Genetic alterations leading to thymic epithelial malignancies

Spyridon Gennatas

Genomic Medicine
National Heart and Lung Institute
Imperial College London

PhD Thesis
Faculty of Medicine, Imperial College London
2017

Declaration of Originality

The original work presented in this thesis is my own. Anyone, who has contributed to any part of the experiments described or the analysis of the data presented, has been explicitly indicated.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

“The smallest oceans still get big, big waves”

EV, 1994

Acknowledgments

I cannot express enough gratitude to Anne Bowcock for her guidance, support and teaching during my PhD. As my primary supervisor she has also been a role model and an immense source of inspiration and motivation.

I would also like to thank my co-supervisors, Sanjay Popat and Andrew Nicholson for their support and input throughout. I would also like to specifically thank both for making me a part of their teams and departments, allowing me to appreciate and approach my research from a clinical and histopathological viewpoint too.

I would like to thank Amit Mandal for his help with data analysis but also his great efforts to teach and guide me.

I would like to thank the members of the lab, and particularly Anastasia, Karl, Ruth, Adam, Hima, Liz, Phil Gillian and Kenny, who all helped me over the last four years with their advice but also created a really happy working environment that I will miss. I hope that our friendships will last in the years to come.

I would like to thank Bill Cookson and Miriam Moffat for the opportunities they gave me as well as their help and guidance.

I would like to thank Angeles, Jan Lucas, Alex, Harshil and their teams for being extremely helpful, hard-working and supportive.

Finally I would like to thank the people, who have always been there for me and have given me all the strength and faith I have ever needed, my parents, my brother and my friends, whether they were near me or really far away, and to Antigoni for being the most beautiful part of my life.

Abstract

Thymic epithelial tumours (TETs) are rare intrathoracic tumours. The key genetic drivers are yet to be discovered. The aim of this research project was to differentiate different TET subsets at the molecular level and identify potential drivers and immunohistochemical targets for each.

Sanger sequencing was used to assess the frequency of the known mutation in *GTF2I*. The prevalence in our cohort of 141 TETs was 34%, being as high as 90% and 69% in type A and AB thymomas respectively. There were no squamous cell carcinomas (SCCs), neuroendocrine thymic tumours (NETTs) or hyperplasias with the mutation. This is a novel finding that could be of diagnostic value in difficult to diagnose mediastinal masses.

Whole exome sequencing (WES) was performed on seventeen TETs and was analysed with the unpublished 'Cancer Genome Atlas' (TCGA) data. 25% of *GTF2I* mutant type A and AB thymomas also had an oncogenic mutation affecting the MAP kinase pathway. The most commonly mutated gene in carcinomas was *TP53* (30%). A recurrent germline insertion in *HLA-A* was seen across the TET subtypes at a frequency of 27%.

Single nucleotide polymorphism (SNP) genotyping on 143 TETs identified multiple genes with copy number gains in chromosome 1q. Very few recurrent copy number losses were seen. A notable exception was *MUTYH* that was shown to have recurrent, homozygous copy number losses in 18% of all TETs and 25% of SCCs.

Eight immunotherapy and two other emerging targets were targeted with immunohistochemistry in 143 TET tissue microarrays. PD-L1, PD-L2 and B7-H3 were the most positive markers. Overall aggressive TETs demonstrated higher staining scores. Clustering of all TETs according to their co-expression of these markers revealed cases in all subtypes that clustered more with other TET subtypes.

These results require validation and expansion. Whole genome sequencing, expression and methylation studies are also needed.

