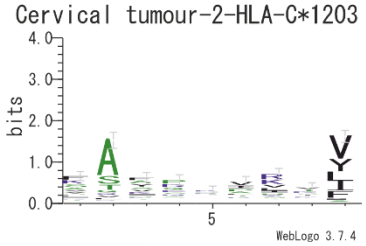




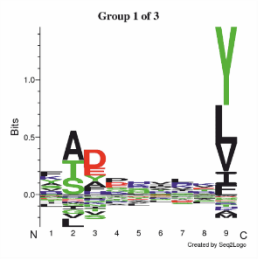


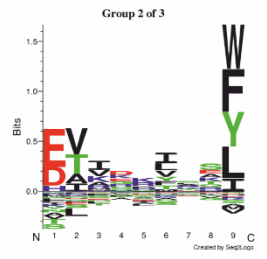


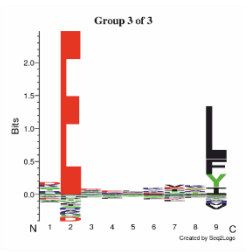




Gibbs Cluster

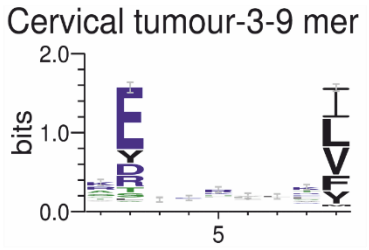




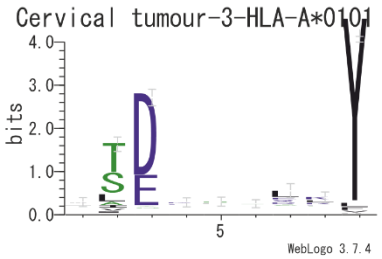


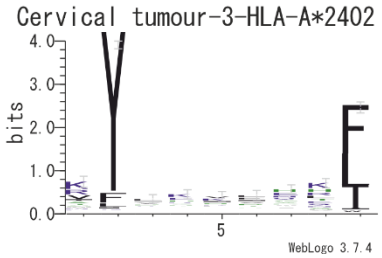
C

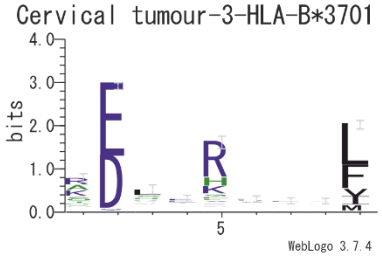
Patient 3

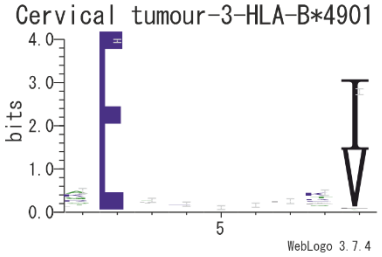


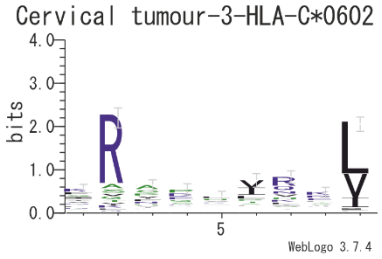
NetMHCpan

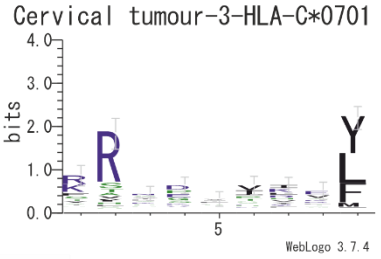




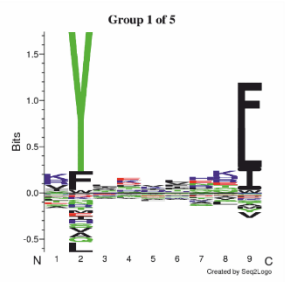


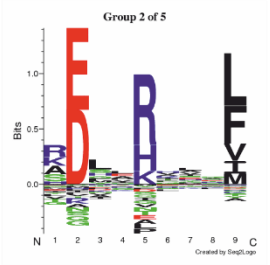


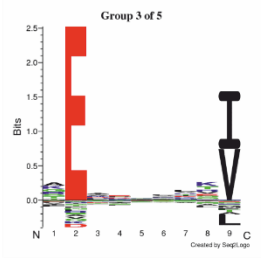


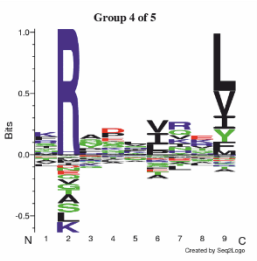


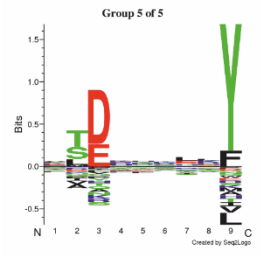
Gibbs Cluster





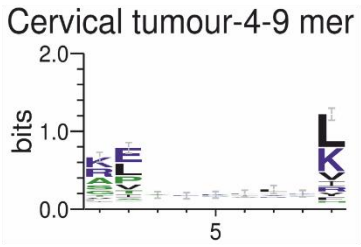




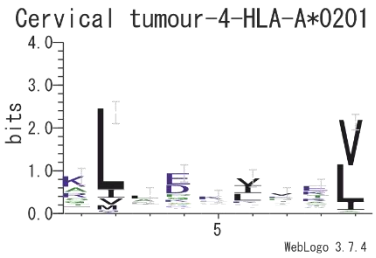


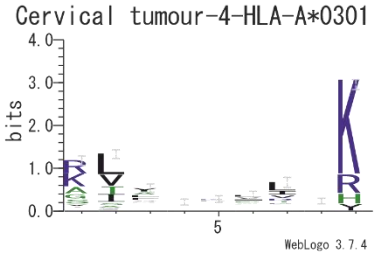
D

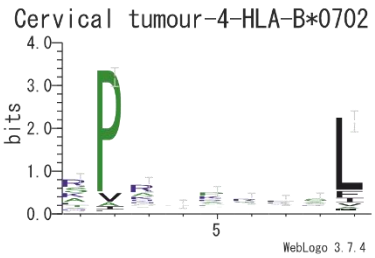
Patient 4

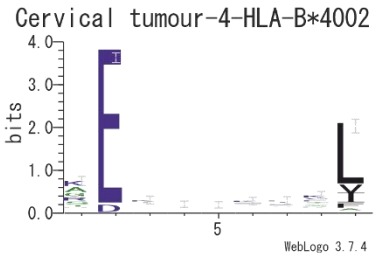


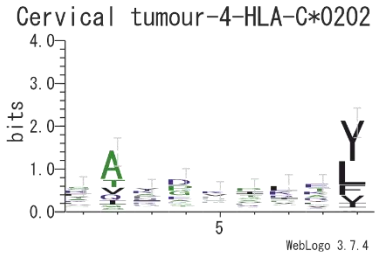
NetMHCpan

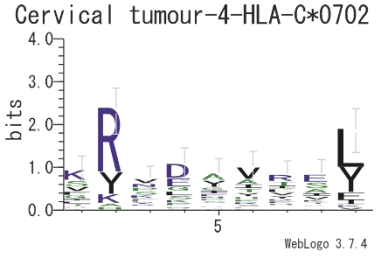




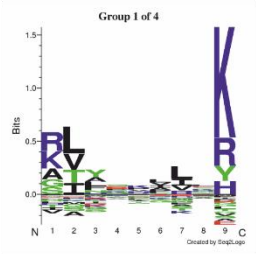


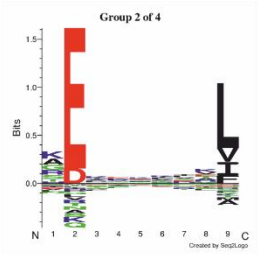


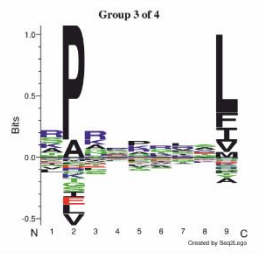


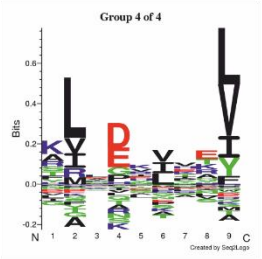


Gibbs Cluster



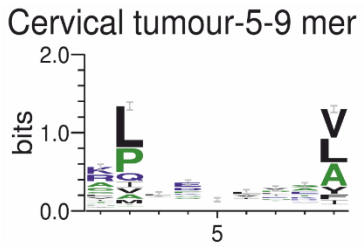




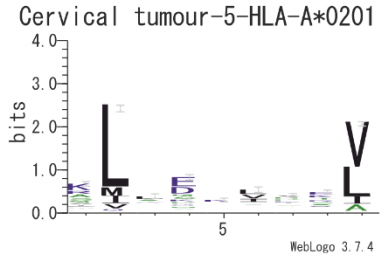


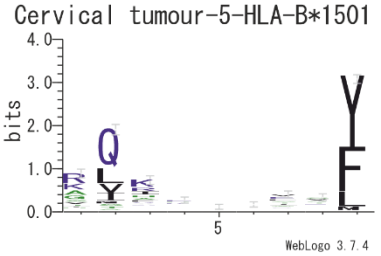
E

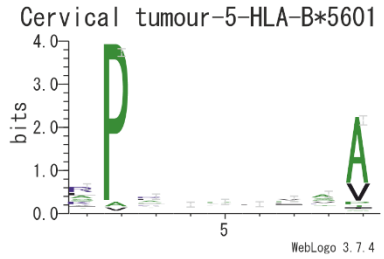
Patient 5

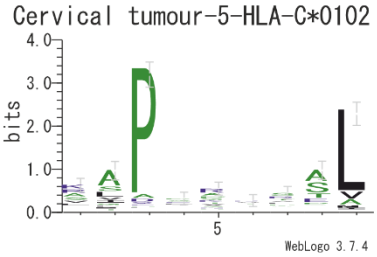


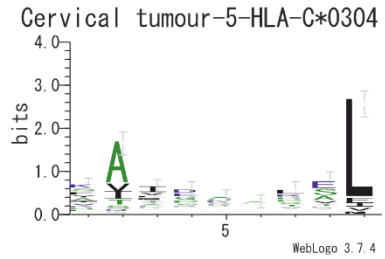
NetMHCpan



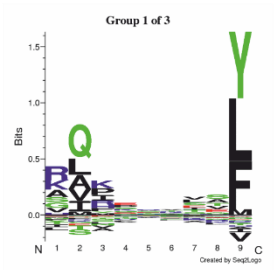


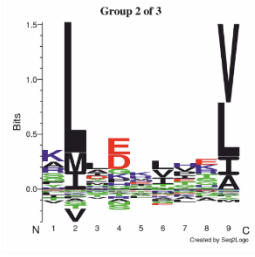


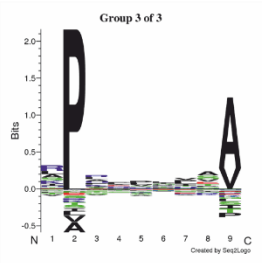




Gibbs Cluster

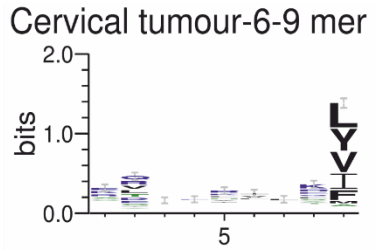




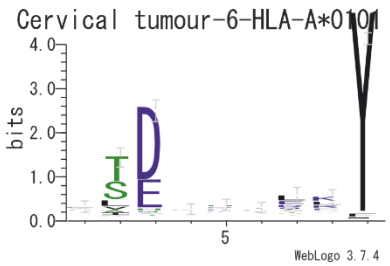


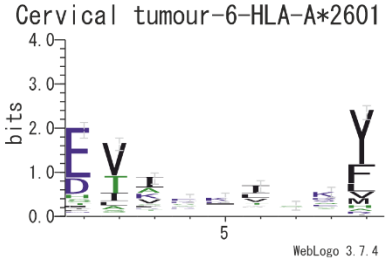
F

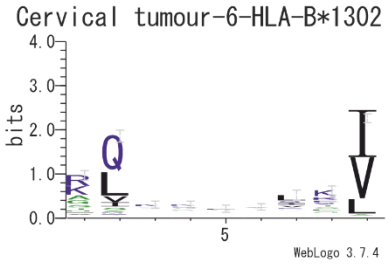
Patient 6

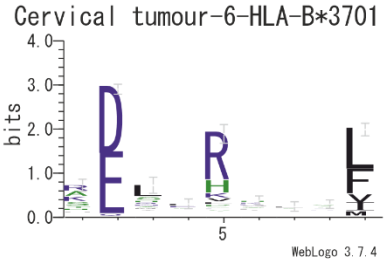


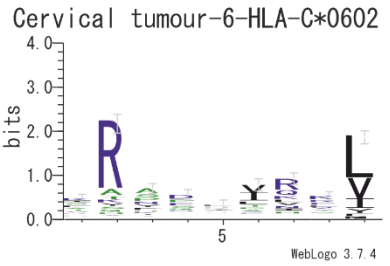
NetMHCpan



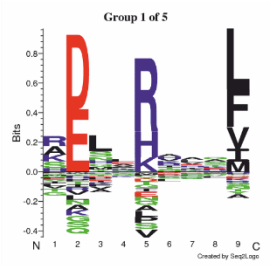


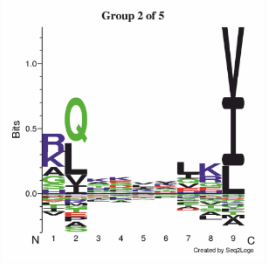


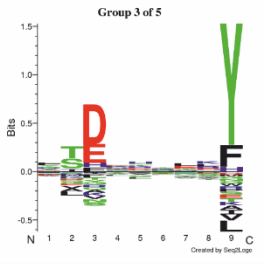


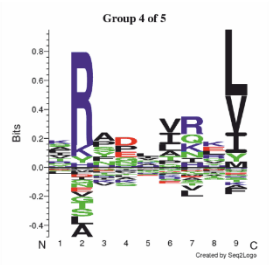


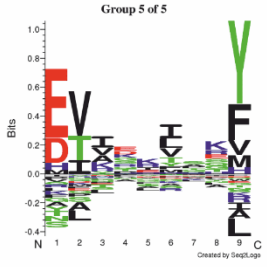
Gibbs Cluster





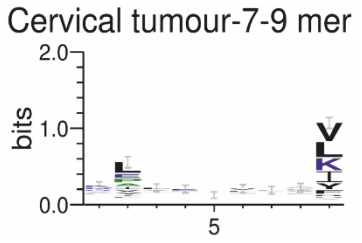




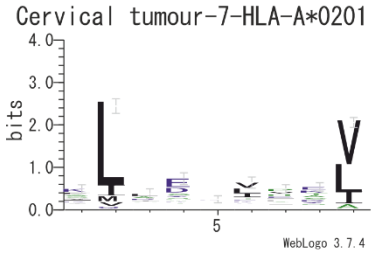


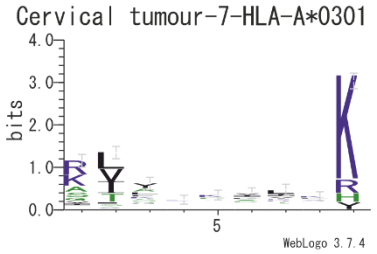
G

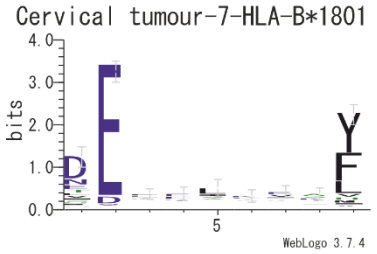
Patient 7

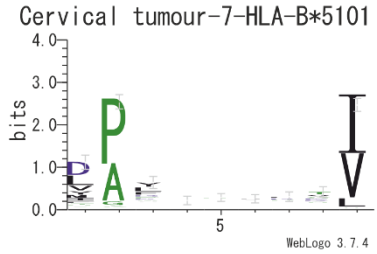


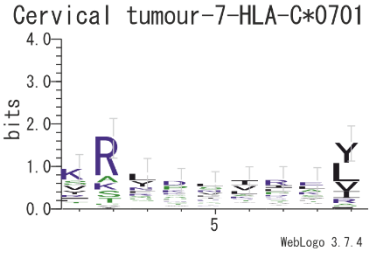
NetMHCpan

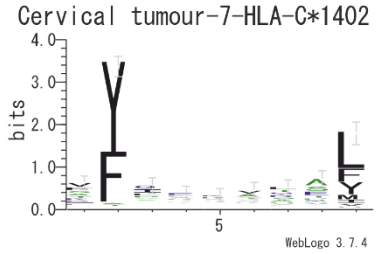




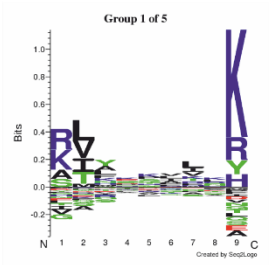


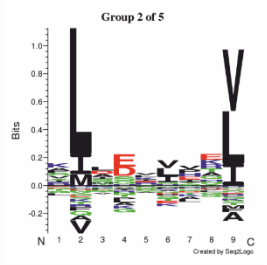


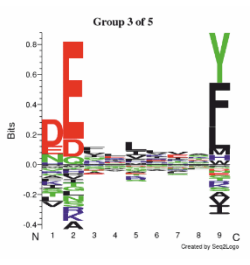


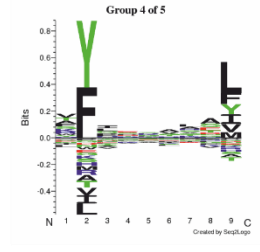


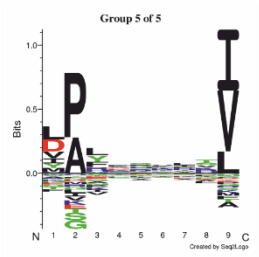
Gibbs Cluster





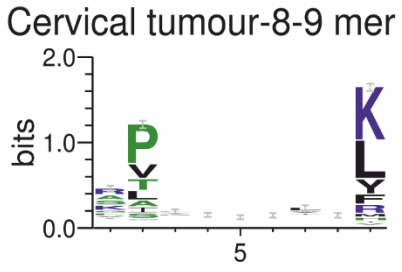




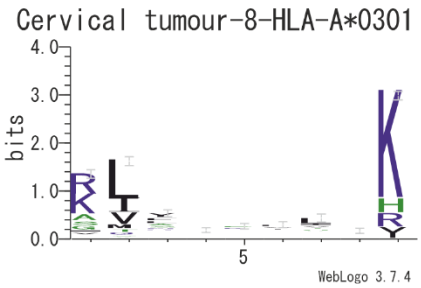


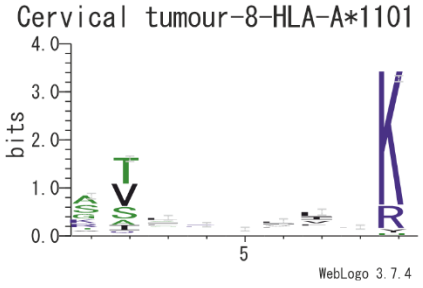
H

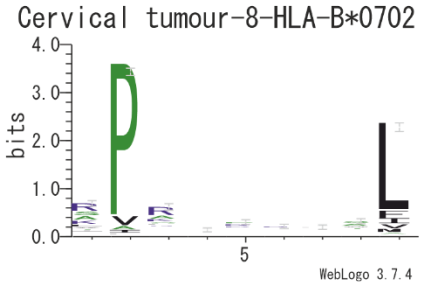
Patient 8

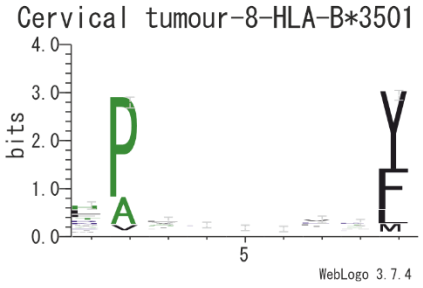


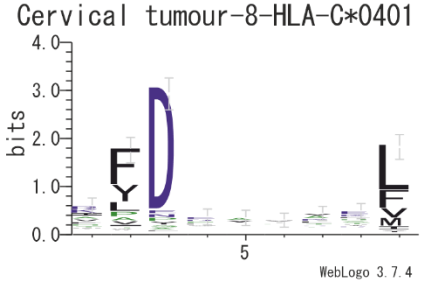
NetMHCpan

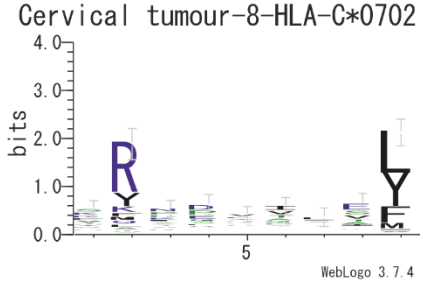




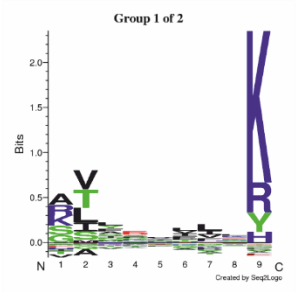


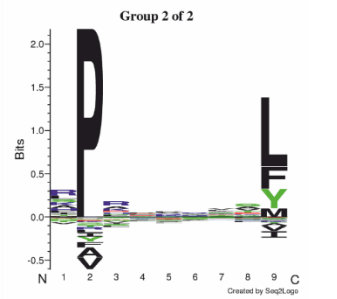




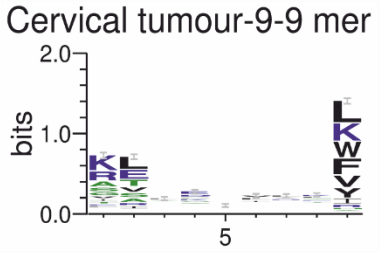


Gibbs Cluster

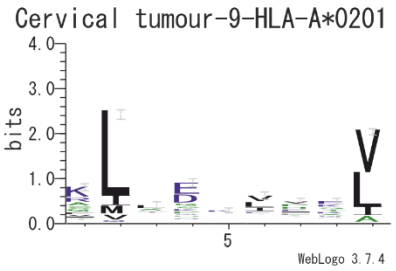


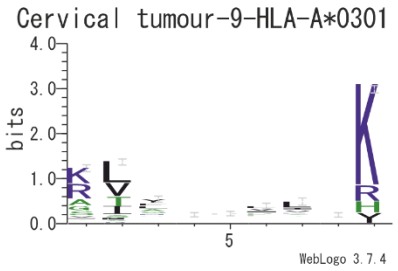


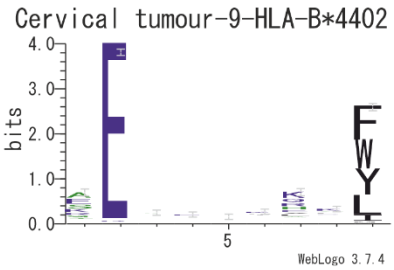
Patient 9

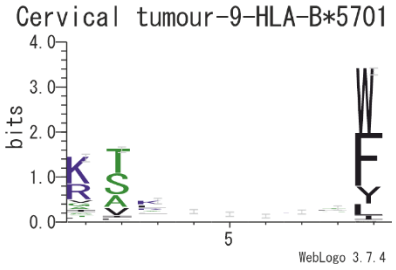


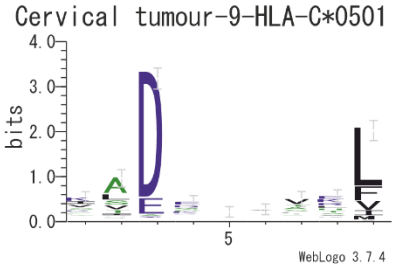
NetMHCpan

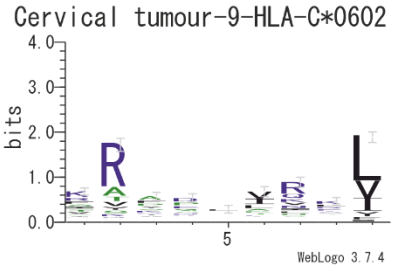




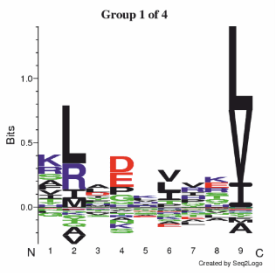


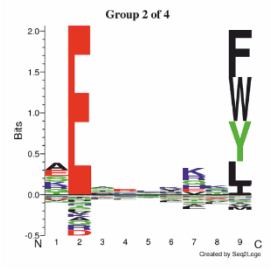


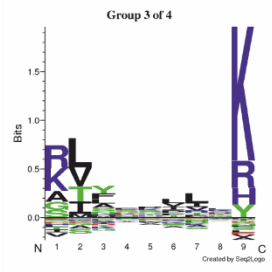


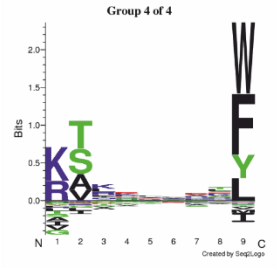


Gibbs Cluster



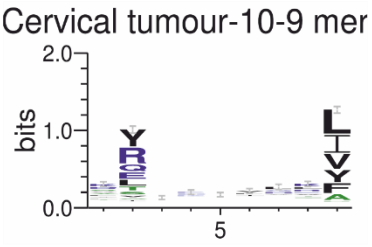




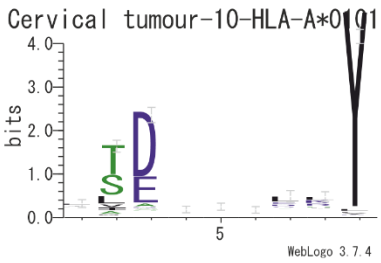


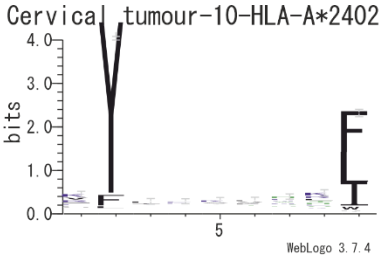
J

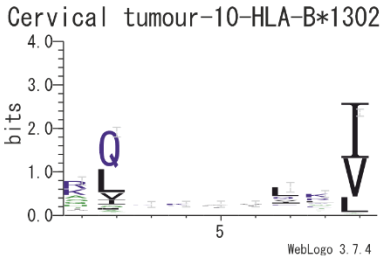
Patient 10

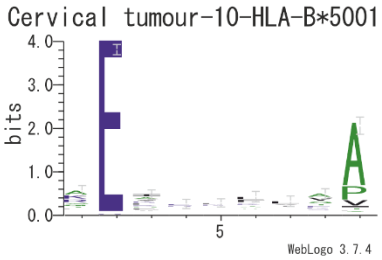


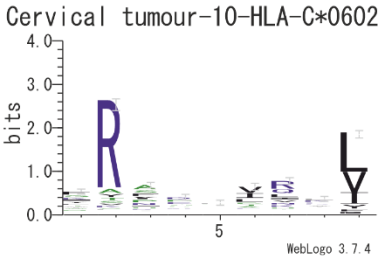
NetMHCpan



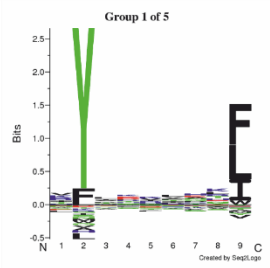


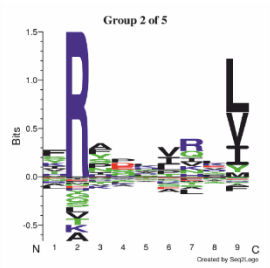


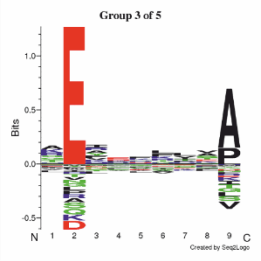


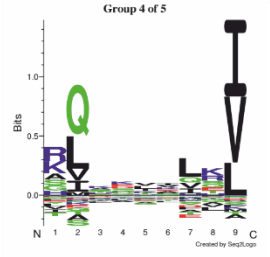


Gibbs Cluster









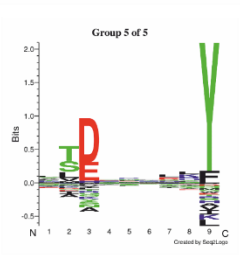


Figure 32. A-J. Amino acid distribution of all eluted 9-mer peptides from primary tumour resections for each patient. The top panel showns the cumulative motif of all 9mers detected for each patient, the middle panel of motifs shows the motifs obtained when peptides are stratified with NetMHCpan to their allele of origin, and the lower panel shows the unsupervised GibbsCluster analysis from each patient, demonstrating the sequence relationships in a prediction-independent fashion.

4.2.3 Viral peptides identified in ten human cervical cancer tumours

Using our RNA-seq data (Trinity database) [224] to build specific databases for each of the patients that included the viral transcript translations, we identified 6 viral peptides from our tumour samples. 50% of them were strong binders to HLA-I (Table 6). However, in patient 3, FEQKGFRLV (FV9) was also share the same sequence with human NME1 and NME2 proteins. Unfortunately, we also could not validate the four other viral peptide sequence candidates by matching to the spectra obtained from synthetic counterparts and had to assume that they were false positive identifications.

Table 6. List of viral peptides identified in ten human cervical tumours. *Trinity represents a novel method for the efficient and robust *de novo* reconstruction of transcriptomes from RNA-seq data. 6Fr is a six-frame translation for the obtained assembled HPV RNA transcript sequences

Patient	Peptide	-10lgP	Mass	Length	m/z	Binder	Rank
2	ICLNGLNIDV	22.91	1072.559	10	537.2853	HLA-A*01:01	53.387
3	FEQKGFRLV	38.45	1122.619	9	562.3163	HLA-B*49:01	0.042
5	APSRLPGV	28.09	795.4603	8	398.7385	HLA-B*56:01	1.857
7	CTCLNCQG	16.19	840.2928	8	421.1523	HLA-A*03:01	87.5
7	TLLQHLGLTL	16.39	1107.665	10	554.8403	HLA-A*02:01	0.652

4.2.4 Several important CT,Raven and protein atlas antigens were identified in the ten human cervical cancer tumours

In order to shortlist known testis antigens that are presented by HLA in the interrogated tumour tissues, we utilized three different open-source databases: the Cancer-testis antigens database (CT database [239], http://www.cta.lncc.br/) and RAVEN[240] (Rich Analysis of Variable gene Expressions in Numerous tissues; https://github.com/JSGerke/RAVENsoftware), and The Human Protein Atlas [241] (HPA, https://www.proteinatlas.org/) [240] as we had done before for interrogation the cell line data.

We identified a total of 236 peptides from 47 proteins from CT database across the cell lines (Fig. 33). Lysine-specific demethylase 5B (KDM5B) (Table 7) was

presented by all ten cervical tumour samples. In addition, we identified 212 antigens mapping to the RAVEN database (Fig. 33), and one antigen (HIST1H1C) was presented by all ten cervical tumour samples. KDM5B was a tumour-associated antigen that stood out as it was also detected to be presented by all 5 cervical tumour cell lines (see Chapter 3).

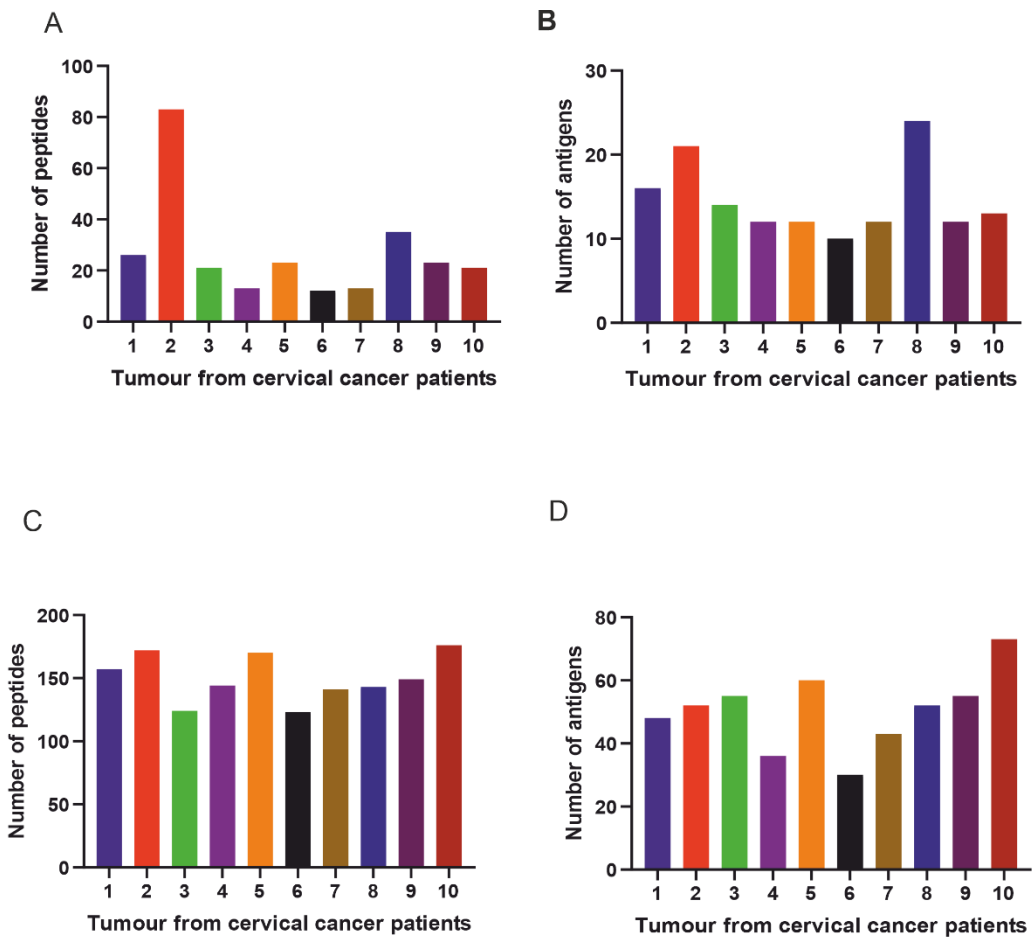


Figure 33: **Identification of CT and Raven antigens in ten cervical cancer resections.** (A) Number of peptides defined by the CT database across the

ten cervical cancer tumours. (B) Number of source proteins defined by the CT database across the ten cervical cancer tumours. (C) Number of peptides defined by the Raven database across the ten cervical cancer tumours. (D) Number of source proteins defined by the Raven database across the ten cervical cancer tumours.

Table 7. List of peptides from KDM5B which was presented in all five human cancer cell lines and ten human cervical tumours

Name/Peptide	Score	Mass	Length	Antigen	Allele_Rank	Allele
C33						
RPRHSLNSL	24.47	1078.6	9	Q9UGL1|KDM5B_HUMAN	0.0039	HLA-B*07:02
GADIASKEF	22.21	936.4552	9	Q9UGL1|KDM5B_HUMAN	0.2403	HLA-C*05:01
RLIDLGVGL	16.99	954.5862	9	Q9UGL1|KDM5B_HUMAN	0.183	HLA-A*02:01
CaSki						
NDLLRHLRL	33.13	1148.678	9	Q9UGL1|KDM5B_HUMAN	0.0896	HLA-B*37:01
KLKEPLPNGKK	29.43	1250.771	11	Q9UGL1|KDM5B_HUMAN	0.308	HLA-A*03:01
RLIDLGVGL	22.09	954.5862	9	Q9UGL1|KDM5B_HUMAN	0.183	HLA-A*02:01
RPRHSLNSL	20.04	1078.6	9	Q9UGL1|KDM5B_HUMAN	0.0039	HLA-B*07:02
HPGPRPAL	17.75	843.4715	8	Q9UGL1|KDM5B_HUMAN	0.1128	HLA-B*07:02
DoTc2						
KLKEPLPNGKK	43.48	1250.771	11	Q9UGL1|KDM5B_HUMAN	0.308	HLA-A*03:01
KLKEPLPNGK	29.7	1122.676	10	Q9UGL1|KDM5B_HUMAN	0.059	HLA-A*03:01
KILDLFQLNK	35.14	1230.734	10	Q9UGL1|KDM5B_HUMAN	0.0577	HLA-A*03:01
HeLa						
GKRDGINSL	23.58	958.5195	9	Q9UGL1|KDM5B_HUMAN	3.1334	HLA-B*15:03
ETASAMATL	22.98	893.4164	9	Q9UGL1|KDM5B_HUMAN	0.0195	HLA-A*68:02
SiHa						
DQIAKYWEL	32.68	1164.582	9	Q9UGL1|KDM5B_HUMAN	1.9612	HLA-B*40:02
HEVPVYRT	28.26	999.5137	8	Q9UGL1|KDM5B_HUMAN	4.379	HLA-B*40:02
REMEALQSL	22.13	1075.533	9	Q9UGL1|KDM5B_HUMAN	0.0063	HLA-B*40:02
1						
AQTRVKLNF	17.05	1075.614	9	Q9UGL1|KDM5B_HUMAN	0.464	HLA-B*1302
RLIDLGVGL	15.32	954.5862	9	Q9UGL1|KDM5B_HUMAN	0.045	HLA-A*02:01
2						
DDWDNRTSY	18.88	1170.458	9	Q9UGL1|KDM5B_HUMAN	0.061	HLA-B*1801
LVTIMNPNTL	15.79	1114.606	10	Q9UGL1|KDM5B_HUMAN	2.999	HLA-C*0304
3						
NDLLRHLRL	31.46	1148.678	9	Q9UGL1|KDM5B_HUMAN	0.075	HLA-B*3701
RETVRKLGV	25.4	1056.64	9	Q9UGL1|KDM5B_HUMAN	0.374	HLA-B*4901
REMEALQSL	23.1	1075.533	9	Q9UGL1|KDM5B_HUMAN	0.004	HLA-B*3701
4						
KLKEPLPNGK	28.81	1122.676	10	Q9UGL1|KDM5B_HUMAN	0.011	HLA-A*0301
RPRHSLNSL	19.62	1078.6	9	Q9UGL1|KDM5B_HUMAN	0.001	HLA-B*0702
5						
APYSAVEKAMA	50.59	1136.554	11	Q9UGL1|KDM5B_HUMAN	0.182	HLA-B*5601
APYSAVEKA	42.71	934.476	9	Q9UGL1|KDM5B_HUMAN	0.012	HLA-B*5601
ALKDSVQRA	34.16	986.5508	9	Q9UGL1|KDM5B_HUMAN	0.46	HLA-A*0201
SIPVHLNSL	16.83	978.5498	9	Q9UGL1|KDM5B_HUMAN	0.005	HLA-C*0102
RLIDLGVGL	15.77	954.5862	9	Q9UGL1|KDM5B_HUMAN	0.053	HLA-A*0201
6						
QCDGSCNQWF	18.81	1186.417	10	sp|Q9UGL1|KDM5B_HUMAN	4.008	HLA-A*01:01
7						
KLKEPLPNGK	28.99	1122.676	10	Q9UGL1|KDM5B_HUMAN	0.011	HLA-A*0301
8						
FPVRDGKIKL	28.41	1171.708	10	Q9UGL1|KDM5B_HUMAN	0.075	HLA-B*0702
HPGPRPAL	18.9	843.4715	8	Q9UGL1|KDM5B_HUMAN	0.023	HLA-B*0702
9						
KLKEPLPNGK	34.55	1122.676	10	Q9UGL1|KDM5B_HUMAN	0.011	HLA-A*0301
KLKEPLPNGKK	32.3	1250.771	11	Q9UGL1|KDM5B_HUMAN	0.046	HLA-A*0301
SVAQQLLNGK	31.62	1056.593	10	Q9UGL1|KDM5B_HUMAN	0.133	HLA-A*0301
RIRLEQARW	30.32	1226.7	9	Q9UGL1|KDM5B_HUMAN	0.199	HLA-B*5701
REMEALQSL	26.82	1075.533	9	Q9UGL1|KDM5B_HUMAN	0.059	HLA-B*4402
LHWGEPKTW	18.76	1152.572	9	Q9UGL1|KDM5B_HUMAN:P	1.287	HLA-B*4402
10						
RLIDLGVGL	17.39	954.5862	9	Q9UGL1|KDM5B_HUMAN	0.135	HLA-B*1302

4.2.5 Long non-coding RNA in ten human cervical cancer tumours

Long ncRNAs were found to be deregulated in several human cancers and show tissue-specific expression. As a potential source for tumour-associated antigens, we performed an analysis aiming to identify lncRNA-derived peptides in our 10 human cervical tumour immunopeptidomics datasets.

For this, we utilised a novel bioinformatics approach developed by the group, which uses a database-independent approach to identify peptide sequences from LC-MS/MS data. Peptide sequences candidates reported by the Peaks *de novo* search (Bioinformatics solutions) are re-ranked using this tool to select the most likely correct peptide sequence candidate for each spectrum. MAR-S is a linear model that integrates both HLA binding prediction and predicted chromatographic retention time for this re-ranking, and reports the highest scoring peptide sequence as the PSM for each spectrum. Spectra that were matching against the canonincal sequence database were excluded from this assessment, and peptide sequences that were found to have a single amino acid difference to the canoncial proteome were also excluded from the list of sequences that were then searched against two lncRNA databases: Lncipedia [327] and LNCRNAK [231].

We used highly scoring peptide candidates (MAR-S score > 90), and excluded sequences that had more than 3 leucine or isoleucine residues in the sequence, as both are isomers and cannot be distinguished due to their identical mass. We further excluded peptides with cysteine residues as these are likely to be false positives due to the high reactivity of cysteine and their likely chemical modification. We then shortlisted sequences that were either present in a minimum of two patients, or two different sequences in a minimum of two patients mapping to the same lncRNA entry.

This analysis resulted in a list of 21 peptides from 11 lncRNAs (Table 8). For validation of the correct annotation of peptide sequence from the LC-MS/MS spectra, we performed spectral matching analysis using spectra from synthetic peptide standards that were acquired under identical conditions (Fig.34). 12 peptides were validated in spectral matching (57%), 6 did not validate (29%), and 3 were undetermined (14%) in our experiment.

For further validation, we also used NCBI T-BLASTn [225] searched against the paired, fully assembled Trinity cDNA sequences to confirm whether the cDNA sequence as defined in the lncRNA databases was present in the transcriptome of the interrogated tumours. We found that there was sequence evidence in Trinity assemblies for peptides IPHGEIPDTSA (IPA11), MTMSTILSK (MTK9) and MTMSTILSKK (MTK10), for which the cDNA entry

from lncipedia [230] fully matched a transcript sequence in our trinity assembly. Partial cDNA sequence matches in Trinity were obtained for peptides ALRAVTLTAK (ALK10), KVHVFLVKK (KVK9), KTKIITGSK (KTK9) and SLSILSLKV (SLV9), in which the cDNA was only partially matched in Trinity but the peptide sequence matched an ORF in a transcript of the respective sample. To note that for the latter two peptides, sequence evidence as found only in a different patient but not the patient the peptide had been identified for. The LRKEIEERI (LI9) could not be found in the trinity assemblies, and may be explained by a peptide fusion event. LI9 may came from cis-spliced when searched against SwissProt human with isoforms (UniProtKB 2021_03). It may be generated from same human protein, such as Golgin subfamily A member 8 or Filensin (Table 9). However, validation of the actual source of this peptide antigen is still outstanding.

Table 8. LncRNA-derived peptide sequence candidates identified in cervical tumour tissue resections

Peptide	Sample	Spectral Match	LNC-RNA ID	cDNA match in Trinity	Trinity hits present in Sample
MRSLALTQV	1,9	Yes	lnckb.10057	Not found	
TRHSPPGVL	4	Maybe	lnckb.10057	Not found	
RQAPALRHL	9	No	lnckb.14510	Not found	
VIFSGIRSL	5	Yes	lnckb.14510	Not found	
IPHGEIPDTSA	5	Yes	lnckb.31346	Yes	1,4,5,9
LPNFARII	4,8	No	lnckb.31346	Yes	8
LRKEIEERI	1,10	Yes	lnckb.31640	Not found	
ALRAVTLTAK	8	Yes	lnckb.36918	Partial	1,2,3,4,5,6,7,8,9,10
SASSLFIYK	4,9	Yes	lnckb.36918	Not found	

VPNFKIISH	2,3,4,5,6,10	Maybe	lnckb.40028	Not found	
MTMSTILSK	8	Yes	lnckb.42318	Yes	1,2,3,4,5,6,8,9,10
MTMSTILSKK	4,8,9	Yes	lnckb.42318	Yes	1,2,3,4,5,6,8,9,11
LPYIKWTI	7	No	lnckb.44609	Not found	
LSSTLNKQI	1	No	lnckb.44609	Not found	
IRNVKIYLI	10	No	lnckb.45851	Partial	1,2,3,4,5,7,8,10
KVHVFLVKK	1,8	Yes	lnckb.45851	Partial	1,4,5,6,7,8
KQLNKQLIK	4,9	Yes	lnckb.53745	Not found	
VTIFQNRVK	8	No	lnckb.53745	Not found	
KLGSITTPK	9	Yes	lnckb.58303	Not found	
KTKIITGSK	8	Yes	lnckb.58303	Partial	3,4,5
SLSILSLKV	1	Maybe	lnckb.58303	Partial	9

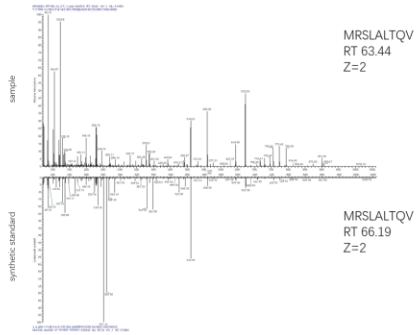
Table 9. Potential protein spliced source of LI9

FullPeptide	ProteinAccession	AA_Sequence	FirstPeptide	SecondPeptide
LRKELEERI	sp|Q5VZ66|JKIP3_HUMAN	LRKELEERL	LR	KELEERI
LRKELEERL	sp|Q86SQ7|SDCG8_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q86SQ7-3|SDCG8_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q86SQ7-4|SDCG8_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q0D2H9|GOG8D_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|Q0D2H9|GOG8D_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERI	sp|H3BQL2|GOG8T_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|H3BQL2|GOG8T_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERI	sp|H3BSY2|GOG8M_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|H3BSY2|GOG8M_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERI	sp|H3BSY2-2|GOG8M_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|H3BSY2-2|GOG8M_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERI	sp|H3BSY2-3|GOG8M_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|H3BSY2-3|GOG8M_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERL	sp|A6NEY3|GG6L3_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NEY3|GG6L3_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|A6NEF3|GG6L4_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NEF3|GG6L4_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|A6NEM1|GG6L9_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NEM1|GG6L9_HUMAN	LRKELEERL	LRKELE	ERL

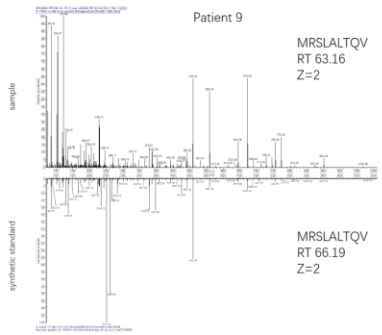
LRKELEERL	sp|A6NEM1-2|GG6L9_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NEM1-2|GG6L9_HUMAN	LRKELEERL	LRKELE	ERL
LRKEIEERL	sp|Q12934|BFSP1_HUMAN	LRKELEERL	LRKEIEE	RL
LRKEIEERL	sp|Q12934-2|BFSP1_HUMAN	LRKELEERL	LRKEIEE	RL
LRKEIEERL	sp|Q12934-3|BFSP1_HUMAN	LRKELEERL	LRKEIEE	RL
LRKELEERI	sp|A6NMD2|GOG8J_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|A6NMD2|GOG8J_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERI	sp|D6RF30|GOG8K_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|D6RF30|GOG8K_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERL	sp|I6L899|GOG8R_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|I6L899|GOG8R_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|I6L899-2|GOG8R_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|I6L899-2|GOG8R_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|A7E2F4|GOG8A_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A7E2F4|GOG8A_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|A7E2F4-3|GOG8A_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A7E2F4-3|GOG8A_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERI	sp|P0CJ92|GOG8H_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|P0CJ92|GOG8H_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERL	sp|A6NCC3|GOG8O_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NCC3|GOG8O_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|A6NCC3-1|GOG8O_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NCC3-1|GOG8O_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|F8WBI6|GOG8N_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|F8WBI6|GOG8N_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q08379|GOGA2_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|Q08379|GOGA2_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q08379-2|GOGA2_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|Q08379-2|GOGA2_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|H0YKK7|GG6LS_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|H0YKK7|GG6LS_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERI	sp|A6NC78|GOG8I_HUMAN	LRKELEERL	LRKEL	EERI
LRKELEERI	sp|A6NC78|GOG8I_HUMAN	LRKELEERL	LRKELE	ERI
LRKELEERL	sp|Q9H095|DRC9_HUMAN	LRKELEERL	LRKE	LEERL
LRKELEERL	sp|Q9H095-2|DRC9_HUMAN	LRKELEERL	LRKE	LEERL
LRKELEERL	sp|A6NI86|GG6LA_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NI86|GG6LA_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q8IWJ2|GCC2_HUMAN	LRKELEERL	LRKELEE	RL
LRKELEERL	sp|A8MQT2|GOG8B_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A8MQT2|GOG8B_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|A8MQT2-2|GOG8B_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A8MQT2-2|GOG8B_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|H3BV12|GOG8Q_HUMAN	LRKELEERL	LRKEL	EERL

LRKELEERL	sp|H3BV12|GOG8Q_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|H3BPF8|GOG8S_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|H3BPF8|GOG8S_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|A6NN73|GOG8C_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|A6NN73|GOG8C_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q08AF8|GOG8F_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|Q08AF8|GOG8F_HUMAN	LRKELEERL	LRKELE	ERL
LRKELEERL	sp|Q08AF8-3|GOG8F_HUMAN	LRKELEERL	LRKEL	EERL
LRKELEERL	sp|Q08AF8-3|GOG8F_HUMAN	LRKELEERL	LRKELE	ERL

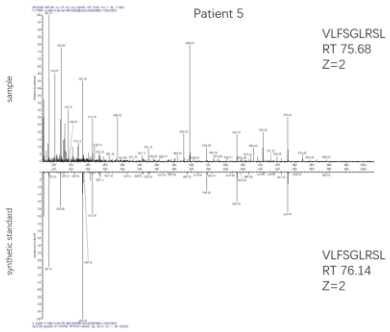
A


sample
synthetic standard
MRSLALTQV
RT 63.44
Z=2
MRSLALTQV
RT 66.19
Z=2

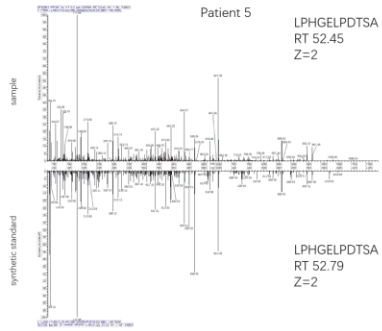
B


Patient 9
sample
synthetic standard
MRSLALTQV
RT 63.16
Z=2
MRSLALTQV
RT 66.19
Z=2

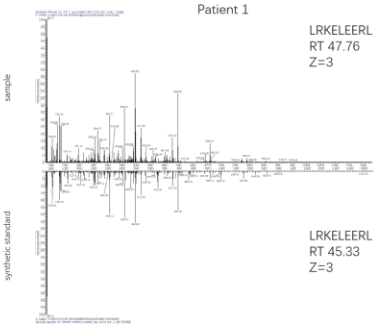
C


Patient 5
sample
synthetic standard
VLFSGLRSL
RT 75.68
Z=2
VLFSGLRSL
RT 76.14
Z=2

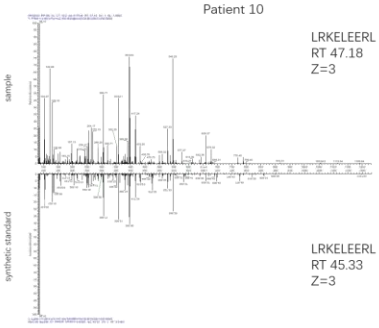
D


Patient 5
sample
synthetic standard
LPHGELPDTSA
RT 52.45
Z=2
LPHGELPDTSA
RT 52.79
Z=2

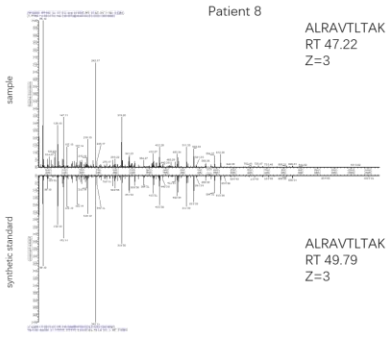
E


Patient 1
sample
synthetic standard
LRKELEERL
RT 47.76
Z=3
LRKELEERL
RT 45.33
Z=3

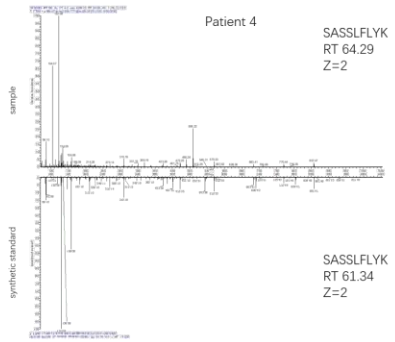
F


Patient 10
sample
synthetic standard
LRKELEERL
RT 47.18
Z=3
LRKELEERL
RT 45.33
Z=3

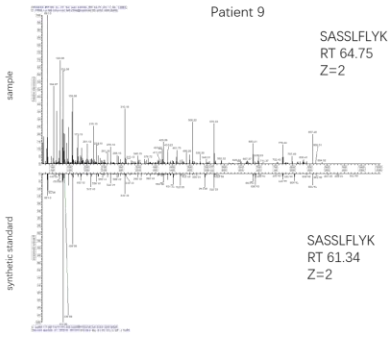
G


Patient 8
ALRAVTLTAK
RT 47.22
Z=3
sample
synthetic standard
ALRAVTLTAK
RT 49.79
Z=3

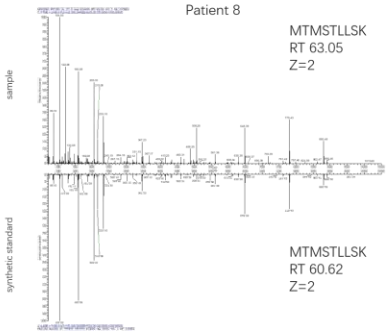
H


Patient 4
SASSLFLYK
RT 64.29
Z=2
sample
synthetic standard
SASSLFLYK
RT 61.34
Z=2

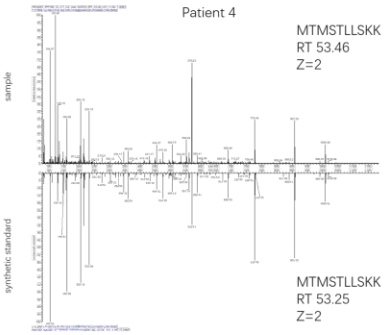
I


Patient 9
SASSLFLYK
RT 64.75
Z=2
sample
synthetic standard
SASSLFLYK
RT 61.34
Z=2

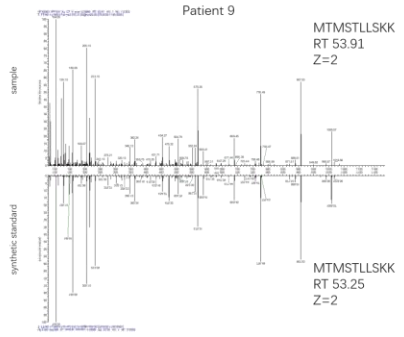
J


Patient 8
MTMSTLLSK
RT 63.05
Z=2
sample
synthetic standard
MTMSTLLSK
RT 60.62
Z=2

K


Patient 4
MTMSTLLSKK
RT 53.46
Z=2
sample
synthetic standard
MTMSTLLSKK
RT 53.25
Z=2

L


Patient 9
MTMSTLLSKK
RT 53.91
Z=2
sample
synthetic standard
MTMSTLLSKK
RT 53.25
Z=2

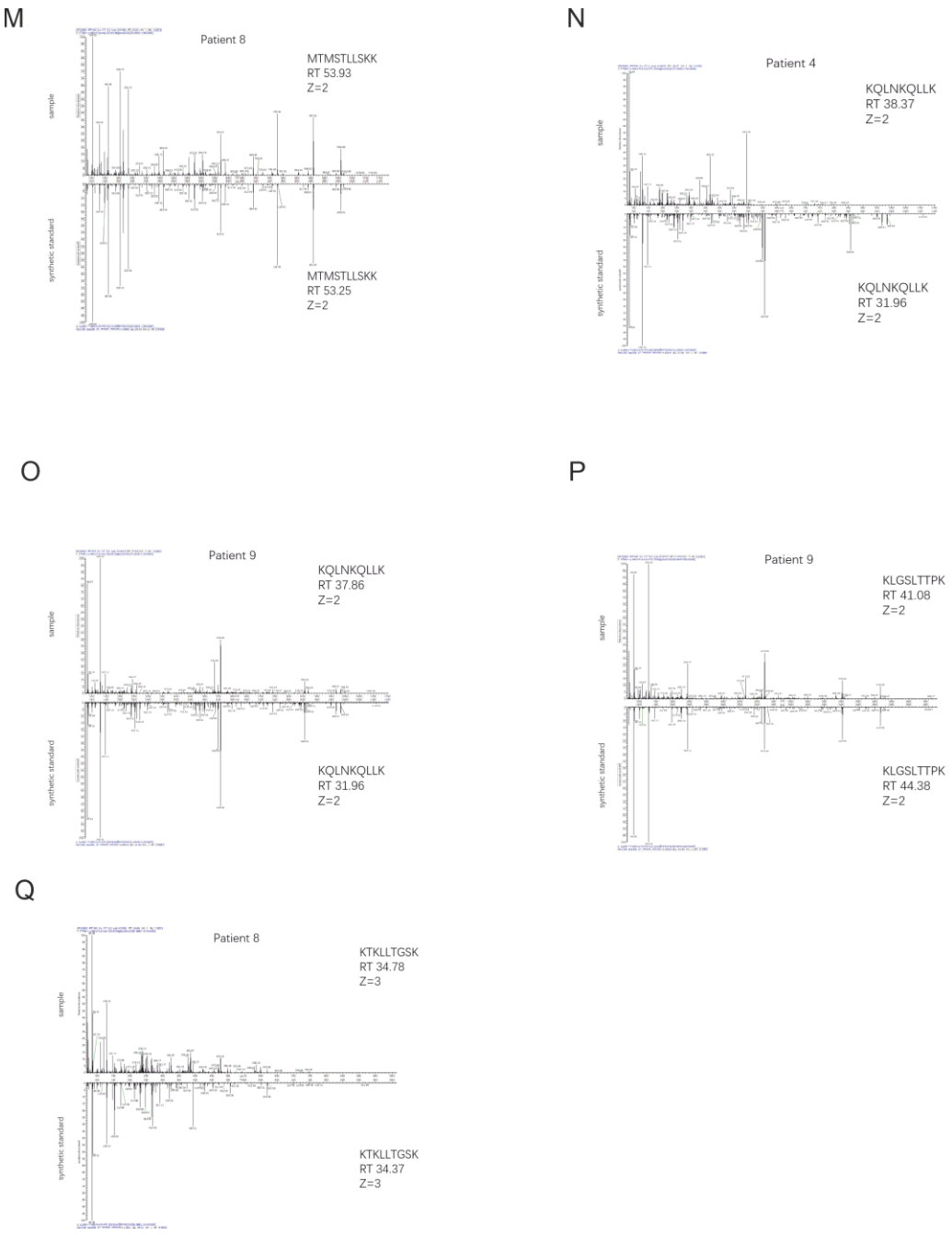


Figure 34. Spectral matching analysis for lncRNA peptide sequences. A-Q. LC-MS fragment spectra are shown as measured in the experiment (sample) and measured from the equivalent synthetic peptide sequence counterpart (synthetic standard).

4.3 Discussion

The information of HLA class -I and class-II peptides presented on tumour cells to T-cells is crucial for designing immunotherapeutic approaches against cancer. Nowadays, with immune checkpoint blockade therapies having shown remarkable clinical efficacy in cancer, the identification of immunogenic T-cell epitopes which are present in disease have become the next major challenge for scientists. Those peptide antigens, that are derived from cancer-specific, mutated proteins and pathogens, could be very suitable clinical targets for not only T-cell based therapy approaches, but also for therapeutic vaccination. As shown in the previous chapter, MS has matured to be able to identify clinically relevant, immunogenic antigens in human tumours. We here expanded the immunopeptidomics pipeline to immunoaffinity purification (IP) of HLA complexes from clinical samples, and subsequent HLA-peptide elution and identification by LC-MS.

The depth of the identified HLA-I peptidomes was outstanding with overall a total of 78,618 peptides identified. A major difficultly of immunopeptidomics is that still a relatively large number of cells or tissue material is needed, which is difficult to achieve for clinical samples. Here, we showed that our methodology enables us to identify an average of 7,861 peptides from 255 mg median weight tumour sample.

We also performed RNA sequencing on all ten cervical tumours and used software OptiType [326] to genotype the patient's alleles. OptiType is a novel HLA genotyping algorithm based on integer linear programming, which capable of producing accurate predictions (overall accuracy of 97%) from RNAseq data [326]. To confirm the accuracy of this software, we also applied it to the RNAseq data for our previous five cervical cancer lines. The result was the same as our conventional genotyping result, which confirmed the reliability of this approach.

We could, however, interrogate our data further, and identified shared CT antigens between all 10 tumours. Our findings demonstrated that a total of 227 peptides from 99 proteins from the CT antigens lists are presented by the 10 tumour biopsies. KDM5B was shared with all cervical cancer lines and all 10 tumours analysed. There was one shared antigen identified in Raven [301] which was found in our tumour samples, which was also found presented in all five cervical cancer cell lines.

It is estimated more than 70% of the genome is transcribed but only 2% of the human genome can be translated to proteins. Recent studies identified lncRNAs to play many important functional roles in cancer and other diseases, making them a novel class of therapeutic targets. However, the knowledge of their specific function in the cell, and their role in diseases are still broadly

unknown. Studies have suggested that lncRNAs are regulating gene expression and interact with RNA binding proteins [187, 328, 329]. As we know, tumour suppressor genes play an important role in inhibition of cellular growth and proliferation. Cancer cells can sustain proliferative signaling by evading growth suppression, such as interfering with tumour suppressor genes like p53 and RB, such as the HPV oncogenes E6 and E7. In addition, cancer cells may also use lncRNAs to inhibit tumour suppressor functions. It was found that there are five human ncRNA fragments that interact with the tumour suppressor PSF which led to an enhanced tumourigenic phenotype [330]. In addition, tumour suppressor p21 is also regulated by lncRNA epigenetic gene regulation [331]. We here identified through a *de novo* sequencing approach (MAR-S) 21 peptide candidates from 11 lncRNAs which we selected for further validation. We could match 12 peptides using spectral matching to spectra from synthetic peptide standards. For three of the peptides, we could match the Lncipedia cDNA entry [230], to our full Trinity assemblies, which provided evidence for active transcription of the lncRNA in the selected tumour. Peptide LRKEIEERI (LI9) (Table 9) was designated a highly likely protein *cis*-spliced peptide candidate, generated by proteasomal fusion from several potential precursor proteins: Serologically defined colon cancer antigen 8 SDCG8, human Golgin subfamily A member 8 (GOG8), Dynein regulatory complex protein 9 (DRC9), Filensin (BFSP1) and others [110, 111].

In future, such non-canonical antigens derived from i.e. lncRNA and protein-spliced peptides may form a new targetable antigen space, and form a new area of tumour immunotherapic developments.

However, efforts are outstanding to validate and test such antigens in the clinic in order to evaluate their therapeutic potential.

Chapter 5: General discussion

The idea of exploiting the host's immune system to treat cancer is based on the immune system's ability to suppress tumour development by elimination of malignant cells. Due to the genetic and epigenetic changes, normal human cells may transform into malignant tumour cells and this process can give rise to foreign antigens that are only present in the tumour, but not healthy cells. If such tumour-specific antigens are presented by HLA on the cell surface, they become detectable by CD8-positive T cells which can in turn initiate the destruction of the antigen presenting cell.

This thesis aimed to discover novel targets for immunotherapy of cervical cancer and high-risk human papillomavirus infection. To discover suitable antigens in HPV-driven cancer, I first utilized five human cervical cancer lines as a research model. I sequenced the exome and mRNA transcriptome of those cell lines to build personalized protein databases for each cell line based on their unique mutational landscapes and transcriptomes. I performed immunopeptidomics analyses for all 5 cell lines in order to directly identify the mutated and viral peptides presented by these tumour cells, based on the databases constructed from exome and mRNA sequence data. Specifically, HPV E1 which is first been reported in immunotherapy research of cervical cancer. We further validated immunogenicity of identified viral sequences in

HPV-positive women with and without CIN. I also explored how to use *in silico* binding prediction and clustering tools on the set of identified HLA peptides, and shortlisted cancer-associated antigens by integration publicly available information on tumour-specificity (CT-database, Raven) with a view to select clinically relevant therapeutic target antigens.

I then proceeded to apply this pipeline to clinical samples. Although we could not validate viral peptides we identified through the integration of RNA sequencing results, we found two tumour-associated antigens presented by all cell lines were also presented by the primary tumour cells: KDM5B, and H1.2.

Finally, we found 12 lncRNA-related peptide candidates presented in cervical tumour samples, with sequence evidence for expression on RNA level for three out of the 12 candidates.

This study provides novel viral antigen, *neo*antigens, and shared tumour-associated antigens that can be integrated in the generation of novel therapeutic approaches for HPV-driven cancers, and validates immunopeptidomics as a powerful discovery platform for clinically relevant antigens for cancer immunotherapies.

6.0 References

1. Franco, E.L., *Cancer Causes Revisited - Human Papillomavirus and Cervical Neoplasia.* Jnci-Journal of the National Cancer Institute, 1995. **87**(11): p. 779-780.
2. Bzhalava, D., C. Eklund, and J. Dillner, *International standardization and classification of human papillomavirus types.* Virology, 2015. **476**: p. 341-344.
3. Rajkumar, R., *Introductory chapter: Human Papillomavirus (HPV) Infections, Associated Diseases and Cervical Cancer Prevention and Control Initiate Countdown Using "The Raj's Cancer Control Clock".* Human Papillomavirus - Research in a Global Perspective, 2016: p. 3-11.
4. States, C.f.D.C.C.i.U., *Number of HPV-Associated and Estimated Number of HPV-Attributable Cancer Cases in United States.* 2021.
5. Brentjens, M.H., et al., *Human papillomavirus: a review.* Dermatologic Clinics, 2002. **20**(2): p. 315-+.
6. Crow, J.M., *Hpv: The Global Burden.* Nature, 2012. **488**(7413): p. 52-53.
7. Accardi, R. and T. Gheit, *Cutaneous HPV and skin cancer.* Presse Medicale, 2014. **43**(12): p. E435-E443.
8. Viarisio, D., L. Gissmann, and M. Tommasino, *Human papillomaviruses and carcinogenesis: well-established and novel models.* Current Opinion in Virology, 2017. **26**: p. 56-62.
9. Tommasino, M., *The biology of beta human papillomaviruses.* Virus Research, 2017. **231**: p. 128-138.
10. Chouhy, D., et al., *Analysis of the genetic diversity and phylogenetic relationships of putative human papillomavirus types.* Journal of General Virology, 2013. **94**: p. 2480-2488.
11. Graham, S.V., *The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review.* Clinical Science, 2017. **131**(17): p. 2201-2221.
12. de Villiers, E.M., et al., *Classification of papillomaviruses.* Virology, 2004. **324**(1): p. 17-27.
13. Bouvard, V., et al., *A review of human carcinogens-Part B: biological agents.* Lancet Oncology, 2009. **10**(4): p. 321-322.
14. Small, W., Jr., et al., *Cervical cancer: A global health crisis.* Cancer, 2017. **123**(13): p. 2404-2412.
15. Bosch, F.X., et al., *The causal relation between human papillomavirus and cervical cancer.* Journal of Clinical Pathology, 2002. **55**(4): p. 244-265.
16. WHO, *Cervical cancer.* 2021.
17. Jaisamrarn, U., et al., *Natural History of Progression of HPV Infection to Cervical Lesion or Clearance: Analysis of the Control Arm of the Large, Randomised PATRICIA Study.* Plos One, 2013. **8**(11).
18. Tainio, K., et al., *Clinical course of untreated cervical intraepithelial neoplasia grade 2 under active surveillance: systematic review and meta-analysis.* Bmj-British Medical Journal, 2018. **360**.

19. Tsikouras, P., et al., *Cervical cancer: screening, diagnosis and staging.* Journal of Buon, 2016. **21**(2): p. 320-325.
20. Lie, A.K. and G. Kristensen, *Human papillomavirus E6/E7 mRNA testing as a predictive marker for cervical carcinoma.* Expert Review of Molecular Diagnostics, 2008. **8**(4): p. 405-415.
21. Koliopoulos, G., et al., *Cytology versus HPV testing for cervical cancer screening in the general population.* Cochrane Database of Systematic Reviews, 2017(8).
22. *Cervical Screening Programme, England - 2019-20 [NS].* 2020; Available from: https://digital.nhs.uk/data-and-information/publications/statistical/cervical-screening-annual/england---2019-20.
23. NHS, *Cervical screening: programme overview.* 2021.
24. Vaccarella, S., et al., *Preventable fractions of cervical cancer via effective screening in six Baltic, central, and eastern European countries 2017-40: a population-based study.* Lancet Oncology, 2016. **17**(10): p. 1445-1452.
25. NHS, *NHS Treatment for cervical cancer.* 2022.
26. Colombo, N., et al., *Pembrolizumab for Persistent, Recurrent, or Metastatic Cervical Cancer.* New England Journal of Medicine, 2021.
27. Stevanovic, S., et al., *Complete Regression of Metastatic Cervical Cancer After Treatment With Human Papillomavirus-Targeted Tumor-Infiltrating T Cells.* Journal of Clinical Oncology, 2015. **33**(14): p. 1543-+.
28. Draper, L.M., et al., *Targeting of HPV-16(+) Epithelial Cancer Cells by TCR Gene Engineered T Cells Directed against E6.* Clinical Cancer Research, 2015. **21**(19): p. 4431-4439.
29. Nagarsheth, N.B., et al., *TCR-engineered T cells targeting E7 for patients with metastatic HPV-associated epithelial cancers.* Nature Medicine, 2021. **27**(3): p. 419-+.
30. Petrosky, E., et al., *Use of 9-Valent Human Papillomavirus (HPV) Vaccine: Updated HPV Vaccination Recommendations of the Advisory Committee on Immunization Practices.* Mmwr-Morbidity and Mortality Weekly Report, 2015. **64**(11): p. 300-304.
31. Munger, K. and P.M. Howley, *Human papillomavirus immortalization and transformation functions.* Virus Research, 2002. **89**(2): p. 213-228.
32. Ashrafi, G.H., et al., *E5 protein of human papillomavirus 16 downregulates HLA class I and interacts with the heavy chain via its first hydrophobic domain.* International Journal of Cancer, 2006. **119**(9): p. 2105-2112.
33. Longworth, M.S. and L.A. Laimins, *Pathogenesis of human papillomaviruses in differentiating epithelia.* Microbiology and Molecular Biology Reviews, 2004. **68**(2): p. 362-+.
34. Zhang, B., et al., *The E5 protein of human papillomavirus type 16 perturbs MHC class II antigen maturation in human foreskin keratinocytes treated with interferon-gamma.* Virology, 2003. **310**(1): p. 100-8.
35. Riemer, A.B., et al., *A conserved E7-derived cytotoxic T lymphocyte epitope expressed on human papillomavirus 16-transformed HLA-A2+ epithelial cancers.* J Biol Chem, 2010. **285**(38): p. 29608-22.
36. Burk, R.D., et al., *Integrated genomic and molecular characterization of cervical cancer.* Nature, 2017. **543**(7645): p. 378-+.

37. Horvath, C.A.J., et al., *Mechanisms of cell entry by human papillomaviruses: an overview.* Virology Journal, 2010. **7**.
38. Doorbar, J., *The papillomavirus life cycle.* Journal of Clinical Virology, 2005. **32**: p. S7-S15.
39. Pappa, K.I., et al., *Novel structural approaches concerning HPV proteins: Insight into targeted therapies for cervical cancer (Review).* Oncology Reports, 2018. **39**(4): p. 1547-1554.
40. Bergvall, M., T. Melendy, and J. Archambault, *The E1 proteins.* Virology, 2013. **445**(1-2): p. 35-56.
41. Swindle, C.S. and J.A. Engler, *Association of the human papillomavirus type 11 E1 protein with histone H1.* Journal of Virology, 1998. **72**(3): p. 1994-2001.
42. Lee, D., et al., *Interaction of E1 and hSNF5 proteins stimulates replication of human papillomavirus DNA.* Nature, 1999. **399**(6735): p. 487-491.
43. Das, D., et al., *The Deacetylase SIRT1 Regulates the Replication Properties of Human Papillomavirus 16 E1 and E2.* Journal of Virology, 2017. **91**(10).
44. Castro-Munoz, L.J., et al., *The Human Papillomavirus (HPV) E1 protein regulates the expression of cellular genes involved in immune response.* Scientific Reports, 2019. **9**.
45. Schoggins, J.W. and C.M. Rice, *Interferon-stimulated genes and their antiviral effector functions.* Current Opinion in Virology, 2011. **1**(6): p. 519-525.
46. Zuniga, E.I., et al., *Innate and Adaptive Immune Regulation During Chronic Viral Infections.* Annual Review of Virology, Vol 2, 2015. **2**: p. 573-597.
47. Schoggins, J.W., *Recent advances in antiviral interferon-stimulated gene biology.* F1000Res, 2018. **7**: p. 309.
48. Ham, J., et al., *The Papillomavirus E2 Protein - a Factor with Many Talents.* Trends in Biochemical Sciences, 1991. **16**(11): p. 440-444.
49. Dell, G., et al., *Comparison of the structure and DNA-binding properties of the E2 proteins from an oncogenic and a non-oncogenic human papillomavirus.* Journal of Molecular Biology, 2003. **334**(5): p. 979-991.
50. Doorbar, J., *Molecular biology of human papillomavirus infection and cervical cancer.* Clinical Science, 2006. **110**(5): p. 525-541.
51. You, J., et al., *Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes.* Cell, 2004. **117**(3): p. 349-360.
52. Stubenrauch, F., H.B. Lim, and L.A. Laimins, *Differential requirements for conserved E2 binding sites in the life cycle of oncogenic human papillomavirus type 31.* Journal of Virology, 1998. **72**(2): p. 1071-1077.
53. Frattini, M.G., H.B. Lim, and L.A. Laimins, *In vitro synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(7): p. 3062-3067.
54. Fehrmann, F., D.J. Klumpp, and L.A. Laimins, *Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation.* Journal of Virology, 2003. **77**(5): p. 2819-2831.
55. Crusius, K., I. Rodriguez, and A. Alonso, *The human papillomavirus type 16 E5 protein modulates ERK1/2 and p38 MAP kinase activation by an EGFR-Independent process in stressed human keratinocytes.* Virus Genes, 2000. **20**(1): p. 65-69.

56. Doorbar, J., *The E4 protein; structure, function and patterns of expression.* Virology, 2013. **445**(1-2): p. 80-98.
57. Nakahara, T., et al., *Human papillomavirus type 16 E1 approximate to E4 contributes to multiple facets of the papillomavirus life cycle.* Journal of Virology, 2005. **79**(20): p. 13150-13165.
58. Wilson, R., F. Fehrmann, and L.A. Laimins, *Role of the E1(boolean AND)E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31.* Journal of Virology, 2005. **79**(11): p. 6732-6740.
59. Davy, C. and J. Doorbar, *G2/M cell cycle arrest in the life cycle of viruses.* Virology, 2007. **368**(2): p. 219-226.
60. Doorbar, J., et al., *Identification of the Human Papilloma Virus-1a E4 Gene-Products.* Embo Journal, 1986. **5**(2): p. 355-362.
61. Sherman, L., et al., *Inhibition of serum- and calcium-induced differentiation of human keratinocytes by HPV16 E6 oncoprotein: Role of p53 inactivation.* Virology, 1997. **237**(2): p. 296-306.
62. Munger, K., et al., *Biological activities and molecular targets of the human papillomavirus E7 oncoprotein.* Oncogene, 2001. **20**(54): p. 7888-7898.
63. Peh, W.L., et al., *The viral E4 protein is required for the completion of the cottontail rabbit papillomavirus productive cycle in vivo.* Journal of Virology, 2004. **78**(4): p. 2142-2151.
64. Ozbun, M.A. and C. Meyers, *Temporal usage of multiple promoters during the life cycle of human papillomavirus type 31b.* Journal of Virology, 1998. **72**(4): p. 2715-2722.
65. Becker, K.A., et al., *Nuclear localization but not PML protein is required for incorporation of the papillomavirus minor capsid protein L2 into virus-like particles.* Journal of Virology, 2004. **78**(3): p. 1121-1128.
66. Day, P.M., et al., *The papillomavirus minor capsid protein, L2, induces localization of the major capsid protein, L1, and the viral transcription/replication protein, E2, to PML oncogenic domains.* Journal of Virology, 1998. **72**(1): p. 142-150.
67. Swindle, C.S., et al., *Human papillomavirus DNA replication compartments in a transient DNA replication system.* Journal of Virology, 1999. **73**(2): p. 1001-1009.
68. Longworth, M.S. and L.A. Laimins, *The binding of histone deacetylases and the integrity of zinc finger-like motifs of the E7 protein are essential for the life cycle of human papillomavirus type 31.* Journal of Virology, 2004. **78**(7): p. 3533-3541.
69. Doorbar, J., et al., *Specific Interaction between Hpv-16 E1-E4 and Cytokeratins Results in Collapse of the Epithelial-Cell Intermediate Filament Network.* Nature, 1991. **352**(6338): p. 824-827.
70. Wang, Q., et al., *Functional analysis of the human papillomavirus type 16 E1(boolean AND)E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization.* Journal of Virology, 2004. **78**(2): p. 821-833.
71. Bryan, J.T. and D.R. Brown, *Association of the human papillomavirus type 11 E1-E4 protein with cornified cell envelopes derived from infected genital epithelium.* Virology, 2000. **277**(2): p. 262-269.
72. Lehr, E., et al., *Infection with human papillomavirus alters expression of the small proline rich proteins 2 and 3.* Journal of Medical Virology, 2004. **72**(3): p. 478-483.

73. Bryan, J.T. and D.R. Brown, *Transmission of human papillomavirus type 11 infection by desquamated cornified cells.* Virology, 2001. **281**(1): p. 35-42.
74. Chen, C.H., et al., *Gene gun-mediated DNA vaccination induces antitumor immunity against human papillomavirus type 16 E7-expressing murine tumor metastases in the liver and lungs.* Gene Therapy, 1999. **6**(12): p. 1972-1981.
75. Crook, T., J.A. Tidy, and K.H. Vousden, *Degradation of P53 Can Be Targeted by Hpv E6 Sequences Distinct from Those Required for P53 Binding and Transactivation.* Cell, 1991. **67**(3): p. 547-556.
76. DiMaio, D. and D. Mattoon, *Mechanisms of cell transformation by papillomavirus E5 proteins.* Oncogene, 2001. **20**(54): p. 7866-7873.
77. Boyer, S.N., D.E. Wazer, and V. Band, *E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway.* Cancer Research, 1996. **56**(20): p. 4620-4624.
78. Gage, J.R., C. Meyers, and F.O. Wettstein, *The E7 Proteins of the Nononcogenic Human Papillomavirus Type-6b (Hpv-6b) and of the Oncogenic Hpv-16 Differ in Retinoblastoma Protein-Binding and Other Properties.* Journal of Virology, 1990. **64**(2): p. 723-730.
79. Jeon, S. and P.F. Lambert, *Integration of Human Papillomavirus Type-16 DNA into the Human Genome Leads to Increased Stability of E6 and E7 Messenger-Rnas - Implications for Cervical Carcinogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(5): p. 1654-1658.
80. Wallace, N.A. and D.A. Galloway, *Novel Functions of the Human Papillomavirus E6 Oncoproteins.* Annual Review of Virology, Vol 2, 2015. **2**: p. 403-423.
81. Pim, D., M. Collins, and L. Banks, *Human Papillomavirus Type-16 E5 Gene Stimulates the Transforming Activity of the Epidermal Growth-Factor Receptor.* Oncogene, 1992. **7**(1): p. 27-32.
82. Ashrafi, G.H., et al., *E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I.* International Journal of Cancer, 2005. **113**(2): p. 276-283.
83. Campo, M.S., et al., *HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells.* Virology, 2010. **407**(1): p. 137-42.
84. Oh, J.M., et al., *Human papillomavirus type 16 E5 protein inhibits hydrogen peroxide-induced apoptosis by stimulating ubiquitin-proteasome-mediated degradation of Bax in human cervical cancer cells.* Carcinogenesis, 2010. **31**(3): p. 402-410.
85. Kabsch, K., et al., *The HPV-16 E5 protein inhibits TRAIL- and FasL-mediated apoptosis in human keratinocyte raft cultures.* Intervirology, 2004. **47**(1): p. 48-56.
86. McBride, A.A. and A. Warburton, *The role of integration in oncogenic progression of HPV-associated cancers.* Plos Pathogens, 2017. **13**(4).
87. Lukusa, T. and J.P. Fryns, *Human chromosome fragility.* Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms, 2008. **1779**(1): p. 3-16.
88. Durkin, S.G. and T.W. Glover, *Chromosome fragile sites.* Annual Review of Genetics, 2007. **41**: p. 169-192.
89. Thorland, E.C., et al., *Common fragile sites are preferential targets for HPV16 integrations in cervical tumors.* Oncogene, 2003. **22**(8): p. 1225-1237.

90. Ferber, M.J., et al., *Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma.* Oncogene, 2003. **22**(46): p. 7233-7242.
91. Yu, T.X., et al., *The role of viral integration in the development of cervical cancer.* Cancer Genetics and Cytogenetics, 2005. **158**(1): p. 27-34.
92. Gao, G., et al., *Common fragile sites (CFS) and extremely large CFS genes are targets for human papillomavirus integrations and chromosome rearrangements in oropharyngeal squamous cell carcinoma.* Genes Chromosomes & Cancer, 2017. **56**(1): p. 59-74.
93. Kadaja, M., et al., *Mechanism of Genomic Instability in Cells Infected with the High-Risk Human Papillomaviruses.* Plos Pathogens, 2009. **5**(4).
94. Chaiwongkot, A., et al., *Differential methylation of E2 binding sites in episomal and integrated HPV 16 genomes in preinvasive and invasive cervical lesions.* International Journal of Cancer, 2013. **132**(9): p. 2087-2094.
95. Akagi, K., et al., *Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability.* Genome Research, 2014. **24**(2): p. 185-199.
96. Pett, M.R., et al., *Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes.* Cancer Research, 2004. **64**(4): p. 1359-1368.
97. Thierry, F. and M. Yaniv, *The Bpv1-E2 Trans-Acting Protein Can Be Either an Activator or a Repressor of the Hpv18 Regulatory Region.* Embo Journal, 1987. **6**(11): p. 3391-3397.
98. Neefjes, J.J. and H.L. Ploegh, *Allele and Locus-Specific Differences in Cell-Surface Expression and the Association of Hla Class-I Heavy-Chain with Beta-2-Microglobulin - Differential-Effects of Inhibition of Glycosylation on Class-I Subunit Association.* European Journal of Immunology, 1988. **18**(5): p. 801-810.
99. Kulkarni, S., et al., *Differential microRNA regulation of HLA-C expression and its association with HIV control.* Nature, 2011. **472**(7344): p. 495-U548.
100. Robinson, J., et al., *The IPD and IMGT/HLA database: allele variant databases.* Nucleic Acids Res, 2015. **43**(Database issue): p. D423-31.
101. Neisig, A., et al., *Allele-specific differences in the interaction of MHC class I molecules with transporters associated with antigen processing.* Journal of Immunology, 1996. **156**(9): p. 3196-3206.
102. Peh, C.A., et al., *HLA-B27-restricted antigen presentation in the absence of tapasin reveals polymorphism in mechanisms of HLA class I peptide loading.* Immunity, 1998. **8**(5): p. 531-542.
103. Fernando, M.M.A., et al., *Defining the role of the MHC in autoimmunity: A review and pooled analysis.* Plos Genetics, 2008. **4**(4).
104. Hughes, E.A., C. Hammond, and P. Cresswell, *Misfolded major histocompatibility complex class I heavy chains are translocated into the cytoplasm and degraded by the proteasome.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(5): p. 1896-1901.
105. Yewdell, J.W., L.C. Anton, and J.R. Bennink, *Defective ribosomal products (DRiPs) - A major source of antigenic peptides for MHC class I molecules?* Journal of Immunology, 1996. **157**(5): p. 1823-1826.
106. Schubert, U., et al., *Rapid degradation of a large fraction of newly synthesized proteins by proteasomes.* Nature, 2000. **404**(6779): p. 770-774.

107. Falk, K., et al., *Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules.* Nature, 1991. **351**(6324): p. 290-6.
108. Trowsdale, J., *HLA genomics in the third millennium.* Current Opinion in Immunology, 2005. **17**(5): p. 498-504.
109. Paes, W., et al., *Contribution of proteasome-catalyzed peptide cis-splicing to viral targeting by CD8(+) T cells in HIV-1 infection.* Proceedings of the National Academy of Sciences of the United States of America, 2019. **116**(49): p. 24748-24759.
110. Liepe, J., et al., *A large fraction of HLA class I ligands are proteasome-generated spliced peptides.* Science, 2016. **354**(6310): p. 354-358.
111. Faridi, P., et al., *A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands.* Science Immunology, 2018. **3**(28).
112. Mylonas, R., et al., *Estimating the Contribution of Proteasomal Spliced Peptides to the HLA-1 Ligandome.* Molecular & Cellular Proteomics, 2018. **17**(12): p. 2347-2357.
113. Roche, P.A. and P. Cresswell, *Invariant Chain Association with Hla-Dr Molecules Inhibits Immunogenic Peptide Binding.* Nature, 1990. **345**(6276): p. 615-618.
114. Denzin, L.K., et al., *Right place, right time, right peptide: DO keeps DM focused.* Immunological Reviews, 2005. **207**: p. 279-292.
115. Fenn, J.B., et al., *Electrospray Ionization for Mass-Spectrometry of Large Biomolecules.* Science, 1989. **246**(4926): p. 64-71.
116. Karas, M. and F. Hillenkamp, *Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10000 Daltons.* Analytical Chemistry, 1988. **60**(20): p. 2299-2301.
117. Senko, M.W., et al., *External accumulation of ions for enhanced electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry.* Journal of the American Society for Mass Spectrometry, 1997. **8**(9): p. 970-976.
118. Marshall, A.G., C.L. Hendrickson, and S.D.H. Shi, *Scaling MS plateaus with high-resolution FT-ICRMS.* Analytical Chemistry, 2002. **74**(9): p. 253a-259a.
119. Sharma, P. and J.P. Allison, *The future of immune checkpoint therapy.* Science, 2015. **348**(6230): p. 56-61.
120. Ishida, Y., et al., *Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death.* EMBO J, 1992. **11**(11): p. 3887-95.
121. Keir, M.E., et al., *PD-1 and its ligands in tolerance and immunity.* Annu Rev Immunol, 2008. **26**: p. 677-704.
122. Okazaki, T., et al., *A rheostat for immune responses: the unique properties of PD-1 and their advantages for clinical application.* Nature Immunology, 2013. **14**(12): p. 1212-1218.
123. Leach, D.R., M.F. Krummel, and J.P. Allison, *Enhancement of antitumor immunity by CTLA-4 blockade.* Science, 1996. **271**(5256): p. 1734-1736.
124. Phan, G.Q., et al., *Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(14): p. 8372-8377.
125. Chambers, C.A., et al., *CTLA-4-mediated inhibition in regulation of T cell responses: Mechanisms and manipulation in tumor immunotherapy.* Annual Review of Immunology, 2001. **19**: p. 565-594.

126. Egen, J.G., M.S. Kuhns, and J.P. Allison, *CTLA-4: new insights into its biological function and use in tumor immunotherapy.* Nature Immunology, 2002. **3**(7): p. 611-618.
127. Parry, R.V., et al., *CTLA-4 and PD-1 receptors inhibit T-cell activation by distinct mechanisms.* Molecular and Cellular Biology, 2005. **25**(21): p. 9543-9553.
128. Chen, D.S., B.A. Irving, and F.S. Hodi, *Molecular Pathways: Next-Generation Immunotherapy-Inhibiting Programmed Death-Ligand 1 and Programmed Death-1.* Clinical Cancer Research, 2012. **18**(24): p. 6580-6587.
129. Chen, D.S. and I. Mellman, *Oncology Meets Immunology: The Cancer-Immunity Cycle.* Immunity, 2013. **39**(1): p. 1-10.
130. Das, R., et al., *Combination Therapy with Anti-CTLA-4 and Anti-PD-1 Leads to Distinct Immunologic Changes In Vivo.* Journal of Immunology, 2015. **194**(3): p. 950-959.
131. Crittenden, M., et al., *Current Clinical Trials Testing Combinations of Immunotherapy and Radiation.* Seminars in Radiation Oncology, 2015. **25**(1): p. 54-64.
132. Tang, C., et al., *Combining Radiation and Immunotherapy: A New Systemic Therapy for Solid Tumors?* Cancer Immunology Research, 2014. **2**(9): p. 831-838.
133. Galluzzi, L., et al., *The secret ally: immunostimulation by anticancer drugs.* Nature Reviews Drug Discovery, 2012. **11**(3): p. 215-233.
134. Burnet, F.M., *Immunological Aspects of Malignant Disease.* Lancet, 1967. **1**(7501): p. 1171-&.
135. Kaplan, D.H., et al., *Demonstration of an interferon gamma-dependent tumor surveillance system in immunocompetent mice.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(13): p. 7556-7561.
136. Shankaran, V., et al., *IFN gamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity.* Nature, 2001. **410**(6832): p. 1107-1111.
137. Linsley, P.S., et al., *CTLA-4 is a second receptor for the B cell activation antigen B7.* J Exp Med, 1991. **174**(3): p. 561-9.
138. Mellman, I., G. Coukos, and G. Dranoff, *Cancer immunotherapy comes of age.* Nature, 2011. **480**(7378): p. 480-9.
139. Okazaki, T., et al., *A rheostat for immune responses: the unique properties of PD-1 and their advantages for clinical application.* Nat Immunol, 2013. **14**(12): p. 1212-8.
140. Phan, G.Q., et al., *Cancer regression and autoimmunity induced by cytotoxic T lymphocyte-associated antigen 4 blockade in patients with metastatic melanoma.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8372-7.
141. Ribas, A., *Releasing the Brakes on Cancer Immunotherapy.* N Engl J Med, 2015. **373**(16): p. 1490-2.
142. Melief, C.J.M., et al., *Therapeutic cancer vaccines.* Journal of Clinical Investigation, 2015. **125**(9): p. 3401-3412.
143. Rosenberg, S.A. and N.P. Restifo, *Adoptive cell transfer as personalized immunotherapy for human cancer.* Science, 2015. **348**(6230): p. 62-68.
144. Morgan, R.A., et al., *Cancer regression in patients after transfer of genetically engineered lymphocytes.* Science, 2006. **314**(5796): p. 126-129.
145. Porter, D.L., et al., *Chimeric Antigen Receptor-Modified T Cells in Chronic Lymphoid Leukemia.* New England Journal of Medicine, 2011. **365**(8): p. 725-733.

146. Efremova, M., et al., *Neoantigens Generated by Individual Mutations and Their Role in Cancer Immunity and Immunotherapy.* Frontiers in Immunology, 2017. **8**.
147. Trimble, C.L., et al., *Safety, efficacy, and immunogenicity of VGX-3100, a therapeutic synthetic DNA vaccine targeting human papillomavirus 16 and 18 E6 and E7 proteins for cervical intraepithelial neoplasia 2/3: a randomised, double-blind, placebo-controlled phase 2b trial.* Lancet, 2015. **386**(10008): p. 2078-2088.
148. Gubin, M.M., et al., *Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens.* Nature, 2014. **515**(7528): p. 577-+.
149. Yadav, M., et al., *Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing.* Nature, 2014. **515**(7528): p. 572-6.
150. Mellman, I., G. Coukos, and G. Dranoff, *Cancer immunotherapy comes of age.* Nature, 2011. **480**(7378): p. 480-489.
151. Rosenberg, S.A., J.C. Yang, and N.P. Restifo, *Cancer immunotherapy: moving beyond current vaccines.* Nature Medicine, 2004. **10**(9): p. 909-915.
152. Seliktar-Ofir, S., et al., *Selection of Shared and Neoantigen-Reactive T Cells for Adoptive Cell Therapy Based on CD137 Separation.* Frontiers in Immunology, 2017. **8**.
153. Inozume, T., et al., *Selection of CD8(+)PD-1(+) Lymphocytes in Fresh Human Melanomas Enriches for Tumor-reactive T Cells.* Journal of Immunotherapy, 2010. **33**(9): p. 956-964.
154. Beane, J.D., et al., *Clinical Scale Zinc Finger Nuclease-mediated Gene Editing of PD-1 in Tumor Infiltrating Lymphocytes for the Treatment of Metastatic Melanoma.* Molecular Therapy, 2015. **23**(8): p. 1380-1390.
155. Veatch, J.R., et al., *Tumor-infiltrating BRAF(V600E)-specific CD4(+) T cells correlated with complete clinical response in melanoma.* Journal of Clinical Investigation, 2018. **128**(4): p. 1563-1568.
156. Scheuermann, R.H. and E. Racila, *Cd19 Antigen in Leukemia and Lymphoma Diagnosis and Immunotherapy.* Leukemia & Lymphoma, 1995. **18**(5-6): p. 385-+.
157. Buchbinder, E.I. and A. Desai, *CTLA-4 and PD-1 Pathways Similarities, Differences, and Implications of Their Inhibition.* American Journal of Clinical Oncology-Cancer Clinical Trials, 2016. **39**(1): p. 98-106.
158. Coulie, P.G., et al., *Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy.* Nature Reviews Cancer, 2014. **14**(2): p. 135-146.
159. Jurtz, V., et al., *NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data.* Journal of Immunology, 2017. **199**(9): p. 3360-3368.
160. Iyer, J.G., et al., *Merkel Cell Polyomavirus-Specific CD8(+) and CD4(+) T-cell Responses Identified in Merkel Cell Carcinomas and Blood.* Clinical Cancer Research, 2011. **17**(21): p. 6671-6680.
161. Lee, S.P., et al., *CTL control of EBV in nasopharyngeal carcinoma (NPC): EBV-specific CTL responses in the blood and tumors of NPC patients and the antigen-processing function of the tumor cells.* Journal of Immunology, 2000. **165**(1): p. 573-582.
162. Coulie, P.G., et al., *A Mutated Intron Sequence Codes for an Antigenic Peptide Recognized by Cytolytic T-Lymphocytes on a Human-Melanoma.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(17): p. 7976-7980.

163. Robbins, P.F., et al., *The intronic region of an incompletely spliced gp100 gene transcript encodes an epitope recognized by melanoma-reactive tumor-infiltrating lymphocytes*. Journal of Immunology, 1997. **159**(1): p. 303-308.
164. Smart, A.C., et al., *Intron retention is a source o neoepitopes in cancer.* Nature Biotechnology, 2018. **36**(11): p. 1056-+.
165. Margolis, C.A., et al., *RNA-based tumor neoantigens from intron retention.* Cancer Research, 2018. **78**(13).
166. Kahles, A., et al., *Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients*. Cancer Cell, 2018. **34**(2): p. 211-+.
167. Jayasinghe, R.G., et al., *Systematic Analysis of Splice-Site-Creating Mutations in Cancer.* Cell Reports, 2018. **23**(1): p. 270-+.
168. Wang, R.F., et al., *Utilization of an alternative open reading frame of a normal gene in generating a novel human cancer antigen.* Journal of Experimental Medicine, 1996. **183**(3): p. 1131-1140.
169. Aarnoudse, C.A., et al., *Interleukin-2-induced, melanoma-specific T cells recognize camel, an unexpected translation product of LAGE-1.* International Journal of Cancer, 1999. **82**(3): p. 442-448.
170. Probst-Kepper, M., et al., *An alternative open reading frame of the human macrophage colony-stimulating factor gene is independently translated and codes for an antigenic peptide of 14 amino acids recognized by tumor-infiltrating CD8 T lymphocytes.* Journal of Experimental Medicine, 2001. **193**(10): p. 1189-1198.
171. Shichijo, S., et al., *A gene encoding antigenic peptides of human squamous cell carcinoma recognized by cytotoxic T lymphocytes.* Journal of Experimental Medicine, 1998. **187**(3): p. 277-288.
172. Wang, R.F., et al., *A breast and melanoma-shared tumor antigen: T cell responses to antigenic peptides translated from different open reading frames*. Journal of Immunology, 1998. **161**(7): p. 3596-3606.
173. Hanada, K., J.W. Yewdell, and J.C. Yang, *Immune recognition of a human renal cancer antigen through post-translational protein splicing.* Nature, 2004. **427**(6971): p. 252-256.
174. Michaux, A., et al., *A Spliced Antigenic Peptide Comprising a Single Spliced Amino Acid Is Produced in the Proteasome by Reverse Splicing of a Longer Peptide Fragment followed by Trimming.* Journal of Immunology, 2014. **192**(4): p. 1962-1971.
175. Warren, E.H., et al., *An antigen produced by splicing of noncontiguous peptides in the reverse order.* Science, 2006. **313**(5792): p. 1444-1447.
176. Dalet, A., et al., *An antigenic peptide produced by reverse splicing and double asparagine deamidation*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(29): p. E323-E331.
177. Djebali, S., et al., *Landscape of transcription in human cells*. Nature, 2012. **489**(7414): p. 101-108.
178. Emst, C. and C.C. Morton, *Identification and function of long non-coding RNA*. Frontiers in Cellular Neuroscience, 2013. **7**.
179. Uszczynska-Ratajczak, B., et al., *Towards a complete map of the human long non-coding RNA transcriptome.* Nature Reviews Genetics, 2018. **19**(9): p. 535-548.

180. Fang, S.S., et al., *NONCODEV5: a comprehensive annotation database for long non-coding RNAs.* Nucleic Acids Research, 2018. **46**(D1): p. D308-D314.
181. Statello, L., et al., *Gene regulation by long non-coding RNAs and its biological functions.* Nature Reviews Molecular Cell Biology, 2021. **22**(2): p. 96-118.
182. Derrien, T., et al., *The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression.* Genome Research, 2012. **22**(9): p. 1775-1789.
183. Tian, B. and J.L. Manley, *Alternative polyadenylation of mRNA precursors.* Nature Reviews Molecular Cell Biology, 2017. **18**(1): p. 18-30.
184. Guo, C.J., et al., *Distinct Processing of lncRNAs Contributes to Non-conserved Functions in Stem Cells.* Cell, 2020. **181**(3): p. 621-+.
185. Mele, M., et al., *Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs.* Genome Research, 2017. **27**(1): p. 27-37.
186. Rosenberg, A.B., et al., *Learning the Sequence Determinants of Alternative Splicing from Millions of Random Sequences.* Cell, 2015. **163**(3): p. 698-711.
187. Mercer, T.R., M.E. Dinger, and J.S. Mattick, *Long non-coding RNAs: insights into functions.* Nature Reviews Genetics, 2009. **10**(3): p. 155-159.
188. Guttman, M. and J.L. Rinn, *Modular regulatory principles of large non-coding RNAs.* Nature, 2012. **482**(7385): p. 339-346.
189. Liu, S.K., et al., *MALAT1-miR-124-RBG2 axis is involved in growth and invasion of HR-HPV-positive cervical cancer cells.* Tumor Biology, 2016. **37**(1): p. 633-640.
190. Guo, X.G., et al., *Long noncoding RNA HOTAIR knockdown inhibits autophagy and epithelial-mesenchymal transition through the Wnt signaling pathway in radioresistant human cervical cancer HeLa cells.* Journal of Cellular Physiology, 2019. **234**(4): p. 3478-3489.
191. Li, J., et al., *HOTAIR enhanced aggressive biological behaviors and induced radio-resistance via inhibiting p21 in cervical cancer.* Tumor Biology, 2015. **36**(5): p. 3611-3619.
192. Li, N., et al., *Overexpression of HOTAIR leads to radioresistance of human cervical cancer via promoting HIF-1 alpha expression.* Radiation Oncology, 2018. **13**.
193. Liu, M.Z., et al., *Long non-coding RNA HOTAIR promotes cervical cancer progression through regulating BCL2 via targeting miR-143-3p.* Cancer Biology & Therapy, 2018. **19**(5): p. 391-399.
194. Sharma, S., et al., *Bridging Links between Long Noncoding RNA HOTAIR and HPV Oncoprotein E7 in Cervical Cancer Pathogenesis.* Scientific Reports, 2015. **5**.
195. Sun, J.B., et al., *Long non-coding RNA HOTAIR modulates HLA-G expression by absorbing miR-148a in human cervical cancer.* International Journal of Oncology, 2016. **49**(3): p. 943-952.
196. Zhang, M., Y.H. Song, and F.R. Zhai, *ARFHPV E7 oncogene, lncRNA HOTAIR, miR-331-3p and its target, NRP2, form a negative feedback loop to regulate the apoptosis in the tumorigenesis in HPV positive cervical cancer.* Journal of Cellular Biochemistry, 2018. **119**(6): p. 4397-4407.

197. Zheng, P., et al., *LncRNA HOTAIR promotes cell migration and invasion by regulating MKL1 via inhibition miR206 expression in HeLa cells.* Cell Communication and Signaling, 2018. **16**.
198. Ou, L., et al., *Decreased Expression of miR-138-5p by lncRNA H19 in Cervical Cancer Promotes Tumor Proliferation.* Oncology Research, 2018. **26**(3): p. 401-410.
199. Han, X.T., et al., *Long non-coding RNA metastasis-associated lung adenocarcinoma transcript 1/microRNA-202-3p/periostin axis modulates invasion and epithelial-mesenchymal transition in human cervical cancer.* Journal of Cellular Physiology, 2019. **234**(8): p. 14170-14180.
200. Lu, H.Z., et al., *Long non-coding RNA MALAT1 modulates radiosensitivity of HR-HPV plus cervical cancer via sponging miR-145.* Tumor Biology, 2016. **37**(2): p. 1683-1691.
201. Zhu, L.H., et al., *Interference of the long noncoding RNA CDKN2B-AS1 upregulates miR-181a-5p/TGF beta I axis to restrain the metastasis and promote apoptosis and senescence of cervical cancer cells.* Cancer Medicine, 2019. **8**(4): p. 1721-1730.
202. Yan, S.P., et al., *LncRNA LINC01305 silencing inhibits cell epithelial-mesenchymal transition in cervical cancer by inhibiting TNXB-mediated PI3K/Akt signalling pathway.* Journal of Cellular and Molecular Medicine, 2019. **23**(4): p. 2656-2666.
203. Wilhelm, M., et al., *Mass-spectrometry-based draft of the human proteome.* Nature, 2014. **509**(7502): p. 582-+.
204. Kim, M.S., et al., *A draft map of the human proteome.* Nature, 2014. **509**(7502): p. 575-+.
205. Slavoff, S.A., et al., *Peptidomic discovery of short open reading frame-encoded peptides in human cells.* Nature Chemical Biology, 2013. **9**(1): p. 59-+.
206. Rinn, J.L., et al., *Functional demarcation of active and silent chromatin domains in human HOX loci by Noncoding RNAs.* Cell, 2007. **129**(7): p. 1311-1323.
207. Zhang, D., et al., *Propofol promotes cell apoptosis via inhibiting HOTAIR mediated mTOR pathway in cervical cancer.* Biochemical and Biophysical Research Communications, 2015. **468**(4): p. 561-567.
208. Lee, M., et al., *The long non-coding RNA HOTAIR increases tumour growth and invasion in cervical cancer by targeting the Notch pathway.* Oncotarget, 2016. **7**(28): p. 44558-44571.
209. Kim, H.J., et al., *Long non-coding RNA HOTAIR is associated with human cervical cancer progression.* International Journal of Oncology, 2015. **46**(2): p. 521-530.
210. Sun, R.L., et al., *Down-regulation of MALAT1 inhibits cervical cancer cell invasion and metastasis by inhibition of epithelial-mesenchymal transition.* Molecular Biosystems, 2016. **12**(3): p. 952-962.
211. Zhu, J., et al., *Long non-coding RNA TUG1 promotes cervical cancer progression by regulating the miR-138-5p-SIRT1 axis.* Oncotarget, 2017. **8**(39): p. 65253-65264.
212. Lin, K.Y., et al., *Treatment of established tumors with a novel vaccine that enhances major histocompatibility class II presentation of tumor antigen.* Cancer Research, 1996. **56**(1): p. 21-26.
213. Hancock, G., et al., *A multi-genotype therapeutic human papillomavirus vaccine elicits potent T cell responses to conserved regions of early proteins.* Scientific Reports, 2019. **9**.
214. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nature Methods, 2012. **9**(4): p. 357-U54.

215. Li, H., *A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data.* Bioinformatics, 2011. **27**(21): p. 2987-2993.
216. Chen, X.Y., et al., *Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications*. Bioinformatics, 2016. **32**(8): p. 1220-1222.
217. Bjerregaard, A.M., et al., *MuPeXI: prediction of neo-epitopes from tumor sequencing data.* Cancer Immunology Immunotherapy, 2017. **66**(9): p. 1123-1130.
218. McLaren, W., et al., *The Ensembl Variant Effect Predictor*. Genome Biology, 2016. **17**.
219. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
220. Li, H., et al., *The Sequence Alignment/Map format and SAMtools*. Bioinformatics, 2009. **25**(16): p. 2078-2079.
221. McKenna, A., et al., *The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data*. Genome Research, 2010. **20**(9): p. 1297-1303.
222. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-930.
223. Inkman, M.J., et al., *HPV-EM: an accurate HPV detection and genotyping EM algorithm.* Scientific Reports, 2020. **10**(1).
224. Grabherr, M.G., et al., *Full-length transcriptome assembly from RNA-Seq data without a reference genome.* Nature Biotechnology, 2011. **29**(7): p. 644-U130.
225. Camacho, C., et al., *BLAST plus : architecture and applications.* Bmc Bioinformatics, 2009. **10**.
226. Guy, L., J. Roat Kultima, and S.G.E. Andersson, *genoPlotR: comparative gene and genome visualization in R.* Bioinformatics, 2010. **26**(18): p. 2334-2335.
227. Charif, D., et al., *Online synonymous codon usage analyses with the ade4 and seqinR packages.* Bioinformatics, 2005. **21**(4): p. 545-547.
228. Andreatta, M., et al., *MS-Rescue: A Computational Pipeline to Increase the Quality and Yield of Immunopeptidomics Experiments*. Proteomics, 2019. **19**(4): p. e1800357.
229. Ma, C.W., et al., *Improved Peptide Retention Time Prediction in Liquid Chromatography through Deep Learning.* Analytical Chemistry, 2018. **90**(18): p. 10881-10888.
230. Volders, P.J., et al., *LNCipedia 5: towards a reference set of human long non-coding RNAs.* Nucleic Acids Research, 2019. **47**(D1): p. D135-D139.
231. Seifuddin, F., et al., *lncRNAKB, a knowledgebase of tissue-specific functional annotation and trait association of long noncoding RNA.* Scientific Data, 2020. **7**(1).
232. Andreatta, M., B. Alvarez, and M. Nielsen, *GibbsCluster: unsupervised clustering and alignment of peptide sequences.* Nucleic Acids Research, 2017. **45**(W1): p. W458-W463.
233. Jurtz, V., et al., *NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data.* J Immunol, 2017. **199**(9): p. 3360-3368.
234. Neisig, A., C.J.M. Melief, and J. Neefjes, *Reduced cell surface expression of HLA-C molecules correlates with restricted peptide binding and stable TAP interaction.* Journal of Immunology, 1998. **160**(1): p. 171-179.

235. Stevanovic, S., et al., *Landscape of immunogenic tumor antigens in successful immunotherapy of virally induced epithelial cancer.* Science, 2017. **356**(6334): p. 200-205.
236. Ojesina, A.I., et al., *Landscape of genomic alterations in cervical carcinomas.* Nature, 2014. **506**(7488): p. 371-+.
237. Qin, Y., et al., *Cervical Cancer Neoantigen Landscape and Immune Activity is Associated with Human Papillomavirus Master Regulator.* Frontiers in Immunology, 2017. **8**.
238. Smith, S.P., et al., *Identification of host transcriptional networks showing concentration-dependent regulation by HPV16 E6 and E7 proteins in basal cervical squamous epithelial cells.* Scientific Reports, 2016. **6**.
239. Almeida, L.G., et al., *CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens.* Nucleic Acids Res, 2009. **37**(Database issue): p. D816-9.
240. Baldauf, M.C., et al., *Systematic identification of cancer-specific MHC-binding peptides with RAVEN.* Oncoimmunology, 2018. **7**(9): p. e1481558.
241. Uhlen, M., et al., *Proteomics. Tissue-based map of the human proteome.* Science, 2015. **347**(6220): p. 1260419.
242. Ressing, M.E., et al., *Occasional memory cytotoxic T-cell responses of patients with human papillomavirus type 16-positive cervical lesions against a human leukocyte antigen-A *0201-restricted E7-encoded epitope.* Cancer Res, 1996. **56**(3): p. 582-8.
243. Xu, B., et al., *Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas.* PLoS One, 2013. **8**(6): p. e66693.
244. Yu, L.L., A.; Majerciak, V.; Mirza, S.; Band, V.; Liu, H.; Cam, M.; Zheng, Z., *Expression of HPV oncogenes from a single integration site in cervical cancer cell lines.* BioRxiv, 2020.
245. Bekker-Jensen, D.B., et al., *An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes.* Cell Systems, 2017. **4**(6): p. 587-+.
246. Granados, D.P., et al., *Impact of genomic polymorphisms on the repertoire of human MHC class I-associated peptides.* Nature Communications, 2014. **5**.
247. Berlin, C., et al., *Mapping the HLA ligandome landscape of acute myeloid leukemia: a targeted approach toward peptide-based immunotherapy.* Leukemia, 2015. **29**(3): p. 647-659.
248. Bassani-Sternberg, M., et al., *Mass Spectrometry of Human Leukocyte Antigen Class I Peptidomes Reveals Strong Effects of Protein Abundance and Turnover on Antigen Presentation.* Molecular & Cellular Proteomics, 2015. **14**(3): p. 658-673.
249. Ressing, M.E., et al., *Human Ctl Epitopes Encoded by Human Papillomavirus Type-16 E6 and E7 Identified through in-Vivo and in-Vitro Immunogenicity Studies of Hla-a-Asterisk-0201-Binding Peptides.* Journal of Immunology, 1995. **154**(11): p. 5934-5943.
250. Villada, I.B., et al., *Identification in humans of HPV-16 E6 and E7 protein epitopes recognized by cytolytic T lymphocytes in association with HLA-B18 and determination of the HLA-B18-specific binding motif.* European Journal of Immunology, 2000. **30**(8): p. 2281-2289.
251. Keskin, D.B., et al., *Direct identification of an HPV-16 tumor antigen from cervical cancer biopsy specimens.* Front Immunol, 2011. **2**: p. 75.

252. Bergvall, M., T. Melendy, and J. Archambault, *The E1 proteins.* Virology, 2013. **445**(1-2): p. 35-56.
253. Unterholzner, L., et al., *IFI16 is an innate immune sensor for intracellular DNA.* Nature Immunology, 2010. **11**(11): p. 997-U42.
254. Huang, Y., et al., *A novel human hydroxysteroid dehydrogenase like 1 gene (HSDL1) is highly expressed in reproductive tissues.* Molecular Biology Reports, 2001. **28**(4): p. 185-191.
255. Meier, M., et al., *Human and zebrafish hydroxysteroid dehydrogenase like 1 (HSDL1) proteins are inactive enzymes but conserved among species.* Chemico-Biological Interactions, 2009. **178**(1-3): p. 197-205.
256. Brooks, S.P., et al., *Identification of the gene for Nance-Horan syndrome (NHS).* Journal of Medical Genetics, 2004. **41**(10): p. 768-771.
257. Kloor, M., et al., *Identification and characterization of UEV3, a human cDNA with similarities to inactive E2 ubiquitin-conjugating enzymes.* Biochimica Et Biophysica Acta-Gene Structure and Expression, 2002. **1579**(2-3): p. 219-224.
258. Deng, Z.H., et al., *Nuclear FAM21 participates in NF-kappa B-dependent gene regulation in pancreatic cancer cells.* Journal of Cell Science, 2015. **128**(2): p. 373-384.
259. Lee, S., J. Chang, and C. Blackstone, *FAM21 directs SNX27-retromer cargoes to the plasma membrane by preventing transport to the Golgi apparatus.* Nature Communications, 2016. **7**.
260. Fabbro, M., et al., *Cdk1/Erk2- and Plk1-dependent phosphorylation of a centrosome protein, Cep55, is required for its recruitment to midbody and cytokinesis.* Developmental Cell, 2005. **9**(4): p. 477-488.
261. Zhao, W.M., A. Seki, and G.W. Fang, *Cep55, a microtubule-bundling protein, associates with centralspindlin to control the midbody integrity and cell abscission during cytokinesis.* Molecular Biology of the Cell, 2006. **17**(9): p. 3881-3896.
262. Frosk, P., et al., *A truncating mutation in CEP55 is the likely cause of MARCH, a novel syndrome affecting neuronal mitosis.* Journal of Medical Genetics, 2017. **54**(7): p. 490-501.
263. Cheeseman, I.M., et al., *KNL1 and the CENP-H/I/K complex coordinately direct kinetochore assembly in vertebrates.* Molecular Biology of the Cell, 2008. **19**(2): p. 587-594.
264. Obuse, C., et al., *A conserved Mis12 centromere complex is linked to heterochromatic HP1 and outer kinetochore protein Zwint-1.* Nature Cell Biology, 2004. **6**(11): p. 1135-U37.
265. Liu, D., et al., *Human NUF2 interacts with centromere-associated protein E and is essential for a stable spindle microtubule-kinetochore attachment.* Journal of Biological Chemistry, 2007. **282**(29): p. 21415-21424.
266. DeLuca, J.G., et al., *hNuf2 inhibition blocks stable kinetochore-microtubule attachment and induces mitotic cell death in HeLa cells.* Journal of Cell Biology, 2002. **159**(4): p. 549-555.
267. Wigge, P.A. and J.V. Kilmartin, *The Ndc80p complex from Saccharomyces cerevisiae contains conserved centromere components and has a function in chromosome segregation.* Journal of Cell Biology, 2001. **152**(2): p. 349-360.

268. Martin-Lluesma, S., V.M. Stucke, and E.A. Nigg, *Role of Hec1 in spindle checkpoint signaling and kinetochore recruitment of Mad1/Mad2.* Science, 2002. **297**(5590): p. 2267-2270.
269. He, X.W., et al., *Molecular analysis of kinetochore-microtubule attachment in budding yeast.* Cell, 2001. **106**(2): p. 195-206.
270. Hori, T., et al., *Dynamic behavior of Nuf2-Hec1 complex that localizes to the centrosome and centromere and is essential for mitotic progression in vertebrate cells.* Journal of Cell Science, 2003. **116**(16): p. 3347-3362.
271. DeLuca, J.G., et al., *Hec1 and Nuf2 are core components of the kinetochore outer plate essential for organizing microtubule attachment sites.* Molecular Biology of the Cell, 2005. **16**(2): p. 519-531.
272. Tanenbaum, M.E., et al., *A Complex of Kif18b and MCAK Promotes Microtubule Depolymerization and Is Negatively Regulated by Aurora Kinases.* Current Biology, 2011. **21**(16): p. 1356-1365.
273. Bakhoum, S.F., et al., *Genome stability is ensured by temporal control of kinetochore-microtubule dynamics.* Nature Cell Biology, 2009. **11**(1): p. 27-U51.
274. Hewitt, S.C., J.C. Harrell, and K.S. Korach, *Lessons in estrogen biology from knockout and transgenic animals.* Annual Review of Physiology, 2005. **67**: p. 285-308.
275. Dahlman-Wright, K., et al., *International Union of Pharmacology. LXIV. Estrogen receptors.* Pharmacological Reviews, 2006. **58**(4): p. 773-781.
276. Wehner, K.A., J.E.G. Gallagher, and S.J. Baserga, *Components of an interdependent unit within the SSU processome regulate and mediate its activity.* Molecular and Cellular Biology, 2002. **22**(20): p. 7258-7267.
277. Hughes, J.M.X. and M. Ares, *Depletion of U3 Small Nucleolar Rna Inhibits Cleavage in the 5' External Transcribed Spacer of Yeast Pre-Ribosomal Rna and Impairs Formation of 18s Ribosomal-Rna.* Embo Journal, 1991. **10**(13): p. 4231-4239.
278. Kass, S., et al., *The U3 Small Nucleolar Ribonucleoprotein Functions in the 1st Step of Preribosomal Rna Processing.* Cell, 1990. **60**(6): p. 897-908.
279. Savino, R. and S.A. Gerbi, *Invivo Disruption of Xenopus U3 Snrna Affects Ribosomal-Rna Processing.* Embo Journal, 1990. **9**(7): p. 2299-2308.
280. Matsumoto, S., et al., *Characterization of a MAPKK-like protein kinase TOPK.* Biochemical and Biophysical Research Communications, 2004. **325**(3): p. 997-1004.
281. Park, J.H., et al., *PDZ-binding kinase/T-LAK cell-originated protein kinase, a putative cancer/testis antigen with an oncogenic activity in breast cancer.* Cancer Research, 2006. **66**(18): p. 9186-9195.
282. Nandi, A.K., et al., *Attenuation of DNA damage checkpoint by PBK, a novel mitotic kinase, involves protein-protein interaction with tumor suppressor p53.* Biochemical and Biophysical Research Communications, 2007. **358**(1): p. 181-188.
283. Bird, A., *Perceptions of epigenetics.* Nature, 2007. **447**(7143): p. 396-398.
284. Martin, C. and Y. Zhang, *The diverse functions of histone lysine methylation.* Nature Reviews Molecular Cell Biology, 2005. **6**(11): p. 838-849.
285. Mosammaparast, N. and Y. Shi, *Reversal of Histone Methylation: Biochemical and Molecular Mechanisms of Histone Demethylases.* Annual Review of Biochemistry, Vol 79, 2010. **79**: p. 155-179.

286. Klose, R.J., E.M. Kallin, and Y. Zhang, *JmjC-domain-containing proteins and histone demethylation.* Nature Reviews Genetics, 2006. **7**(9): p. 715-727.
287. Christensen, J., et al., *RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3.* Cell, 2007. **128**(6): p. 1063-1076.
288. Iwase, S., et al., *The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases.* Cell, 2007. **128**(6): p. 1077-1088.
289. Klose, R.J., et al., *The retinoblastoma binding protein RBP2 is an H3K4 demethylase.* Cell, 2007. **128**(5): p. 889-900.
290. Yamane, K., et al., *PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation.* Mol Cell, 2007. **25**(6): p. 801-12.
291. Schmitz, S.U., et al., *Jarid1b targets genes regulating development and is involved in neural differentiation.* Embo Journal, 2011. **30**(22): p. 4586-4600.
292. Barrett, A., et al., *PLU-1 nuclear protein, which is upregulated in breast cancer, shows restricted expression in normal human adult tissues: A new cancer/testis antigen?* International Journal of Cancer, 2002. **101**(6): p. 581-588.
293. Lu, P.J., et al., *A novel gene (PLU-1) containing highly conserved putative DNA chromatin binding motifs is specifically up-regulated in breast cancer.* Journal of Biological Chemistry, 1999. **274**(22): p. 15633-15645.
294. Wang, L.S., et al., *Overexpression of JARID1B is associated with poor prognosis and chemotherapy resistance in epithelial ovarian cancer.* Tumor Biology, 2015. **36**(4): p. 2465-2472.
295. Wang, Z.R., et al., *KDM5B is overexpressed in gastric cancer and is required for gastric cancer cell proliferation and metastasis.* American Journal of Cancer Research, 2015. **5**(1): p. 87-100.
296. Hayami, S., et al., *Overexpression of the JmjC histone demethylase KDM5B in human carcinogenesis: involvement in the proliferation of cancer cells through the E2F/RB pathway.* Molecular Cancer, 2010. **9**.
297. Wang, D., et al., *Depletion of histone demethylase KDM5B inhibits cell proliferation of hepatocellular carcinoma by regulation of cell cycle checkpoint proteins p15 and p27.* Journal of Experimental & Clinical Cancer Research, 2016. **35**.
298. Ohta, K., et al., *Depletion of JARID1B induces cellular senescence in human colorectal cancer.* International Journal of Oncology, 2013. **42**(4): p. 1212-1218.
299. Dai, B., et al., *Overexpressed KDM5B is associated with the progression of glioma and promotes glioma cell growth via downregulating p21.* Biochemical and Biophysical Research Communications, 2014. **454**(1): p. 221-227.
300. Roesch, A., et al., *RBP2-H1/JARID1B is a transcriptional regulator with a tumor suppressive potential in melanoma cells.* International Journal of Cancer, 2008. **122**(5): p. 1047-1057.
301. Baldauf, M.C., et al., *Systematic identification of cancer-specific MHC-binding peptides with RAVEN.* Oncoimmunology, 2018. **7**(9).
302. Ellgaard, L. and A. Helenius, *Quality control in the endoplasmic reticulum.* Nature Reviews Molecular Cell Biology, 2003. **4**(3): p. 181-191.
303. Romisch, K., *Endoplasmic reticulum-associated degradation.* Annual Review of Cell and Developmental Biology, 2005. **21**: p. 435-456.

304. Alcock, F. and E. Swanton, *Mammalian OS-9 Is Upregulated in Response to Endoplasmic Reticulum Stress and Facilitates Ubiquitination of Misfolded Glycoproteins.* Journal of Molecular Biology, 2009. **385**(4): p. 1032-1042.
305. Wang, Y., et al., *OS-9 regulates the transit and polyubiquitination of TRPV4 in the endoplasmic reticulum.* Journal of Biological Chemistry, 2007. **282**(50): p. 36561-36570.
306. Feng, Y., B. Press, and A. Wandingerness, *Rab-7 - an Important Regulator of Late Endocytic Membrane Traffic.* Journal of Cell Biology, 1995. **131**(6): p. 1435-1452.
307. Press, B., et al., *Mutant Rab7 causes the accumulation of cathepsin D and cation-independent mannose 6-phosphate receptor in an early endocytic compartment.* Journal of Cell Biology, 1998. **140**(5): p. 1075-1089.
308. Christoforidis, S., et al., *Phosphatidylinositol-3-OH kinases are Rab5 effectors.* Nature Cell Biology, 1999. **1**(4): p. 249-252.
309. Patki, V., et al., *Identification of an early endosomal protein regulated by phosphatidylinositol 3-kinase.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(14): p. 7326-7330.
310. Futter, C.E., et al., *Human VPS34 is required for internal vesicle formation within multivesicular endosomes.* Journal of Cell Biology, 2001. **155**(7): p. 1251-1263.
311. Bucci, C., et al., *Rab7: A key to lysosome biogenesis.* Molecular Biology of the Cell, 2000. **11**(2): p. 467-480.
312. Eick, S., et al., *Human H1 histones: conserved and varied sequence elements in two H1 subtype genes.* Eur J Cell Biol, 1989. **49**(1): p. 110-5.
313. Marzluff, W.F., et al., *The human and mouse replication-dependent histone genes.* Genomics, 2002. **80**(5): p. 487-98.
314. Vasu, S.K. and D.J. Forbes, *Nuclear pores and nuclear assembly.* Current Opinion in Cell Biology, 2001. **13**(3): p. 363-375.
315. Allen, T.D., et al., *The nuclear pore complex: mediator of translocation between nucleus and cytoplasm.* Journal of Cell Science, 2000. **113**(10): p. 1651-1659.
316. Ryan, K.J. and S.R. Wente, *The nuclear pore complex: a protein machine bridging the nucleus and cytoplasm.* Current Opinion in Cell Biology, 2000. **12**(3): p. 361-371.
317. Rout, M.P. and J.D. Aitchison, *The nuclear pore complex as a transport machine.* Journal of Biological Chemistry, 2001. **276**(20): p. 16593-16596.
318. Cronshaw, J.A., et al., *Proteomic analysis of the mammalian nuclear pore complex.* Journal of Cell Biology, 2002. **158**(5): p. 915-927.
319. Boehmer, T., et al., *Depletion of a single nucleoporin, Nup107, prevents the assembly of a subset of nucleoporins into the nuclear pore complex.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(3): p. 981-985.
320. Braun, D.A., et al., *Mutations in multiple components of the nuclear pore complex cause nephrotic syndrome.* Journal of Clinical Investigation, 2018. **128**(10): p. 4313-4328.
321. Chua, K.L. and A. Hjerpe, *Persistence of human papillomavirus (HPV) infections preceding cervical carcinoma.* Cancer, 1996. **77**(1): p. 121-127.
322. Huibregtse, J.M., M. Scheffner, and P.M. Howley, *A Cellular Protein Mediates Association of P53 with the E6 Oncoprotein of Human Papillomavirus Type-16 or Type-18.* Embo Journal, 1991. **10**(13): p. 4129-4135.

323. Zhang, J., et al., *PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification.* Mol Cell Proteomics, 2012. **11**(4): p. M111 010587.
324. Kotecha, M.T., et al., *Enhanced tumour growth after DNA vaccination against human papilloma virus E7 oncoprotein: evidence for tumour-induced immune deviation.* Vaccine, 2003. **21**(19-20): p. 2506-2515.
325. Crooks, G.E., et al., *WebLogo: A sequence logo generator.* Genome Research, 2004. **14**(6): p. 1188-1190.
326. Szolek, A., et al., *OptiType: precision HLA typing from next-generation sequencing data.* Bioinformatics, 2014. **30**(23): p. 3310-3316.
327. Volders, P.J., et al., *LNCipedia: a database for annotated human lncRNA transcript sequences and structures*. Nucleic Acids Research, 2013. **41**(D1): p. D246-D251.
328. Feng, J.C., et al., *The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator.* Genes & Development, 2006. **20**(11): p. 1470-1484.
329. Wilusz, J.E., H. Sunwoo, and D.L. Spector, *Long noncoding RNAs: functional surprises from the RNA world.* Genes & Development, 2009. **23**(13): p. 1494-1504.
330. Li, L., et al., *Role of human noncoding RNAs in the control of tumorigenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(31): p. 12956-12961.
331. Morris, K.V., et al., *Bidirectional Transcription Directs Both Transcriptional Gene Activation and Suppression in Human Cells.* Plos Genetics, 2008. **4**(11).

7.0 Appendix

Table 10. List of peptides defined by the CT database across the ten cervical cancer tumours.

Patient/Peptide	-10lgP	Mass	Length	m/z	CT_Record
1					
YLQEKLHAL	28.55	1113.618	9	557.8175	O15078|CE290_HUMAN:CEP290
HQHNVSLQL	18.98	1074.557	9	538.286	O15078|CE290_HUMAN:CEP290
VTNNVVLEAP	15.08	1054.566	10	528.2889	O43482|MS18B_HUMAN:OIP5
PGSVIRV	22.47	726.4388	7	364.2262	O60271|JIP4_HUMAN:SPAG9
KLKDSILSI	15.83	1015.628	9	508.8212	O60271|JIP4_HUMAN:SPAG9
SVPLSDALLNK	27.68	1155.65	11	578.8327	P33981|TTK_HUMAN:TTK
NMLEAVHTI	21.15	1026.517	9	514.2659	P33981|TTK_HUMAN:TTK
GVYDGREHTV	38.42	1131.531	10	566.7735	P43358|MAGA4_HUMAN:MAGEA4
KVLEHVVRV	36.35	1077.666	9	539.8402	P43358|MAGA4_HUMAN:MAGEA4
KILDLETQL	27.94	1071.618	9	536.8169	Q5BJF6|ODFP2_HUMAN:ODF2
SRYNQVVKEK	26.08	1249.678	10	417.5658	Q5BJF6|ODFP2_HUMAN:ODF2
RQKDLIMKV	28.92	1129.664	9	377.5617	Q5H9L4|TAF7L_HUMAN:TAF7L
APCDIDHASA	15.56	998.4127	10	500.2119	Q6NUI2|GPAT2_HUMAN:GPAT2
SQNAIDHKI	34.74	1024.53	9	342.5171	Q6PL18|ATAD2_HUMAN:ATAD2
GLSNHIAAL	21.29	894.4923	9	448.2539	Q6PL18|ATAD2_HUMAN:ATAD2
VLINRINEV	23	1068.629	9	535.3222	Q8N0S2|SYCE1_HUMAN:SYCE1
KQVIQTHV	24.94	951.5502	8	476.7827	Q8NG31|KNL1_HUMAN:CASC5

KQDQSNVQI	24.56	1058.536	9	530.2756	Q8NG31|KNL1_HUMAN:CASC5
HTANIQTLI	17.65	1009.556	9	505.7861	Q8NG31|KNL1_HUMAN:CASC5
GRMHGKFSL	28.2	1031.533	9	344.8521	Q99661|KIF2C_HUMAN:MCAK
ATAQFKINKK	34.87	1147.671	10	383.5626	Q9BZD4|NUF2_HUMAN:CDCA1
ALLDKLYAL	29.88	1018.606	9	510.3111	Q9NV31|IMP3_HUMAN:IMP-3
LLDKLYAL	22.31	947.5692	8	474.7919	Q9NV31|IMP3_HUMAN:IMP-3
LSAVTPELI	21	941.5433	9	471.778	Q9NXZ1|SAGE1_HUMAN:SAGE1
AQTRVKLNF	17.05	1075.614	9	538.8156	Q9UGL1|KDM5B_HUMAN:JARID1B
RLIDLGVGL	15.32	954.5862	9	478.3003	Q9UGL1|KDM5B_HUMAN:JARID1B
2					
AELKEKIVL	30.02	1041.643	9	521.8308	O43482|MS18B_HUMAN:OIP5
YVHSSVAQW	25.11	1075.509	9	538.7628	O60271|JIP4_HUMAN:SPAG9
TESSLLAKL	25.38	960.5491	9	481.2815	P33981|TTK_HUMAN:TTK
EETDRFVVF	17.67	1140.545	9	571.2799	Q14667|K0100_HUMAN:KIAA0100
AELESKTNTL	32.64	1104.566	10	553.2917	Q53EZ4|CEP55_HUMAN:CEP55
REENDIARGKL	30.21	1299.69	11	434.2372	Q53EZ4|CEP55_HUMAN:CEP55
AETEHENTVL	17.04	1141.525	10	571.7701	Q5BJF6|ODFP2_HUMAN:ODF2
EHGIITNWDDMEKI	41.46	1699.788	14	567.6028	Q6S8J3|POTEE_HUMAN:POTEE
AASSSSLEKSYELPDGQVI	60.9	1979.969	19	990.9951	Q6S8J3|POTEE_HUMAN:POTEE
TAASSSSLEKSYELPDGQVI	58.19	2081.016	20	1041.518	Q6S8J3|POTEE_HUMAN:POTEE
LLTEAPLNPKANREKMTQI	48.13	2166.183	19	723.0688	Q6S8J3|POTEE_HUMAN:POTEE
LLTEAPLNPKANREKMTQ	46.04	2053.099	18	514.2838	Q6S8J3|POTEE_HUMAN:POTEE

LTEAPLNPKANREKMTQI	37.43	2053.099	18	514.2804	Q6S8J3|POTEE_HUMAN:POTEE
LLTEAPLNPKANREKM	33.51	1823.993	16	457.0044	Q6S8J3|POTEE_HUMAN:POTEE
LLTEAPLNPKANREKMTQIM	30.48	2297.224	20	575.3141	Q6S8J3|POTEE_HUMAN:POTEE
TEAPLNPKANREKMTQI	25.69	1940.015	17	647.6774	Q6S8J3|POTEE_HUMAN:POTEE
EKIWHHTFYNELRVAP	62.16	2039.038	16	680.687	Q6S8J3|POTEE_HUMAN:POTEE
SSSSLEKSYELPDGQVI	57.75	1837.895	17	919.9569	Q6S8J3|POTEE_HUMAN:POTEE
ASSSSLEKSYELPDGQVI	57.05	1908.932	18	955.476	Q6S8J3|POTEE_HUMAN:POTEE
HHTFYNELRVAP	55.75	1482.737	12	495.2511	Q6S8J3|POTEE_HUMAN:POTEE
MEKIWHHTF	55.74	1227.586	9	614.798	Q6S8J3|POTEE_HUMAN:POTEE
EKIWHHTFYNELRV	55.23	1870.948	14	624.6577	Q6S8J3|POTEE_HUMAN:POTEE
WHHTFYNELRVAP	54.44	1668.816	13	557.2779	Q6S8J3|POTEE_HUMAN:POTEE
SLEKSYELPDGQVI	52.28	1576.798	14	789.4081	Q6S8J3|POTEE_HUMAN:POTEE
WHHTFYNELR	50.96	1401.658	10	468.225	Q6S8J3|POTEE_HUMAN:POTEE
SSLEKSYELPDGQVI	48.82	1663.83	15	832.9241	Q6S8J3|POTEE_HUMAN:POTEE
LEKSYELPDGQVI	48.77	1489.766	13	745.8925	Q6S8J3|POTEE_HUMAN:POTEE
SSSLEKSYELPDGQVI	48.41	1750.862	16	876.4406	Q6S8J3|POTEE_HUMAN:POTEE
HTFYNELRVAP	44.95	1345.678	11	673.8442	Q6S8J3|POTEE_HUMAN:POTEE
EKIWHHTFYNELR	42.57	1771.879	13	591.6342	Q6S8J3|POTEE_HUMAN:POTEE
REIVRDIKEKL	41.31	1397.835	11	350.4655	Q6S8J3|POTEE_HUMAN:POTEE
HHTFYNELR	38.92	1215.579	9	406.1988	Q6S8J3|POTEE_HUMAN:POTEE
AEREIVRDI	38.44	1099.599	9	550.8073	Q6S8J3|POTEE_HUMAN:POTEE
EKIWHHTFYNEL	35.4	1615.	12	539.6	Q6S8J3|POTEE_HUMAN:POTEE

	4	778		005	
WHHTFYNEL	32.81	1245.557	9	623.7862	Q6S8J3|POTEE_HUMAN:POTEE
HHTFYNEL	30.96	1059.477	8	530.7468	Q6S8J3|POTEE_HUMAN:POTEE
FAGDDAPRA	30.73	918.4195	9	460.2162	Q6S8J3|POTEE_HUMAN:POTEE
EAQSKRGILTL	26.81	1214.698	11	608.3578	Q6S8J3|POTEE_HUMAN:POTEE
DIRKDLY	25.67	921.4919	7	461.7518	Q6S8J3|POTEE_HUMAN:POTEE
AGDDAPRAVFPSIV	23.32	1413.725	14	707.869	Q6S8J3|POTEE_HUMAN:POTEE
WHHTFYNE	23.29	1132.473	8	567.244	Q6S8J3|POTEE_HUMAN:POTEE
SSSSLEKSYELPD	22.21	1440.662	13	721.3384	Q6S8J3|POTEE_HUMAN:POTEE
APRAVFPSIV	17.44	1055.613	10	528.8132	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIV	15.03	789.4344	8	395.7248	Q6S8J3|POTEE_HUMAN:POTEE
WISKQEYDESGPSIV	66.25	1736.826	15	869.4218	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIVMDSGDGV	59.63	1450.672	15	726.3428	Q6S8J3|POTEE_HUMAN:POTEE
ISKQEYDESGPSIV	57.31	1550.746	14	776.3793	Q6S8J3|POTEE_HUMAN:POTEE
WISKQEYDESGPSIVHRK	57.13	2158.081	18	540.5264	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIVMDSGD	55.36	1294.582	13	648.2969	Q6S8J3|POTEE_HUMAN:POTEE
SKQEYDESGPSIV	55.27	1437.662	13	719.8394	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIVMDSGDGVTHTVPIYEG	51.6	2448.159	24	817.0626	Q6S8J3|POTEE_HUMAN:POTEE
KQEYDESGPSIV	43.44	1350.63	12	676.321	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIVMDS	39.32	1122.534	11	562.2744	Q6S8J3|POTEE_HUMAN:POTEE
DESGPSIVH	33.23	939.4297	9	470.7225	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIVMDSGDGVTHTVPIYE	33.07	2391.138	23	798.0509	Q6S8J3|POTEE_HUMAN:POTEE

HETTFNSI	31.7	947.4349	8	474.7246	Q6S8J3|POTEE_HUMAN:POTEE
ESGPSIVHRK	31.69	1108.599	10	370.5395	Q6S8J3|POTEE_HUMAN:POTEE
GDGVTHTVPIYEG	30.78	1343.636	13	672.8258	Q6S8J3|POTEE_HUMAN:POTEE
GIHETTFNSI	29.94	1117.54	10	559.7764	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIVMD	27.14	1035.502	10	518.7585	Q6S8J3|POTEE_HUMAN:POTEE
ISKQEYDESGPSIVHRK	22.97	1972.001	17	494.0086	Q6S8J3|POTEE_HUMAN:POTEE
GIHETTF	17.49	803.3813	7	402.6967	Q6S8J3|POTEE_HUMAN:POTEE
SGRTTGIVMDSG	15.97	1179.555	12	590.7851	Q6S8J3|POTEE_HUMAN:POTEE
HETTFNSIM	15.57	1078.475	9	540.2457	Q6S8J3|POTEE_HUMAN:POTEE
LDNKKRTAL	17.29	1057.624	9	353.5488	Q6S8J3|POTEE_HUMAN:POTEE
KMDIAWVI	17.67	974.5259	8	488.2722	Q6UXF1|TM108_HUMAN:TMEM108
EVVVPMPSW	16.15	1042.516	9	522.2674	Q86VD1|MORC1_HUMAN:MORC1
LSAPIHTQM	26.78	996.5062	9	499.2603	Q8NG31|KNL1_HUMAN:CASC5
SAAMDEKVI	23.22	962.4742	9	482.2444	Q8NG31|KNL1_HUMAN:CASC5
ASEFQILKY	41.52	1097.576	9	549.7961	Q96PC5|MIA2_HUMAN:CTAGE5
SEVDKTKEEL	40.65	1176.587	10	589.3015	Q96PC5|MIA2_HUMAN:CTAGE5
SEFQILKY	33.32	1026.539	8	514.2773	Q96PC5|MIA2_HUMAN:CTAGE5
READFKETLL	28.71	1220.64	10	611.3257	Q96Q89|KI20B_HUMAN:MPHOSPH1
READFKETL	24.04	1107.556	9	554.7861	Q96Q89|KI20B_HUMAN:MPHOSPH1
AEIEDIRVL	24.02	1056.582	9	529.2974	Q96Q89|KI20B_HUMAN:MPHOSPH1
HEPKLKVDL	41.08	1077.618	9	360.2131	Q99661|KIF2C_HUMAN:MCAK
RELEEKAMEEL	21.6	1375.	11	688.8	Q99661|KIF2C_HUMAN:MCAK

	1	665		404	
EVVEKYEIY	27.7	1170.581	9	586.2992	Q9BZD4|NUF2_HUMAN:CDCA1
KLMEKWLG	16.7	1003.552	8	502.784	Q9C093|SPEF2_HUMAN:SPEF2
EVTDHNLHEL	45.32	1205.568	10	603.7917	Q9NV31|IMP3_HUMAN:IMP-3
VTDHNLHEL	39.71	1076.525	9	539.2711	Q9NV31|IMP3_HUMAN:IMP-3
DDWDNRTSY	18.88	1170.458	9	586.2352	Q9UGL1|KDM5B_HUMAN:JARID1B
LVTIMNPNTL	15.79	1114.606	10	558.3085	Q9UGL1|KDM5B_HUMAN:JARID1B
3					
KEDGRVQAF	25.13	1048.53	9	525.2717	O60271|JIP4_HUMAN:SPAG9
KSEVQAII	15.43	886.5123	8	444.2639	O60271|JIP4_HUMAN:SPAG9
VEERAQAII	18.1	1027.566	9	514.7908	O60732|MAGC1_HUMAN:MAGEC1
AQAIIDT	15.03	730.3861	7	366.2014	O60732|MAGC1_HUMAN:Q8TD91|MAGC3_HUMAN:
GELLPARKI	16.65	995.6127	9	498.8141	Q14667|K0100_HUMAN:KIAA0100
RDSLVERL	17.64	986.5508	8	494.2828	Q5BJF6|ODFP2_HUMAN:ODF2
AYAIIKEEL	31.6	1048.58	9	525.2972	Q6PL18|ATAD2_HUMAN:ATAD2
GESERQLRL	22.95	1086.578	9	544.2968	Q6PL18|ATAD2_HUMAN:ATAD2
KYLTVKDYL	21.29	1141.638	9	571.826	Q6PL18|ATAD2_HUMAN:ATAD2
VYVPHIHVW	19.38	1148.613	9	383.8782	Q6PL18|ATAD2_HUMAN:ATAD2
AEQEASRVI	24.54	1001.514	9	501.7645	Q6Q759|SPG17_HUMAN:SPAG17
WQDILQQI	18.8	1042.545	8	522.2797	Q6Q759|SPG17_HUMAN:SPAG17
GEILLLPEL	15.94	995.5903	9	498.8016	Q8TC59|PIWL2_HUMAN:PIWIL2
MPKNCLLIL	16.54	1043.587	9	522.801	Q8TD91|MAGC3_HUMAN:MAGEC3
SYQKVIELF	15.3	1125.	9	563.8	Q96KB5|TOPK_HUMAN:PBK

	9	607		112	
DRIPEVHV	16.67	963.5137	8	482.7649	Q96PC5|MIA2_HUMAN:CTAGE5
AENIRNKEM	21.01	1103.539	9	552.777	Q96Q89|KI20B_HUMAN:MPHOSPH1
LEEQASRQI	29.36	1072.551	9	537.2828	Q99661|KIF2C_HUMAN:MCAK
NDLLRHLRL	31.46	1148.678	9	383.9001	Q9UGL1|KDM5B_HUMAN:JARID1B
RETVRKLGV	25.4	1056.64	9	353.2209	Q9UGL1|KDM5B_HUMAN:JARID1B
REMEALQSL	23.1	1075.533	9	538.7743	Q9UGL1|KDM5B_HUMAN:JARID1B
4					
LKHVPEGAETEQGLKRE	15.79	1920.007	17	481.0083	O15078|CE290_HUMAN:CEP290
ESKGSFSI	15.25	853.4181	8	427.7164	P04279|SEMG1_HUMAN:SEMG1
KLEPGSRHQL	21.77	1163.641	10	582.8283	Q02246|CNTN2_HUMAN:TAG
KILDLETQL	26.42	1071.618	9	536.8167	Q5BJF6|ODFP2_HUMAN:ODF2
GYVLGAPLVK	15.56	1015.607	10	508.8101	Q86UG4|SO6A1_HUMAN:SLCO6A1
RLYVGREK	26.14	1019.588	8	340.8703	Q96PC5|MIA2_HUMAN:CTAGE5
PAKKGSIHVSSAITEDQKK	15.19	2023.106	19	506.7817	Q96Q89|KI20B_HUMAN:MPHOSPH1
QDFVRLTY	15.58	1040.529	8	521.2737	Q9BZ81|MAGB5_HUMAN:MAGEB5
RAESTTPEL	16.51	1002.498	9	502.2558	Q9C093|SPEF2_HUMAN:SPEF2
MLEILFELL	16.93	1119.625	9	374.2171	Q9H568|ACTL8_HUMAN:ACTL8
ALLDKLYAL	26.71	1018.606	9	510.3102	Q9NV31|IMP3_HUMAN:IMP-3
KLKEPLPNGK	28.81	1122.676	10	375.2321	Q9UGL1|KDM5B_HUMAN:JARID1B
RPRHSLNSL	19.62	1078.6	9	360.5404	Q9UGL1|KDM5B_HUMAN:JARID1B
5					
YLQEKLHAL	28.6	1113.618	9	372.2137	O15078|CE290_HUMAN:CEP290

SPNSILKAA	42.25	899.5076	9	450.7628	P33981|TTK_HUMAN:TTK
GVYDGREHTV	46.27	1131.531	10	566.7737	P43358|MAGA4_HUMAN:MAGEA4
KVLEHVVRV	33.37	1077.666	9	360.2289	P43358|MAGA4_HUMAN:MAGEA4
YDGREHTV	31.52	975.441	8	488.7289	P43358|MAGA4_HUMAN:MAGEA4
RPSMVWLSA	24.27	1045.538	9	523.7766	P78395|PRAME_HUMAN:PRAME
VPWMPPGKSSA	45.89	1155.575	11	578.7964	Q5BJF6|ODFP2_HUMAN:ODF2
KILDLETQL	34.33	1071.618	9	536.816	Q5BJF6|ODFP2_HUMAN:ODF2
SPIRSRSPPA	23.88	1066.588	10	356.5366	Q5BJF6|ODFP2_HUMAN:ODF2
LAPAVIHAL	25.23	903.5541	9	452.7853	Q6PL18|ATAD2_HUMAN:ATAD2
GLSNHIAAL	21.88	894.4923	9	448.2533	Q6PL18|ATAD2_HUMAN:ATAD2
RPSAAHIVEA	20.74	1049.562	10	525.7902	Q96KB5|TOPK_HUMAN:PBK
KIIETEDQID	20.15	1202.603	10	602.3059	Q96PC5|MIA2_HUMAN:CTAGE5
ALLRQIKEV	24.69	1068.666	9	535.3396	Q96Q89|KI20B_HUMAN:MPHOSPH1
ALLDKLYAL	36.05	1018.606	9	510.3109	Q9NV31|IMP3_HUMAN:IMP-3
LLDKLYAL	30.06	947.5692	8	474.7919	Q9NV31|IMP3_HUMAN:IMP-3
AQHLQAAVAF	26.33	1054.556	10	528.2867	Q9NV31|IMP3_HUMAN:IMP-3
LDKLYAL	20.53	834.485	7	418.2489	Q9NV31|IMP3_HUMAN:IMP-3
APYSAVEKAMA	50.59	1136.554	11	569.2828	Q9UGL1|KDM5B_HUMAN:JARID1B
APYSAVEKA	42.71	934.476	9	468.2467	Q9UGL1|KDM5B_HUMAN:JARID1B
ALKDSVQRA	34.16	986.5508	9	329.8569	Q9UGL1|KDM5B_HUMAN:JARID1B
SIPVHLNSL	16.83	978.5498	9	490.283	Q9UGL1|KDM5B_HUMAN:JARID1B
RLIDLGVGL	15.7	954.5	9	478.3	Q9UGL1|KDM5B_HUMAN:JARID

	7	862		005	1B
6					
LERSRTVI	17.49	972.5716	8	487.2939	O15078|CE290_HUMAN:CEP290
GVYDGREHTV	32.58	1131.531	10	566.772	P43358|MAGA4_HUMAN:MAGEA4
ALAPESRAL	18.03	926.5185	9	464.2667	Q07617|SPAG1_HUMAN:SPAG1
KLEMAREK	15.34	1003.548	8	502.7829	Q5BJF6|ODFP2_HUMAN:ODF2
DVIRALAKY	32.36	1047.608	9	524.8126	Q5TZF3|ANR45_HUMAN:ANKRD45
SQNAIDHKI	21.32	1024.53	9	513.2732	Q6PL18|ATAD2_HUMAN:ATAD2
SVISKIDLHKY	20.75	1301.734	11	434.9189	Q6PL18|ATAD2_HUMAN:ATAD2
EVKNENAKL	17.29	1043.561	9	522.7874	Q8TC20|CAGE1_HUMAN:CAGE1
KEFFQGCIM	19.18	1101.499	9	551.7579	Q96Q89|KI20B_HUMAN:MPHOSPH1
ICIQNKIPTL	17.64	1141.653	10	571.836	Q9BXT5|TEX15_HUMAN:TEX15
PGSLLPLENP	17.47	1035.56	10	346.1942	Q9BXT5|TEX15_HUMAN:TEX15
EVVEKYEIY	33.69	1170.581	9	586.2991	Q9BZD4|NUF2_HUMAN:CDCA1
7					
LSLNRGDEDEDDHEDHDDSE	17.64	2340.885	20	586.2312	O94818|NOL4_HUMAN:NOL4
NMLEAVHTI	21.7	1026.517	9	514.2656	P33981|TTK_HUMAN:TTK
KAEMLESVIKN	17.55	1260.675	11	631.3434	P43355|MAGA1_HUMAN:MAGEA1
VPKALEPYV	39.4	1014.575	9	508.2951	Q5H9I0|TFDP3_HUMAN:TFDP3
KKPAATTAK	20.84	914.5549	9	458.2834	Q6PL18|ATAD2_HUMAN:ATAD2
KNKGGLTPLLLAVHGQK	16.27	1773.063	17	887.5417	Q6S8J7|POTEA_HUMAN:POTEA
STNGKKIEV	16.66	974.5396	9	488.278	Q7Z7J5|DPPA2_HUMAN:DPPA2
SAAMDEKVI	21.93	962.4742	9	482.2398	Q8NG31|KNL1_HUMAN:CASC5

VPHVSKERI	15.6	1063.614	9	355.5445	Q8NG31|KNL1_HUMAN:CASC5
RLYVGREK	24.23	1019.588	8	340.8701	Q96PC5|MIA2_HUMAN:CTAGE5
KTYTFQGT	17.75	944.4603	8	473.2381	Q96Q89|KI20B_HUMAN:MPHOSPH1
ELPLCFAL	16.43	904.4728	8	453.2452	Q9NXZ2|DDX43_HUMAN:DDX43
KLKEPLPNGK	28.99	1122.676	10	375.2326	Q9UGL1|KDM5B_HUMAN:JARID1B
8					
GLEDAVVEIKNC	21.32	1288.633	12	645.3223	O15078|CE290_HUMAN:CEP290
SILSIVHVK	21.34	994.6175	9	498.3157	O60271|JIP4_HUMAN:SPAG9
SVPLSDALLNK	32.98	1155.65	11	578.8326	P33981|TTK_HUMAN:TTK
GPREALSRL	35.26	997.5668	9	333.5293	P49910|ZN165_HUMAN:ZNF165
RVSSHLIRH	33.58	1103.631	9	368.884	P49910|ZN165_HUMAN:ZNF165
RTLDEAVGVQK	33.12	1214.662	11	608.3394	Q14667|K0100_HUMAN:KIAA0100
RLYGNTPMR	20.45	1106.565	9	554.2911	Q14667|K0100_HUMAN:KIAA0100
LPRVWTDANL	20.69	1183.635	10	592.8259	Q17RY6|LY6K_HUMAN:LY6K
KPSNSKSETTL	25.61	1190.614	11	596.3157	Q53EZ4|CEP55_HUMAN:CEP55
RLMEQQGALLK	32.51	1285.718	11	643.8672	Q5BJF6|ODFP2_HUMAN:ODF2
STDPNIVRK	38.91	1028.561	9	515.2887	Q5H9L4|TAF7L_HUMAN:TAF7L
KPRLAGENL	21.7	996.5716	9	499.2936	Q5T6F0|DCA12_HUMAN:DCAF12
ATFGDGTTIIAK	40.6	1193.629	12	597.8223	Q6Q759|SPG17_HUMAN:SPAG17
GSISTILPEK	29.21	1043.586	10	522.801	Q6Q759|SPG17_HUMAN:SPAG17
RPRPTWEEV	25.75	1168.599	9	585.3071	Q8N9E0|F133A_HUMAN:FAM133A
SVQKPKFPK	34.05	1057.628	9	353.551	Q8NG31|KNL1_HUMAN:CASC5

RPMDKTVVF	32.68	1091.58	9	546.7973	Q8NG31|KNL1_HUMAN:CASC5
SPIEKSGVL	26.25	928.5229	9	465.2689	Q8NG31|KNL1_HUMAN:CASC5
RPMDKTVLF	24.48	1105.595	9	553.8049	Q8NG31|KNL1_HUMAN:CASC5
GLLDNPISEK	16.91	1084.576	10	543.2964	Q8NG31|KNL1_HUMAN:CASC5
TVATFILQK	25.5	1019.602	9	510.8085	Q92600|CNOT9_HUMAN:RQCD1
RLMDEAKILK	32.85	1215.701	10	406.2419	Q96KB5|TOPK_HUMAN:PBK
RLYVGREK	20.95	1019.588	8	340.87	Q96PC5|MIA2_HUMAN:CTAGE5
RPSSKKTYSL	39.83	1165.646	10	389.556	Q96Q89|KI20B_HUMAN:MPHOSPH1
AIFKDLVGK	31.11	989.5909	9	495.8033	Q96Q89|KI20B_HUMAN:MPHOSPH1
VQKDNEIEQLK	20.03	1342.709	11	448.5775	Q96Q89|KI20B_HUMAN:MPHOSPH1
GPLRLSPLLPR	20.04	1217.761	11	406.9276	Q96RT6|CTGE2_HUMAN:CTAGE1__CTAGE-2
ATAQFKINKK	43.92	1147.671	10	574.8441	Q9BZD4|NUF2_HUMAN:CDCA1
ATAQFKINK	39.9	1019.576	9	510.7963	Q9BZD4|NUF2_HUMAN:CDCA1
KIDEKTAELK	33.64	1173.66	10	392.2272	Q9BZD4|NUF2_HUMAN:CDCA1
RVRSQGQQL	21.62	1070.595	9	536.3046	Q9NPE6|SPAG4_HUMAN:SPAG4
RLHGMSSKNIK	34.33	1269.698	11	424.2394	Q9NW75|GPTC2_HUMAN:GPATCH2
RPKGLGLGF	31.52	943.5603	9	472.7879	Q9NW75|GPTC2_HUMAN:GPATCH2
MPVHMVPTEL	35.14	1152.567	10	577.2913	Q9UGL1|KDM5B_HUMAN:JARID1B
FPVRDGKIKL	28.41	1171.708	10	391.5768	Q9UGL1|KDM5B_HUMAN:JARID1B
HPGPRPAL	18.9	843.4715	8	422.7429	Q9UGL1|KDM5B_HUMAN:JARID1B
9					
REFERLIGRY	36.34	1337.72	10	446.914	O60271|JIP4_HUMAN:SPAG9

RTLDEAVGVQKW	44.71	1400.741	12	701.3787	Q14667|K0100_HUMAN:KIAA0100
RLYGNTPMRR	39.48	1262.667	10	421.8963	Q14667|K0100_HUMAN:KIAA0100
RLYGNTPMR	33.96	1106.565	9	554.2902	Q14667|K0100_HUMAN:KIAA0100
DEAVGVQKW	27.17	1030.508	9	516.2609	Q14667|K0100_HUMAN:KIAA0100
KILDLETQL	33.21	1071.618	9	536.8163	Q5BJF6|ODFP2_HUMAN:ODF2
RLMEQQGALLK	29.27	1285.718	11	643.8664	Q5BJF6|ODFP2_HUMAN:ODF2
RTKTELSQL	20.67	1074.603	9	538.3098	Q5BJF6|ODFP2_HUMAN:ODF2
RSMQNYVQF	17.82	1171.544	9	586.7786	Q5BJF6|ODFP2_HUMAN:ODF2
AENTLSKLL	30.52	987.5601	9	494.787	Q5T6F0|DCA12_HUMAN:DCAF12
KSHTVVIGF	31.2	986.5549	9	494.2844	Q8NG31|KNL1_HUMAN:CASC5
RLYVGREK	24.8	1019.588	8	340.8697	Q96PC5|MIA2_HUMAN:CTAGE5
RLYTKMNLK	40.78	1165.664	9	389.5619	Q96Q89|KI20B_HUMAN:MPHOSPH1
RLFPGLAIK	16.21	1013.639	9	507.8259	Q99661|KIF2C_HUMAN:MCAK
KIDEKTAELK	36.85	1173.66	10	392.2267	Q9BZD4|NUF2_HUMAN:CDCA1
KIDEKTAEL	30.93	1045.566	9	523.7904	Q9BZD4|NUF2_HUMAN:CDCA1
ALLDKLYAL	30.2	1018.606	9	510.3104	Q9NV31|IMP3_HUMAN:IMP-3
KLKEPLPNGK	34.55	1122.676	10	375.2318	Q9UGL1|KDM5B_HUMAN:JARID1B
KLKEPLPNGKK	32.3	1250.771	11	417.9307	Q9UGL1|KDM5B_HUMAN:JARID1B
SVAQQLLNGK	31.62	1056.593	10	529.3056	Q9UGL1|KDM5B_HUMAN:JARID1B
RIRLEQARW	30.32	1226.7	9	409.9072	Q9UGL1|KDM5B_HUMAN:JARID1B
REMEALQSL	26.82	1075.533	9	538.7739	Q9UGL1|KDM5B_HUMAN:JARID1B
LHWGEPKTW	18.7	1152.	9	577.2	Q9UGL1|KDM5B_HUMAN:JARID

	6	572		95	1B
10					
ELDETKHQNQL	16.99	1353.653	11	677.8315	A0A0A6YYL3|POTEB_HUMAN:POTEB
MGPNPGMDGDCVYEN	19.99	1597.585	15	799.8005	O00295|TULP2_HUMAN:TULP2
HQHNVSLQL	19.79	1074.557	9	538.2852	O15078|CE290_HUMAN:CEP290
KNRELEEEL	20.64	1158.588	9	580.3016	O60271|JIP4_HUMAN:SPAG9
IRKERPISL	19.78	1110.687	9	371.237	O60271|JIP4_HUMAN:SPAG9
SFYEHIITV	30.4	1107.56	9	554.7859	Q5T6F0|DCA12_HUMAN:DCAF12
VQTSQITKI	24.86	1016.587	9	509.3053	Q5T6F0|DCA12_HUMAN:DCAF12
KYLTVKDYL	40.25	1141.638	9	571.8258	Q6PL18|ATAD2_HUMAN:ATAD2
VYVPHIHVW	37.7	1148.613	9	383.8779	Q6PL18|ATAD2_HUMAN:ATAD2
AYAIIKEEL	31.34	1048.58	9	525.2971	Q6PL18|ATAD2_HUMAN:ATAD2
RYSGVNQSMLF	30.99	1300.623	11	651.3194	Q6PL18|ATAD2_HUMAN:ATAD2
ITVEKALAI	24.41	956.5906	9	479.3028	Q6PL18|ATAD2_HUMAN:ATAD2
TVTPVEVHI	16.6	993.5495	9	497.7825	Q6Q759|SPG17_HUMAN:SPAG17
TSDVLSWNEVE	16.21	1277.578	11	639.7968	Q6Q759|SPG17_HUMAN:SPAG17
SYQKVIELF	30.37	1125.607	9	563.8116	Q96KB5|TOPK_HUMAN:PBK
SRSHSIFTV	25.35	1032.535	9	517.2795	Q96Q89|KI20B_HUMAN:MPHOSPH1
LDLTEARVQV	15.21	1142.63	10	572.3262	Q96QS3|ARX_HUMAN:ARX
TRMSTVSEL	29.89	1022.507	9	512.2617	Q99661|KIF2C_HUMAN:MCAK
MKTIEHMKMI	22.7	1260.639	10	631.3291	Q9BXT5|TEX15_HUMAN:TEX15
ICIQNKIPTL	15.73	1141.653	10	571.8363	Q9BXT5|TEX15_HUMAN:TEX15
RLIDLGVGL	17.3	954.5	9	478.3	Q9UGL1|KDM5B_HUMAN:JARID

	9	862		001	1B

Table 11. List of peptides defined by the RAVEN database across the ten cervical cancer tumours.

Patient/Peptide	-10lgP	Mass	Length	m/z	Raven_Record
1					
TEEIDKDLEEA	15.7	1290.583	11	646.2976	A6H8Y1|BDP1_HUMAN:BDP1
RLIEPIPDSHI	23.83	1288.714	11	645.3652	O15344|TRI18_HUMAN:MID1
RQIIGTKI	17.51	927.5865	8	464.8003	O15344|TRI18_HUMAN:MID1
TDRQLFDTI	15.22	1107.556	9	554.7862	O43497|CAC1G_HUMAN:CACNA1G
KLLEKAFSI	27.06	1047.633	9	524.8247	O75881|CP7B1_HUMAN:CYP7B1
LQNKVVVTI	27.79	1012.628	9	507.3219	O95199|RCBT2_HUMAN:RCBTB2
TKPPYEI	17.56	846.4487	7	424.2333	O95619|YETS4_HUMAN:YEATS4
TLKESGPTL	16.82	944.5178	9	473.2644	P01817|HV205_HUMAN:IGHV2-5
LLPKKTESHHKAKGK	51.12	1701.005	15	568.0098	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
SRAGLQFPVGR	32.25	1186.657	11	396.5596	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
RAGLQFPV	18.23	886.5024	8	444.259	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
SGVSLAALKK	39.51	972.5967	10	487.3057	P16403|H12_HUMAN:HIST1H1C
ASGPPVSELITK	38.27	1197.66	12	599.8376	P16403|H12_HUMAN:HIST1H1C
GTLVQTKGTGASGSFKLNK	59.65	1893.032	19	632.019	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNK	57.29	2407.344	24	602.8435	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKG	53.2	1316.766	13	659.3912	P16403|H12_HUMAN:HIST1H1C

SLVSKGTLVQTK	50.81	1259.745	12	630.8798	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASG	46.89	1689.926	18	564.3157	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	46.07	1165.609	12	583.8118	P16403|H12_HUMAN:HIST1H1C
GASGSFKLNK	39.33	1007.54	10	504.7772	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTG	38.48	1474.835	15	492.6191	P16403|H12_HUMAN:HIST1H1C
GTLVQTKGTGASGSFK	38.02	1537.81	16	513.6103	P16403|H12_HUMAN:HIST1H1C
ASGSFKLNK	33.22	950.5185	9	476.2669	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQT	27.29	1131.65	11	566.8323	P16403|H12_HUMAN:HIST1H1C
SRIKLGLKSLVSK	47.2	1427.919	13	357.9875	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	26.02	1036.676	10	346.5661	P16403|H12_HUMAN:HIST1H1C
KAVKPKAAKPK	25.64	1164.771	11	389.2646	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGG	72.41	2174.337	23	725.7856	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAG	70.39	2117.316	22	706.7792	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTK	66.58	2478.537	24	620.6428	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEAKPKV	65.92	1946.095	19	487.5312	P16403|H12_HUMAN:HIST1H1C
GATPKKSAKKTPKKAKKPAAATVTK	65.29	2535.559	25	634.8958	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTKK	59.87	2081.305	20	521.3331	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	58.06	2017.289	20	673.4373	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAA	54.92	2060.294	21	687.7733	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTKK	54.73	1923.236	18	481.8167	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	54.1	1918.22	19	640.414	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	53.8	1579.	15	790.5	P16403|H12_HUMAN:HIST1H1

	3	03		223	C
KKAASGEAKPKVKKAGGTKPK	53.79	2108.279	21	528.0779	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	53.51	1662.03	17	555.0176	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	53.36	1790.125	18	597.7159	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTK	51.89	1953.21	19	489.3105	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	51.75	1495.956	15	499.6591	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	50.92	1795.141	17	449.7925	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	50.89	1424.919	14	713.4672	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	50.7	2145.384	21	537.3535	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKA	50.04	1806.193	17	603.0717	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	49.56	1478.966	14	370.7484	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAA	49.5	1450.935	14	363.7412	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	49.09	1745.177	16	582.7322	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	48.65	1607.061	15	536.6939	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKA	48.56	1678.098	16	560.3735	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	48.31	1667.046	16	417.769	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	48.08	1749.135	17	584.0529	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	48	1296.824	13	433.2816	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	47.93	1877.23	18	626.7507	P16403|H12_HUMAN:HIST1H1C
TKPKKPVGAAKKPK	47.35	1476.95	14	370.2451	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAG	47.27	1989.221	21	498.313	P16403|H12_HUMAN:HIST1H1C
KPKAAKPKVVKPKKAAPKKK	45.94	2169.456	20	543.3719	P16403|H12_HUMAN:HIST1H1C

KKPAAATVTKKVAKSPK	45.65	1752.099	17	439.032	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	44.78	1521.972	15	381.4998	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	44.64	1816.214	17	455.0598	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	43.3	1624.051	16	407.0199	P16403|H12_HUMAN:HIST1H1C
SAAKAVKPKAAKPKVVKPKKAAPKK	43.26	2568.668	25	1285.3452	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAKA	43.21	1695.088	17	424.7787	P16403|H12_HUMAN:HIST1H1C
TKPKKPVGAAKKPKK	42.76	1605.045	15	402.2687	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	42.51	1617.082	15	405.2784	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPK	41.77	1308.861	12	437.2935	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTKK	40.69	1340.85	13	447.9575	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPKA	40.58	1367.861	14	456.9617	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKP	40.13	1662.03	17	416.5154	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	39.99	1436.883	15	360.2276	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	39.94	1166.714	12	584.3645	P16403|H12_HUMAN:HIST1H1C
AKSPKKAKVAKPK	39.79	1379.898	13	460.9733	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKKK	38.41	1674.14	15	419.5423	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	37.94	1707.125	16	427.7887	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGA	37.07	1237.751	13	413.5909	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAK	35.68	1877.23	18	470.3149	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	35.38	1588.055	15	398.0214	P16403|H12_HUMAN:HIST1H1C
AKKPAAATVTKK	33.97	1212.755	12	405.26	P16403|H12_HUMAN:HIST1H1C
AKKPAAATVTKKVAKS	33.9	2249.	22	563.3	P16403|H12_HUMAN:HIST1H1

PKKAKV	7	431		643	C
KSPKKAKVAKPKKA	33.96	1507.993	14	503.6716	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	33.16	1064.707	10	355.9093	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKK	32.83	1688.119	16	423.0365	P16403|H12_HUMAN:HIST1H1C
KPKVVKPK	32.58	922.6327	8	462.3242	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKK	32.49	1417.95	13	473.6569	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	31.71	1320.897	12	331.2311	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKKK	31.31	1546.045	14	516.3562	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	30.83	1192.802	11	398.6081	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKKAAPK	30.38	1688.119	16	423.0361	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPK	29.66	993.6698	9	332.2306	P16403|H12_HUMAN:HIST1H1C
KAAGGATPKKSAKKTPKKA	29.53	1867.137	19	467.7913	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKK	29.49	1546.045	14	387.5182	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAA	29.43	1308.788	14	437.2701	P16403|H12_HUMAN:HIST1H1C
PKKAAKSAAKAVKPKAAKP	27.93	1889.194	19	473.3061	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	27.18	1735.156	16	434.7969	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKK	25.93	1790.125	18	448.5381	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKP	24.72	1180.766	11	394.5957	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKA	24.64	1989.257	20	498.322	P16403|H12_HUMAN:HIST1H1C
AKVAKPKKAAKSAAK	24.39	1495.956	15	374.996	P16403|H12_HUMAN:HIST1H1C
SFKLNKKAASGEAKPKVK	21.11	1930.136	18	483.5399	P16403|H12_HUMAN:HIST1H1C
KKPAAATVTK	21.05	1013.623	10	507.8185	P16403|H12_HUMAN:HIST1H1C

KKTPKKAKKPAAATVTK	19.16	1795.141	17	449.7913	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTK	16.27	1212.755	12	607.3855	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVK	15.57	1071.676	11	536.8441	P16403|H12_HUMAN:HIST1H1C
AKPKKAAKSAAKAVKPKAAK	15.22	1991.273	20	664.7623	P16403|H12_HUMAN:HIST1H1C
GQFDHPNIIRL	34.04	1308.694	11	437.2382	P29322|EPHA8_HUMAN:P29320|EPHA3_HUMAN:
GQFDHPNII	24.23	1039.509	9	520.762	P29322|EPHA8_HUMAN:P29320|EPHA3_HUMAN:
GVYDGREHTV	38.42	1131.531	10	566.7735	P43358|MAGA4_HUMAN:MAGEA4
KVLEHVVRV	36.35	1077.666	9	539.8402	P43358|MAGA4_HUMAN:MAGEA4
KGLAAALLV	15.5	854.5589	9	428.2868	P55774|CCL18_HUMAN:CCL18
AQKRVLLQA	30.59	1025.635	9	513.8253	P57740|NU107_HUMAN:NUP107
IFLSETWSL	15.02	1094.565	9	548.2921	P59045|NAL11_HUMAN:NLRP11
TIMPKDIQLAR	39.9	1284.722	11	429.248	P68431|H31_HUMAN:HIST1H3A
RKSTGGKAPRKQLATK	38.77	1726.033	16	432.5165	P68431|H31_HUMAN:HIST1H3A
FQRLVREI	31.93	1059.619	8	354.213	P68431|H31_HUMAN:HIST1H3A
KAPRKQLATK	24.37	1139.714	10	380.9119	P68431|H31_HUMAN:HIST1H3A
LRFQSSAVM	22.8	1037.533	9	519.7741	P68431|H31_HUMAN:HIST1H3A
YQKSTELLI	18.25	1093.602	9	547.8091	P68431|H31_HUMAN:HIST1H3A
HRYRPGTVAL	15.51	1168.647	10	390.5561	P68431|H31_HUMAN:HIST1H3A
KRNYIRLSF	31.22	1195.683	9	399.5681	Q07864|DPOE1_HUMAN:POLE
WLLQKNPQL	30.1	1138.65	9	570.3323	Q07864|DPOE1_HUMAN:POLE
ALYSAEHGL	29.23	959.4712	9	480.7433	Q07864|DPOE1_HUMAN:POLE
REYDVPYHI	28.5	1190.	9	596.2	Q07864|DPOE1_HUMAN:POL

	9	572		937	E
KVFMRALF	16.1	1010.574	8	506.2959	Q13368|MPP3_HUMAN:MPP3
GRHIQQYHM	46.01	1168.556	9	390.5261	Q13438|OS9_HUMAN:OS9
ARFNDLRFVGR	47.61	1349.732	11	338.4399	Q13950|RUNX2_HUMAN:RUNX2
ALNEDIVRV	26.09	1027.566	9	514.7908	Q2M1P5|KIF7_HUMAN:KIF7
TLVDQRFLI	34.47	1103.634	9	552.8254	Q5HYA8|MKS3_HUMAN:TMEM67
ITIAYSDIDI	20.66	1122.581	10	562.2968	Q5HYA8|MKS3_HUMAN:TMEM67
GLLKKINSV	27.86	970.6175	9	486.3167	Q5XXA6|ANO1_HUMAN:ANO1
KLHSVTTKL	29.69	1025.623	9	342.8821	Q68BL8|OLM2B_HUMAN:OLFML2B
QLLEEPKL	15.95	968.5542	8	485.2834	Q6UXZ4|UNC5D_HUMAN:UNC5D
HRNPLIRKY	41.57	1195.694	9	399.5706	Q6ZUX3|TGRM2_HUMAN:FAM179A
FIQRAAGQSL	17.23	1089.593	10	545.8039	Q6ZUX3|TGRM2_HUMAN:FAM179A
ATVITTEKR	28.49	1017.582	9	509.7975	Q7Z5N4|SDK1_HUMAN:SDK1
HLEGNRLVL	15.56	1049.598	9	525.8066	Q7Z5N4|SDK1_HUMAN:SDK1
TVSSYNLSQL	15.83	1110.556	10	556.2872	Q8IVU1|IGDC3_HUMAN:IGDCC3
PVMKKLVP	22.36	910.5674	8	456.2897	Q8N6L0|KASH5_HUMAN:CCDC155
ALQPLLHTV	24.63	990.5862	9	496.3004	Q8NFM7|I17RD_HUMAN:IL17RD
GQVVHLTLV	23.21	964.5706	9	483.2921	Q8NI35|INADL_HUMAN:PATJ
RLFDDEASV	22.68	1050.498	9	526.2561	Q8NI35|INADL_HUMAN:PATJ
KQSPPIVKI	32.05	1008.633	9	505.3248	Q96GZ6|S41A3_HUMAN:SLC41A3
GQSALPQRL	23.08	968.5403	9	485.2771	Q96L91|EP400_HUMAN:EP400
ESVVQCLHI	18.38	1026.517	9	514.2633	Q96N23|CFA54_HUMAN:CFAP54

IKCIVVLQ	15.99	914.5623	8	458.2872	Q96N23|CFA54_HUMAN:CFAP54
TRLRLEEA	19.54	986.5508	8	494.2849	Q96PX9|PKH4B_HUMAN:PLEKHG4B
RQGFKLAYQ	16.85	1109.598	9	370.8748	Q96PZ7|CSMD1_HUMAN:CSMD1
VKPSHVILL	16.94	1004.638	9	503.327	Q99665|I12R2_HUMAN:IL12RB2
IRANATNKL	29.77	999.5825	9	334.2009	Q9BV19|CA050_HUMAN:C1orf50
STKGNLKVHF	17.71	1129.624	10	565.8185	Q9BXA9|SALL3_HUMAN:SALL3
GRLIGQHV	17.25	878.5086	8	440.2622	Q9HCF6|TRPM3_HUMAN:TRPM3
SGATIRIFK	30	991.5814	9	496.7977	Q9NUV7|SPTC3_HUMAN:SPTLC3
SLDSRVREV	17.07	1059.567	9	530.7918	Q9UGC6|RGS17_HUMAN:RGS17
SLRDLLAEV	23.61	1014.571	9	508.2936	Q9ULI4|KI26A_HUMAN:KIF26A
IQGGILNLNK	17.27	1068.629	10	535.3221	Q9UNK4|PA2GD_HUMAN:PLA2G2D
RQADLKNWSL	30.87	1229.652	10	615.8338	Q9UPW0|FOXJ3_HUMAN:FOXJ3
2					
EVSSHFSHF	29.22	1075.472	9	359.4983	A6H8Y1|BDP1_HUMAN:BDP1
DETLVHSSF	34.05	1033.472	9	517.7423	O14595|CTDS2_HUMAN:CTDSP2
DVVPKTAENFRQF	45.86	1549.789	13	775.9036	O43447|PPIH_HUMAN:PPIH
EVFTWGHNAY	39.44	1222.541	10	612.2788	O95199|RCBT2_HUMAN:RCBTB2
YTDVPFHNQY	47.95	1282.562	10	642.2891	O95897|NOE2_HUMAN:OLFM2
HRLLRKGNYAERV	19.08	1610.912	13	403.7349	P0C0S8|H2A1_HUMAN:HIST1H2AL
RNDEELNKLLGKVTI	43.74	1740.973	15	581.3327	P0C0S8|H2A1_HUMAN:HIST1H2AL
RNDEELNKLLGKVTIA	42.35	1812.011	16	605.0099	P0C0S8|H2A1_HUMAN:HIST1H2AL
KVTIAQGGVLPNIQAV	37.5	1606.941	16	804.4779	P0C0S8|H2A1_HUMAN:HIST1H2AL

IPRHLQLAI	22.76	1059.655	9	530.8337	P0C0S8|H2A1_HUMAN:HIST1H2AL
LLPKKTESHHKAKGK	64.51	1701.005	15	851.5096	P0C0S8|H2A1_HUMAN:HIST1H2AL
LPKKTESHHKAKGK	58.45	1587.921	14	397.9879	P0C0S8|H2A1_HUMAN:HIST1H2AL
PKKTESHHKAKGK	56.75	1474.837	13	492.6194	P0C0S8|H2A1_HUMAN:HIST1H2AL
LPKKTESHHKA	46.01	1274.71	11	425.9098	P0C0S8|H2A1_HUMAN:HIST1H2AL
LLPKKTESHHKAK	40.62	1515.889	13	506.3032	P0C0S8|H2A1_HUMAN:HIST1H2AL
KAKTRSSRAGLQFPV	22.21	1644.942	15	549.3211	P0C0S8|H2A1_HUMAN:HIST1H2AL
FEHPNVVRL	42.72	1109.598	9	370.8735	P11802|CDK4_HUMAN:CDK4
GAAKKPKKAAGGATPKKSAKKTPK	84.44	2348.438	24	588.1162	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKKAASGEAKPKVK	73.64	2360.317	24	787.781	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEAKPKVK	69.06	2074.19	20	519.5538	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAAK	66.19	1936.231	19	485.0646	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	65.17	1918.22	19	640.4142	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTK	64.31	1953.21	19	652.0777	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	62.55	1790.125	18	597.7166	P16403|H12_HUMAN:HIST1H1C
TKKVAKSPKKAKVAKPKKAAK	59.27	2234.468	21	559.6239	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTKK	59.18	2081.305	20	521.3336	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	58.74	1795.141	17	449.7926	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	57.92	2017.289	20	673.437	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAKSAAK	57.14	2234.431	22	559.6155	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAKAVKPK	55.18	2147.399	21	537.8574	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGG	53.6	1804.	19	602.3	P16403|H12_HUMAN:HIST1H1

ATPK	7	105		765	C
VAKSPKKAKVAKPKKAAK	53.25	1877.23	18	470.315	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	52.71	1424.919	14	357.2367	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	52.25	1478.966	14	493.9961	P16403|H12_HUMAN:HIST1H1C
TPKKAKKPAAATVTKK	51.58	1667.046	16	417.7684	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	51.55	1495.956	15	374.9965	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	50.83	1662.03	17	416.5149	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAK	50.55	2005.325	19	669.4488	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	50.48	1607.061	15	536.6944	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	50.35	1667.046	16	417.7684	P16403|H12_HUMAN:HIST1H1C
KAAGGATPKKSAKKTPKKAKKPA	50.14	2291.417	23	573.8606	P16403|H12_HUMAN:HIST1H1C
TKKVAKSPKKAKVAKPKK	49.48	1964.299	18	492.0815	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	48.95	1745.177	16	437.3012	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	48.36	1624.051	16	407.0199	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	48.07	2145.384	21	537.3529	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	47.78	1707.125	16	427.7885	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	46.63	1436.883	15	360.2283	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	46.48	1296.824	13	433.2821	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	46.44	1816.214	17	909.1138	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKA	45.65	1989.257	20	498.322	P16403|H12_HUMAN:HIST1H1C
TKKVAKSPKKAKVAKPKKAA	44.96	2106.373	20	527.6009	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTKK	44.83	1340.85	13	447.9575	P16403|H12_HUMAN:HIST1H1C

AKSPKKAKVAKPKKAAK	44.33	1778.162	17	445.5476	P16403|H12_HUMAN:HIST1H1C
GAAKKPKKAAGGATPK	43.91	1479.888	16	494.3037	P16403|H12_HUMAN:HIST1H1C
SAAKAVKPKAAKPK	43.45	1393.877	14	465.6328	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPKKAAPKKK	41.66	2339.562	22	585.8978	P16403|H12_HUMAN:HIST1H1C
KAVKPKAAKPKVVKPKKAAPKKK	41.03	2467.657	23	617.9224	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKKK	41.01	1674.14	15	419.5422	P16403|H12_HUMAN:HIST1H1C
KPKAAKPKVVKPKKAAPKKK	40.18	2169.456	20	724.1595	P16403|H12_HUMAN:HIST1H1C
SAAKAVKPKAAKPKVVKPK	39.58	1945.256	19	487.3218	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	39.51	1320.897	12	331.2309	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPKKAAPK	39.4	2083.372	20	521.8495	P16403|H12_HUMAN:HIST1H1C
SFKLNKKAASGEAKPKVK	39.31	1930.136	18	483.5398	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	37.89	1579.03	15	395.7645	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	36.78	1735.156	16	434.7957	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	35.95	1588.055	15	398.0205	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPV	35.32	1109.692	11	555.8535	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	31.64	1064.707	10	355.9096	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	30.6	1166.714	12	389.9116	P16403|H12_HUMAN:HIST1H1C
AKSPKKAKVAKPK	30.43	1379.898	13	460.9733	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKKAAPK	29.91	1688.119	16	423.0371	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKKK	29.73	1546.045	14	516.3554	P16403|H12_HUMAN:HIST1H1C
KAVKPKAAKPKVVKPKKAAPK	28.42	2211.467	21	553.8735	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKK	27.6	1749.	17	438.2	P16403|H12_HUMAN:HIST1H1

AA		135		914	C
KPKVVKPK	26.87	922.6327	8	462.3239	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKKAAPKK	26.38	1816.214	17	455.0594	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGA	25.58	1237.751	13	413.5905	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKK	25.51	1688.119	16	423.0365	P16403|H12_HUMAN:HIST1H1C
KAGGTPRKASGPPVSELI	25.17	1763.989	18	589.0018	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	24.02	1192.802	11	398.6082	P16403|H12_HUMAN:HIST1H1C
KSAKKTPKKAKKPAAATV	20.44	1852.162	18	618.3944	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	20.17	1617.082	15	405.2782	P16403|H12_HUMAN:HIST1H1C
PAAATVTKKVAKSPKKAKVAK	17.93	2121.336	21	531.3392	P16403|H12_HUMAN:HIST1H1C
PKKAAKSAAK	17.08	998.6236	10	500.3188	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPV	16.68	1336.855	13	335.2212	P16403|H12_HUMAN:HIST1H1C
KALAAAGYDVEKNNSRIKLGLK	63.69	2358.338	22	590.593	P16403|H12_HUMAN:HIST1H1C
ALAAAGYDVEKNNSRIKLGLK	46.41	2230.243	21	558.5693	P16403|H12_HUMAN:HIST1H1C
AGYDVEKNNSRIKLGLK	38.19	1904.048	17	477.0202	P16403|H12_HUMAN:HIST1H1C
TKAVAASKERSGV	33.26	1302.726	13	435.2498	P16403|H12_HUMAN:HIST1H1C
RKASGPPVSELI	19.74	1252.714	12	627.3622	P16403|H12_HUMAN:HIST1H1C
KASGPPVSELI	19.21	1096.613	11	549.3143	P16403|H12_HUMAN:HIST1H1C
TKAVAASKERSGVSLAALK	16.46	1886.095	19	472.5312	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNKK	80.9	2535.439	25	634.8668	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNK	73.41	2407.344	24	602.8442	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTK	52.56	1259.745	12	630.8812	P16403|H12_HUMAN:HIST1H1C

GTGASGSFKLNKK	48.8	1293.704	13	647.8597	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	45.56	1165.609	12	583.8125	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFK	42.36	2052.122	21	685.0504	P16403|H12_HUMAN:HIST1H1C
TGASGSFKLNK	41.72	1108.588	11	555.3011	P16403|H12_HUMAN:HIST1H1C
SLAALKKALAA	40.05	1055.67	11	528.8406	P16403|H12_HUMAN:HIST1H1C
GTLVQTKGTGASGSFKLNKK	39.53	2021.127	20	506.2906	P16403|H12_HUMAN:HIST1H1C
VSKGTLVQTK	36.1	1059.629	10	530.8231	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLN	35.21	2279.249	23	570.8197	P16403|H12_HUMAN:HIST1H1C
SKGTLVQTKGTGASGSFKLNKK	33.74	2236.254	22	560.0707	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQT	25.71	1131.65	11	566.8311	P16403|H12_HUMAN:HIST1H1C
SLAALKK	24.64	729.4749	7	365.7448	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLV	20.1	902.5436	9	452.2793	P16403|H12_HUMAN:HIST1H1C
TGASGSFKLNKK	16.07	1236.683	12	413.2355	P16403|H12_HUMAN:HIST1H1C
KGTGASGSFKLNKKAA	44.36	1563.873	16	391.9748	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAA	38.68	1149.651	11	384.2224	P16403|H12_HUMAN:HIST1H1C
NNSRIKLGLK	36.35	1141.693	10	381.5724	P16403|H12_HUMAN:HIST1H1C
SKGTLVQTKGTGASGSFKLNKKAA	35.87	2378.328	24	595.5906	P16403|H12_HUMAN:HIST1H1C
EKNNSRIKLGLK	31.94	1398.831	12	467.2844	P16403|H12_HUMAN:HIST1H1C
NSRIKLGLK	30.77	1027.65	9	343.5576	P16403|H12_HUMAN:HIST1H1C
GLKSLVSKGTLV	28.08	1200.744	12	601.3784	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKKAA	24.69	1435.778	15	479.6003	P16403|H12_HUMAN:HIST1H1C
KLGLKSLVSKGTLV	24.5	1441.	14	481.6	P16403|H12_HUMAN:HIST1H1

	3	923		477	C
KPKAAKPK	20.81	866.5701	8	434.2921	P16403|H12_HUMAN:HIST1H1C
KLGLKSLV	20.65	856.5746	8	429.2926	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKKAASGEAKPK	44.03	2133.154	22	534.2954	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	35.23	1036.676	10	346.5659	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEAKPK	21.87	1847.026	18	462.7638	P16403|H12_HUMAN:HIST1H1C
AKAVKPKAAKPK	16.21	1235.808	12	412.9433	P16403|H12_HUMAN:HIST1H1C
DWIRSTLRR	15.2	1201.668	9	401.5623	P24158|PRTN3_HUMAN:PRTN3
DVFGGASKW	17.75	965.4606	9	483.7373	P47972|NPTX2_HUMAN:NPTX2
NEFPEPIKL	36.11	1085.576	9	543.7953	P51149|RAB7A_HUMAN:RAB7A
KEVMVDDRL	30.7	1103.565	9	552.7892	P51149|RAB7A_HUMAN:RAB7A
DEVRLLKY	29.09	1034.576	8	518.2965	P57740|NU107_HUMAN:NUP107
AEDDNAIREHL	24.31	1281.595	11	428.2056	P57740|NU107_HUMAN:NUP107
KEEVSIEVL	17.01	1044.57	9	523.2914	P57740|NU107_HUMAN:NUP107
DVTDAVHEW	32.45	1070.467	9	536.242	P61812|TGFB2_HUMAN:TGFB2
DEKQQHIVY	41.12	1158.567	9	580.2916	Q00987|MDM2_HUMAN:MDM2
DEIILPRGL	29.71	1024.592	9	513.3028	Q02223|TNR17_HUMAN:TNFRSF17
DVKELIKKW	30.87	1157.681	9	386.8995	Q07325|CXCL9_HUMAN:CXCL9
DLKIHVAHW	37.96	1117.603	9	373.5421	Q07864|DPOE1_HUMAN:POLE
YADNQVMHFY	23.41	1286.539	10	644.2776	Q07864|DPOE1_HUMAN:POLE
KKHKKKRVVPKKPPPSPQPT	39.19	2304.427	20	577.1136	Q13438|OS9_HUMAN:OS9
IEGLVKELL	28.16	1012.617	9	507.3153	Q13438|OS9_HUMAN:OS9

SEIKGEVL	26.54	873.4807	8	437.7477	Q13438|OS9_HUMAN:OS9
KELLEREGL	19.78	1085.608	9	543.8099	Q13438|OS9_HUMAN:OS9
DESVWRPY	15.83	1050.477	8	526.2466	Q13950|RUNX2_HUMAN:RUNX2
YTDPEVFKY	35.64	1160.539	9	581.2779	Q16647|PTGIS_HUMAN:PTGIS
IEWEKPEHY	27.61	1229.572	9	410.8645	Q17RS7|GEN_HUMAN:GEN1
FMEKETPLRY	38.41	1312.649	10	438.5572	Q2PZI1|D19L1_HUMAN:DPY19L1P1
YTDENQHQY	39.34	1196.473	9	599.245	Q5HYA8|MKS3_HUMAN:TMEM67
DDFHIYVY	24.49	1070.471	8	536.2432	Q5JUX0|SPIN3_HUMAN:SPIN3
ICSFASTDSI	18.47	1042.464	10	522.2372	Q5M8T2|S35D3_HUMAN:SLC35D3
RSEEGRLYY	28.88	1171.562	9	586.7875	Q5PSV4|BRM1L_HUMAN:BRMS1L
QVQGTSDLSKL	16.29	1174.619	11	588.3196	Q5TAH2|SL9C2_HUMAN:SLC9C2
EVIGLLGGRY	21.13	1075.603	10	538.8097	Q5VVJ2|MYSM1_HUMAN:MYSM1
VEFRNLENF	34.88	1166.572	9	584.2919	Q68BL8|OLM2B_HUMAN:OLFML2B
DERIYVTNY	34.02	1171.551	9	586.7839	Q68BL8|OLM2B_HUMAN:OLFML2B
HASVGPTL	26.05	780.413	8	391.2128	Q68BL8|OLM2B_HUMAN:OLFML2B
YGNTLVEFR	23.68	1097.551	9	549.7814	Q68BL8|OLM2B_HUMAN:OLFML2B
TIDEVKGTIY	22.39	1137.592	10	569.8044	Q6V0I7|FAT4_HUMAN:FAT4
RELEDYIDNL	18.68	1278.609	10	640.3123	Q7L804|RFIP2_HUMAN:RAB11FIP2
VYHLTWLRN	15.62	1200.64	9	401.2204	Q7Z2Y5|NRK_HUMAN:NRK
VSDRHVYKY	23.53	1165.588	9	389.5367	Q86UP3|ZFHX4_HUMAN:ZFHX4
EVVVPMPSW	16.15	1042.516	9	522.2674	Q86VD1|MORC1_HUMAN:MORC1
EEVLLTFK	15.1	977.5	8	489.7	Q8IZP9|AGRG2_HUMAN:ADG

		433		785	RG2
HEFLTPRL	19.76	1011.55	8	506.7812	Q8NI35|INADL_HUMAN:PATJ
NTDFGDSGRLYY	40.13	1406.61	12	704.3141	Q96JM2|ZN462_HUMAN:ZNF462
CSHADGCHPT	15.88	1026.365	10	514.1909	Q96KG7|MEG10_HUMAN:MEGF10
KLMEKWLG	16.7	1003.552	8	502.784	Q9C093|SPEF2_HUMAN:SPEF2
LEQKVVELY	23.16	1119.618	9	560.8151	Q9HAI6|CX021_HUMAN:CXorf21
SEDPVVGHL	25.36	951.4661	9	476.7411	Q9UBP6|TRMB_HUMAN:METTL1
YEIDDVERL	19.29	1150.551	9	576.2831	Q9ULI4|KI26A_HUMAN:KIF26A
FAEDGARTV	22.51	964.4614	9	483.2388	Q9UP38|FZD1_HUMAN:FZD1
NTQKPTGIADVY	36.78	1305.657	12	653.8367	Q9Y6H5|SNCAP_HUMAN:SNCAIP
IESVIEKF	30.3	963.5276	8	482.7718	Q9Y6Q5|AP1M2_HUMAN:AP1M2
SEIVGTIKL	23.97	958.5699	9	480.2917	Q9Y6Q5|AP1M2_HUMAN:AP1M2
DVAMSKIEHF	19.42	1175.565	10	588.7905	Q9Y6Q5|AP1M2_HUMAN:AP1M2
3					
SESHPLSTI	17.4	969.4767	9	485.7454	A6H8Y1|BDP1_HUMAN:BDP1
KEAEFLQKL	31.81	1104.618	9	553.3165	O14986|PI51B_HUMAN:PIP5K1B
RDHQVAAL	31.34	908.4828	8	455.2492	O15344|TRI18_HUMAN:MID1
VEERAQAII	18.1	1027.566	9	514.7908	O60732|MAGC1_HUMAN:MAGEC1
AQAIIDT	15.03	730.3861	7	366.2014	O60732|MAGC1_HUMAN:MAGEC1
MDHSISPF	26.8	932.4062	8	467.2103	O75897|ST1C4_HUMAN:SULT1C4
YTDVPFHNQY	41.55	1282.562	10	642.2878	O95897|NOE2_HUMAN:OLFM2
LLPKKTESHHKAKGK	33.01	1701.005	15	426.2581	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
APRESPLPY	24.0	1028.	9	515.2	P09172|DOPO_HUMAN:DBH

	2	529		732	
KKNTESVK	17.61	932.5291	8	467.2715	P11055|MYH3_HUMAN:MYH3
CPTLAGILDI	15.79	1014.542	10	508.2769	P15248|IL9_HUMAN:IL9
KAGGTKPKKPVGAAKKPKKAAG	71.62	2117.316	22	706.7784	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGG	70.34	2174.337	23	725.7859	P16403|H12_HUMAN:HIST1H1C
KSAKKTPKKAKKPAAATVTKK	62.01	2209.4	21	553.3574	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTKK	61.63	2081.305	20	521.3319	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAKSAAK	60.5	2234.431	22	745.8156	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNK	59.93	2407.344	24	602.8437	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAKSAAK	58.99	2362.526	23	788.516	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAG	57.1	1989.221	21	498.3127	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAAK	55.91	1936.231	19	485.0649	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	55.71	1790.125	18	597.7161	P16403|H12_HUMAN:HIST1H1C
GATPKKSAKKTPKKAKKPAAATVTK	53.25	2535.559	25	634.8972	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	51.96	1918.22	19	480.5621	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAK	51.61	1877.23	18	626.7505	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTK	50.85	1259.745	12	630.8801	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAK	49.45	2005.325	19	669.4485	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	48.99	2017.289	20	505.3297	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	48.85	1667.046	16	417.769	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKG	47.87	1316.766	13	659.3907	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	47.29	1795.141	17	449.7923	P16403|H12_HUMAN:HIST1H1C

KVAKSPKKAKVAKPK	47.04	1607.061	15	402.7719	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTKK	46.91	1923.236	18	481.8165	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	46.91	1877.23	18	626.7502	P16403|H12_HUMAN:HIST1H1C
TKKVAKSPKKAKVAKPKKAAK	46.26	2234.468	21	559.6239	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	46.24	1296.824	13	433.2823	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	46.2	1436.883	15	360.2277	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTK	45.7	1795.141	17	599.3853	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	45.38	1816.214	17	606.4118	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASG	45.31	1689.926	18	564.3165	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	44.88	1745.177	16	437.3008	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	44.3	1165.609	12	583.8109	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPKKAAPKKK	44.13	2339.562	22	585.8995	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	43.99	1478.966	14	493.9962	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAGG	43.66	2046.242	22	512.5678	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	43.41	1749.135	17	584.0521	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	43.14	1424.919	14	357.2373	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGGA	42.84	2245.375	24	562.3524	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTK	41.46	2478.537	24	620.6409	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	40.71	1166.714	12	584.3636	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	40.35	1624.051	16	407.0202	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAKSAA	39.95	2234.431	22	559.6143	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKKK	39.6	1674.	15	419.5	P16403|H12_HUMAN:HIST1H1

	7	14		423	C
KKVAKSPKKAKVAKP K	37.2 7	1735. 156	16	434.7 961	P16403|H12_HUMAN:HIST1H1 C
KVAKSPKKAKVAKPK K	36.7 9	1735. 156	16	434.7 953	P16403|H12_HUMAN:HIST1H1 C
VKKAGGTKPKKPVGA AKKPKK	35.5	2145. 384	21	537.3 527	P16403|H12_HUMAN:HIST1H1 C
KAGGTKPKKPVGA	35.4 9	1237. 751	13	413.5 907	P16403|H12_HUMAN:HIST1H1 C
KSPKKAKVAKPKKAA KSAAK	35.3 8	2064. 326	20	517.0 894	P16403|H12_HUMAN:HIST1H1 C
SLVSKGTLVQTKGTG	33.5 8	1474. 835	15	492.6 175	P16403|H12_HUMAN:HIST1H1 C
AVKPKAAKPK	32.4 5	1036. 676	10	346.5 649	P16403|H12_HUMAN:HIST1H1 C
AKSPKKAKVAKPKKA AK	31.9 6	1778. 162	17	445.5 479	P16403|H12_HUMAN:HIST1H1 C
SELITKAV	31.2	859.5 015	8	430.7 581	P16403|H12_HUMAN:HIST1H1 C
ATPKKSAKKTPKKAK KPAAATVTKK	29.2 9	2606. 632	25	652.6 658	P16403|H12_HUMAN:HIST1H1 C
SFKLNKKAASGEAKP KVK	28.0 1	1930. 136	18	483.5 417	P16403|H12_HUMAN:HIST1H1 C
SLVSKGTLVQTKGTG A	27.6	1545. 873	16	516.2 98	P16403|H12_HUMAN:HIST1H1 C
KAAKPKVVKPKKAAP KK	26.9 6	1816. 214	17	455.0 609	P16403|H12_HUMAN:HIST1H1 C
KVVKPKKAAPKK	25.2 1	1320. 897	12	331.2 306	P16403|H12_HUMAN:HIST1H1 C
KKPAAATVTKK	24.1 3	1141. 718	11	381.5 797	P16403|H12_HUMAN:HIST1H1 C
PKVVKPKKAAPKKK	23.3 3	1546. 045	14	387.5 182	P16403|H12_HUMAN:HIST1H1 C
AKPKKAAKSAAKAVK PKAAK	21.3	1991. 273	20	498.8 253	P16403|H12_HUMAN:HIST1H1 C
KERSGVSL	20.5 3	874.4 872	8	438.2 504	P16403|H12_HUMAN:HIST1H1 C
AKPKVVKPKKAAPKK	19.7 3	1617. 082	15	405.2 771	P16403|H12_HUMAN:HIST1H1 C
AASKERSGVSL	18.3 4	1103. 594	11	552.8 036	P16403|H12_HUMAN:HIST1H1 C
SAKKTPKKAKKPAAA TVTK	17.3 9	1953. 21	19	489.3 11	P16403|H12_HUMAN:HIST1H1 C

KPVGAAKKPKKAAG	17.38	1349.851	14	338.4702	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKA	17.38	1989.257	20	498.3215	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	17.2	1707.125	16	427.7888	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAK	16.69	901.5709	9	451.7917	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKK	16.47	1293.704	13	432.2417	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	15.95	1495.956	15	374.9955	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPKAAKPK	15.51	1792.141	18	449.0436	P16403|H12_HUMAN:HIST1H1C
LTTFRMTG	15.08	925.4691	8	463.7438	P20849|CO9A1_HUMAN:COL9A1
DRNNRPKFE	15.95	1174.584	9	392.5344	P29320|EPHA3_HUMAN:EPHA3
RESVPARTV	22.41	1013.562	9	507.79	P31644|GBRA5_HUMAN:GABRA5
YFVIQTYL	18.38	1045.549	8	523.7808	P31644|GBRA5_HUMAN:GABRA5
YEKGIMQKV	25.52	1094.579	9	548.2965	P43268|ETV4_HUMAN:P41161|ETV5_HUMAN:
AYLEAHETF	34.52	1079.492	9	540.7538	P57740|NU107_HUMAN:NUP107
FEILTPKI	25.8	959.5692	8	480.7922	Q02108|GCYA1_HUMAN:GUCY1A3
KETFHSLTI	22.29	1074.571	9	538.2938	Q07864|DPOE1_HUMAN:POLE
YEFKGLHKV	16.36	1119.608	9	374.2094	Q07864|DPOE1_HUMAN:POLE
NKQLVRHL	15.44	1006.604	8	336.5409	Q07864|DPOE1_HUMAN:POLE
GEIAPRTVV	17.79	940.5342	9	471.2741	Q12904|AIMP1_HUMAN:AIMP1
RENPQLKQI	16.88	1124.63	9	563.3223	Q13438|OS9_HUMAN:OS9
TESLRALK	18.37	916.5342	8	459.2741	Q13515|BFSP2_HUMAN:BFSP2
KEFEKASKI	18.35	1078.602	9	540.3094	Q15554|TERF2_HUMAN:TERF2
YTDPEVFKY	41.3	1160.	9	581.2	Q16647|PTGIS_HUMAN:PTGIS

	1	539		772	
TLTGFSLVSL	16.69	1036.58	10	519.2962	Q5FYB1|ARSI_HUMAN:ARSI
YTDENQHQY	40.97	1196.473	9	599.2438	Q5HYA8|MKS3_HUMAN:TMEM67
KWDELAV	16.09	859.444	7	430.7274	Q5JU67|CF157_HUMAN:CFAP157
ICSFASTDSI	26.61	1042.464	10	522.2394	Q5M8T2|S35D3_HUMAN:SLC35D3
RSEEGRLYY	30.16	1171.562	9	586.7893	Q5PSV4|BRM1L_HUMAN:BRMS1L
NKQDAVKEL	22.28	1043.561	9	522.7877	Q5TAH2|SL9C2_HUMAN:SLC9C2
VYGKTSHL	18.5	903.4814	8	452.7469	Q6BEB4|SP5_HUMAN:SP5
AEQEASRVI	24.54	1001.514	9	501.7645	Q6Q759|SPG17_HUMAN:SPAG17
WQDILQQI	18.8	1042.545	8	522.2797	Q6Q759|SPG17_HUMAN:SPAG17
IEAERAAHL	19.26	1008.535	9	505.2755	Q6TFL3|CC171_HUMAN:CCDC171
LPHGEGSL	17.2	808.4079	8	405.2113	Q6ZUX3|TGRM2_HUMAN:FAM179A
IYVPSSYHL	15.23	1077.549	9	539.7822	Q8IX21|SLF2_HUMAN:SLF2
VCSSWAGSTQL	15.79	1137.513	11	569.7651	Q8IY50|S35F3_HUMAN:SLC35F3
REFDDPGQQV	26.03	1189.536	10	595.7759	Q8NF99|ZN397_HUMAN:ZNF397
HEFLTPRL	24.42	1011.55	8	506.7831	Q8NI35|INADL_HUMAN:PATJ
GEILLLPEL	15.94	995.5903	9	498.8016	Q8TC59|PIWL2_HUMAN:PIWIL2
GEQGYIRLI	24.01	1047.571	9	524.793	Q96FA3|PELI1_HUMAN:PELI1
REHAAPYF	24.75	989.4719	8	495.7433	Q96L91|EP400_HUMAN:EP400
KDLIARAF	20.04	932.5443	8	467.2814	Q96MC2|DRC1_HUMAN:DRC1
ETMMKLVL	18.52	963.5133	8	482.7656	Q96PX9|PKH4B_HUMAN:PLEKHG4B
QLEEGYDFL	17.56	1112.503	9	557.2576	Q96PZ7|CSMD1_HUMAN:CSMD1

IRANATNKL	26.86	999.5825	9	334.2023	Q9BV19|CA050_HUMAN:C1orf50
GEHLVRTI	20.58	923.5189	8	462.7666	Q9H6P5|TASP1_HUMAN:TASP1
SDSQVRAF	36.24	908.4352	8	455.2244	Q9UJ98|STAG3_HUMAN:STAG3
FRGVFVHRY	28.87	1179.63	9	394.217	Q9UJ98|STAG3_HUMAN:STAG3
RDLARRFAL	15.95	1116.652	9	373.2242	Q9UJ98|STAG3_HUMAN:STAG3
RDSLHTYKL	27.88	1131.604	9	378.2087	Q9Y2H5|PKHA6_HUMAN:PLEKHA6
NTQKPTGIADVY	29.74	1305.657	12	653.8362	Q9Y6H5|SNCAP_HUMAN:SNCAIP
IESVIEKF	26.38	963.5276	8	482.7711	Q9Y6Q5|AP1M2_HUMAN:AP1M2
4					
RVFSNKLLEK	31.89	1232.724	10	411.9161	O75881|CP7B1_HUMAN:CYP7B1
IASLSVKEP	19.24	942.5386	9	472.2764	O94779|CNTN5_HUMAN:CNTN5
RVKGVTIVK	23.75	998.66	9	333.8936	O95619|YETS4_HUMAN:YEATS4
KLFQSDTNAMLGK	15.4	1451.744	13	484.9218	O95619|YETS4_HUMAN:YEATS4
LLPKKTESHHKAKGK	47.08	1701.005	15	568.0093	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
GIAAPFLRK	25.55	971.5916	9	486.8032	P11055|MYH3_HUMAN:MYH3
KAGGTKPKKPVGAAKKPKKAAGG	70.11	2174.337	23	725.7865	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAG	64.7	2117.316	22	706.7798	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKG	57.72	1316.766	13	659.3903	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	57.29	2017.289	20	505.3294	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	56.46	1918.22	19	640.4135	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	56.44	1662.03	17	555.0176	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAAK	53.78	1936.231	19	485.0643	P16403|H12_HUMAN:HIST1H1C

GTLVQTKGTGASGSFK	53.67	1537.81	16	769.9125	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	53.48	1579.03	15	790.5214	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTKK	53.07	1923.236	18	481.816	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKA	52.49	1806.193	17	603.0723	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	52.38	1424.919	14	357.2367	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	51.51	1795.141	17	449.7922	P16403|H12_HUMAN:HIST1H1C
KALAAAGYDVEKNNSRIKLGLK	51.13	2358.338	22	590.5908	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTK	50.97	1953.21	19	489.3096	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	50.52	1749.135	17	584.0521	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	50.32	1790.125	18	448.5375	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAA	49.77	1450.935	14	363.7404	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	49.3	1478.966	14	493.996	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	49.13	1667.046	16	556.6899	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTK	48.9	1259.745	12	630.8801	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	48.86	1877.23	18	626.7507	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAA	47.07	2060.294	21	516.0808	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	46.39	1607.061	15	402.7726	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	45.66	2145.384	21	537.3529	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	45.44	1745.177	16	582.7324	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	44.54	1165.609	12	583.8124	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGGATPK	44.28	1804.105	19	452.0327	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTG	43.5	2407.	24	602.8	P16403|H12_HUMAN:HIST1H1

ASGSFKLNK	2	344		425	C
SLVSKGTLVQTKGTG	43.24	1474.835	15	738.4252	P16403|H12_HUMAN:HIST1H1C
AASKERSGVSL	43.18	1103.594	11	552.8049	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	42.94	1495.956	15	374.9965	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNKK	42.8	2535.439	25	634.8676	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	42.6	1296.824	13	433.2817	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKKK	42.6	1674.14	15	419.5408	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPK	42.58	1308.861	12	437.2944	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	42.43	1816.214	17	455.0605	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	41.58	1707.125	16	427.7882	P16403|H12_HUMAN:HIST1H1C
ALAAAGYDVEKNNSRIKLGLK	41.26	2230.243	21	558.5682	P16403|H12_HUMAN:HIST1H1C
TPKKAKKPAAATVTKK	40.97	1667.046	16	417.7688	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	40.71	1036.676	10	519.3452	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	40.5	1588.055	15	398.0211	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAA	40.48	1808.136	18	453.0412	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFK	40.22	2052.122	21	685.0458	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPKA	39.58	1367.861	14	456.9608	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASG	39.36	1689.926	18	564.316	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPK	38.56	1180.766	11	394.5961	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEAKPKV	38.21	1946.095	19	487.5317	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	38.07	1064.707	10	533.3606	P16403|H12_HUMAN:HIST1H1C
VSKGTLVQTK	38.04	1059.629	10	530.8215	P16403|H12_HUMAN:HIST1H1C

KAGGTKPKKPVGAAK	37.82	1436.883	15	360.2276	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	37.8	1735.156	16	434.7962	P16403|H12_HUMAN:HIST1H1C
GASGSFKLNK	37.45	1007.54	10	504.7772	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	37.24	1617.082	15	405.2773	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	36.69	1624.051	16	407.0185	P16403|H12_HUMAN:HIST1H1C
KPKVVKPK	36.55	922.6327	8	462.3229	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTKK	35.27	2606.632	25	652.6652	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	34.47	1166.714	12	389.9121	P16403|H12_HUMAN:HIST1H1C
ASGSFKLNK	33.62	950.5185	9	476.2675	P16403|H12_HUMAN:HIST1H1C
KLGLKSLVSK	33.6	1071.702	10	358.2413	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	33.32	1192.802	11	398.6075	P16403|H12_HUMAN:HIST1H1C
AKAVKPKAAKPK	33.26	1235.808	12	412.9436	P16403|H12_HUMAN:HIST1H1C
AKSPKKAKVAKPK	33.05	1379.898	13	460.9728	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAKA	32.57	1695.088	17	424.7793	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGA	32.04	1545.873	16	516.2985	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAA	31.82	1138.755	11	570.3848	P16403|H12_HUMAN:HIST1H1C
KPKAAKPK	31.17	866.5701	8	434.292	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTK	31.05	2478.537	24	620.6403	P16403|H12_HUMAN:HIST1H1C
VKPKAAKPK	30.23	965.6385	9	483.8265	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	30.09	1521.972	15	381.5	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAKSAAK	29.89	2064.326	20	517.0888	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTG	29.7	2279.	23	570.8	P16403|H12_HUMAN:HIST1H1

ASGSFKLN	3	249		199	C
KKTPKKAKKPAAATVTK	29.04	1795.141	17	599.3878	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKK	28.92	1417.95	13	473.6565	P16403|H12_HUMAN:HIST1H1C
KERSGVSL	28.73	874.4872	8	438.2508	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAGG	27.79	2046.242	22	512.5685	P16403|H12_HUMAN:HIST1H1C
GVSLAALKK	26.55	885.5647	9	443.7894	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAKSAA	26.18	1936.231	19	485.0652	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKKK	25.8	1546.045	14	387.5177	P16403|H12_HUMAN:HIST1H1C
SAAKAVKPKAAKPKVVKPKKAAPKK	25.4	2568.668	25	643.1747	P16403|H12_HUMAN:HIST1H1C
AKSAAKAVKPKAAKPK	25.1	1593.009	16	399.2592	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKP	25.06	1662.03	17	416.515	P16403|H12_HUMAN:HIST1H1C
SELITKAV	25.02	859.5015	8	430.7584	P16403|H12_HUMAN:HIST1H1C
KKPAAATVTKK	24.83	1141.718	11	571.8673	P16403|H12_HUMAN:HIST1H1C
KERSGVSLA	24.28	945.5243	9	473.7698	P16403|H12_HUMAN:HIST1H1C
TKPKKPVGAAKKPKK	24.04	1605.045	15	402.2683	P16403|H12_HUMAN:HIST1H1C
PKKAAKSAAKAVKPKAAKPK	23.29	2017.289	20	505.3315	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAK	22.79	1877.23	18	470.313	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPKAAKPK	21.98	1920.236	19	481.0657	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAG	21.89	1989.221	21	498.3126	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTK	21.4	1212.755	12	405.2581	P16403|H12_HUMAN:HIST1H1C
KERSGVSLAA	20.62	1016.561	10	509.2899	P16403|H12_HUMAN:HIST1H1C
PKAAKPK	18.98	738.4752	7	370.2457	P16403|H12_HUMAN:HIST1H1C

SLVSKGTLVQT	18.32	1131.65	11	566.8342	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	17.8	1320.897	12	331.2316	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGAS	17.5	1632.905	17	545.3086	P16403|H12_HUMAN:HIST1H1C
KKPAAATVTKKVAK	17.37	1439.919	14	480.9799	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPKAAKPK	16.81	1792.141	18	449.0431	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKK	16.06	1320.897	12	441.3062	P16403|H12_HUMAN:HIST1H1C
KPKAAKPKVVKPKKAAPK	15.93	1913.267	18	479.3221	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKA	15.53	1989.257	20	498.3195	P16403|H12_HUMAN:HIST1H1C
GPAPAPHSL	20.59	845.4395	9	423.727	P41161|ETV5_HUMAN:ETV5
LRCVWVPK	15.25	999.5688	8	500.7937	P43146|DCC_HUMAN:DCC
EGLGVDRRI	15.55	1013.562	9	507.7874	P48664|EAA4_HUMAN:SLC1A6
REIAQDFKTDL	45.25	1334.683	11	668.349	P68431|H31_HUMAN:HIST1H3A
RKSTGGKAPRKQLATK	44.51	1726.033	16	432.5159	P68431|H31_HUMAN:HIST1H3A
KSTELLIRK	33.09	1086.676	9	363.2325	P68431|H31_HUMAN:HIST1H3A
KSTGGKAPRKQLATK	32.39	1569.931	15	393.4896	P68431|H31_HUMAN:HIST1H3A
TELLIRKL	31.64	984.6331	8	493.3241	P68431|H31_HUMAN:HIST1H3A
KAPRKQLATKAARK	29.62	1565.984	14	392.5035	P68431|H31_HUMAN:HIST1H3A
REIAQDFKT	24.3	1106.572	9	554.2938	P68431|H31_HUMAN:HIST1H3A
ELLIRKL	23.16	883.5854	7	442.8003	P68431|H31_HUMAN:HIST1H3A
KAPRKQLATK	16.17	1139.714	10	380.9117	P68431|H31_HUMAN:HIST1H3A
RPREAEVRF	40.08	1158.626	9	387.2155	Q13438|OS9_HUMAN:OS9
RPREAEVRFL	36.2	1271.	10	424.9	Q13438|OS9_HUMAN:OS9

	6	71		106	
SEIKGEVL	23.35	873.4807	8	437.7475	Q13438|OS9_HUMAN:OS9
TLNKLIKR	16.59	984.6443	8	329.2217	Q13438|OS9_HUMAN:OS9
LAQLTFNQSIL	16.94	1246.692	11	624.3516	Q4G0P3|HYDIN_HUMAN:HYDIN
ELLGCLAAEL	16.25	1030.537	10	516.2775	Q56UN5|M3K19_HUMAN:MAP3K19
ICSFASTDSI	17.59	1042.464	10	522.2396	Q5M8T2|S35D3_HUMAN:SLC35D3
TLEISNIQI	16.5	1029.571	9	515.7923	Q6UXK5|LRRN1_HUMAN:LRRN1
KNAASNILI	18.19	942.5498	9	472.2822	Q6ZMW3|EMAL6_HUMAN:EML6
KENDCWCH	15.42	1033.375	8	517.6942	Q76B58|BRNP3_HUMAN:BRINP3
EAVTVVTIAP	16.44	998.5648	10	500.2888	Q7Z5N4|SDK1_HUMAN:SDK1
LLRYYPPGI	18.69	1090.617	9	546.316	Q8N136|DAW1_HUMAN:DAW1
VLGPTPVQK	25.8	937.5596	9	469.7873	Q8WU20|FRS2_HUMAN:FRS2
MMKKHRSMVK	18.3	1274.677	10	638.3451	Q96JM2|ZN462_HUMAN:ZNF462
NCHNGAFCSAYDGE	18.63	1486.524	14	744.2685	Q96KG7|MEG10_HUMAN:MEGF10
PAARPAAAKLKASRLF	15.41	1666.999	16	834.5051	Q96L91|EP400_HUMAN:EP400
TVLVPTRHP	18.6	1018.592	9	510.304	Q96M63|CC114_HUMAN:CCDC114
DISLPSIFNL	16.56	1117.602	10	559.8077	Q96N23|CFA54_HUMAN:CFAP54
LYQYFVSGI	16.49	1088.554	9	545.2863	Q96N23|CFA54_HUMAN:CFAP54
KETLRDFIF	15.45	1167.629	9	584.8214	Q96N23|CFA54_HUMAN:CFAP54
GSLLGAAPGL	18.78	854.4861	10	428.2503	Q9BXA9|SALL3_HUMAN:SALL3
RAESTTPEL	16.51	1002.498	9	502.2558	Q9C093|SPEF2_HUMAN:SPEF2
ESQRLRTL	15.03	1001.562	8	501.7892	Q9UJ90|KCNE5_HUMAN:KCNE5

GECAPGKPTLVCSS	16.85	1347.616	14	674.8128	Q9Y5H5|PCDA9_HUMAN:PCDHA9
SEIVGTIKL	19.29	958.5699	9	480.292	Q9Y6Q5|AP1M2_HUMAN:AP1M2
5					
RPIDTVISQQA	34.26	1226.662	11	614.3396	A2RUB1|MEIOC_HUMAN:MEIOC
SPREKTPEV	22.34	1041.545	9	348.1895	A6H8Y1|BDP1_HUMAN:BDP1
KPDDYLYSI	18.78	1112.539	9	557.2786	O14986|PI51B_HUMAN:PIP5K1B
KLLEKAFSI	32.01	1047.633	9	524.826	O75881|CP7B1_HUMAN:CYP7B1
FILDPFQYQ	19.89	1169.576	9	585.7954	O75881|CP7B1_HUMAN:CYP7B1
KPYRNEDMSA	28.25	1209.545	10	605.7799	O95619|YETS4_HUMAN:YEATS4
LPKKTESHHKA	58.39	1274.71	11	638.3611	P0C0S8|H2A1_HUMAN:HIST1H2AL
LPKKTESHHKAKGK	46.9	1587.921	14	397.987	P0C0S8|H2A1_HUMAN:HIST1H2AL
LLPKKTESHHKAKGK	46.21	1701.005	15	426.2584	P0C0S8|H2A1_HUMAN:HIST1H2AL
PKKTESHHKA	40.23	1161.625	10	388.2157	P0C0S8|H2A1_HUMAN:HIST1H2AL
LQFPVGRVH	34.52	1051.593	9	351.5386	P0C0S8|H2A1_HUMAN:HIST1H2AL
KLLGKVTIA	30.82	941.6273	9	471.8204	P0C0S8|H2A1_HUMAN:HIST1H2AL
RIIPRHLQL	23.24	1144.719	9	382.5811	P0C0S8|H2A1_HUMAN:HIST1H2AL
TLWYRAPEV	30.85	1133.587	9	567.801	P11802|CDK4_HUMAN:CDK4
KAGGTKPKKPVGAAKKPKKAAG	70.47	2117.316	22	706.7797	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGGATPKKSA	65.85	2090.269	22	697.7641	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAA	65.23	1808.136	18	603.7191	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	62.78	1662.03	17	555.0175	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	59.89	2017.289	20	673.4377	P16403|H12_HUMAN:HIST1H1C

SPKKAKVAKPKKAA	59.84	1450.935	14	363.7399	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	59.31	2145.384	21	716.1353	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTKK	58.94	1923.236	18	481.816	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGG	58.57	2174.337	23	544.592	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	56.93	1795.141	17	449.792	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	56.56	1790.125	18	896.0699	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	55.38	1918.22	19	480.562	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	54.79	1579.03	15	790.5216	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	54.78	1877.23	18	626.7512	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAAK	53.31	1936.231	19	485.0652	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKA	53.14	1806.193	17	603.0717	P16403|H12_HUMAN:HIST1H1C
KPAAATVTKKVA	52.17	1183.729	12	395.5833	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	51.83	1424.919	14	357.2369	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	50.15	1478.966	14	493.9965	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGGATPKK	50.01	1932.2	20	484.0574	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	50.01	1607.061	15	536.694	P16403|H12_HUMAN:HIST1H1C
TPKKAKKPAAATVTKK	49.94	1667.046	16	417.769	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGGATPK	49.16	1804.105	19	452.0333	P16403|H12_HUMAN:HIST1H1C
KKAGGTKPKKPVGAAKKPK	48.69	1918.22	19	480.562	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	48.4	1667.046	16	417.769	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	47.71	1745.177	16	582.7326	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	47.5	2362.	23	591.6	P16403|H12_HUMAN:HIST1H1

KAAKSAAK	5	526		386	C
KSPKKAKVAKPKKAAK	47.5	1707.125	16	427.7886	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPK	47.05	1180.766	11	394.5954	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPK	46.74	1308.861	12	655.4379	P16403|H12_HUMAN:HIST1H1C
TKPKKPVGAAKKPK	46.54	1476.95	14	370.2452	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	45.89	1617.082	15	809.5492	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	45.07	1436.883	15	360.2273	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	45	1495.956	15	374.9961	P16403|H12_HUMAN:HIST1H1C
VAKPKKAAKSAA	44.72	1168.729	12	390.5831	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	44.34	1588.055	15	398.0212	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	44.28	1816.214	17	455.0608	P16403|H12_HUMAN:HIST1H1C
KKAASGEAKPKVKKAGGTKPK	43.69	2108.279	21	528.0765	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTK	42.87	2478.537	24	620.642	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKK	42.43	1790.125	18	448.5377	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	41.03	1521.972	15	508.3311	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAA	40.51	1138.755	11	380.5921	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	39.76	1296.824	13	433.2819	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAK	39.21	1664.046	17	417.0184	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKA	36.8	1507.993	14	378.0053	P16403|H12_HUMAN:HIST1H1C
AKAVKPKAAKPK	36.72	1235.808	12	412.9439	P16403|H12_HUMAN:HIST1H1C
AALKKALAA	35.61	855.5541	9	428.785	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQT	35.38	1131.65	11	566.831	P16403|H12_HUMAN:HIST1H1C

ALKKALAAA	34.78	855.5541	9	428.7844	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	34.7	1064.707	10	355.9091	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	34.45	1735.156	16	434.7956	P16403|H12_HUMAN:HIST1H1C
KKPAAATVTKKVAKSPK	34.33	1752.099	17	439.0302	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	33.61	1166.714	12	389.9108	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	32.95	1036.676	10	346.5649	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	32.51	1320.897	12	661.4554	P16403|H12_HUMAN:HIST1H1C
KPKVVKPK	32.44	922.6327	8	462.3241	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKP	32.2	1180.766	11	394.5959	P16403|H12_HUMAN:HIST1H1C
KPKKPVGAA	31.98	894.5651	9	448.2902	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	31.34	1624.051	16	407.0199	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKK	31.18	1417.95	13	473.6574	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	31.12	1192.802	11	398.6073	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGA	30.85	1237.751	13	413.5905	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTK	30.66	1953.21	19	489.3104	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAK	29.99	1579.03	15	395.7647	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKK	28.71	1607.061	15	402.7726	P16403|H12_HUMAN:HIST1H1C
GATPKKSAKKTPKKAKKPAAATVTK	27.76	2535.559	25	634.8973	P16403|H12_HUMAN:HIST1H1C
LITKAVAA	27.17	785.501	8	393.7596	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	27.09	1749.135	17	438.2915	P16403|H12_HUMAN:HIST1H1C
VKPKKAAPKKK	25.96	1221.829	11	408.2838	P16403|H12_HUMAN:HIST1H1C
KVVKPKKAAPKK	25.4	1320.	12	331.2	P16403|H12_HUMAN:HIST1H1

	6	897		311	C
SGSFKLNKKAASGEAKPK	24.87	1847.026	18	462.7638	P16403|H12_HUMAN:HIST1H1C
PKVVKPK	24.48	794.5378	7	398.2763	P16403|H12_HUMAN:HIST1H1C
KAVKPKAAKPKVVKP	24.13	1588.055	15	398.0207	P16403|H12_HUMAN:HIST1H1C
GASGSFKL	23.46	765.4021	8	383.708	P16403|H12_HUMAN:HIST1H1C
SAAKAVKPKAAKPKVVKPK	22.37	1945.256	19	487.3214	P16403|H12_HUMAN:HIST1H1C
KAVKPKAAKPK	21.62	1164.771	11	389.2641	P16403|H12_HUMAN:HIST1H1C
KKPAAATVTKKVAK	20.89	1439.919	14	720.9649	P16403|H12_HUMAN:HIST1H1C
KVAKPKKAAKSAA	20.72	1296.824	13	433.2817	P16403|H12_HUMAN:HIST1H1C
KKVAKSPKKAKVAKPK	20.43	1735.156	16	434.796	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLN	19.77	2279.249	23	570.8193	P16403|H12_HUMAN:HIST1H1C
GPPVSELITKAVA	18.73	1280.734	13	641.3743	P16403|H12_HUMAN:HIST1H1C
KPKKAAKSAAKAVKPKA	17.88	1721.104	17	431.2825	P16403|H12_HUMAN:HIST1H1C
TGASGSFKLNKKAASGEAKPKVK	17.03	2303.296	23	576.8311	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLV	15.26	902.5436	9	452.2785	P16403|H12_HUMAN:HIST1H1C
KQRRIKLGF	21.18	1144.719	9	382.5813	P20265|PO3F2_HUMAN:POU3F2
GVYDGREHTV	46.27	1131.531	10	566.7737	P43358|MAGA4_HUMAN:MAGEA4
KVLEHVVRV	33.37	1077.666	9	360.2289	P43358|MAGA4_HUMAN:MAGEA4
YDGREHTV	31.52	975.441	8	488.7289	P43358|MAGA4_HUMAN:MAGEA4
SPQRVLVQA	18.78	996.5716	9	499.2933	P48378|RFX2_HUMAN:RFX2
SLIEKLWQT	26.67	1116.618	9	559.3162	P48380|RFX3_HUMAN:RFX3
RPIPSALTQA	24.83	1052.598	10	527.305	P48380|RFX3_HUMAN:RFX3

ALSAVLALPA	19.55	924.5644	10	463.2902	P49683|PRLHR_HUMAN:PRLHR
SLLYHLTAV	28.65	1015.57	9	508.7917	P55899|FCGRN_HUMAN:FCGRT
KQRQLTSY	23.33	1022.551	8	512.2828	P57740|NU107_HUMAN:NUP107
APRKQLATKAA	45.28	1153.693	11	385.5717	P68431|H31_HUMAN:HIST1H3A
RKSTGGKAPRKQLATK	43	1726.033	16	432.5156	P68431|H31_HUMAN:HIST1H3A
KSTGGKAPRKQLATK	27.35	1569.931	15	393.4898	P68431|H31_HUMAN:HIST1H3A
APRKQLATKA	20.91	1082.656	10	361.8919	P68431|H31_HUMAN:HIST1H3A
RPSMVWLSA	24.27	1045.538	9	523.7766	P78395|PRAME_HUMAN:PRAME
SPTRPVSREHE	17.38	1293.643	11	432.2233	Q01101|INSM1_HUMAN:INSM1
SLAELKGFEV	36.03	1091.586	10	546.8002	Q07864|DPOE1_HUMAN:POLE
ALYSAEHGL	26.87	959.4712	9	480.7432	Q07864|DPOE1_HUMAN:POLE
WLLQKNPQL	24.09	1138.65	9	570.3325	Q07864|DPOE1_HUMAN:POLE
HPLLRPPPSA	42.51	1083.619	10	542.8187	Q13438|OS9_HUMAN:OS9
RQKELESNY	38.4	1165.573	9	583.793	Q13438|OS9_HUMAN:OS9
VPKKPPPSP	35.64	945.5647	9	473.7895	Q13438|OS9_HUMAN:OS9
REREEETPAY	34.18	1278.584	10	640.3004	Q13438|OS9_HUMAN:OS9
LLRPPPSA	34.07	849.5072	8	425.7595	Q13438|OS9_HUMAN:OS9
RPWAEGTEEGA	33.05	1201.536	11	601.7781	Q13438|OS9_HUMAN:OS9
GLAHSAPLPA	21.26	932.5079	10	467.2635	Q13516|OLIG2_HUMAN:OLIG2
LPVAFKVVA	32.97	942.5902	9	472.3015	Q13950|RUNX2_HUMAN:RUNX2
AAGEAEAMCAYL	20.7	1198.5	12	600.2561	Q17RS7|GEN_HUMAN:GEN1
FVMEGEPPKL	16.0	1145.	10	573.7	Q17RS7|GEN_HUMAN:GEN1

	2	579		988	
ALNEDIVRV	29.18	1027.566	9	514.7899	Q2M1P5|KIF7_HUMAN:KIF7
APSRLPRPA	25.13	963.5613	9	322.1943	Q2M1P5|KIF7_HUMAN:KIF7
EKVVQHALDK	20.59	1165.646	10	389.5563	Q2M3G0|ABCB5_HUMAN:ABCB5
SPASAFRVA	28.89	904.4766	9	453.2456	Q5U5Z8|CBPC2_HUMAN:AGBL2
GLLKKINSV	26.12	970.6175	9	486.3151	Q5XXA6|ANO1_HUMAN:ANO1
KLMDKISLV	23.08	1045.621	9	523.8187	Q6TFL3|CC171_HUMAN:CCDC171
ALEDRVWEL	17.26	1129.577	9	565.7957	Q6ZP82|CC141_HUMAN:CCDC141
RPRVTYQV	17.48	1017.572	8	509.792	Q6ZTQ4|CDHR3_HUMAN:CDHR3
AGPGARTPL	28.09	838.4661	9	420.2418	Q7Z5N4|SDK1_HUMAN:SDK1
APVFTQRP	26.48	914.4974	8	458.2551	Q7Z5N4|SDK1_HUMAN:SDK1
NLIGKIENV	21.08	998.576	9	500.2965	Q86X45|TILB_HUMAN:LRRC6
SPAHFVNAA	30.73	912.4453	9	457.2309	Q8N635|MEIOB_HUMAN:MEIOB
MPRETVIFA	18.43	1062.553	9	532.2852	Q8N635|MEIOB_HUMAN:MEIOB
ALQPLLHTV	24.34	990.5862	9	496.3011	Q8NFM7|I17RD_HUMAN:IL17RD
RLMDNLPQL	17.88	1098.586	9	550.3	Q8NFM7|I17RD_HUMAN:IL17RD
YPLSHIQEA	35.39	1056.524	9	529.2688	Q8NI35|INADL_HUMAN:PATJ
GLLPGDRLVSV	24.85	1124.655	11	563.3338	Q8NI35|INADL_HUMAN:PATJ
RLIQGDQILSV	23.61	1240.714	11	621.3641	Q8NI35|INADL_HUMAN:PATJ
VPVVLVGNKA	37.18	994.6175	10	498.3148	Q8TAI7|REBL1_HUMAN:RHEBL1
KSPEGTQAL	26.14	929.4818	9	465.7472	Q8WX77|IBPL1_HUMAN:IGFBPL1
GVKLPKPDYND	17.38	1244.64	11	623.3302	Q8WXX0|DYH7_HUMAN:DNAH7

QMEKQLSEF	24.08	1138.533	9	570.2737	Q8WZ74|CTTB2_HUMAN:CTTNBP2
KLTDIQIEL	27.65	1071.618	9	536.8173	Q96AA8|JKIP2_HUMAN:JAKMIP2
LLAVRHSVF	18.51	1040.613	9	521.3141	Q96DT5|DYH11_HUMAN:DNAH11
RPFGHLEEV	32.47	1082.551	9	542.2842	Q96JM2|ZN462_HUMAN:ZNF462
KPLPPIQVA	36.45	961.596	9	481.8053	Q96L91|EP400_HUMAN:EP400
ALAKPPVVSV	28.88	979.6066	10	490.8111	Q96L91|EP400_HUMAN:EP400
SLQDVIDRV	27.31	1043.561	9	522.7871	Q96L91|EP400_HUMAN:EP400
KPPVVSVPAA	26.12	963.5753	10	482.795	Q96L91|EP400_HUMAN:EP400
YLLRSLDYV	19.6	1140.618	9	571.315	Q9NS86|LANC2_HUMAN:LANCL2
RPFVIVTA	28.52	901.5385	8	451.7778	Q9NYQ6|CELR1_HUMAN:CELSR1
KLAENEITL	23.78	1029.571	9	515.793	Q9UGC7|RF1ML_HUMAN:MTRF1L
RPVKEPRGLSA	42.19	1208.699	11	403.9072	Q9UHG2|PCSK1_HUMAN:PCSK1
RPVDVLVPSVSL	24.63	1279.75	12	640.8821	Q9UI42|CBPA4_HUMAN:CPA4
RPVDVLVPSV	22.69	1079.634	10	540.8246	Q9UI42|CBPA4_HUMAN:CPA4
KLMTSLVKV	32.88	1017.626	9	509.8206	Q9UJ98|STAG3_HUMAN:STAG3
VPHQVILPAL	17.9	1085.66	10	543.8366	Q9UJ98|STAG3_HUMAN:STAG3
VAINTWGVSL	17.33	1058.576	10	530.2974	Q9UKG4|S13A4_HUMAN:SLC13A4
SLRDLLAEV	25.47	1014.571	9	508.292	Q9ULI4|KI26A_HUMAN:KIF26A
APLLLGVRA	23.08	908.5807	9	455.2984	Q9UP38|FZD1_HUMAN:FZD1
GLGRVWAV	20.33	856.4919	8	429.2531	Q9Y289|SC5A6_HUMAN:SLC5A6
YLVTKVVAV	32.82	990.6113	9	496.3138	Q9Y5E3|PCDB6_HUMAN:Q9Y5E7
ATATLHVL	25.1	824.4	8	413.2	Q9Y5E3|PCDB6_HUMAN:Q9Y

	3	756		442	5E7
6					
VPIGYKGSTFHRVIK	49.58	1700.973	15	426.2501	O43447|PPIH_HUMAN:PPIH
DVVPKTAENFRQF	46.24	1549.789	13	775.9031	O43447|PPIH_HUMAN:PPIH
VPKTAENFRQF	41.5	1335.694	11	668.8549	O43447|PPIH_HUMAN:PPIH
PKTAENFRQF	38.1	1236.625	10	619.3215	O43447|PPIH_HUMAN:PPIH
DFVNGDG	20.23	722.2871	7	362.1504	O43447|PPIH_HUMAN:PPIH
SIYRGPFADE	15.63	1153.54	10	577.7749	O43447|PPIH_HUMAN:PPIH
EVSAATGRF	28.87	936.4664	9	469.2405	O75881|CP7B1_HUMAN:CYP7B1
KNGGVVLFPPGGGC	17.04	1300.66	14	651.3373	O94933|SLIK3_HUMAN:SLITRK3
YTDVPFHNQY	44.51	1282.562	10	642.2899	O95897|NOE2_HUMAN:OLFM2
IAKDFVVKV	34.11	1017.622	9	340.2133	P12109|CO6A1_HUMAN:COL6A1
KAGGTKPKKPVGAAKKPKKAAGG	67.62	2174.337	23	725.7865	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	63.85	1918.22	19	640.4141	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAG	61.3	2117.316	22	706.7795	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNK	59.3	2407.344	24	602.8428	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAKSAAK	55.79	2362.526	23	591.6395	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	54.92	1662.03	17	555.0171	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	54.56	2017.289	20	505.3293	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAAK	53.71	1936.231	19	485.0649	P16403|H12_HUMAN:HIST1H1C
TPKKAKKPAAATVTKK	53.2	1667.046	16	417.7683	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAA	52.6	1450.935	14	363.7397	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	51.8	1790.	18	896.0	P16403|H12_HUMAN:HIST1H1

KPK	1	125		705	C
KVAKSPKKAKVAKPKKAA	51.69	1877.23	18	626.7496	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	51.59	1579.03	15	790.5227	P16403|H12_HUMAN:HIST1H1C
GATPKKSAKKTPKKAKKPAAATVTK	51.38	2535.559	25	634.8973	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAA	51.26	2060.294	21	516.0807	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTKK	50.41	1923.236	18	481.817	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	50.22	1424.919	14	357.2373	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	49.56	1795.141	17	449.7919	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAKSAA	49.09	2234.431	22	559.6151	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	48.78	1478.966	14	493.9959	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTK	48.72	2478.537	24	620.6428	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNKK	48.1	2535.439	25	634.8663	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	47.97	1607.061	15	402.7724	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKA	47.59	1806.193	17	452.5549	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	47.55	1667.046	16	556.6887	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	46.89	2145.384	21	716.1351	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	46.82	1707.125	16	427.7889	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	46.65	1745.177	16	582.7325	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEAKPK	46.1	1847.026	18	462.764	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	45.8	1296.824	13	433.2819	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	45.33	1749.135	17	584.05	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	45.03	1436.883	15	360.2282	P16403|H12_HUMAN:HIST1H1C

SLVSKGTLVQTKG	44.86	1316.766	13	439.9291	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	44.42	1495.956	15	499.6597	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPK	44.28	1308.861	12	655.4382	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTK	44.26	1259.745	12	630.8798	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	43.79	1816.214	17	455.0602	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGGATPK	43.17	1804.105	19	452.0336	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAG	42.39	1989.221	21	498.3129	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEAKPKV	41.97	1946.095	19	487.5322	P16403|H12_HUMAN:HIST1H1C
KKAASGEAKPKVKKAGGTKPK	41.37	2108.279	21	528.0768	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAK	41.26	1877.23	18	470.3135	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	40.18	1166.714	12	584.364	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTKK	40.06	2606.632	25	652.666	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASG	39.76	1689.926	18	564.3155	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	39.27	1624.051	16	407.0203	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPK	39.06	1180.766	11	394.5956	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	38.72	1735.156	16	579.3934	P16403|H12_HUMAN:HIST1H1C
SRIKLGLKSLVSK	38.34	1427.919	13	357.9874	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKKK	37.94	1674.14	15	419.5423	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	37.92	1521.972	15	381.4996	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGGA	37.75	2245.375	24	562.3528	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKK	37.48	1790.125	18	448.5377	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	37.0	1588.	15	398.0	P16403|H12_HUMAN:HIST1H1

	9	055		215	C
AKPKVVKPKKAAPKK	36.18	1617.082	15	405.2775	P16403|H12_HUMAN:HIST1H1C
AKVAKPKKAAKSAAK	35.92	1495.956	15	499.6591	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	35.44	1320.897	12	331.2313	P16403|H12_HUMAN:HIST1H1C
TKPKKPVGAAKKPKK	35.17	1605.045	15	402.2684	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTK	34.3	1795.141	17	599.388	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTG	34.2	1474.835	15	492.6182	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	33.31	1192.802	11	398.6078	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	32.86	1036.676	10	346.5654	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKKAAPK	32.59	1688.119	16	423.0371	P16403|H12_HUMAN:HIST1H1C
TGASGSFKLNKKAASGEAKPKVK	32.51	2303.296	23	576.8313	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTKK	32.36	2081.305	20	521.3334	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTK	30.81	1953.21	19	489.3098	P16403|H12_HUMAN:HIST1H1C
AKAVKPKAAKPK	30.05	1235.808	12	412.9431	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAA	29.54	1138.755	11	380.5926	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKKAAPKK	29.39	1816.214	17	455.0609	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKKK	28.85	1546.045	14	387.5175	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	27.77	1064.707	10	355.9086	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVG	27.71	1038.619	11	347.2135	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKK	26.63	1607.061	15	402.7723	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGA	26.04	1237.751	13	413.5903	P16403|H12_HUMAN:HIST1H1C
SELITKAV	25.83	859.5015	8	430.7576	P16403|H12_HUMAN:HIST1H1C

KSPKKAKVAKP	25.55	1180.766	11	394.5958	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	25.28	1165.609	12	583.8118	P16403|H12_HUMAN:HIST1H1C
PKKAAKSAAKAVKPKAAKPK	24.22	2017.289	20	505.3271	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKK	24.21	1688.119	16	423.0374	P16403|H12_HUMAN:HIST1H1C
AKPKKAAKSAAKAVKPKAAK	23.93	1991.273	20	664.7629	P16403|H12_HUMAN:HIST1H1C
KAAGGATPKKSAKKTPKKA	23.23	1867.137	19	467.7913	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAGG	22.6	2046.242	22	512.5676	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKK	21.51	1320.897	12	441.3062	P16403|H12_HUMAN:HIST1H1C
KKAGGTKPKKPVGAAKKPK	19.51	1918.22	19	480.5624	P16403|H12_HUMAN:HIST1H1C
AAKKPKKAAGGATPKKSAKK	19.49	1965.221	20	492.3126	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAKSAAK	17.54	2064.326	20	689.1162	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPK	16.27	1192.802	11	597.4061	P16403|H12_HUMAN:HIST1H1C
KPKAAKPKVVKPKK	15.31	1546.045	14	387.5181	P16403|H12_HUMAN:HIST1H1C
GVYDGREHTV	32.58	1131.531	10	566.772	P43358|MAGA4_HUMAN:MAGEA4
AQKRVLLQA	18.19	1025.635	9	513.8248	P57740|NU107_HUMAN:NUP107
RKSTGGKAPRKQLATK	38.53	1726.033	16	432.5153	P68431|H31_HUMAN:HIST1H3A
KAPRKQLATKAARK	29.8	1565.984	14	392.5035	P68431|H31_HUMAN:HIST1H3A
KSTGGKAPRKQLATK	25.79	1569.931	15	393.4901	P68431|H31_HUMAN:HIST1H3A
YQKSTELLI	23.16	1093.602	9	547.808	P68431|H31_HUMAN:HIST1H3A
EVYGSVAKV	27.23	950.5073	9	476.2608	Q07864|DPOE1_HUMAN:POLE
YTDPEVFKY	38.86	1160.539	9	581.2765	Q16647|PTGIS_HUMAN:PTGIS
ASSGPPTSPEPGELE	18.0	2153.	21	718.9	Q2M329|CCD96_HUMAN:CCD

SEPEEE	7	912		77	C96
ELKPKTRIF	15.47	1130.681	9	566.3467	Q4G0P3|HYDIN_HUMAN:HYDIN
DVIRALAKY	32.36	1047.608	9	524.8126	Q5TZF3|ANR45_HUMAN:ANKRD45
EVIGLLGGR	26.45	912.5392	9	457.2767	Q5VVJ2|MYSM1_HUMAN:MYSM1
EVIGLLGGRY	17.19	1075.603	10	538.8083	Q5VVJ2|MYSM1_HUMAN:MYSM1
KLHSVTTKL	22.15	1025.623	9	342.8811	Q68BL8|OLM2B_HUMAN:OLFML2B
LQDKNIIQY	18.4	1133.608	9	567.8109	Q6TDU7|CASC1_HUMAN:CASC1
EIFNKQILKY	20.27	1294.729	10	648.3725	Q6V0I7|FAT4_HUMAN:FAT4
PNQGPFTYYLLSTGPATSYF	15.56	2223.053	20	742.028	Q6V0I7|FAT4_HUMAN:FAT4
APVFTQRP	17.71	914.4974	8	458.2556	Q7Z5N4|SDK1_HUMAN:SDK1
KVIEGHGVT	15.56	938.5185	9	470.2674	Q8N3T6|T132C_HUMAN:TMEM132C
KENISVADTL	16.62	1088.571	10	545.2948	Q8N7X0|ADGB_HUMAN:ADGB
LKEEAEKKL	16.93	1086.628	9	363.2173	Q8NEH6|MNS1_HUMAN:MNS1
GQSALPQRL	16.13	968.5403	9	485.2782	Q96L91|EP400_HUMAN:EP400
KDLIARAF	20.58	932.5443	8	467.2804	Q96MC2|DRC1_HUMAN:DRC1
IRANATNKL	26.64	999.5825	9	334.2007	Q9BV19|CA050_HUMAN:C1orf50
RQADLKNWSL	27.87	1229.652	10	615.8339	Q9UPW0|FOXJ3_HUMAN:FOXJ3
7					
EADTYTKLF	16.18	1086.523	9	544.2679	A2RUB1|MEIOC_HUMAN:MEIOC
LCLRNSSTYY	20.82	1218.57	10	610.2925	A8MT70|ZBBX_HUMAN:ZBBX
RLPLRVVKH	24.03	1116.724	9	373.248	O00626|CCL22_HUMAN:CCL22
DETLVHSSF	26.56	1033.472	9	517.7431	O14595|CTDS2_HUMAN:CTDSP2
DEFIIKTV	26.8	963.5	8	482.7	O14986|PI51B_HUMAN:PIP5K

	9	276		712	1B
YYPWPYNY	16.77	1164.492	8	583.2515	O75829|CNMD_HUMAN:LECT1
KLLEKAFSI	22.75	1047.633	9	524.8187	O75881|CP7B1_HUMAN:CYP7B1
PKKTESHHKAKGK	47.81	1474.837	13	492.6187	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
LLPKKTESHHKAKGK	45.3	1701.005	15	568.0089	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
LPKKTESHHKAKGK	29.37	1587.921	14	397.9869	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
IPRHLQLAI	18.52	1059.655	9	354.2224	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
KAGGTKPKKPVGAAKKPKKAAGG	82.38	2174.337	23	725.7864	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAG	72.37	2117.316	22	706.7807	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTKK	60.03	1923.236	18	481.8163	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAA	58.65	1808.136	18	453.0408	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAA	58.2	2060.294	21	516.0808	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAAK	56.59	1936.231	19	485.064	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTKK	55.77	2081.305	20	521.3315	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	55.62	1790.125	18	597.7159	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	55.45	1918.22	19	640.4135	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAKSAAK	54.05	2362.526	23	788.5158	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	53.53	1667.046	16	417.7691	P16403|H12_HUMAN:HIST1H1C
TKKVAKSPKKAKVAKPKKAA	52.97	2106.373	20	527.6006	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAKSAAK	51.57	2234.431	22	745.8181	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKK	51.37	1790.125	18	597.7159	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAG	51.35	1989.221	21	498.3125	P16403|H12_HUMAN:HIST1H1C

AGGTKPKKPVGAAKKPK	51.27	1662.03	17	555.0173	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAKSAA	51.21	2234.431	22	559.6152	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	50.14	1607.061	15	536.6947	P16403|H12_HUMAN:HIST1H1C
SAAKAVKPKAAKPKVVKPKKAAPKK	50.03	2568.668	25	643.1739	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTK	49.39	1953.21	19	652.0774	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAA	49.24	1450.935	14	363.7409	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	49.18	2017.289	20	505.3291	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAKAVKPK	49.02	2147.399	21	537.858	P16403|H12_HUMAN:HIST1H1C
KALAAAGYDVEKNNSRIKLGLK	48.98	2358.338	22	590.5859	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAK	48.96	1877.23	18	626.7505	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	48.3	2145.384	21	716.1349	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	48.09	1795.141	17	449.7923	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	47.95	1424.919	14	357.2364	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	47.62	1478.966	14	370.7486	P16403|H12_HUMAN:HIST1H1C
TPKKAKKPAAATVTKK	47.3	1667.046	16	417.7675	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	46.34	1745.177	16	437.3015	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	46.16	1495.956	15	499.6595	P16403|H12_HUMAN:HIST1H1C
TKKVAKSPKKAKVAKPKKAAK	46.11	2234.468	21	559.6233	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	45.88	1436.883	15	360.2274	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	45.6	1165.609	12	583.8077	P16403|H12_HUMAN:HIST1H1C
TKKVAKSPKKAKVAKPKK	45.36	1964.299	18	492.082	P16403|H12_HUMAN:HIST1H1C
GTLVQTKGTGASGSF	45.2	1537.	16	769.9	P16403|H12_HUMAN:HIST1H1

K	8	81		045	C
KSPKKAKVAKPKKAA	45.13	1579.03	15	527.3502	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	44.85	1877.23	18	626.75	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	44.63	1749.135	17	584.0515	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAAK	44.06	2005.325	19	669.4487	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	43.15	1707.125	16	570.0486	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	41.06	1816.214	17	455.0604	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPKKAAPKKK	40.62	2339.562	22	585.898	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPKKAAGG	40.43	2046.242	22	512.5677	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTK	39.96	1212.755	12	405.2588	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGGATPK	39.9	1804.105	19	602.376	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	39.22	1296.824	13	433.2817	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTK	39.22	2478.537	24	620.6423	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	39.18	1036.676	10	519.3454	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTKK	39.17	1340.85	13	447.9574	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTKK	38.88	2606.632	25	652.6652	P16403|H12_HUMAN:HIST1H1C
KAAGGATPKKSAKKTPKKA	38.16	1867.137	19	467.7913	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	38.14	1624.051	16	407.0194	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPK	37.76	1180.766	11	394.5959	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKKK	37.7	1674.14	15	419.5418	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKK	37.41	1688.119	16	423.036	P16403|H12_HUMAN:HIST1H1C
KAASGEAKPKVKKAG	37.07	1468.872	15	368.2253	P16403|H12_HUMAN:HIST1H1C

KAGGTKPKKPVG	37	1166.714	12	584.3647	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPKA	36.83	1367.861	14	456.9609	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	36.04	1588.055	15	398.0205	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	35.94	1521.972	15	508.3314	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	35.29	1735.156	16	434.7959	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGA	34.85	1237.751	13	413.5903	P16403|H12_HUMAN:HIST1H1C
VKPKAAKPK	34.39	965.6385	9	483.8257	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	34.29	1064.707	10	355.9091	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAKA	33.92	1695.088	17	424.7792	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	33.75	1320.897	12	331.2309	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	33.58	1617.082	15	405.2775	P16403|H12_HUMAN:HIST1H1C
TGASGSFKLNKKAASGEAKPKVK	32.86	2303.296	23	576.8315	P16403|H12_HUMAN:HIST1H1C
TGASGSFKLNKKAASGEAKPK	31.83	2076.133	21	520.0408	P16403|H12_HUMAN:HIST1H1C
KKPKKAAGGATPKKSAKKTPK	31.65	2149.342	21	538.3419	P16403|H12_HUMAN:HIST1H1C
KPKAAKPK	30.45	866.5701	8	434.2918	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	29.97	1192.802	11	398.6077	P16403|H12_HUMAN:HIST1H1C
KKPAAATVTKK	29.52	1141.718	11	381.5799	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKKAAPKK	29.44	1816.214	17	455.0609	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPV	26.11	1109.692	11	370.9044	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLN	25.54	2279.249	23	570.8141	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAGG	25.49	1406.872	15	352.7253	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEA	25.2	2074.	20	519.5	P16403|H12_HUMAN:HIST1H1

KPKVK	8	19		569	C
VKPKKAAPK	25.14	965.6385	9	483.826	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKKK	22.5	1546.045	14	516.3552	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKA	22.34	1989.257	20	498.3203	P16403|H12_HUMAN:HIST1H1C
SGSFKLNKKAASGEAKPK	20.97	1847.026	18	462.7638	P16403|H12_HUMAN:HIST1H1C
PKAAKPK	20.46	738.4752	7	370.2456	P16403|H12_HUMAN:HIST1H1C
AKAVKPKAAKPK	19.01	1235.808	12	412.943	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKK	17.28	1320.897	12	441.3051	P16403|H12_HUMAN:HIST1H1C
AKSPKKAKVAKPK	17.1	1379.898	13	460.973	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTK	16.12	1795.141	17	449.7912	P16403|H12_HUMAN:HIST1H1C
KPVGAAKKPKKAAG	15.42	1349.851	14	450.9573	P16403|H12_HUMAN:HIST1H1C
KAVKPKAAKP	15.23	1036.676	10	346.5652	P16403|H12_HUMAN:HIST1H1C
KAASGEAKPKVKKAGGTK	15.19	1755.037	18	439.7652	P16403|H12_HUMAN:HIST1H1C
KVAKPKKAAKSAA	15.07	1296.824	13	433.2816	P16403|H12_HUMAN:HIST1H1C
KPKAAKPKVVKPKKAA	15.02	1688.119	16	423.0367	P16403|H12_HUMAN:HIST1H1C
AYLEAHETF	36.68	1079.492	9	540.7539	P57740|NU107_HUMAN:NUP107
DEVRLLKY	27.47	1034.576	8	518.2946	P57740|NU107_HUMAN:NUP107
RKSTGGKAPRKQLATK	34.99	1726.033	16	432.515	P68431|H31_HUMAN:HIST1H3A
KSTGGKAPRKQLATK	30.82	1569.931	15	393.4898	P68431|H31_HUMAN:HIST1H3A
VPVRYRI	19.6	901.5497	7	451.7803	Q16647|PTGIS_HUMAN:PTGIS
EIDEEKNLL	15.98	1101.555	9	551.7814	Q4G0P3|HYDIN_HUMAN:HYDIN
PFARIVDI	15.03	929.5334	8	465.7707	Q4G0P3|HYDIN_HUMAN:HYDIN

DIDEVTVNVTVLDANDPPI	15.2	2038.011	19	510.5075	Q6V0I7|FAT4_HUMAN:FAT4
GELINWVAL	15.9	1013.555	9	507.7834	Q6V1P9|PCD23_HUMAN:DCHS2
RLASVGLDAK	15.22	1028.598	10	515.3027	Q6ZMW3|EMAL6_HUMAN:EML6
PTLTWWVAY	15.22	1135.57	9	568.7939	Q6ZP82|CC141_HUMAN:CCDC141
CLVTDLIIW	18.12	1074.578	9	538.2982	Q7Z5N4|SDK1_HUMAN:SDK1
FTQRPVDTT	16.53	1063.53	9	532.7713	Q7Z5N4|SDK1_HUMAN:SDK1
SSSSSTETS	18.37	871.3407	9	436.6763	Q86X95|CIR1_HUMAN:CIR1
TFISAASKF	19.63	970.5123	9	486.2635	Q8N755|PQLC3_HUMAN:PQLC3
RSRQILTTTAVRKI	15.98	1642	14	411.5063	Q8NDX2|VGLU3_HUMAN:SLC17A8
ALQPLLHTV	23.07	990.5862	9	496.3002	Q8NFM7|I17RD_HUMAN:IL17RD
FWKPRNL	17.11	959.5341	7	480.7725	Q8NFM7|I17RD_HUMAN:IL17RD
RLIQGDQILSV	16.74	1240.714	11	621.3649	Q8NI35|INADL_HUMAN:PATJ
VLGPTPVQK	16.89	937.5596	9	469.7824	Q8WU20|FRS2_HUMAN:FRS2
PAATFQSINK	15.92	1075.566	10	538.7878	Q96L91|EP400_HUMAN:EP400
IKCIVVLQ	17.12	914.5623	8	458.2901	Q96N23|CFA54_HUMAN:CFAP54
KDKQRELGDK	15.32	1215.657	10	608.8341	Q96PX9|PKH4B_HUMAN:PLEKHG4B
ELPTTLTF	17.74	920.4855	8	461.2501	Q99645|EPYC_HUMAN:EPYC
LEQKVVELY	24.48	1119.618	9	560.8163	Q9HAI6|CX021_HUMAN:CXorf21
ELPLCFAL	16.43	904.4728	8	453.2452	Q9NXZ2|DDX43_HUMAN:DDX43
AKKPKKEQDNDDENKPK	19.38	2011.033	17	503.7637	Q9NY46|SCN3A_HUMAN:SCN3A
KLGSKKPQKP	16.51	1109.692	10	370.9041	Q9NY46|SCN3A_HUMAN:SCN3A
PLERPVLV	17.9	921.5	8	461.7	Q9NYQ6|CELR1_HUMAN:CEL

	3	647		899	SR1
AILLHYIY	15.27	1004.57	8	503.2897	Q9NYQ6|CELR1_HUMAN:CELSR1
DELGRHVL	17.54	937.4981	8	469.7566	Q9UBX1|CATF_HUMAN:CTSF
SLRDLLAEV	18.49	1014.571	9	508.2919	Q9ULI4|KI26A_HUMAN:KIF26A
FAEDGARTV	16.58	964.4614	9	483.2334	Q9UP38|FZD1_HUMAN:FZD1
TYQFPDGTV	15.98	1026.466	9	514.2367	Q9UQP3|TENN_HUMAN:TNN
LPNQTVLKI	25.43	1024.628	9	513.3216	Q9Y2I9|TBC30_HUMAN:TBC1D30
YLVTKVVAV	28.95	990.6113	9	496.313	Q9Y5E9|PCDBE_HUMAN:Q9Y5E8|
8					
GTKYIFMVK	43.86	1085.594	9	543.8052	O15344|TRI18_HUMAN:MID1
VPIGYKGSTFHRVIK	39.55	1700.973	15	426.2507	O43447|PPIH_HUMAN:PPIH
KVSDYILQH	35.06	1101.582	9	551.7991	O75129|ASTN2_HUMAN:ASTN2
SVSLVPPSK	17.15	912.528	9	457.272	O95156|NXPH2_HUMAN:NXPH2
KLFQSDTNAMLGK	54.38	1451.744	13	484.9228	O95619|YETS4_HUMAN:YEATS4
KPTQTLTL	21.73	900.528	8	451.2715	P01817|HV205_HUMAN:IGHV2-5
VVFPFPVNK	21.87	1045.596	9	523.8064	P04198|MYCN_HUMAN:MYCN
FPVGRVHRL	26.35	1079.635	9	360.8857	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
IQAVLLPK	19.39	880.5746	8	441.295	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
AGLQFPVGR	17.9	943.5239	9	472.7694	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
ALFSSKDFAGK	41.66	1169.608	11	585.8127	P08842|STS_HUMAN:STS
HVSTVNPNY	16.57	1029.488	9	515.7524	P09017|HXC4_HUMAN:HOXC4
GTLVQTKGTGASGSFKLNK	68.14	1893.032	19	632.0208	P16403|H12_HUMAN:HIST1H1C
GLKSLVSKGTLVQTK	64.9	2350.	24	588.5	P16403|H12_HUMAN:HIST1H1

GTGASGSFK	6	322		876	C
KALAAAGYDVEKNNSRIKLGLK	63.44	2358.338	22	590.5923	P16403|H12_HUMAN:HIST1H1C
GTLVQTKGTGASGSFK	61.94	1537.81	16	769.9143	P16403|H12_HUMAN:HIST1H1C
VQTKGTGASGSFKLNK	60.94	1621.879	16	541.6355	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNK	59.96	2407.344	24	602.8444	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	57.56	1790.125	18	597.7157	P16403|H12_HUMAN:HIST1H1C
LVQTKGTGASGSFKLNK	56.6	1734.963	17	579.3281	P16403|H12_HUMAN:HIST1H1C
GLKSLVSKGTLVQTK	53.91	1557.945	15	520.3228	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	53.84	2017.289	20	673.4375	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTK	52.94	1259.745	12	630.88	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	52.67	1165.609	12	583.8128	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	52.3	1918.22	19	640.4136	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	51.23	1662.03	17	416.5147	P16403|H12_HUMAN:HIST1H1C
KASGPPVSELITK	50.61	1325.755	13	663.8855	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGG	49.87	2174.337	23	544.5913	P16403|H12_HUMAN:HIST1H1C
RIKLGLKSLVSK	49.65	1340.887	12	336.2291	P16403|H12_HUMAN:HIST1H1C
KLGLKSLVSKGTLVQTK	49.44	1799.124	17	450.7887	P16403|H12_HUMAN:HIST1H1C
SGPPVSELITK	48.52	1126.623	11	564.3195	P16403|H12_HUMAN:HIST1H1C
GASGSFKLNK	47.29	1007.54	10	504.7777	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTK	47.07	1212.755	12	405.2594	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKK	46.79	1293.704	13	432.2422	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	46.59	1478.966	14	370.7488	P16403|H12_HUMAN:HIST1H1C

KVAKSPKKAKVAKPK	46.07	1607.061	15	402.7724	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	45.09	1745.177	16	437.3017	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAG	44.94	2117.316	22	530.3366	P16403|H12_HUMAN:HIST1H1C
ALAAAGYDVEKNNSRIKLGLK	44.28	2230.243	21	558.5679	P16403|H12_HUMAN:HIST1H1C
AVAASKERSGVSLAALK	43.97	1656.952	17	553.3264	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	43.95	1036.676	10	346.5659	P16403|H12_HUMAN:HIST1H1C
AASKERSGVSL	43.87	1103.594	11	552.8048	P16403|H12_HUMAN:HIST1H1C
SGVSLAALKK	43.79	972.5967	10	487.3059	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLN	43.54	1037.514	11	519.7649	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	43.36	1816.214	17	455.06	P16403|H12_HUMAN:HIST1H1C
ALAAAGYDVEK	43.3	1106.561	11	554.288	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	42.83	1624.051	16	407.0186	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	42.53	1877.23	18	626.7504	P16403|H12_HUMAN:HIST1H1C
ASGPPVSELITK	42.44	1197.66	12	599.8381	P16403|H12_HUMAN:HIST1H1C
VSKGTLVQTK	42.07	1059.629	10	530.8221	P16403|H12_HUMAN:HIST1H1C
ASGSFKLNK	41.17	950.5185	9	476.2668	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	41.11	1064.707	10	533.3608	P16403|H12_HUMAN:HIST1H1C
GTKPKKPVGAAK	40.97	1180.729	12	394.5836	P16403|H12_HUMAN:HIST1H1C
TGASGSFKLNK	39.7	1108.588	11	370.5367	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKK	39.13	1688.119	16	423.0372	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	39	1588.055	15	398.0212	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	38.8	1579.	15	395.7	P16403|H12_HUMAN:HIST1H1

	8	03		646	C
KAAKSAAKAVKPK	38.87	1296.824	13	433.282	P16403|H12_HUMAN:HIST1H1C
SGSFKLNK	38.64	879.4814	8	440.7482	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPK	38.49	1308.861	12	437.2938	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	37.68	2145.384	21	537.3533	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	36.87	1320.897	12	331.2312	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	36.29	1436.883	15	360.2278	P16403|H12_HUMAN:HIST1H1C
SGVSLAALK	36.09	844.5018	9	423.2582	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	35.58	1617.082	15	405.2782	P16403|H12_HUMAN:HIST1H1C
KLGLKSLVSK	35.29	1071.702	10	358.242	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	34.86	1707.125	16	427.7885	P16403|H12_HUMAN:HIST1H1C
VKPKAAKPK	34.71	965.6385	9	483.8268	P16403|H12_HUMAN:HIST1H1C
ASGPPVSELITKAVAASK	33.55	1724.967	18	575.9972	P16403|H12_HUMAN:HIST1H1C
VKPKKAAPK	33.27	965.6385	9	322.8867	P16403|H12_HUMAN:HIST1H1C
GVSLAALKK	31.93	885.5647	9	443.7903	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	31.06	1667.046	16	417.7687	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	30.7	1735.156	16	434.7959	P16403|H12_HUMAN:HIST1H1C
KPKKAAPK	30.25	866.5701	8	434.292	P16403|H12_HUMAN:HIST1H1C
KPKAAKPK	29.63	866.5701	8	434.292	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	29.45	1192.802	11	398.608	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPK	28.06	1180.766	11	394.5961	P16403|H12_HUMAN:HIST1H1C
KPKVVKPKKAAPKKK	27.49	1674.14	15	419.5411	P16403|H12_HUMAN:HIST1H1C

LGLKSLVSK	25.71	943.6066	9	472.8105	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPKKAAPKK	24.03	2211.467	21	553.8749	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAA	22.12	1308.788	14	437.2707	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	21.78	1521.972	15	381.5006	P16403|H12_HUMAN:HIST1H1C
PKKAAPK	21.69	738.4752	7	370.2445	P16403|H12_HUMAN:HIST1H1C
PKAAKPK	21.18	738.4752	7	370.2451	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPK	20.55	993.6698	9	332.2308	P16403|H12_HUMAN:HIST1H1C
KKPAAATVTK	17.48	1013.623	10	507.8203	P16403|H12_HUMAN:HIST1H1C
KAAKPKVVKPKK	15.71	1320.897	12	441.3064	P16403|H12_HUMAN:HIST1H1C
NNSRIKLGL	15.62	1013.598	9	338.8735	P16403|H12_HUMAN:HIST1H1C
KVIRDFNTNK	20.16	1233.683	10	412.2343	P26378|ELAV4_HUMAN:ELAVL4
KMVDPGEEQPT	16.89	1229.56	11	615.7886	P29120|NEC1_HUMAN:PCSK1
SPSLHKLVL	33.27	992.6018	9	497.3084	P43897|EFTS_HUMAN:TSFM
SPRDPENFPFV	22.49	1303.62	11	652.818	P51149|RAB7A_HUMAN:RAB7A
IPQKFIVDY	40.8	1121.612	9	561.8126	P55774|CCL18_HUMAN:CCL18
VPRPQPWAL	22.47	1062.597	9	532.308	P55899|FCGRN_HUMAN:FCGRT
SVRPLVTEL	25.81	1012.592	9	507.3034	P57740|NU107_HUMAN:NUP107
TIMPKDIQLAR	45.95	1284.722	11	643.3685	P68431|H31_HUMAN:HIST1H3A
APRKQLATKAA	45.07	1153.693	11	385.5715	P68431|H31_HUMAN:HIST1H3A
KSTGGKAPRKQLATK	41.78	1569.931	15	393.4911	P68431|H31_HUMAN:HIST1H3A
MPKDIQLAR	40.71	1070.591	9	536.3025	P68431|H31_HUMAN:HIST1H3A
APRKQLATK	35.4	1011.	9	338.2	P68431|H31_HUMAN:HIST1H3

	5	619		143	A
KSTELLIRK	35.43	1086.676	9	363.2326	P68431|H31_HUMAN:HIST1H3A
PKDIQLAR	29.45	939.5501	8	470.7824	P68431|H31_HUMAN:HIST1H3A
KAPRKQLATK	24.03	1139.714	10	380.9125	P68431|H31_HUMAN:HIST1H3A
YRPGTVAL	20.88	875.4865	8	438.7514	P68431|H31_HUMAN:HIST1H3A
KPRTIYSSF	28.97	1097.587	9	549.8013	Q07687|DLX2_HUMAN:DLX2
FPVLPGSHL	35.1	965.5334	9	483.774	Q07864|DPOE1_HUMAN:POLE
RPREAEVRF	47.01	1158.626	9	387.2167	Q13438|OS9_HUMAN:OS9
RPREAEVRFL	43.37	1271.71	10	424.9114	Q13438|OS9_HUMAN:OS9
HPSLQPEEY	23.57	1098.498	9	550.2573	Q13438|OS9_HUMAN:OS9
NVRPPNTAL	23.15	980.5403	9	491.2778	Q17RS7|GEN_HUMAN:GEN1
SPFARGNVF	29.48	993.5032	9	497.7591	Q1MSJ5|CSPP1_HUMAN:CSPP1
VPRPTKSSF	36.25	1017.561	9	509.7889	Q5VVJ2|MYSM1_HUMAN:MYSM1
PRPTKSSF	25.58	918.4923	8	460.2535	Q5VVJ2|MYSM1_HUMAN:MYSM1
YPLHDGDYNGENVEF	40.52	1767.738	15	884.8779	Q5XXA6|ANO1_HUMAN:ANO1
KSFEERLIFK	16.94	1295.724	10	432.9153	Q5XXA6|ANO1_HUMAN:ANO1
SPDPTRESVL	26.66	1099.551	10	550.7845	Q68BL8|OLM2B_HUMAN:OLFML2B
MPDIRSAL	18.81	901.4691	8	451.7418	Q68BL8|OLM2B_HUMAN:OLFML2B
ATFGDGTTIIAK	40.6	1193.629	12	597.8223	Q6Q759|SPG17_HUMAN:SPAG17
GSISTILPEK	29.21	1043.586	10	522.801	Q6Q759|SPG17_HUMAN:SPAG17
RPRPTWEEV	25.75	1168.599	9	585.3071	Q8N9E0|F133A_HUMAN:FAM133A
GPREALSRL	35.26	997.5668	9	333.5293	Q8NF99|ZN397_HUMAN:ZNF397

RIHTGEKPYK	32.84	1227.672	10	410.232	Q8NF99|ZN397_HUMAN:ZNF397
ASQETVATILK	42.06	1159.645	11	580.8307	Q8NI35|INADL_HUMAN:PATJ
SPAGKTNAL	30.08	857.4606	9	429.7381	Q8NI35|INADL_HUMAN:PATJ
AVIDIERTAAMSNLQK	42.61	1758.93	16	587.3173	Q8WU20|FRS2_HUMAN:FRS2
VLGPTPVQK	23.82	937.5596	9	469.7877	Q8WU20|FRS2_HUMAN:FRS2
SSSPQLPPK	20.18	939.5025	9	470.7589	Q8WZ74|CTTB2_HUMAN:CTTNBP2
DAGPPTHAF	28.92	911.4137	9	456.7144	Q96FA3|PELI1_HUMAN:PELI1
APRPKSHWDYL	16.77	1368.694	11	457.2389	Q96L91|EP400_HUMAN:EP400
APRPGALL	15.88	793.481	8	397.7482	Q96L91|EP400_HUMAN:EP400
HVLGAIFSK	33.64	970.5599	9	486.288	Q96MM7|H6ST2_HUMAN:HS6ST2
LPRALLSSL	32.88	968.6018	9	485.309	Q96RQ9|OXLA_HUMAN:IL4I1
RPFWREEHI	31.63	1268.641	9	423.8883	Q96RQ9|OXLA_HUMAN:IL4I1
MPSSHRILHKL	22.17	1317.734	11	330.441	Q96RQ9|OXLA_HUMAN:IL4I1
KTIEESLIYRK	45.21	1378.782	11	460.601	Q9BX26|SYCP2_HUMAN:SYCP2
ASLEITNVTQK	32.39	1202.651	11	602.334	Q9BX26|SYCP2_HUMAN:SYCP2
SPYPKTSSL	23.7	978.5022	9	490.2584	Q9BX26|SYCP2_HUMAN:SYCP2
RPIHPINGL	35.34	1015.593	9	508.8038	Q9NYQ6|CELR1_HUMAN:CELSR1
ATSAPLVSR	28.69	900.5029	9	451.2595	Q9NYQ6|CELR1_HUMAN:CELSR1
SPGAQRTFFQL	34.45	1250.641	11	626.3287	Q9UJ98|STAG3_HUMAN:STAG3
IQGGILNLNK	29.73	1068.629	10	535.3223	Q9UNK4|PA2GD_HUMAN:PLA2G2D
SPRSSLNNSL	24.56	1073.547	10	537.7809	Q9UPW0|FOXJ3_HUMAN:FOXJ3
GTYVPGYLLK	28.9	1109.	10	555.8	Q9Y653|AGRG1_HUMAN:ADG

	7	612		139	RG1
9					
SEISKRLGAEW	35.33	1274.662	11	638.3386	O00570|SOX1_HUMAN:SOX1
ISKRLGAEW	31.97	1058.587	9	530.3004	O00570|SOX1_HUMAN:SOX1
RLPLRVVKH	24.21	1116.724	9	373.2486	O00626|CCL22_HUMAN:CCL22
KTAENFRQF	32.73	1139.572	9	570.7922	O43447|PPIH_HUMAN:PPIH
KVSDYILQH	36.61	1101.582	9	551.7973	O75129|ASTN2_HUMAN:ASTN2
KLLEKAFSI	30.46	1047.633	9	524.8234	O75881|CP7B1_HUMAN:CYP7B1
KLFQSDTNAMLGK	51.84	1451.744	13	484.9221	O95619|YETS4_HUMAN:YEATS4
REEDGHTHQW	44.8	1293.549	10	647.7821	O95619|YETS4_HUMAN:YEATS4
EEDGHTHQW	37.53	1137.448	9	569.7307	O95619|YETS4_HUMAN:YEATS4
KVKEEIGSSW	37.99	1161.603	10	581.809	O95936|EOMES_HUMAN:EOMES
LLPKKTESHHKAKGK	39.51	1701.005	15	426.2574	P0C0S8|H2A1_HUMAN:HIST1H2AL
IQAVLLPK	25.19	880.5746	8	441.2937	P0C0S8|H2A1_HUMAN:HIST1H2AL
KLLGKVTIA	23.13	941.6273	9	471.8209	P0C0S8|H2A1_HUMAN:HIST1H2AL
GVFHQTVSRK	33.26	1157.631	10	386.8844	P12109|CO6A1_HUMAN:COL6A1
IAKDFVVKV	32.29	1017.622	9	340.2144	P12109|CO6A1_HUMAN:COL6A1
KEAIKSLQW	32.07	1101.618	9	551.8166	P12109|CO6A1_HUMAN:COL6A1
GTLVQTKGTGASGSFKLNK	70.99	1893.032	19	947.5231	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNK	66.24	2407.344	24	602.844	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLN	63.42	2279.249	23	760.7573	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNKK	62.61	2535.439	25	634.8676	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	58.6	1918.	19	640.4	P16403|H12_HUMAN:HIST1H1

KPKK	3	22		136	C
GTLVQTKGTGASGSFK	58.31	1537.81	16	769.913	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	57.99	1662.03	17	416.5143	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	57.28	1667.046	16	417.7686	P16403|H12_HUMAN:HIST1H1C
KALAAAGYDVEKNNSRIKLGLK	56.53	2358.338	22	590.5916	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	55.29	1795.141	17	449.7917	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	54.08	1877.23	18	626.7496	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFK	53.4	2052.122	21	685.0468	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	52.26	2017.289	20	505.3289	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	51.03	1749.135	17	438.2901	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTK	50.5	1259.745	12	630.8807	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	50.19	1478.966	14	493.9949	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKA	49.37	1806.193	17	452.5545	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPK	48.22	1790.125	18	448.5389	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	47.97	1165.609	12	583.8114	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	47.63	1579.03	15	527.351	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	47.62	1816.214	17	455.0594	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	47.55	1436.883	15	360.2277	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	47.36	2145.384	21	537.3523	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAAKSAAK	47.33	1936.231	19	485.0653	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	46.22	1607.061	15	402.7718	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	46.09	1424.919	14	475.9801	P16403|H12_HUMAN:HIST1H1C

ALAAAGYDVEKNNSRIKLGLK	45.37	2230.243	21	558.5686	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	45.32	1745.177	16	437.3005	P16403|H12_HUMAN:HIST1H1C
VSKGTLVQTK	45.17	1059.629	10	530.8225	P16403|H12_HUMAN:HIST1H1C
GTKPKKPVGAAK	45.14	1180.729	12	394.5841	P16403|H12_HUMAN:HIST1H1C
GLKSLVSKGTLVQTK	45.06	1557.945	15	520.322	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	44.33	1296.824	13	433.2818	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	44.21	1707.125	16	427.7883	P16403|H12_HUMAN:HIST1H1C
AVAASKERSGVSLAALK	43.61	1656.952	17	553.3248	P16403|H12_HUMAN:HIST1H1C
GASGSFKLNK	43.47	1007.54	10	504.7776	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	43.43	1036.676	10	519.344	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	43.06	1588.055	15	398.0215	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKK	42.76	1688.119	16	423.0375	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTK	42.6	1212.755	12	405.2592	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAA	41.82	1450.935	14	363.7409	P16403|H12_HUMAN:HIST1H1C
LVQTKGTGASGSFKLNK	41.17	1734.963	17	579.3286	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPK	40.6	1308.861	12	437.2946	P16403|H12_HUMAN:HIST1H1C
ASGSFKLNK	40.5	950.5185	9	476.2673	P16403|H12_HUMAN:HIST1H1C
RIKLGLKSLVSK	40.39	1340.887	12	336.2287	P16403|H12_HUMAN:HIST1H1C
SAAKAVKPKAAKPK	39.95	1393.877	14	465.6336	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	39.3	1064.707	10	533.3593	P16403|H12_HUMAN:HIST1H1C
KAKKPAAATVTKK	39.18	1340.85	13	447.9575	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	38.6	1617.	15	405.2	P16403|H12_HUMAN:HIST1H1

		082		778	C
KAKVAKPKKAAKSAA	38.3	1495.956	15	374.9967	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLN	38.07	1037.514	11	519.7651	P16403|H12_HUMAN:HIST1H1C
AKAVKPKAAKPK	37.87	1235.808	12	412.9436	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKK	36.47	1293.704	13	432.2422	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	36.15	1735.156	16	434.7953	P16403|H12_HUMAN:HIST1H1C
VKPKAAKPK	36.15	965.6385	9	483.8256	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKKK	36.11	1320.897	12	331.2314	P16403|H12_HUMAN:HIST1H1C
KLGLKSLVSK	35.66	1071.702	10	358.2401	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	35.15	1624.051	16	407.0202	P16403|H12_HUMAN:HIST1H1C
ALAAAGYDVEKNNSRIK	34.75	1818.959	17	607.3294	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQT	34.72	1131.65	11	566.8319	P16403|H12_HUMAN:HIST1H1C
KPKVVKPK	34.66	922.6327	8	462.3237	P16403|H12_HUMAN:HIST1H1C
SGSFKLNK	33.39	879.4814	8	440.749	P16403|H12_HUMAN:HIST1H1C
KSLVSKGTLVQTK	33.32	1387.84	13	463.6212	P16403|H12_HUMAN:HIST1H1C
LGLKSLVSK	31.69	943.6066	9	472.8098	P16403|H12_HUMAN:HIST1H1C
VKPKKAAPK	31.49	965.6385	9	483.8261	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	30	1192.802	11	398.608	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	29.67	1521.972	15	508.3308	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPK	29.41	1180.766	11	394.5956	P16403|H12_HUMAN:HIST1H1C
GTLVQTKGTGASGSFKLNKK	28.64	2021.127	20	506.2876	P16403|H12_HUMAN:HIST1H1C
KPKAAKPK	28.3	866.5701	8	434.2926	P16403|H12_HUMAN:HIST1H1C

KPKKAAPK	28.11	866.5701	8	434.292	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPK	27.85	1064.707	10	355.9094	P16403|H12_HUMAN:HIST1H1C
GVSLAALKK	26.56	885.5647	9	443.7896	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPKKAAPKKK	25.91	2339.562	22	585.8991	P16403|H12_HUMAN:HIST1H1C
PKKAAPK	22.4	738.4752	7	370.2441	P16403|H12_HUMAN:HIST1H1C
PKAAKPK	22.06	738.4752	7	370.2447	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	22.05	1166.714	12	389.9115	P16403|H12_HUMAN:HIST1H1C
KKVAKSPKKAKVAKPK	18.02	1735.156	16	434.796	P16403|H12_HUMAN:HIST1H1C
KAASGEAKPK	17.87	985.5556	10	493.7857	P16403|H12_HUMAN:HIST1H1C
KPKKAAKSAAKAVKPKAAKP	15.65	2017.289	20	505.3287	P16403|H12_HUMAN:HIST1H1C
GGTKPKKPVGAAKKPKK	15.34	1719.088	17	430.7791	P16403|H12_HUMAN:HIST1H1C
KLGLKSLV	15.08	856.5746	8	429.2951	P16403|H12_HUMAN:HIST1H1C
RLWGAARKPSGAGWEK	40.02	1768.949	16	443.2449	P22083|FUT4_HUMAN:FUT4
RILGQQVPYATK	35.48	1372.783	12	687.3985	P36222|CH3L1_HUMAN:CHI3L1
LRFPLTNAI	23.22	1043.613	9	522.8141	P36222|CH3L1_HUMAN:CHI3L1
ALSAVLALPA	15.12	924.5644	10	463.289	P49683|PRLHR_HUMAN:PRLHR
ATKRAQAW	35.7	930.5035	8	466.2596	P51149|RAB7A_HUMAN:RAB7A
KARPSSPGF	34.33	945.5032	9	473.7584	P55899|FCGRN_HUMAN:FCGRT
AEDELFNRY	32.86	1155.52	9	578.7665	P57740|NU107_HUMAN:NUP107
LLEKAQQLL	21	1054.639	9	528.3286	P60508|SYCY2_HUMAN:ERVFRD-1
RKSTGGKAPRKQLATK	47.36	1726.033	16	432.5154	P68431|H31_HUMAN:HIST1H3A
KSTELLIRKL	40.7	1199.	10	400.9	P68431|H31_HUMAN:HIST1H3

	6	76		277	A
KSTGGKAPRKQLATK	38.31	1569.931	15	393.49	P68431|H31_HUMAN:HIST1H3A
TIMPKDIQLAR	36.11	1284.722	11	429.2484	P68431|H31_HUMAN:HIST1H3A
VALREIRRY	26.62	1174.694	9	392.5714	P68431|H31_HUMAN:HIST1H3A
KAPRKQLATK	22.79	1139.714	10	380.9117	P68431|H31_HUMAN:HIST1H3A
ITSSIHSKETF	34.65	1248.635	11	417.2184	Q07864|DPOE1_HUMAN:POLE
ALYSAEHGL	26.06	959.4712	9	480.7431	Q07864|DPOE1_HUMAN:POLE
NEPDEAHLIQRW	20.62	1506.722	12	503.2468	Q07864|DPOE1_HUMAN:POLE
KTRDELERSF	23.51	1279.652	10	427.5588	Q12999|TSN31_HUMAN:TSPAN31
AEGTEEGARW	36.95	1104.484	10	553.2487	Q13438|OS9_HUMAN:OS9
RENPQLKQI	31.26	1124.63	9	563.3226	Q13438|OS9_HUMAN:OS9
KTLPVAFKV	25.7	1001.627	9	501.8209	Q13950|RUNX2_HUMAN:RUNX2
VRNPKILIL	15.61	1064.707	9	355.9094	Q2M3G0|ABCB5_HUMAN:ABCB5
RIHETLIRK	22.54	1164.709	9	389.2435	Q5HYA8|MKS3_HUMAN:TMEM67
KTRWIIEKY	16.45	1235.703	9	412.9087	Q5VVJ2|MYSM1_HUMAN:MYSM1
YGNTLVEFR	30.68	1097.551	9	549.783	Q68BL8|OLM2B_HUMAN:OLFML2B
KRVGICKGASSYITHID	16.22	1846.972	17	462.7515	Q6ZMW3|EMAL6_HUMAN:EML6
KVKDLVARI	24.31	1040.671	9	347.8969	Q8IX21|SLF2_HUMAN:SLF2
TSAYHYVQCP	18.31	1167.502	10	584.758	Q8IX21|SLF2_HUMAN:SLF2
KTRDSGTVSALTW	21.74	1420.731	13	474.5844	Q8N755|PQLC3_HUMAN:PQLC3
GIVGLLLTTL	17.71	998.6376	10	500.325	Q8NEN0|ARMC2_HUMAN:ARMC2
KSLGKRDLY	18.6	1078.613	9	360.5447	Q8NF99|ZN397_HUMAN:ZNF397

RIHTGEKPYK	17.51	1227.672	10	410.2317	Q8NF99|ZN397_HUMAN:ZNF397
KEELVDEPFLDLGKS	42.11	1717.877	15	573.6327	Q8NI35|INADL_HUMAN:PATJ
KSRREAVSF	23.68	1078.588	9	360.5359	Q8NI35|INADL_HUMAN:PATJ
VLGPTPVQK	25.46	937.5596	9	469.7875	Q8WU20|FRS2_HUMAN:FRS2
KLTDIQIEL	22.25	1071.618	9	536.8168	Q96AA8|JKIP2_HUMAN:JAKMIP2
LAALNHTLL	16.2	964.5706	9	483.2921	Q96MS0|ROBO3_HUMAN:ROBO3
RQPNFARKL	18.33	1128.652	9	377.2241	Q96P67|GPR82_HUMAN:GPR82
GEHTAYPHGW	39.86	1153.494	10	577.7543	Q96RQ9|OXLA_HUMAN:IL4I1
HTAYPHGW	39.81	967.4301	8	484.7209	Q96RQ9|OXLA_HUMAN:IL4I1
TLSTHLLIH	31.88	1033.592	9	345.5382	Q99684|GFI1_HUMAN:GFI1
GLYERPTAA	29.9	976.4977	9	489.2584	Q99684|GFI1_HUMAN:GFI1
KTFGHAVSL	29.56	958.5236	9	480.2692	Q99684|GFI1_HUMAN:GFI1
IRANATNKL	22.54	999.5825	9	334.2007	Q9BV19|CA050_HUMAN:C1orf50
KELETKQSY	30.44	1124.571	9	563.2924	Q9H6P5|TASP1_HUMAN:TASP1
RTFKGNAQKMFSPK	27.77	1638.866	14	410.7239	Q9H8H2|DDX31_HUMAN:DDX31
NETDRLVHW	36.23	1168.563	9	585.2892	Q9NS86|LANC2_HUMAN:LANCL2
GTFSPPKQW	25.86	1046.519	9	524.2654	Q9NYF0|DACT1_HUMAN:DACT1
VRAQKIQAL	16.71	1025.635	9	513.8246	Q9UBX1|CATF_HUMAN:CTSF
SLRDLLAEV	21.34	1014.571	9	508.2922	Q9ULI4|KI26A_HUMAN:KIF26A
RSKDDPGKGSYW	44.81	1394.658	12	465.8928	Q9UPW0|FOXJ3_HUMAN:FOXJ3
MALPYHIF	16.75	990.4997	8	496.2591	Q9UQF0|SYCY1_HUMAN:ERVW-1
10					

SYDTITVHW	31.05	1120.519	9	561.2647	O15344|TRI18_HUMAN:MID1
TFRLMRVL	15.12	1034.606	8	518.3087	O43497|CAC1G_HUMAN:CACNA1G
KLLEKAFSI	31.88	1047.633	9	524.8239	O75881|CP7B1_HUMAN:CYP7B1
VPLPDAGVYI	17.02	1042.57	10	522.2968	O94779|CNTN5_HUMAN:CNTN5
AGEKLGKEARSLYER	43.58	1705.911	15	427.4856	O94923|GLCE_HUMAN:GLCE
TAGEKLGKEARSLYER	41.29	1806.959	16	452.7472	O94923|GLCE_HUMAN:GLCE
TAGEKLGKEARSLYERG	40.11	1863.98	17	467.0031	O94923|GLCE_HUMAN:GLCE
LQNKVVVTI	30.83	1012.628	9	507.3206	O95199|RCBT2_HUMAN:RCBTB2
RLKASRETI	22.84	1072.635	9	358.5518	O95619|YETS4_HUMAN:YEATS4
YYLPNKNQMVF	35.98	1415.691	11	708.8524	O95672|ECEL1_HUMAN:ECEL1
LTDKHAPEHY	35.31	1209.578	10	605.7965	O95672|ECEL1_HUMAN:ECEL1
KRRSQSIYL	16.31	1149.662	9	384.228	O95672|ECEL1_HUMAN:ECEL1
YTDVPFHNQY	41.43	1282.562	10	642.2887	O95897|NOE2_HUMAN:OLFM2
PKKTESHHKAKGK	51.25	1474.837	13	492.6192	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
LLPKKTESHHKAKGK	49.41	1701.005	15	426.2583	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
RIIPRHLQL	27.34	1144.719	9	382.5801	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
LLPKKTESHHKAKG	26.1	1572.91	14	394.2348	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
IQAVLLPK	25.71	880.5746	8	441.2941	P04908|H2A1B_HUMAN:P0C0S8|H2A1_HUMAN:
MSDKQFHEAAYLADE	20.17	1753.762	15	585.595	P04920|B3A2_HUMAN:SLC4A2
IAKDFVVKV	31.58	1017.622	9	509.8179	P12109|CO6A1_HUMAN:COL6A1
AKDFVVKV	29.03	904.5381	8	453.2763	P12109|CO6A1_HUMAN:COL6A1
KAGGTKPKKPVGAAK	60.3	2117.	22	706.7	P16403|H12_HUMAN:HIST1H1

KPKKAAG	1	316		801	C
KAGGTKPKKPVGAAKKPK	58.92	1790.125	18	597.7162	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGG	55.17	2174.337	23	725.787	P16403|H12_HUMAN:HIST1H1C
GATPKKSAKKTPKKAKKPAAATVTK	54.74	2535.559	25	634.9	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTK	54.38	2478.537	24	620.6451	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKK	54.34	1918.22	19	640.4142	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPK	53.78	2017.289	20	505.3295	P16403|H12_HUMAN:HIST1H1C
AGGTKPKKPVGAAKKPK	53.12	1662.03	17	416.5147	P16403|H12_HUMAN:HIST1H1C
KSAKKTPKKAKKPAAATVTK	53.04	2081.305	20	521.3337	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAA	52.81	1579.03	15	790.5215	P16403|H12_HUMAN:HIST1H1C
KKTPKKAKKPAAATVTKK	52.21	1923.236	18	481.8157	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAA	51.79	1749.135	17	584.052	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAA	51.38	1495.956	15	374.9962	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPKKAA	50.07	1450.935	14	363.7405	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKAA	49.88	1877.23	18	626.7506	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPK	49.77	1478.966	14	493.9959	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKKK	47.85	1745.177	16	437.3015	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPK	47.29	1607.061	15	402.7732	P16403|H12_HUMAN:HIST1H1C
KKAAKSAAKAVKPK	47.23	1424.919	14	357.2368	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTKK	46.25	1795.141	17	449.7922	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTK	45.68	1259.745	12	630.8796	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAA	45.62	2060.294	21	516.0808	P16403|H12_HUMAN:HIST1H1C

KKPAAATVTKKVAKSPK	45.45	1752.099	17	439.032	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKKK	45.11	1816.214	17	455.0602	P16403|H12_HUMAN:HIST1H1C
AKPKVVKPKKAAPKK	44.92	1617.082	15	405.2774	P16403|H12_HUMAN:HIST1H1C
KTPKKAKKPAAATVTK	44.52	1667.046	16	417.7682	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKPKKAAGGA	44.27	2245.375	24	562.3549	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKG	44.27	1316.766	13	659.3964	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNK	43.27	1165.609	12	583.8129	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPK	43.21	1296.824	13	433.2816	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKKA	42.77	1806.193	17	603.0707	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPKVVKPK	41.64	1588.055	15	398.0204	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPK	41.28	1308.861	12	437.2938	P16403|H12_HUMAN:HIST1H1C
TPKKSAKKTPKKAKKPAAATVTK	40.92	2407.5	23	602.8875	P16403|H12_HUMAN:HIST1H1C
AVKPKAAKPK	38.7	1036.676	10	346.5653	P16403|H12_HUMAN:HIST1H1C
SPKKAKVAKPK	37.16	1180.766	11	394.5951	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKA	37.13	1678.098	16	420.5309	P16403|H12_HUMAN:HIST1H1C
TPKKAKKPAAATVTKK	36.96	1667.046	16	417.7715	P16403|H12_HUMAN:HIST1H1C
AAKPKVVKPKKAAPKK	36.52	1688.119	16	423.0389	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASGSFKLNKK	35.6	2535.439	25	634.8673	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVG	35.15	1166.714	12	584.3636	P16403|H12_HUMAN:HIST1H1C
VAKSPKKAKVAKPKKAAK	35.08	1877.23	18	470.3142	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTG	35.05	1474.835	15	492.6232	P16403|H12_HUMAN:HIST1H1C
KPKVVKPK	35.0	922.6	8	462.3	P16403|H12_HUMAN:HIST1H1

	3	327		235	C
KAGGTKPKKPVGA	34.99	1237.751	13	413.5928	P16403|H12_HUMAN:HIST1H1C
KVAKSPKKAKVAKPKK	34.09	1735.156	16	434.7957	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTKK	33.84	2081.305	20	521.3379	P16403|H12_HUMAN:HIST1H1C
SELITKAVA	33.69	930.5386	9	466.2759	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAA	32.25	1138.755	11	380.592	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPK	32.09	1064.707	10	355.909	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAKA	31.72	1695.088	17	424.7783	P16403|H12_HUMAN:HIST1H1C
KSPKKAKVAKPKKAAK	31.52	1707.125	16	427.7888	P16403|H12_HUMAN:HIST1H1C
KAKVAKPKKAAKSAAK	31.48	1624.051	16	407.0197	P16403|H12_HUMAN:HIST1H1C
SAKKTPKKAKKPAAATVTK	29.8	1953.21	19	489.313	P16403|H12_HUMAN:HIST1H1C
VKKAGGTKPKKPVGAAKKPKK	29.62	2145.384	21	537.3527	P16403|H12_HUMAN:HIST1H1C
PKVVKPKKAAPKKK	28.4	1546.045	14	516.3552	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGA	27.9	1545.873	16	516.3033	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKKAASG	27.17	1579.832	17	527.6187	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAA	26.84	1308.788	14	437.2701	P16403|H12_HUMAN:HIST1H1C
AAKKPKKAAGGATPKKSAKK	26.46	1965.221	20	492.3143	P16403|H12_HUMAN:HIST1H1C
VVKPKKAAPKK	25.26	1192.802	11	398.6073	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAK	23.49	1436.883	15	360.228	P16403|H12_HUMAN:HIST1H1C
KPKKAAKSAAKAVKPKA	23.08	1721.104	17	431.2832	P16403|H12_HUMAN:HIST1H1C
KSAAKAVKPKAAKPK	23.07	1521.972	15	508.3315	P16403|H12_HUMAN:HIST1H1C
KAAKSAAKAVKPKA	22.91	1367.861	14	456.9609	P16403|H12_HUMAN:HIST1H1C

AKKPAAATVTKK	22.08	1212.755	12	405.2621	P16403|H12_HUMAN:HIST1H1C
KAVKPKAAKPK	21.86	1164.771	11	583.3947	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKKAA	21.23	1435.778	15	479.6011	P16403|H12_HUMAN:HIST1H1C
SLVSKGTLVQTKGTGASG	20.91	1689.926	18	564.3168	P16403|H12_HUMAN:HIST1H1C
ATPKKSAKKTPKKAKKPAAATVTKK	20.14	2606.632	25	652.6692	P16403|H12_HUMAN:HIST1H1C
KPKKAAKSAAKAVKPK	18.57	1650.067	16	413.5236	P16403|H12_HUMAN:HIST1H1C
GTGASGSFKLNKKAAS	18.1	1522.81	16	508.6098	P16403|H12_HUMAN:HIST1H1C
KPKAAKPKVVKPK	16.93	1417.95	13	473.6606	P16403|H12_HUMAN:HIST1H1C
KAGGTKPKKPVGAAKKP	15.62	1662.03	17	416.5141	P16403|H12_HUMAN:HIST1H1C
GQFDHPNIIRL	30.37	1308.694	11	437.2378	P29320|EPHA3_HUMAN:P29322|EPHA8_HUMAN:
RNFAKSL	16.11	834.4711	7	418.2435	P48378|RFX2_HUMAN:RFX2
AYLEAHETF	33.11	1079.492	9	540.7537	P57740|NU107_HUMAN:NUP107
NRYERAIYA	24.09	1154.583	9	578.2997	P57740|NU107_HUMAN:NUP107
YYIGKTPKI	35.49	1081.617	9	541.8166	P61812|TGFB2_HUMAN:TGFB2
YYAKEVYKI	31.75	1175.623	9	588.8239	P61812|TGFB2_HUMAN:TGFB2
RKSTGGKAPRKQLATK	44.07	1726.033	16	432.5151	P68431|H31_HUMAN:HIST1H3A
RYRPGTVAL	31.04	1031.588	9	516.8019	P68431|H31_HUMAN:HIST1H3A
ARTKQTARKSTGGKAPRKQLATK	25.57	2482.457	23	621.6201	P68431|H31_HUMAN:HIST1H3A
RYQKSTELL	22.6	1136.619	9	569.3174	P68431|H31_HUMAN:HIST1H3A
YQKSTELLI	17.9	1093.602	9	547.8127	P68431|H31_HUMAN:HIST1H3A
YRPGTVAL	16.99	875.4865	8	438.7502	P68431|H31_HUMAN:HIST1H3A
IMPKDIQLARRIRG	15.9	1665.	14	417.5	P68431|H31_HUMAN:HIST1H3

	8	982		028	A
IYTMIYRNL	17.18	1185.622	9	593.818	Q00987|MDM2_HUMAN:MDM2
TEIEKSISA	28.93	976.5077	9	489.2611	Q02223|TNR17_HUMAN:TNFRSF17
YADNQVMHFY	46.32	1286.539	10	644.2769	Q07864|DPOE1_HUMAN:POLE
IYADNQVMHF	41.74	1236.56	10	619.2866	Q07864|DPOE1_HUMAN:POLE
FYVDTVRAF	31.34	1116.56	9	559.2872	Q07864|DPOE1_HUMAN:POLE
LYMKYVHPF	31.26	1196.605	9	399.8759	Q07864|DPOE1_HUMAN:POLE
LEREGLTAA	22.83	958.5083	9	480.2653	Q13438|OS9_HUMAN:OS9
KRFEIQKY	17.37	1110.619	8	556.3177	Q14774|HLX_HUMAN:HLX
LSDDHDEKY	29.64	1120.467	9	374.4968	Q16526|CRY1_HUMAN:CRY1
YTDPEVFKY	32.7	1160.539	9	581.2763	Q16647|PTGIS_HUMAN:PTGIS
KYFDKVVTL	22.77	1111.628	9	556.8221	Q2M1P5|KIF7_HUMAN:KIF7
FMEKETPLRY	37.41	1312.649	10	438.5561	Q2PZI1|D19L1_HUMAN:DPY19L1P1
YTDENQHQY	41.61	1196.473	9	599.2433	Q5HYA8|MKS3_HUMAN:TMEM67
RSEEGRLYY	27.43	1171.562	9	586.7886	Q5PSV4|BRM1L_HUMAN:BRMS1L
TDVTPRPLY	17.23	1060.555	9	531.2852	Q5VTH9|WDR78_HUMAN:WDR78
TRSTIVYEI	26.13	1080.582	9	541.3027	Q5XXA6|ANO1_HUMAN:ANO1
PYDPALGSPSRLF	30.64	1418.719	13	710.3676	Q6BEB4|SP5_HUMAN:SP5
VYGKTSHL	25.07	903.4814	8	452.7483	Q6BEB4|SP5_HUMAN:SP5
TVTPVEVHI	16.6	993.5495	9	497.7825	Q6Q759|SPG17_HUMAN:SPAG17
TSDVLSWNEVE	16.21	1277.578	11	639.7968	Q6Q759|SPG17_HUMAN:SPAG17
TIDEVKGTIY	19	1137.592	10	569.8026	Q6V0I7|FAT4_HUMAN:FAT4

YEWEPFAEVKGN	38.93	1467.667	12	734.8411	Q6ZMP0|THSD4_HUMAN:THSD4
KGKTLQKRKPK	16.85	1310.851	11	656.4351	Q6ZN17|LN28B_HUMAN:LIN28B
YSIECTCNFL	19.27	1191.494	10	596.7585	Q86VV8|RTTN_HUMAN:RTTN
IYVPSSYHL	18.3	1077.549	9	539.7858	Q8IX21|SLF2_HUMAN:SLF2
EETTSHLQRI	15.22	1212.61	10	405.2088	Q8N3T6|T132C_HUMAN:TMEM132C
SYLIIHQRI	35.02	1141.661	9	571.8383	Q8NF99|ZN397_HUMAN:ZNF397
KAYECSECGK	18.02	1116.458	10	559.2382	Q8NF99|ZN397_HUMAN:ZNF397
KEELVDEPFLDLGKS	17.66	1717.877	15	573.6342	Q8NI35|INADL_HUMAN:PATJ
HYSTLVHMF	40.66	1133.533	9	567.7743	Q96DT5|DYH11_HUMAN:DNAH11
TAATTWLTW	17.97	1049.518	9	525.7668	Q96DT5|DYH11_HUMAN:DNAH11
KPVWNDLNPK	15.29	1209.651	10	605.8356	Q96DT5|DYH11_HUMAN:DNAH11
KQSPPIVKI	31.81	1008.633	9	505.3244	Q96GZ6|S41A3_HUMAN:SLC41A3
YRHKDSRYL	30.11	1236.636	9	413.2195	Q96GZ6|S41A3_HUMAN:SLC41A3
NTDFGDSGRLYY	35.09	1406.61	12	704.3139	Q96JM2|ZN462_HUMAN:ZNF462
NTDFGDSGRLY	16.45	1243.547	11	622.7809	Q96JM2|ZN462_HUMAN:ZNF462
IRANATNKL	32.02	999.5825	9	334.2014	Q9BV19|CA050_HUMAN:C1orf50
HRTTGPRFL	30.79	1083.594	9	362.2044	Q9H4F8|SMOC1_HUMAN:SMOC1
LYIDEAHSI	28.55	1059.524	9	530.7691	Q9NUV7|SPTC3_HUMAN:SPTLC3
LYMRIRDNW	16.09	1265.634	9	422.8856	Q9NUV7|SPTC3_HUMAN:SPTLC3
IYMSTFAWTL	18.64	1231.595	10	616.8036	Q9NYQ6|CELR1_HUMAN:CELSR1
IPYLDDDIPLL	16.16	1285.681	11	643.8489	Q9P2K9|DISP3_HUMAN:DISP3
DENEDLRKL	16.6	1130.	9	377.8	Q9UGC7|RF1ML_HUMAN:MTR

	9	557		608	F1L
GIYKKAQHPQKTHH	15.55	1671.896	14	418.981	Q9ULW2|FZD10_HUMAN:FZD10
STLFTVLTYL	19.92	1156.638	10	579.3326	Q9UP38|FZD1_HUMAN:FZD1
RQADLKNWSL	24.64	1229.652	10	615.8337	Q9UPW0|FOXJ3_HUMAN:FOXJ3
GVCQCHEDFMS	19.03	1254.447	11	628.234	Q9UQP3|TENN_HUMAN:TNN
LRFNKTVRV	25.16	1131.688	9	378.2368	Q9Y289|SC5A6_HUMAN:SLC5A6
NTQKPTGIADVY	43.16	1305.657	12	653.8364	Q9Y6H5|SNCAP_HUMAN:SNCAIP

Table 12. List of selected lncRNA peptides in ten human cervical tumour samples

Peptide	ALC	Length	lnckbID	chr	idents	Samples
LLSASKVLL	62.77778	9	lnckb.9710	chr11	DISC1FP1	CT_7
LLSASKVLL	62.77778	9	lnckb.9710	chr11	DISC1FP1	CT_7
LPYLKWTL	71.75	8	lnckb.44609	chr5	CTC-340A15.2	CT_7
VLLSLFLEHL	83.3	10	lnckb.14510	chr13	LOC101927284	CT_7
HLLCARTLK	77.44444	9	lnckb.31640	chr2	lnckb.31640	CT_7
SLSLLSLKV	71.88889	9	lnckb.58303	chrX	LOC107985664	CT_1
LSSTLNKQL	64.33333	9	lnckb.44609	chr5	CTC-340A15.2	CT_1
KVHVFLVKK	99.66667	9	lnckb.45851	chr6	HLA-F-AS1	CT_1
KVHVFLVKK	99.66667	9	lnckb.45851	chr6	HLA-F-AS1	CT_1
MRSLALTQV	63.66667	9	lnckb.10057	chr11	LOC107987166	CT_1
MRSLALTQV	63.66667	9	lnckb.10057	chr11	LOC107987166	CT_1
LRKELEERL	97.6666	9	lnckb.3164	chr2	lnckb.31640	CT_1

	7		0			
LRKELEERL	97.66667	9	lnckb.31640	chr2	lnckb.31640	CT_1
LRKELEERL	97.66667	9	lnckb.31640	chr2	lnckb.31640	CT_1
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	89.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	80	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5

LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	75.55556	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	84.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.2222	9	lnckb.4460	chr5	CTC-340A15.2	CT_5

	2		9			
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	82.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	83.66667	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5

LLYFCLHKL	81.22222	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LLYFCLHKL	79.88889	9	lnckb.44609	chr5	CTC-340A15.2	CT_5
LPHGELPDTSA	77.18182	11	lnckb.31346	chr2	LOC107985960	CT_5
LPHGELPDTSA	77.18182	11	lnckb.31346	chr2	LOC107985960	CT_5
VLFSGLRSL	82.55556	9	lnckb.14510	chr13	LOC101927284	CT_5
LPHGELPDTSA	74.81818	11	lnckb.31346	chr2	LOC107985960	CT_5
LPHGELPDTSA	74.81818	11	lnckb.31346	chr2	LOC107985960	CT_5
TRHSPPGVL	70.11111	9	lnckb.10057	chr11	LOC107987166	CT_4
MTMSTLLSKK	99.2	10	lnckb.42318	chr5	MIR4458HG	CT_4
MTMSTLLSKK	99.2	10	lnckb.42318	chr5	MIR4458HG	CT_4
MTMSTLLSKK	99.2	10	lnckb.42318	chr5	MIR4458HG	CT_4
MTMSTLLSKK	99.2	10	lnckb.42318	chr5	MIR4458HG	CT_4
MTMSTLLSKK	99.2	10	lnckb.42318	chr5	MIR4458HG	CT_4
MTMSTLLSKK	99.2	10	lnckb.4231	chr5	MIR4458HG	CT_4

			8			
MTMSTLLSKK	99.2	10	lnckb.42318	chr5	MIR4458HG	CT_4
LPNFARLL	69.5	8	lnckb.31346	chr2	LOC107985960	CT_4
LPNFARLL	69.5	8	lnckb.31346	chr2	LOC107985960	CT_4
LPNFARLL	69.5	8	lnckb.31346	chr2	LOC107985960	CT_4
LPNFARLL	69.5	8	lnckb.31346	chr2	LOC107985960	CT_4
LPNFARLL	65.75	8	lnckb.31346	chr2	LOC107985960	CT_4
LPNFARLL	65.75	8	lnckb.31346	chr2	LOC107985960	CT_4
LPNFARLL	65.75	8	lnckb.31346	chr2	LOC107985960	CT_4
LPNFARLL	65.75	8	lnckb.31346	chr2	LOC107985960	CT_4
SASSLFLYK	83.55556	9	lnckb.36918	chr3	LOC101928135	CT_4
SASSLFLYK	83.55556	9	lnckb.36918	chr3	LOC101928135	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
LLYFCLHKL	86.44444	9	lnckb.44609	chr5	CTC-340A15.2	CT_4
KQLNKQLLK	72.33333	9	lnckb.53745	chr8	LOC105375911	CT_4

KQLNKQLLK	72.33333	9	lnckb.53745	chr8	LOC105375911	CT_4
KQLNKQLLK	94.44444	9	lnckb.53745	chr8	LOC105375911	CT_9
KQLNKQLLK	94.44444	9	lnckb.53745	chr8	LOC105375911	CT_9
KLGSLTTPK	86.44444	9	lnckb.58303	chrX	LOC107985664	CT_9
MRSLALTQV	66.66667	9	lnckb.10057	chr11	LOC107987166	CT_9
MRSLALTQV	66.66667	9	lnckb.10057	chr11	LOC107987166	CT_9
LLSASKVLL	71.77778	9	lnckb.9710	chr11	DISC1FP1	CT_9
LLSASKVLL	71.77778	9	lnckb.9710	chr11	DISC1FP1	CT_9
MTMSTLLSKK	99.4	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.4	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.4	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.4	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.4	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.4	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.4	10	lnckb.42318	chr5	MIR4458HG	CT_9
SASSLFLYK	91.22222	9	lnckb.36918	chr3	LOC101928135	CT_9
SASSLFLYK	91.22222	9	lnckb.36918	chr3	LOC101928135	CT_9
RQAPALRHL	68.11111	9	lnckb.14510	chr13	LOC101927284	CT_9
MTMSTLLSKK	99.3	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.3	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.3	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.3	10	lnckb.4231	chr5	MIR4458HG	CT_9

			8			
MTMSTLLSKK	99.3	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.3	10	lnckb.42318	chr5	MIR4458HG	CT_9
MTMSTLLSKK	99.3	10	lnckb.42318	chr5	MIR4458HG	CT_9
ALRAVTLTAK	77.1	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	77.1	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	77.1	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	77.1	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	88.5	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	88.5	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	88.5	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	88.5	10	lnckb.36918	chr3	LOC101928135	CT_8
MTMSTLLSKK	99.8	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.8	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.8	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.8	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.8	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.8	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.8	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.9	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.9	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.9	10	lnckb.42318	chr5	MIR4458HG	CT_8

MTMSTLLSKK	95.9	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.9	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.9	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.9	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSK	98	9	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSK	98	9	lnckb.42318	chr5	MIR4458HG	CT_8
VTLFQNRVK	66.33333	9	lnckb.53745	chr8	LOC105375911	CT_8
LPNFARLL	61.25	8	lnckb.31346	chr2	LOC107985960	CT_8
LPNFARLL	61.25	8	lnckb.31346	chr2	LOC107985960	CT_8
LPNFARLL	61.25	8	lnckb.31346	chr2	LOC107985960	CT_8
LPNFARLL	61.25	8	lnckb.31346	chr2	LOC107985960	CT_8
ALRAVTLTAK	69.6	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	69.6	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	69.6	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	69.6	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	92.2	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	92.2	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	92.2	10	lnckb.36918	chr3	LOC101928135	CT_8
ALRAVTLTAK	92.2	10	lnckb.36918	chr3	LOC101928135	CT_8
MTMSTLLSKK	99.7	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.7	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.7	10	lnckb.4231	chr5	MIR4458HG	CT_8

			8			
MTMSTLLSKK	99.7	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.7	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.7	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	99.7	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.5	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.5	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.5	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.5	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.5	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.5	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSKK	95.5	10	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSK	97.33333	9	lnckb.42318	chr5	MIR4458HG	CT_8
MTMSTLLSK	97.33333	9	lnckb.42318	chr5	MIR4458HG	CT_8
LPNFARLL	64	8	lnckb.31346	chr2	LOC107985960	CT_8
LPNFARLL	64	8	lnckb.31346	chr2	LOC107985960	CT_8
LPNFARLL	64	8	lnckb.31346	chr2	LOC107985960	CT_8
LPNFARLL	64	8	lnckb.31346	chr2	LOC107985960	CT_8
KTKLLTGSK	74.88889	9	lnckb.58303	chrX	LOC107985664	CT_8
KVHVFLVKK	98.66667	9	lnckb.45851	chr6	HLA-F-AS1	CT_8
KVHVFLVKK	98.66667	9	lnckb.45851	chr6	HLA-F-AS1	CT_8
LRKELEERL	98.55556	9	lnckb.31640	chr2	lnckb.31640	CT_10

LRKELEERL	98.55556	9	lnckb.31640	chr2	lnckb.31640	CT_10
LRKELEERL	98.55556	9	lnckb.31640	chr2	lnckb.31640	CT_10
LRKELEERL	98.22222	9	lnckb.31640	chr2	lnckb.31640	CT_10
LRKELEERL	98.22222	9	lnckb.31640	chr2	lnckb.31640	CT_10
LRKELEERL	98.22222	9	lnckb.31640	chr2	lnckb.31640	CT_10
LRNVKLYLL	63.11111	9	lnckb.45851	chr6	HLA-F-AS1	CT_10
TELKAWKL	83.125	8	lnckb.9710	chr11	DISC1FP1	CT_3

Table.13. List of peptides shared RAVAN antigens HIST1H1C identified in five human cancer cell lines and ten human cervical tumours

Name	Score	Mass	Length	Antigen
C33				
AKPKVVKPKKAAPKKK	48.02	1745.18	16	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	36.48	1816.21	17	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	27.07	1617.08	15	P16403|H12_HUMAN
AKAVKPKAAKPKVVKPKKAAPKKK	19.86	2538.69	24	P16403|H12_HUMAN
AASKERSGVSL	46.3	1103.59	11	P16403|H12_HUMAN
KERSGVSL	29.03	874.487	8	P16403|H12_HUMAN
Caski				
AKPKVVKPKKAAPKKK	54.52	1745.18	16	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	54.36	1790.13	18	P16403|H12_HUMAN
KPKKAAKSAAKAVKPKA	50.94	1721.1	17	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	47.71	1707.12	16	P16403|H12_HUMAN

AAKPKVVKPKKAAPKKK	47.36	1816.21	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	46.07	1918.22	19	P16403|H12_HUMAN
KSPKKAKVAKPKKAAKSA	45.93	1865.19	18	P16403|H12_HUMAN
AKPKVVKPKKAAPK	44.82	1488.99	14	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	44.17	1617.08	15	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	41.98	1579.03	15	P16403|H12_HUMAN
SAAKAVKPK	40.89	898.56	9	P16403|H12_HUMAN
VVKPKKAAPK	39.34	1064.71	10	P16403|H12_HUMAN
KSPKKAKVAKPKKA	33.96	1507.99	14	P16403|H12_HUMAN
KSPKKAKVAKPK	32.27	1308.86	12	P16403|H12_HUMAN
VVKPKKAAPKKK	28.04	1320.9	12	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	23.98	1674.14	15	P16403|H12_HUMAN
Dotc2				
KLGLKSLVSK	37.36	1071.7	10	P16403|H12_HUMAN
GTGASGSFK	38.6	810.387	9	P16403|H12_HUMAN
ASGSFKLNK	38.86	950.519	9	P16403|H12_HUMAN
Hela				
AAATVTKKV	38.78	887.544	9	P16403|H12_HUMAN
KAAKPKVV	35.4	839.559	8	P16403|H12_HUMAN
AKKPAAATV	28.54	855.518	9	P16403|H12_HUMAN
SKERSGVSL	32.76	961.519	9	P16403|H12_HUMAN
KERSGVSL	19	874.487	8	P16403|H12_HUMAN
KKALAAAGY	41.55	891.51	9	P16403|H12_HUMA

		8		N
Siha				
AKPKVVKPKKAAPKKK	50.05	1745.18	16	P16403|H12_HUMAN
AKPKVVKPKKAAPK	47.15	1488.99	14	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	45.63	1816.21	17	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	42.36	1617.08	15	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	19.96	1674.14	15	P16403|H12_HUMAN
ASGPPVSEL	35.63	855.434	9	P16403|H12_HUMAN
KERSGVSLAAL	30.89	1129.65	11	P16403|H12_HUMAN
KERSGVSL	28.63	874.487	8	P16403|H12_HUMAN
KERSGVSLA	28.51	945.524	9	P16403|H12_HUMAN
KERSGVSLAA	24.16	1016.56	10	P16403|H12_HUMAN
AALKKALAA	34.21	855.554	9	P16403|H12_HUMAN
SELITKAV	32.44	859.502	8	P16403|H12_HUMAN
SELITKAVA	23.58	930.539	9	P16403|H12_HUMAN
1				
KAGGTKPKKPVGAAKKPKKAAGG	72.41	2174.34	23	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAG	70.39	2117.32	22	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTK	66.58	2478.54	24	P16403|H12_HUMAN
SGSFKLNKKAASGEAKPKV	65.92	1946.09	19	P16403|H12_HUMAN
GATPKKSAKKTPKKAKKPAAATVTK	65.29	2535.56	25	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTKK	59.87	2081.3	20	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	58.06	2017.29	20	P16403|H12_HUMAN

KAGGTKPKKPVGAAKKPKKAA	54.92	2060.29	21	P16403|H12_HUMAN
KKTPKKAKKPAAATVTKK	54.73	1923.24	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	54.1	1918.22	19	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	53.83	1579.03	15	P16403|H12_HUMAN
KKAASGEAKPKVKKAGGTKPK	53.79	2108.28	21	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPK	53.51	1662.03	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	53.36	1790.13	18	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	51.89	1953.21	19	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	51.75	1495.96	15	P16403|H12_HUMAN
KTPKKAKKPAAATVTKK	50.92	1795.14	17	P16403|H12_HUMAN
KKAAKSAAKAVKPK	50.89	1424.92	14	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	50.7	2145.38	21	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKA	50.04	1806.19	17	P16403|H12_HUMAN
VAKSPKKAKVAKPK	49.56	1478.97	14	P16403|H12_HUMAN
SPKKAKVAKPKKAA	49.5	1450.93	14	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	49.09	1745.18	16	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	48.65	1607.06	15	P16403|H12_HUMAN
VAKSPKKAKVAKPKKA	48.56	1678.1	16	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	48.31	1667.05	16	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	48.08	1749.14	17	P16403|H12_HUMAN
KAAKSAAKAVKPK	48	1296.82	13	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	47.93	1877.2	18	P16403|H12_HUMA

		3		N
TKPKKPVGAAKKPK	47.35	1476.95	14	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAG	47.27	1989.22	21	P16403|H12_HUMAN
KPKAAKPKVVKPKKAAPKKK	45.94	2169.46	20	P16403|H12_HUMAN
KKPAAATVTKKVAKSPK	45.65	1752.1	17	P16403|H12_HUMAN
KSAAKAVKPKAAKPK	44.78	1521.97	15	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	44.64	1816.21	17	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	43.3	1624.05	16	P16403|H12_HUMAN
SAAKAVKPKAAKPKVVKPKKAAPKK	43.26	2568.67	25	P16403|H12_HUMAN
KAKVAKPKKAAKSAAKA	43.21	1695.09	17	P16403|H12_HUMAN
TKPKKPVGAAKKPKK	42.76	1605.05	15	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	42.51	1617.08	15	P16403|H12_HUMAN
KSPKKAKVAKPK	41.77	1308.86	12	P16403|H12_HUMAN
KAKKPAAATVTKK	40.69	1340.85	13	P16403|H12_HUMAN
KAAKSAAKAVKPKA	40.58	1367.86	14	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKP	40.13	1662.03	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	39.99	1436.88	15	P16403|H12_HUMAN
KAGGTKPKKPVG	39.94	1166.71	12	P16403|H12_HUMAN
AKSPKKAKVAKPK	39.79	1379.9	13	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	38.41	1674.14	15	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	37.94	1707.12	16	P16403|H12_HUMAN
KAGGTKPKKPVGA	37.07	1237.75	13	P16403|H12_HUMAN

VAKSPKKAKVAKPKKAAK	35.68	1877.23	18	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	35.38	1588.06	15	P16403|H12_HUMAN
AKKPAAATVTKK	33.97	1212.76	12	P16403|H12_HUMAN
AKKPAAATVTKKVAKSPKKAKV	33.97	2249.43	22	P16403|H12_HUMAN
KSPKKAKVAKPKKA	33.96	1507.99	14	P16403|H12_HUMAN
VVKPKKAAPK	33.16	1064.71	10	P16403|H12_HUMAN
AAKPKVVKPKKAAPKK	32.83	1688.12	16	P16403|H12_HUMAN
KPKVVKPK	32.58	922.633	8	P16403|H12_HUMAN
PKVVKPKKAAPKK	32.49	1417.95	13	P16403|H12_HUMAN
VVKPKKAAPKKK	31.71	1320.9	12	P16403|H12_HUMAN
PKVVKPKKAAPKKK	31.31	1546.04	14	P16403|H12_HUMAN
VVKPKKAAPKK	30.83	1192.8	11	P16403|H12_HUMAN
KAAKPKVVKPKKAAPK	30.38	1688.12	16	P16403|H12_HUMAN
AKPKVVKPK	29.66	993.67	9	P16403|H12_HUMAN
KAAGGATPKKSAKKTPKKA	29.53	1867.14	19	P16403|H12_HUMAN
KPKVVKPKKAAPKK	29.49	1546.04	14	P16403|H12_HUMAN
KAGGTKPKKPVGAA	29.43	1308.79	14	P16403|H12_HUMAN
PKKAAKSAAKAVKPKAAKP	27.93	1889.19	19	P16403|H12_HUMAN
KVAKSPKKAKVAKPKK	27.18	1735.16	16	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKK	25.93	1790.13	18	P16403|H12_HUMAN
KSPKKAKVAKP	24.72	1180.77	11	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKA	24.64	1989.2	20	P16403|H12_HUMA

		6		N
AKVAKPKKAAKSAAK	24.39	1495.96	15	P16403|H12_HUMAN
SFKLNKKAASGEAKPKVK	21.11	1930.14	18	P16403|H12_HUMAN
KKPAAATVTK	21.05	1013.62	10	P16403|H12_HUMAN
KKTPKKAKKPAAATVTK	19.16	1795.14	17	P16403|H12_HUMAN
KAKKPAAATVTK	16.27	1212.76	12	P16403|H12_HUMAN
KAAKSAAKAVK	15.57	1071.68	11	P16403|H12_HUMAN
AKPKKAAKSAAKAVKPKAAK	15.22	1991.27	20	P16403|H12_HUMAN
AVKPKAAKPK	26.02	1036.68	10	P16403|H12_HUMAN
KAVKPKAAKPK	25.64	1164.77	11	P16403|H12_HUMAN
SRIKLGLKSLVSK	47.2	1427.92	13	P16403|H12_HUMAN
GTLVQTKGTGASGSFKLNK	59.65	1893.03	19	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLNK	57.29	2407.34	24	P16403|H12_HUMAN
SLVSKGTLVQTKG	53.2	1316.77	13	P16403|H12_HUMAN
SLVSKGTLVQTK	50.81	1259.74	12	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASG	46.89	1689.93	18	P16403|H12_HUMAN
GTGASGSFKLNK	46.07	1165.61	12	P16403|H12_HUMAN
GASGSFKLNK	39.33	1007.54	10	P16403|H12_HUMAN
SLVSKGTLVQTKGTG	38.48	1474.84	15	P16403|H12_HUMAN
GTLVQTKGTGASGSFK	38.02	1537.81	16	P16403|H12_HUMAN
ASGSFKLNK	33.22	950.519	9	P16403|H12_HUMAN
SLVSKGTLVQT	27.29	1131.65	11	P16403|H12_HUMAN

SGVSLAALKK	39.51	972.597	10	P16403|H12_HUMAN
ASGPPVSELITK	38.27	1197.66	12	P16403|H12_HUMAN
2				
GAAKKPKKAAGGATPKKSAKKTPK	84.44	2348.44	24	P16403|H12_HUMAN
GTGASGSFKLNKKAASGEAKPKVK	73.64	2360.32	24	P16403|H12_HUMAN
SGSFKLNKKAASGEAKPKVK	69.06	2074.19	20	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAAK	66.19	1936.23	19	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	65.17	1918.22	19	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	64.31	1953.21	19	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	62.55	1790.13	18	P16403|H12_HUMAN
TKKVAKSPKKAKVAKPKKAAK	59.27	2234.47	21	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTKK	59.18	2081.3	20	P16403|H12_HUMAN
KTPKKAKKPAAATVTKK	58.74	1795.14	17	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	57.92	2017.29	20	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAKSAAK	57.14	2234.43	22	P16403|H12_HUMAN
KAKVAKPKKAAKSAAKAVKPK	55.18	2147.4	21	P16403|H12_HUMAN
KPVGAAKKPKKAAGGATPK	53.67	1804.1	19	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAK	53.25	1877.23	18	P16403|H12_HUMAN
KKAAKSAAKAVKPK	52.71	1424.92	14	P16403|H12_HUMAN
VAKSPKKAKVAKPK	52.25	1478.97	14	P16403|H12_HUMAN
TPKKAKKPAAATVTKK	51.58	1667.05	16	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	51.55	1495.96	15	P16403|H12_HUMAN

AGGTKPKKPVGAAKKPK	50.83	1662.03	17	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAK	50.55	2005.33	19	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	50.48	1607.06	15	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	50.35	1667.05	16	P16403|H12_HUMAN
KAAGGATPKKSAKKTPKKAKKPA	50.14	2291.42	23	P16403|H12_HUMAN
TKKVAKSPKKAKVAKPKK	49.48	1964.3	18	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	48.95	1745.18	16	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	48.36	1624.05	16	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	48.07	2145.38	21	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	47.78	1707.12	16	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	46.63	1436.88	15	P16403|H12_HUMAN
KAAKSAAKAVKPK	46.48	1296.82	13	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	46.44	1816.21	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKA	45.65	1989.26	20	P16403|H12_HUMAN
TKKVAKSPKKAKVAKPKKAA	44.96	2106.37	20	P16403|H12_HUMAN
KAKKPAAATVTKK	44.83	1340.85	13	P16403|H12_HUMAN
AKSPKKAKVAKPKKAAK	44.33	1778.16	17	P16403|H12_HUMAN
GAAKKPKKAAGGATPK	43.91	1479.89	16	P16403|H12_HUMAN
SAAKAVKPKAAKPK	43.45	1393.88	14	P16403|H12_HUMAN
AVKPKAAKPKVVKPKKAAPKKK	41.66	2339.56	22	P16403|H12_HUMAN
KAVKPKAAKPKVVKPKKAAPKKK	41.03	2467.66	23	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	41.01	1674.1	15	P16403|H12_HUMA

		4		N
KPKAAKPKVVKPKKAAPKKK	40.18	2169.46	20	P16403|H12_HUMAN
SAAKAVKPKAAKPKVVKPK	39.58	1945.26	19	P16403|H12_HUMAN
VVKPKKAAPKKK	39.51	1320.9	12	P16403|H12_HUMAN
AVKPKAAKPKVVKPKKAAPK	39.4	2083.37	20	P16403|H12_HUMAN
SFKLNKKAASGEAKPKVK	39.31	1930.14	18	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	37.89	1579.03	15	P16403|H12_HUMAN
KVAKSPKKAKVAKPKK	36.78	1735.16	16	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	35.95	1588.06	15	P16403|H12_HUMAN
KAGGTKPKKPV	35.32	1109.69	11	P16403|H12_HUMAN
VVKPKKAAPK	31.64	1064.71	10	P16403|H12_HUMAN
KAGGTKPKKPVG	30.6	1166.71	12	P16403|H12_HUMAN
AKSPKKAKVAKPK	30.43	1379.9	13	P16403|H12_HUMAN
KAAKPKVVKPKKAAPK	29.91	1688.12	16	P16403|H12_HUMAN
PKVVKPKKAAPKKK	29.73	1546.04	14	P16403|H12_HUMAN
KAVKPKAAKPKVVKPKKAAPK	28.42	2211.47	21	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	27.6	1749.14	17	P16403|H12_HUMAN
KPKVVKPK	26.87	922.633	8	P16403|H12_HUMAN
KAAKPKVVKPKKAAPKK	26.38	1816.21	17	P16403|H12_HUMAN
KAGGTKPKKPVGA	25.58	1237.75	13	P16403|H12_HUMAN
AAKPKVVKPKKAAPKK	25.51	1688.12	16	P16403|H12_HUMAN
KAGGTPRKASGPPVSELI	25.17	1763.99	18	P16403|H12_HUMAN

VVKPKKAAPKK	24.02	1192.8	11	P16403|H12_HUMAN
KSAKKTPKKAKKPAAATV	20.44	1852.16	18	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	20.17	1617.08	15	P16403|H12_HUMAN
PAAATVTKKVAKSPKKAKVAK	17.93	2121.34	21	P16403|H12_HUMAN
PKKAAKSAAK	17.08	998.624	10	P16403|H12_HUMAN
VKKAGGTKPKKPV	16.68	1336.86	13	P16403|H12_HUMAN
3				
KAGGTKPKKPVGAAKKPKKAAG	71.62	2117.32	22	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAGG	70.34	2174.34	23	P16403|H12_HUMAN
KSAKKTPKKAKKPAAATVTKK	62.01	2209.4	21	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTKK	61.63	2081.3	20	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAKSAAK	60.5	2234.43	22	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAKSAAK	58.99	2362.53	23	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAG	57.1	1989.22	21	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAAK	55.91	1936.23	19	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	55.71	1790.13	18	P16403|H12_HUMAN
GATPKKSAKKTPKKAKKPAAATVTK	53.25	2535.56	25	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	51.96	1918.22	19	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAK	51.61	1877.23	18	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAK	49.45	2005.33	19	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	48.99	2017.29	20	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	48.85	1667.05	16	P16403|H12_HUMAN

KTPKKAKKPAAATVTKK	47.29	1795.14	17	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	47.04	1607.06	15	P16403|H12_HUMAN
KKTPKKAKKPAAATVTKK	46.91	1923.24	18	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	46.91	1877.23	18	P16403|H12_HUMAN
TKKVAKSPKKAKVAKPKKAAK	46.26	2234.47	21	P16403|H12_HUMAN
KAAKSAAKAVKPK	46.24	1296.82	13	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	46.2	1436.88	15	P16403|H12_HUMAN
KKTPKKAKKPAAATVTK	45.7	1795.14	17	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	45.38	1816.21	17	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	44.88	1745.18	16	P16403|H12_HUMAN
AVKPKAAKPKVVKPKKAAPKKK	44.13	2339.56	22	P16403|H12_HUMAN
VAKSPKKAKVAKPK	43.99	1478.97	14	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAGG	43.66	2046.24	22	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	43.41	1749.14	17	P16403|H12_HUMAN
KKAAKSAAKAVKPK	43.14	1424.92	14	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAGGA	42.84	2245.37	24	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTK	41.46	2478.54	24	P16403|H12_HUMAN
KAGGTKPKKPVG	40.71	1166.71	12	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	40.35	1624.05	16	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAKSAA	39.95	2234.43	22	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	39.67	1674.14	15	P16403|H12_HUMAN
KKVAKSPKKAKVAKPK	37.27	1735.1	16	P16403|H12_HUMA

		6		N
KVAKSPKKAKVAKPKK	36.79	1735.16	16	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	35.5	2145.38	21	P16403|H12_HUMAN
KAGGTKPKKPVGA	35.49	1237.75	13	P16403|H12_HUMAN
KSPKKAKVAKPKKAAKSAAK	35.38	2064.33	20	P16403|H12_HUMAN
AKSPKKAKVAKPKKAAK	31.96	1778.16	17	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTKK	29.29	2606.63	25	P16403|H12_HUMAN
SFKLNKKAASGEAKPKVK	28.01	1930.14	18	P16403|H12_HUMAN
KAAKPKVVKPKKAAPKK	26.96	1816.21	17	P16403|H12_HUMAN
KVVKPKKAAPKK	25.21	1320.9	12	P16403|H12_HUMAN
KKPAAATVTKK	24.13	1141.72	11	P16403|H12_HUMAN
PKVVKPKKAAPKKK	23.33	1546.04	14	P16403|H12_HUMAN
AKPKKAAKSAAKAVKPKAAK	21.3	1991.27	20	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	19.73	1617.08	15	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	17.39	1953.21	19	P16403|H12_HUMAN
KPVGAAKKPKKAAG	17.38	1349.85	14	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKA	17.38	1989.26	20	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	17.2	1707.12	16	P16403|H12_HUMAN
KKAAKSAAK	16.69	901.571	9	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	15.95	1495.96	15	P16403|H12_HUMAN
KAAKSAAKAVKPKAAKPK	15.51	1792.14	18	P16403|H12_HUMAN
AVKPKAAKPK	32.45	1036.68	10	P16403|H12_HUMAN

SELITKAV	31.2	859.502	8	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLNK	59.93	2407.34	24	P16403|H12_HUMAN
SLVSKGTLVQTK	50.85	1259.74	12	P16403|H12_HUMAN
SLVSKGTLVQTKG	47.87	1316.77	13	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASG	45.31	1689.93	18	P16403|H12_HUMAN
GTGASGSFKLNK	44.3	1165.61	12	P16403|H12_HUMAN
SLVSKGTLVQTKGTG	33.58	1474.84	15	P16403|H12_HUMAN
SLVSKGTLVQTKGTGA	27.6	1545.87	16	P16403|H12_HUMAN
GTGASGSFKLNKK	16.47	1293.7	13	P16403|H12_HUMAN
KERSGVSL	20.53	874.487	8	P16403|H12_HUMAN
AASKERSGVSL	18.34	1103.59	11	P16403|H12_HUMAN
4				
KAGGTKPKKPVGAAKKPKKAAGG	70.11	2174.34	23	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAG	64.7	2117.32	22	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	57.29	2017.29	20	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	56.46	1918.22	19	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPK	56.44	1662.03	17	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAAK	53.78	1936.23	19	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	53.48	1579.03	15	P16403|H12_HUMAN
KKTPKKAKKPAAATVTKK	53.07	1923.24	18	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKA	52.49	1806.19	17	P16403|H12_HUMAN
KKAAKSAAKAVKPK	52.38	1424.92	14	P16403|H12_HUMAN

KTPKKAKKPAAATVTKK	51.51	1795.14	17	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	50.97	1953.21	19	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	50.52	1749.14	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	50.32	1790.13	18	P16403|H12_HUMAN
SPKKAKVAKPKKAA	49.77	1450.93	14	P16403|H12_HUMAN
VAKSPKKAKVAKPK	49.3	1478.97	14	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	49.13	1667.05	16	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	48.86	1877.23	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAA	47.07	2060.29	21	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	46.39	1607.06	15	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	45.66	2145.38	21	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	45.44	1745.18	16	P16403|H12_HUMAN
KPVGAAKKPKKAAGGATPK	44.28	1804.1	19	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	42.94	1495.96	15	P16403|H12_HUMAN
KAAKSAAKAVKPK	42.6	1296.82	13	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	42.6	1674.14	15	P16403|H12_HUMAN
KSPKKAKVAKPK	42.58	1308.86	12	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	42.43	1816.21	17	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	41.58	1707.12	16	P16403|H12_HUMAN
TPKKAKKPAAATVTKK	40.97	1667.05	16	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	40.5	1588.06	15	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAA	40.48	1808.1	18	P16403|H12_HUMA

		4		N
KAAKSAAKAVKPKA	39.58	1367.86	14	P16403|H12_HUMAN
SPKKAKVAKPK	38.56	1180.77	11	P16403|H12_HUMAN
SGSFKLNKKAASGEAKPKV	38.21	1946.09	19	P16403|H12_HUMAN
VVKPKKAAPK	38.07	1064.71	10	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	37.82	1436.88	15	P16403|H12_HUMAN
KVAKSPKKAKVAKPKK	37.8	1735.16	16	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	37.24	1617.08	15	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	36.69	1624.05	16	P16403|H12_HUMAN
KPKVVKPK	36.55	922.633	8	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTKK	35.27	2606.63	25	P16403|H12_HUMAN
KAGGTKPKKPVG	34.47	1166.71	12	P16403|H12_HUMAN
VVKPKKAAPKK	33.32	1192.8	11	P16403|H12_HUMAN
AKSPKKAKVAKPK	33.05	1379.9	13	P16403|H12_HUMAN
KAKVAKPKKAAKSAAKA	32.57	1695.09	17	P16403|H12_HUMAN
KAKVAKPKKAA	31.82	1138.75	11	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTK	31.05	2478.54	24	P16403|H12_HUMAN
KSAAKAVKPKAAKPK	30.09	1521.97	15	P16403|H12_HUMAN
KSPKKAKVAKPKKAAKSAAK	29.89	2064.33	20	P16403|H12_HUMAN
KKTPKKAKKPAAATVTK	29.04	1795.14	17	P16403|H12_HUMAN
PKVVKPKKAAPKK	28.92	1417.95	13	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAGG	27.79	2046.24	22	P16403|H12_HUMAN

KSPKKAKVAKPKKAAKSAA	26.18	1936.23	19	P16403|H12_HUMAN
PKVVKPKKAAPKKK	25.8	1546.04	14	P16403|H12_HUMAN
SAAKAVKPKAAKPKVVKPKKAAPKK	25.4	2568.67	25	P16403|H12_HUMAN
AKSAAKAVKPKAAKPK	25.1	1593.01	16	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKP	25.06	1662.03	17	P16403|H12_HUMAN
KKPAAATVTKK	24.83	1141.72	11	P16403|H12_HUMAN
TKPKKPVGAAKKPKK	24.04	1605.05	15	P16403|H12_HUMAN
PKKAAKSAAKAVKPKAAKPK	23.29	2017.29	20	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAK	22.79	1877.23	18	P16403|H12_HUMAN
KKAAKSAAKAVKPKAAKPK	21.98	1920.24	19	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAG	21.89	1989.22	21	P16403|H12_HUMAN
KAKKPAAATVTK	21.4	1212.76	12	P16403|H12_HUMAN
VVKPKKAAPKKK	17.8	1320.9	12	P16403|H12_HUMAN
KKPAAATVTKKVAK	17.37	1439.92	14	P16403|H12_HUMAN
KAAKSAAKAVKPKAAKPK	16.81	1792.14	18	P16403|H12_HUMAN
KAAKPKVVKPKK	16.06	1320.9	12	P16403|H12_HUMAN
KPKAAKPKVVKPKKAAPK	15.93	1913.27	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKA	15.53	1989.26	20	P16403|H12_HUMAN
5				
KAGGTKPKKPVGAAKKPKKAAG	70.47	2117.32	22	P16403|H12_HUMAN
KPVGAAKKPKKAAGGATPKKSA	65.85	2090.27	22	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAA	65.23	1808.14	18	P16403|H12_HUMAN

AGGTKPKKPVGAAKKPK	62.78	1662.03	17	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	59.89	2017.29	20	P16403|H12_HUMAN
SPKKAKVAKPKKAA	59.84	1450.93	14	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	59.31	2145.38	21	P16403|H12_HUMAN
KKTPKKAKKPAAATVTKK	58.94	1923.24	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAGG	58.57	2174.34	23	P16403|H12_HUMAN
KTPKKAKKPAAATVTKK	56.93	1795.14	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	56.56	1790.13	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	55.38	1918.22	19	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	54.79	1579.03	15	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	54.78	1877.23	18	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAAK	53.31	1936.23	19	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKA	53.14	1806.19	17	P16403|H12_HUMAN
KPAAATVTKKVA	52.17	1183.73	12	P16403|H12_HUMAN
KKAAKSAAKAVKPK	51.83	1424.92	14	P16403|H12_HUMAN
VAKSPKKAKVAKPK	50.15	1478.97	14	P16403|H12_HUMAN
KPVGAAKKPKKAAGGATPKK	50.01	1932.2	20	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	50.01	1607.06	15	P16403|H12_HUMAN
TPKKAKKPAAATVTKK	49.94	1667.05	16	P16403|H12_HUMAN
KPVGAAKKPKKAAGGATPK	49.16	1804.1	19	P16403|H12_HUMAN
KKAGGTKPKKPVGAAKKPK	48.69	1918.22	19	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	48.4	1667.0	16	P16403|H12_HUMA

		5		N
AKPKVVKPKKAAPKKK	47.71	1745.18	16	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAKSAAK	47.55	2362.53	23	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	47.5	1707.12	16	P16403|H12_HUMAN
SPKKAKVAKPK	47.05	1180.77	11	P16403|H12_HUMAN
KSPKKAKVAKPK	46.74	1308.86	12	P16403|H12_HUMAN
TKPKKPVGAAKKPK	46.54	1476.95	14	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	45.89	1617.08	15	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	45.07	1436.88	15	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	45	1495.96	15	P16403|H12_HUMAN
VAKPKKAAKSAA	44.72	1168.73	12	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	44.34	1588.06	15	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	44.28	1816.21	17	P16403|H12_HUMAN
KKAASGEAKPKVKKAGGTKPK	43.69	2108.28	21	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTK	42.87	2478.54	24	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKK	42.43	1790.13	18	P16403|H12_HUMAN
KSAAKAVKPKAAKPK	41.03	1521.97	15	P16403|H12_HUMAN
KAKVAKPKKAA	40.51	1138.75	11	P16403|H12_HUMAN
KAAKSAAKAVKPK	39.76	1296.82	13	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAK	39.21	1664.05	17	P16403|H12_HUMAN
KSPKKAKVAKPKKA	36.8	1507.99	14	P16403|H12_HUMAN
VVKPKKAAPK	34.7	1064.71	10	P16403|H12_HUMAN

KVAKSPKKAKVAKPKK	34.45	1735.16	16	P16403|H12_HUMAN
KKPAAATVTKKVAKSPK	34.33	1752.1	17	P16403|H12_HUMAN
KAGGTKPKKPVG	33.61	1166.71	12	P16403|H12_HUMAN
VVKPKKAAPKKK	32.51	1320.9	12	P16403|H12_HUMAN
KPKVVKPK	32.44	922.633	8	P16403|H12_HUMAN
KSPKKAKVAKP	32.2	1180.77	11	P16403|H12_HUMAN
KPKKPVGAA	31.98	894.565	9	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	31.34	1624.05	16	P16403|H12_HUMAN
PKVVKPKKAAPKK	31.18	1417.95	13	P16403|H12_HUMAN
VVKPKKAAPKK	31.12	1192.8	11	P16403|H12_HUMAN
KAGGTKPKKPVGA	30.85	1237.75	13	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	30.66	1953.21	19	P16403|H12_HUMAN
SPKKAKVAKPKKAAK	29.99	1579.03	15	P16403|H12_HUMAN
VAKSPKKAKVAKPKK	28.71	1607.06	15	P16403|H12_HUMAN
GATPKKSAKKTPKKAKKPAAATVTK	27.76	2535.56	25	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	27.09	1749.14	17	P16403|H12_HUMAN
VKPKKAAPKKK	25.96	1221.83	11	P16403|H12_HUMAN
KVVKPKKAAPKK	25.46	1320.9	12	P16403|H12_HUMAN
PKVVKPK	24.48	794.538	7	P16403|H12_HUMAN
KAVKPKAAKPKVVKP	24.13	1588.06	15	P16403|H12_HUMAN
SAAKAVKPKAAKPKVVKPK	22.37	1945.26	19	P16403|H12_HUMAN
KKPAAATVTKKVAK	20.89	1439.9	14	P16403|H12_HUMA

		2		N
KVAKPKKAAKSAA	20.72	1296.82	13	P16403|H12_HUMAN
KKVAKSPKKAKVAKPK	20.43	1735.16	16	P16403|H12_HUMAN
KPKKAAKSAAKAVKPKA	17.88	1721.1	17	P16403|H12_HUMAN
TGASGSFKLNKKAASGEAKPKVK	17.03	2303.3	23	P16403|H12_HUMAN
6				
KAGGTKPKKPVGAAKKPKKAAGG	67.62	2174.34	23	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	63.85	1918.22	19	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAG	61.3	2117.32	22	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAKSAAK	55.79	2362.53	23	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPK	54.92	1662.03	17	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	54.56	2017.29	20	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAAK	53.71	1936.23	19	P16403|H12_HUMAN
TPKKAKKPAAATVTKK	53.2	1667.05	16	P16403|H12_HUMAN
SPKKAKVAKPKKAA	52.6	1450.93	14	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	51.81	1790.13	18	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	51.69	1877.23	18	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	51.59	1579.03	15	P16403|H12_HUMAN
GATPKKSAKKTPKKAKKPAAATVTK	51.38	2535.56	25	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAA	51.26	2060.29	21	P16403|H12_HUMAN
KKTPKKAKKPAAATVTKK	50.41	1923.24	18	P16403|H12_HUMAN
KKAAKSAAKAVKPK	50.22	1424.92	14	P16403|H12_HUMAN
KTPKKAKKPAAATVTKK	49.56	1795.1	17	P16403|H12_HUMA

		4		N
KVAKSPKKAKVAKPKKAAKSAA	49.09	2234.43	22	P16403|H12_HUMAN
VAKSPKKAKVAKPK	48.78	1478.97	14	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTK	48.72	2478.54	24	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	47.97	1607.06	15	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKA	47.59	1806.19	17	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	47.55	1667.05	16	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	46.89	2145.38	21	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	46.82	1707.12	16	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	46.65	1745.18	16	P16403|H12_HUMAN
KAAKSAAKAVKPK	45.8	1296.82	13	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	45.33	1749.14	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	45.03	1436.88	15	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	44.42	1495.96	15	P16403|H12_HUMAN
KSPKKAKVAKPK	44.28	1308.86	12	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	43.79	1816.21	17	P16403|H12_HUMAN
KPVGAAKKPKKAAGGATPK	43.17	1804.1	19	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAG	42.39	1989.22	21	P16403|H12_HUMAN
SGSFKLNKKAASGEAKPKV	41.97	1946.09	19	P16403|H12_HUMAN
KKAASGEAKPKVKKAGGTKPK	41.37	2108.28	21	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAK	41.26	1877.23	18	P16403|H12_HUMAN
KAGGTKPKKPVG	40.18	1166.71	12	P16403|H12_HUMAN

ATPKKSAKKTPKKAKKPAAATVTKK	40.06	2606.63	25	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	39.27	1624.05	16	P16403|H12_HUMAN
SPKKAKVAKPK	39.06	1180.77	11	P16403|H12_HUMAN
KVAKSPKKAKVAKPKK	38.72	1735.16	16	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	37.94	1674.14	15	P16403|H12_HUMAN
KSAAKAVKPKAAKPK	37.92	1521.97	15	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAGGA	37.75	2245.37	24	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKK	37.48	1790.13	18	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	37.09	1588.06	15	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	36.18	1617.08	15	P16403|H12_HUMAN
AKVAKPKKAAKSAAK	35.92	1495.96	15	P16403|H12_HUMAN
VVKPKKAAPKKK	35.44	1320.9	12	P16403|H12_HUMAN
TKPKKPVGAAKKPKK	35.17	1605.05	15	P16403|H12_HUMAN
KKTPKKAKKPAAATVTK	34.3	1795.14	17	P16403|H12_HUMAN
VVKPKKAAPKK	33.31	1192.8	11	P16403|H12_HUMAN
KAAKPKVVKPKKAAPK	32.59	1688.12	16	P16403|H12_HUMAN
TGASGSFKLNKKAASGEAKPKVK	32.51	2303.3	23	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTKK	32.36	2081.3	20	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	30.81	1953.21	19	P16403|H12_HUMAN
KAKVAKPKKAA	29.54	1138.75	11	P16403|H12_HUMAN
KAAKPKVVKPKKAAPKK	29.39	1816.21	17	P16403|H12_HUMAN
PKVVKPKKAAPKKK	28.85	1546.0	14	P16403|H12_HUMA

		4		N
VVKPKKAAPK	27.77	1064.71	10	P16403|H12_HUMAN
AGGTKPKKPVG	27.71	1038.62	11	P16403|H12_HUMAN
VAKSPKKAKVAKPKK	26.63	1607.06	15	P16403|H12_HUMAN
KAGGTKPKKPVGA	26.04	1237.75	13	P16403|H12_HUMAN
KSPKKAKVAKP	25.55	1180.77	11	P16403|H12_HUMAN
PKKAAKSAAKAVKPKAAKPK	24.22	2017.29	20	P16403|H12_HUMAN
AAKPKVVKPKKAAPKK	24.21	1688.12	16	P16403|H12_HUMAN
AKPKKAAKSAAKAVKPKAAK	23.93	1991.27	20	P16403|H12_HUMAN
KAAGGATPKKSAKKTPKKA	23.23	1867.14	19	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAGG	22.6	2046.24	22	P16403|H12_HUMAN
KAAKPKVVKPKK	21.51	1320.9	12	P16403|H12_HUMAN
KKAGGTKPKKPVGAAKKPK	19.51	1918.22	19	P16403|H12_HUMAN
AAKKPKKAAGGATPKKSAKK	19.49	1965.22	20	P16403|H12_HUMAN
KSPKKAKVAKPKKAAKSAAK	17.54	2064.33	20	P16403|H12_HUMAN
KAAKPKVVKPK	16.27	1192.8	11	P16403|H12_HUMAN
KPKAAKPKVVKPKK	15.31	1546.04	14	P16403|H12_HUMAN
AKAVKPKAAKPK	30.05	1235.81	12	P16403|H12_HUMAN
SGSFKLNKKAASGEAKPK	46.1	1847.03	18	P16403|H12_HUMAN
AVKPKAAKPK	32.86	1036.68	10	P16403|H12_HUMAN
SRIKLGLKSLVSK	38.34	1427.92	13	P16403|H12_HUMAN
SELITKAV	25.83	859.502	8	P16403|H12_HUMAN

SLVSKGTLVQTKGTGASGSFKLNK	59.3	2407.34	24	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLNKK	48.1	2535.44	25	P16403|H12_HUMAN
SLVSKGTLVQTKG	44.86	1316.77	13	P16403|H12_HUMAN
SLVSKGTLVQTK	44.26	1259.74	12	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASG	39.76	1689.93	18	P16403|H12_HUMAN
SLVSKGTLVQTKGTG	34.2	1474.84	15	P16403|H12_HUMAN
GTGASGSFKLNK	25.28	1165.61	12	P16403|H12_HUMAN
7				
KAGGTKPKKPVGAAKKPKKAAGG	82.38	2174.34	23	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAG	72.37	2117.32	22	P16403|H12_HUMAN
KKTPKKAKKPAAATVTKK	60.03	1923.24	18	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAA	58.65	1808.14	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAA	58.2	2060.29	21	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAAK	56.59	1936.23	19	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTKK	55.77	2081.3	20	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	55.62	1790.13	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	55.45	1918.22	19	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAKSAAK	54.05	2362.53	23	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	53.53	1667.05	16	P16403|H12_HUMAN
TKKVAKSPKKAKVAKPKKAA	52.97	2106.37	20	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAKSAAK	51.57	2234.43	22	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKK	51.37	1790.13	18	P16403|H12_HUMAN

AGGTKPKKPVGAAKKPKKAAG	51.35	1989.22	21	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPK	51.27	1662.03	17	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAKSAA	51.21	2234.43	22	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	50.14	1607.06	15	P16403|H12_HUMAN
SAAKAVKPKAAKPKVVKPKKAAPKK	50.03	2568.67	25	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	49.39	1953.21	19	P16403|H12_HUMAN
SPKKAKVAKPKKAA	49.24	1450.93	14	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	49.18	2017.29	20	P16403|H12_HUMAN
KAKVAKPKKAAKSAAKAVKPK	49.02	2147.4	21	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAK	48.96	1877.23	18	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	48.3	2145.38	21	P16403|H12_HUMAN
KTPKKAKKPAAATVTKK	48.09	1795.14	17	P16403|H12_HUMAN
KKAAKSAAKAVKPK	47.95	1424.92	14	P16403|H12_HUMAN
VAKSPKKAKVAKPK	47.62	1478.97	14	P16403|H12_HUMAN
TPKKAKKPAAATVTKK	47.3	1667.05	16	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	46.34	1745.18	16	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	46.16	1495.96	15	P16403|H12_HUMAN
TKKVAKSPKKAKVAKPKKAAK	46.11	2234.47	21	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	45.88	1436.88	15	P16403|H12_HUMAN
TKKVAKSPKKAKVAKPKK	45.36	1964.3	18	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	45.13	1579.03	15	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	44.85	1877.2	18	P16403|H12_HUMA

		3		N
VAKSPKKAKVAKPKKAA	44.63	1749.14	17	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAAK	44.06	2005.33	19	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	43.15	1707.12	16	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	41.06	1816.21	17	P16403|H12_HUMAN
AVKPKAAKPKVVKPKKAAPKKK	40.62	2339.56	22	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPKKAAGG	40.43	2046.24	22	P16403|H12_HUMAN
KAKKPAAATVTK	39.96	1212.76	12	P16403|H12_HUMAN
KPVGAAKKPKKAAGGATPK	39.9	1804.1	19	P16403|H12_HUMAN
KAAKSAAKAVKPK	39.22	1296.82	13	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTK	39.22	2478.54	24	P16403|H12_HUMAN
KAKKPAAATVTKK	39.17	1340.85	13	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTKK	38.88	2606.63	25	P16403|H12_HUMAN
KAAGGATPKKSAKKTPKKA	38.16	1867.14	19	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	38.14	1624.05	16	P16403|H12_HUMAN
SPKKAKVAKPK	37.76	1180.77	11	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	37.7	1674.14	15	P16403|H12_HUMAN
AAKPKVVKPKKAAPKK	37.41	1688.12	16	P16403|H12_HUMAN
KAASGEAKPKVKKAG	37.07	1468.87	15	P16403|H12_HUMAN
KAGGTKPKKPVG	37	1166.71	12	P16403|H12_HUMAN
KAAKSAAKAVKPKA	36.83	1367.86	14	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	36.04	1588.06	15	P16403|H12_HUMAN

KSAAKAVKPKAAKPK	35.94	1521.97	15	P16403|H12_HUMAN
KVAKSPKKAKVAKPKK	35.29	1735.16	16	P16403|H12_HUMAN
KAGGTKPKKPVGA	34.85	1237.75	13	P16403|H12_HUMAN
VVKPKKAAPK	34.29	1064.71	10	P16403|H12_HUMAN
KAKVAKPKKAAKSAAKA	33.92	1695.09	17	P16403|H12_HUMAN
VVKPKKAAPKKK	33.75	1320.9	12	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	33.58	1617.08	15	P16403|H12_HUMAN
TGASGSFKLNKKAASGEAKPKVK	32.86	2303.3	23	P16403|H12_HUMAN
KKPKKAAGGATPKKSAKKTPK	31.65	2149.34	21	P16403|H12_HUMAN
VVKPKKAAPKK	29.97	1192.8	11	P16403|H12_HUMAN
KKPAAATVTKK	29.52	1141.72	11	P16403|H12_HUMAN
KAAKPKVVKPKKAAPKK	29.44	1816.21	17	P16403|H12_HUMAN
KAGGTKPKKPV	26.11	1109.69	11	P16403|H12_HUMAN
KPVGAAKKPKKAAGG	25.49	1406.87	15	P16403|H12_HUMAN
SGSFKLNKKAASGEAKPKVK	25.28	2074.19	20	P16403|H12_HUMAN
VKPKKAAPK	25.14	965.639	9	P16403|H12_HUMAN
PKVVKPKKAAPKKK	22.5	1546.04	14	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKA	22.34	1989.26	20	P16403|H12_HUMAN
KAAKPKVVKPKK	17.28	1320.9	12	P16403|H12_HUMAN
AKSPKKAKVAKPK	17.1	1379.9	13	P16403|H12_HUMAN
KKTPKKAKKPAAATVTK	16.12	1795.14	17	P16403|H12_HUMAN
KPVGAAKKPKKAAG	15.42	1349.8	14	P16403|H12_HUMA

		5		N
KAASGEAKPKVKKAGGTK	15.19	1755.04	18	P16403|H12_HUMAN
KVAKPKKAAKSAA	15.07	1296.82	13	P16403|H12_HUMAN
KPKAAKPKVVKPKKAA	15.02	1688.12	16	P16403|H12_HUMAN
AKAVKPKAAKPK	19.01	1235.81	12	P16403|H12_HUMAN
AVKPKAAKPK	39.18	1036.68	10	P16403|H12_HUMAN
VKPKAAKPK	34.39	965.639	9	P16403|H12_HUMAN
TGASGSFKLNKKAASGEAKPK	31.83	2076.13	21	P16403|H12_HUMAN
SGSFKLNKKAASGEAKPK	20.97	1847.03	18	P16403|H12_HUMAN
KAVKPKAAKP	15.23	1036.68	10	P16403|H12_HUMAN
KALAAAGYDVEKNNSRIKLGLK	48.98	2358.34	22	P16403|H12_HUMAN
KPKAAKPK	30.45	866.57	8	P16403|H12_HUMAN
PKAAKPK	20.46	738.475	7	P16403|H12_HUMAN
GTGASGSFKLNK	45.6	1165.61	12	P16403|H12_HUMAN
GTLVQTKGTGASGSFK	45.28	1537.81	16	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLN	25.54	2279.25	23	P16403|H12_HUMAN
8				
KAGGTKPKKPVGAAKKPK	57.56	1790.13	18	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	53.84	2017.29	20	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	52.3	1918.22	19	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPK	51.23	1662.03	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAGG	49.87	2174.34	23	P16403|H12_HUMAN
KAKKPAAATVTK	47.07	1212.7	12	P16403|H12_HUMA

		6		N
VAKSPKKAKVAKPK	46.59	1478.97	14	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	46.07	1607.06	15	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	45.09	1745.18	16	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAG	44.94	2117.32	22	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	43.36	1816.21	17	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	42.83	1624.05	16	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	42.53	1877.23	18	P16403|H12_HUMAN
VVKPKKAAPK	41.11	1064.71	10	P16403|H12_HUMAN
GTKPKKPVGAAK	40.97	1180.73	12	P16403|H12_HUMAN
AAKPKVVKPKKAAPKK	39.13	1688.12	16	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	39	1588.06	15	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	38.88	1579.03	15	P16403|H12_HUMAN
KAAKSAAKAVKPK	38.87	1296.82	13	P16403|H12_HUMAN
KSPKKAKVAKPK	38.49	1308.86	12	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	37.68	2145.38	21	P16403|H12_HUMAN
VVKPKKAAPKKK	36.87	1320.9	12	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	36.29	1436.88	15	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	35.58	1617.08	15	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	34.86	1707.12	16	P16403|H12_HUMAN
VKPKKAAPK	33.27	965.639	9	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	31.06	1667.05	16	P16403|H12_HUMAN

KVAKSPKKAKVAKPKK	30.7	1735.16	16	P16403|H12_HUMAN
VVKPKKAAPKK	29.45	1192.8	11	P16403|H12_HUMAN
SPKKAKVAKPK	28.06	1180.77	11	P16403|H12_HUMAN
KPKVVKPKKAAPKKK	27.49	1674.14	15	P16403|H12_HUMAN
AVKPKAAKPKVVKPKKAAPKK	24.03	2211.47	21	P16403|H12_HUMAN
KAGGTKPKKPVGAA	22.12	1308.79	14	P16403|H12_HUMAN
KSAAKAVKPKAAKPK	21.78	1521.97	15	P16403|H12_HUMAN
AKPKVVKPK	20.55	993.67	9	P16403|H12_HUMAN
KKPAAATVTK	17.48	1013.62	10	P16403|H12_HUMAN
KAAKPKVVKPKK	15.71	1320.9	12	P16403|H12_HUMAN
AVKPKAAKPK	43.95	1036.68	10	P16403|H12_HUMAN
VKPKAAKPK	34.71	965.639	9	P16403|H12_HUMAN
KALAAAGYDVEKNNSRIKLGLK	63.44	2358.34	22	P16403|H12_HUMAN
KASGPPVSELITK	50.61	1325.76	13	P16403|H12_HUMAN
SGPPVSELITK	48.52	1126.62	11	P16403|H12_HUMAN
ALAAAGYDVEKNNSRIKLGLK	44.28	2230.24	21	P16403|H12_HUMAN
AVAASKERSGVSLAALK	43.97	1656.95	17	P16403|H12_HUMAN
AASKERSGVSL	43.87	1103.59	11	P16403|H12_HUMAN
SGVSLAALKK	43.79	972.597	10	P16403|H12_HUMAN
ASGPPVSELITK	42.44	1197.66	12	P16403|H12_HUMAN
SGVSLAALK	36.09	844.502	9	P16403|H12_HUMAN
ASGPPVSELITKAVAASK	33.55	1724.9	18	P16403|H12_HUMA

		7		N
GLKSLVSKGTLVQTKGTGASGSFK	64.96	2350.32	24	P16403|H12_HUMAN
GLKSLVSKGTLVQTK	53.91	1557.95	15	P16403|H12_HUMAN
RIKLGLKSLVSK	49.65	1340.89	12	P16403|H12_HUMAN
KLGLKSLVSKGTLVQTK	49.44	1799.12	17	P16403|H12_HUMAN
KLGLKSLVSK	35.29	1071.7	10	P16403|H12_HUMAN
KPKAAKPK	29.63	866.57	8	P16403|H12_HUMAN
LGLKSLVSK	25.71	943.607	9	P16403|H12_HUMAN
PKAAKPK	21.18	738.475	7	P16403|H12_HUMAN
NNSRIKLGL	15.62	1013.6	9	P16403|H12_HUMAN
GTLVQTKGTGASGSFKLNK	68.14	1893.03	19	P16403|H12_HUMAN
GTLVQTKGTGASGSFK	61.94	1537.81	16	P16403|H12_HUMAN
VQTKGTGASGSFKLNK	60.94	1621.88	16	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLNK	59.96	2407.34	24	P16403|H12_HUMAN
LVQTKGTGASGSFKLNK	56.6	1734.96	17	P16403|H12_HUMAN
SLVSKGTLVQTK	52.94	1259.74	12	P16403|H12_HUMAN
GTGASGSFKLNK	52.67	1165.61	12	P16403|H12_HUMAN
GASGSFKLNK	47.29	1007.54	10	P16403|H12_HUMAN
GTGASGSFKLNKK	46.79	1293.7	13	P16403|H12_HUMAN
GTGASGSFKLN	43.54	1037.51	11	P16403|H12_HUMAN
VSKGTLVQTK	42.07	1059.63	10	P16403|H12_HUMAN
ASGSFKLNK	41.17	950.519	9	P16403|H12_HUMAN

TGASGSFKLNK	39.7	1108.59	11	P16403|H12_HUMAN
SGSFKLNK	38.64	879.481	8	P16403|H12_HUMAN
ALAAAGYDVEK	43.3	1106.56	11	P16403|H12_HUMAN
GVSLAALKK	31.93	885.565	9	P16403|H12_HUMAN
KPKKAAPK	30.25	866.57	8	P16403|H12_HUMAN
PKKAAPK	21.69	738.475	7	P16403|H12_HUMAN
9				
KAGGTKPKKPVGAAKKPKK	58.63	1918.22	19	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPK	57.99	1662.03	17	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	57.28	1667.05	16	P16403|H12_HUMAN
KTPKKAKKPAAATVTKK	55.29	1795.14	17	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	54.08	1877.23	18	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	52.26	2017.29	20	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	51.03	1749.14	17	P16403|H12_HUMAN
VAKSPKKAKVAKPK	50.19	1478.97	14	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKA	49.37	1806.19	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	48.22	1790.13	18	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	47.63	1579.03	15	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	47.62	1816.21	17	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	47.55	1436.88	15	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	47.36	2145.38	21	P16403|H12_HUMAN
SPKKAKVAKPKKAAKSAAK	47.33	1936.23	19	P16403|H12_HUMAN

KVAKSPKKAKVAKPK	46.22	1607.06	15	P16403|H12_HUMAN
KKAAKSAAKAVKPK	46.09	1424.92	14	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	45.32	1745.18	16	P16403|H12_HUMAN
GTKPKKPVGAAK	45.14	1180.73	12	P16403|H12_HUMAN
KAAKSAAKAVKPK	44.33	1296.82	13	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	44.21	1707.12	16	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	43.06	1588.06	15	P16403|H12_HUMAN
AAKPKVVKPKKAAPKK	42.76	1688.12	16	P16403|H12_HUMAN
KAKKPAAATVTK	42.6	1212.76	12	P16403|H12_HUMAN
SPKKAKVAKPKKAA	41.82	1450.93	14	P16403|H12_HUMAN
KSPKKAKVAKPK	40.6	1308.86	12	P16403|H12_HUMAN
SAAKAVKPKAAKPK	39.95	1393.88	14	P16403|H12_HUMAN
VVKPKKAAPK	39.3	1064.71	10	P16403|H12_HUMAN
KAKKPAAATVTKK	39.18	1340.85	13	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	38.6	1617.08	15	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	38.3	1495.96	15	P16403|H12_HUMAN
KVAKSPKKAKVAKPKK	36.15	1735.16	16	P16403|H12_HUMAN
VVKPKKAAPKKK	36.11	1320.9	12	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	35.15	1624.05	16	P16403|H12_HUMAN
KPKVVKPK	34.66	922.633	8	P16403|H12_HUMAN
VKPKKAAPK	31.49	965.639	9	P16403|H12_HUMAN
VVKPKKAAPKK	30	1192.8	11	P16403|H12_HUMA

				N
KSAAKAVKPKAAKPK	29.67	1521.97	15	P16403|H12_HUMAN
SPKKAKVAKPK	29.41	1180.77	11	P16403|H12_HUMAN
AAKPKVVKPK	27.85	1064.71	10	P16403|H12_HUMAN
AVKPKAAKPKVVKPKKAAPKKK	25.91	2339.56	22	P16403|H12_HUMAN
KAGGTKPKKPVG	22.05	1166.71	12	P16403|H12_HUMAN
KKVAKSPKKAKVAKPK	18.02	1735.16	16	P16403|H12_HUMAN
KPKKAAKSAAKAVKPKAAKP	15.65	2017.29	20	P16403|H12_HUMAN
GGTKPKKPVGAAKKPKK	15.34	1719.09	17	P16403|H12_HUMAN
AKAVKPKAAKPK	37.87	1235.81	12	P16403|H12_HUMAN
AVKPKAAKPK	43.43	1036.68	10	P16403|H12_HUMAN
VKPKAAKPK	36.15	965.639	9	P16403|H12_HUMAN
KAASGEAKPK	17.87	985.556	10	P16403|H12_HUMAN
KALAAAGYDVEKNNSRIKLGLK	56.53	2358.34	22	P16403|H12_HUMAN
ALAAAGYDVEKNNSRIKLGLK	45.37	2230.24	21	P16403|H12_HUMAN
AVAASKERSGVSLAALK	43.61	1656.95	17	P16403|H12_HUMAN
GLKSLVSKGTLVQTK	45.06	1557.95	15	P16403|H12_HUMAN
RIKLGLKSLVSK	40.39	1340.89	12	P16403|H12_HUMAN
KLGLKSLVSK	35.66	1071.7	10	P16403|H12_HUMAN
LGLKSLVSK	31.69	943.607	9	P16403|H12_HUMAN
KPKAAKPK	28.3	866.57	8	P16403|H12_HUMAN
PKAAKPK	22.06	738.475	7	P16403|H12_HUMAN

KLGLKSLV	15.08	856.575	8	P16403|H12_HUMAN
GTLVQTKGTGASGSFKLNK	70.99	1893.03	19	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLNK	66.24	2407.34	24	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLN	63.42	2279.25	23	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLNKK	62.61	2535.44	25	P16403|H12_HUMAN
GTLVQTKGTGASGSFK	58.31	1537.81	16	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFK	53.4	2052.12	21	P16403|H12_HUMAN
SLVSKGTLVQTK	50.5	1259.74	12	P16403|H12_HUMAN
GTGASGSFKLNK	47.97	1165.61	12	P16403|H12_HUMAN
VSKGTLVQTK	45.17	1059.63	10	P16403|H12_HUMAN
GASGSFKLNK	43.47	1007.54	10	P16403|H12_HUMAN
LVQTKGTGASGSFKLNK	41.17	1734.96	17	P16403|H12_HUMAN
ASGSFKLNK	40.5	950.519	9	P16403|H12_HUMAN
GTGASGSFKLN	38.07	1037.51	11	P16403|H12_HUMAN
GTGASGSFKLNKK	36.47	1293.7	13	P16403|H12_HUMAN
SLVSKGTLVQT	34.72	1131.65	11	P16403|H12_HUMAN
SGSFKLNK	33.39	879.481	8	P16403|H12_HUMAN
KSLVSKGTLVQTK	33.32	1387.84	13	P16403|H12_HUMAN
GTLVQTKGTGASGSFKLNKK	28.64	2021.13	20	P16403|H12_HUMAN
ALAAAGYDVEKNNSRIK	34.75	1818.96	17	P16403|H12_HUMAN
GVSLAALKK	26.56	885.565	9	P16403|H12_HUMAN
KPKKAAPK	28.11	866.57	8	P16403|H12_HUMA

				N
PKKAAPK	22.4	738.475	7	P16403|H12_HUMAN
10				
KAGGTKPKKPVGAAKKPKKAAG	60.31	2117.32	22	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPK	58.92	1790.13	18	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAGG	55.17	2174.34	23	P16403|H12_HUMAN
GATPKKSAKKTPKKAKKPAAATVTK	54.74	2535.56	25	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTK	54.38	2478.54	24	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKK	54.34	1918.22	19	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPK	53.78	2017.29	20	P16403|H12_HUMAN
AGGTKPKKPVGAAKKPK	53.12	1662.03	17	P16403|H12_HUMAN
KSAKKTPKKAKKPAAATVTK	53.04	2081.3	20	P16403|H12_HUMAN
KSPKKAKVAKPKKAA	52.81	1579.03	15	P16403|H12_HUMAN
KKTPKKAKKPAAATVTKK	52.21	1923.24	18	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAA	51.79	1749.14	17	P16403|H12_HUMAN
KAKVAKPKKAAKSAA	51.38	1495.96	15	P16403|H12_HUMAN
SPKKAKVAKPKKAA	50.07	1450.93	14	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKAA	49.88	1877.23	18	P16403|H12_HUMAN
VAKSPKKAKVAKPK	49.77	1478.97	14	P16403|H12_HUMAN
AKPKVVKPKKAAPKKK	47.85	1745.18	16	P16403|H12_HUMAN
KVAKSPKKAKVAKPK	47.29	1607.06	15	P16403|H12_HUMAN
KKAAKSAAKAVKPK	47.23	1424.92	14	P16403|H12_HUMAN
KTPKKAKKPAAATVTKK	46.25	1795.1	17	P16403|H12_HUMA

		4		N
KAGGTKPKKPVGAAKKPKKAA	45.62	2060.29	21	P16403|H12_HUMAN
KKPAAATVTKKVAKSPK	45.45	1752.1	17	P16403|H12_HUMAN
AAKPKVVKPKKAAPKKK	45.11	1816.21	17	P16403|H12_HUMAN
AKPKVVKPKKAAPKK	44.92	1617.08	15	P16403|H12_HUMAN
KTPKKAKKPAAATVTK	44.52	1667.05	16	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKPKKAAGGA	44.27	2245.37	24	P16403|H12_HUMAN
KAAKSAAKAVKPK	43.21	1296.82	13	P16403|H12_HUMAN
KVAKSPKKAKVAKPKKA	42.77	1806.19	17	P16403|H12_HUMAN
AVKPKAAKPKVVKPK	41.64	1588.06	15	P16403|H12_HUMAN
KSPKKAKVAKPK	41.28	1308.86	12	P16403|H12_HUMAN
TPKKSAKKTPKKAKKPAAATVTK	40.92	2407.5	23	P16403|H12_HUMAN
SPKKAKVAKPK	37.16	1180.77	11	P16403|H12_HUMAN
VAKSPKKAKVAKPKKA	37.13	1678.1	16	P16403|H12_HUMAN
TPKKAKKPAAATVTKK	36.96	1667.05	16	P16403|H12_HUMAN
AAKPKVVKPKKAAPKK	36.52	1688.12	16	P16403|H12_HUMAN
KAGGTKPKKPVG	35.15	1166.71	12	P16403|H12_HUMAN
VAKSPKKAKVAKPKKAAK	35.08	1877.23	18	P16403|H12_HUMAN
KPKVVKPK	35.03	922.633	8	P16403|H12_HUMAN
KAGGTKPKKPVGA	34.99	1237.75	13	P16403|H12_HUMAN
KVAKSPKKAKVAKPKK	34.09	1735.16	16	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTKK	33.84	2081.3	20	P16403|H12_HUMAN

KAKVAKPKKAA	32.25	1138.75	11	P16403|H12_HUMAN
VVKPKKAAPK	32.09	1064.71	10	P16403|H12_HUMAN
KAKVAKPKKAAKSAAKA	31.72	1695.09	17	P16403|H12_HUMAN
KSPKKAKVAKPKKAAK	31.52	1707.12	16	P16403|H12_HUMAN
KAKVAKPKKAAKSAAK	31.48	1624.05	16	P16403|H12_HUMAN
SAKKTPKKAKKPAAATVTK	29.8	1953.21	19	P16403|H12_HUMAN
VKKAGGTKPKKPVGAAKKPKK	29.62	2145.38	21	P16403|H12_HUMAN
PKVVKPKKAAPKKK	28.4	1546.04	14	P16403|H12_HUMAN
KAGGTKPKKPVGAA	26.84	1308.79	14	P16403|H12_HUMAN
AAKKPKKAAGGATPKKSAKK	26.46	1965.22	20	P16403|H12_HUMAN
VVKPKKAAPKK	25.26	1192.8	11	P16403|H12_HUMAN
KAGGTKPKKPVGAAK	23.49	1436.88	15	P16403|H12_HUMAN
KPKKAAKSAAKAVKPKA	23.08	1721.1	17	P16403|H12_HUMAN
KSAAKAVKPKAAKPK	23.07	1521.97	15	P16403|H12_HUMAN
KAAKSAAKAVKPKA	22.91	1367.86	14	P16403|H12_HUMAN
AKKPAAATVTKK	22.08	1212.76	12	P16403|H12_HUMAN
ATPKKSAKKTPKKAKKPAAATVTKK	20.14	2606.63	25	P16403|H12_HUMAN
KPKKAAKSAAKAVKPK	18.57	1650.07	16	P16403|H12_HUMAN
KPKAAKPKVVKPK	16.93	1417.95	13	P16403|H12_HUMAN
KAGGTKPKKPVGAAKKP	15.62	1662.03	17	P16403|H12_HUMAN
AVKPKAAKPK	38.7	1036.68	10	P16403|H12_HUMAN
KAVKPKAAKPK	21.86	1164.77	11	P16403|H12_HUMA

				N
SELITKAVA	33.69	930.539	9	P16403|H12_HUMAN
GTGASGSFKLNKKAASG	27.17	1579.83	17	P16403|H12_HUMAN
GTGASGSFKLNKKAA	21.23	1435.78	15	P16403|H12_HUMAN
GTGASGSFKLNKKAAS	18.1	1522.81	16	P16403|H12_HUMAN
SLVSKGTLVQTK	45.68	1259.74	12	P16403|H12_HUMAN
SLVSKGTLVQTKG	44.27	1316.77	13	P16403|H12_HUMAN
GTGASGSFKLNK	43.27	1165.61	12	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASGSFKLNKK	35.6	2535.44	25	P16403|H12_HUMAN
SLVSKGTLVQTKGTG	35.05	1474.84	15	P16403|H12_HUMAN
SLVSKGTLVQTKGTGA	27.9	1545.87	16	P16403|H12_HUMAN
SLVSKGTLVQTKGTGASG	20.91	1689.93	18	P16403|H12_HUMAN

Table 14. List of peptides were synthesized by GenScript Biotech (Netherlands) B.V for lncRNA peptides MS matching.

	Peptide	Quantity
1	ALRAVTLTAK	Quantity: 0.2-0.5 mg Purity: Crude
2	HLLCARTIK	Quantity: 0.2-0.5 mg Purity: Crude

3	IISASKVIL	Quantity: 0.2-0.5 mg Purity: Crude
4	IIYFCLHKI	Quantity: 0.2-0.5 mg Purity: Crude
5	IPHGEIPDTSA	Quantity: 0.2-0.5 mg Purity: Crude
6	IRNVKIYLI	Quantity: 0.2-0.5 mg Purity: Crude
7	KLGSITTPK	Quantity: 0.2-0.5 mg Purity: Crude
8	KQLNKQLIK	Quantity: 0.2-0.5 mg Purity: Crude
9	KTKIITGSK	Quantity: 0.2-0.5 mg Purity: Crude
10	KVHVFLVKK	Quantity: 0.2-0.5 mg Purity: Crude
11	LPNFARII	Quantity: 0.2-0.5 mg Purity: Crude
12	LPYIKWTI	Quantity: 0.2-0.5 mg Purity: Crude
13	LRKEIEERI	Quantity: 0.2-0.5 mg Purity: Crude

14	LSSTLNKQI	Quantity: 0.2-0.5 mg Purity: Crude
15	MRSLALTQV	Quantity: 0.2-0.5 mg Purity: Crude
16	MTMSTILSK	Quantity: 0.2-0.5 mg Purity: Crude
17	MTMSTILSKK	Quantity: 0.2-0.5 mg Purity: Crude
18	RQAPALRHL	Quantity: 0.2-0.5 mg Purity: Crude
19	SASSLFIYK	Quantity: 0.2-0.5 mg Purity: Crude
20	SLSILSLKV	Quantity: 0.2-0.5 mg Purity: Crude
21	TELKAWKI	Quantity: 0.2-0.5 mg Purity: Crude
22	TRHSPPGVL	Quantity: 0.2-0.5 mg Purity: Crude
23	VIFSGIRSL	Quantity: 0.2-0.5 mg Purity: Crude
24	VLLSIFIEHL	Quantity: 0.2-0.5 mg Purity: Crude

25	VTIFQNRVK	Quantity: 0.2-0.5 mg Purity: Crude
26	VPNFKIISH	Quantity: 0.2-0.5 mg Purity: Crude

8.0 List of Publications

1. Andreatta, M., Nicastri, A., Peng, X., Hancock, G., Dorrell, L., Ternette, N., Nielsen, M. (2019) MS-Rescue: A Computational Pipeline to Increase the Quality and Yield of Immunopeptidomics Experiments. Proteomics, 2019. 19(4): p. e1800357.

2. Refsgaard, C., Barra, C., Peng, X., Ternette, N., Nielsen, M. (2021). NetMHCphosPan - Pan-specific prediction of MHC class I antigen presentation of phosphorylated ligands. ImmunoInformatics, 1–2 (2021) 100005

3. Parker, R., Tailor, A., Peng, X., Nicastri, A., Zerweck, J., Reimer, U., Wenschuh, H., Schnatbaum, K. and Ternette, N. (2021) The Choice of Search Engine Affects Sequencing Depth and HLA Class I Allele-Specific Peptide Repertoires. Mol Cell Proteomics, 2021. 20: p. 100124.

4. Peng, X; Woodhouse, I; Hancock, G; Parker, R; Muller, J; Salatino, S; Partridge, T; Nicastri, A; Liao, H; Ternette, N; Hellner, K; Dorrell, L. (2021) Beyond E6 and E7: The importance of E1 antigen presentation in human papillomavirus-driven cervical cancer. PLOS Pathogens, In Review