Table of Contents

Acknowledgments **5**

Abstract **6**

Table of Contents **7**

List of Figures **12**

List of Tables **14**

List of abbreviations **16**

Chapter 1 ..... **19**

Introduction **19**

1.1 The aims of the chapter are: **19**

1.2 Background **19**

1.2.1 Anatomy and function of the normal thymus gland 19

1.2.2. Origin of thymic epithelial tumours (TETs) 20

1.2.3 Epidemiology and presentation of TETs 23

1.3 Pathology **24**

1.3.1 Thymomas 24

1.3.2 Thymic Carcinomas 25

1.3.3 Clarifications from the 4th edition of the WHO criteria 26

1.4 Staging **30**

1.5 Prognosis **33**

1.6 Autoimmune paraneoplastic syndromes (APNS) **34**

1.6.1 Myasthenia Gravis 35

1.7 The immune system as a target for treatment in cancer **36**

1.7.1 The role of immunotherapy in the treatment of solid tumours 36

1.7.1.1 CTLA-4 36

1.7.1.2 PD-1 37

1.7.1.3 PD-L1 and other biomarkers of response to immune checkpoint inhibitors 38

1.7.1.4 The role of PD-L2 39

1.7.2 PD-1/ PD-L1 expression in TETs 40

1.7.3 Antibody clones and kits for PD-L1 staining 45

1.8 Treatment of TETs **46**

1.8.1 Standard of care for patients with TETs 46

1.8.2 Targeted therapies for patients with TETs 47

1.8.3 PD-1 inhibitors in TETs: The story so far 48

1.9 Known molecular changes involved in TET formation **49**

1.9.1 Gene mutations 49

1.9.2 Copy number alterations/aberrations (CNA) 55
1.9.3 DNA methylation 60
1.9.4 Gene expression profiling 60
1.9.5 MicroRNA (miRNA) expression 62
1.10 Conclusions 62
1.11 Hypothesis 63
1.12 Aims 63
Chapter 2 64
Materials & Methods 64
2.1 Materials 64
2.2 Methods 66
2.2.1 Thymic epithelial tissue/ DNA ascertainment: 66
2.2.2 Patient data 66
2.2.3 Histologic assessment of tumour samples 67
2.2.4 Polymerase chain reaction (PCR) 67
2.2.5 Sanger sequencing 69
2.2.6 DNA extraction from FFPEs 70
2.2.7 Simultaneous purification of genomic DNA and total RNA, including miRNA, from frozen tissues 71
2.2.8 DNA quantification and quality control 73
2.2.9 Whole exome sequencing (WES) 74
2.2.9.1 Data analysis 74
2.2.9.2 Variant filtering and identification of potential driver mutations 74
2.2.10 SNP genotyping of TETs 80
2.2.10.1 Preparation of LRR and BAF files 80
2.2.10.2 Segmentation and Somatic Copy number calling 81
2.2.10.3 Custom thresholding on ASCAT copy number output 83
2.2.10.4 Gene annotation of Copy number segments and recurrent aberrations 83
2.2.11 Immunohistochemistry (IHC) 84
2.2.11.1 Construction of Tissue Microarray Blocks (TMAs) 84
2.2.11.2 Microtomy 85
2.2.11.3 IHC staining 85
2.2.11.4 Scoring 88
2.2.11.5 Statistical analyses 88
Chapter 3 89
TET samples and clinical data 89
3.1 The aims of this chapter were the following: 89

3.2 Clinicopathological characteristics **89**
3.3 Prevalence of myasthenia gravis (MG) in this patient cohort **97**
3.4 Nature of collected samples **98**
3.4.1. Suitability of FFPE for NGS 98
3.4.2 Normal tissue for germ line DNA sequence analysis 98
Chapter 4 ..... **100**
GTF2I mutation in TETs **100**
4.1 The aims of the research described in this chapter were the following: **100**
4.2 Background **100**
4.3 Methods **101**
4.4 Results **102**
4.5 Discussion **104**
Chapter 5 ..... **109**
Whole exome sequencing of TETs **109**
5.1 The aims of the research described in this chapter were the following: **109**
5.2 Introduction **109**
5.3 Materials & Methods **111**
5.3.1 Samples and DNA extraction 111
5.3.2 Whole exome sequencing 111
5.4 Results **111**
5.4.1 Patient demographics and sample characteristics 111
5.4.2 Somatic mutations 112
5.4.2.1 Overview of somatic mutations 112
5.4.2.2 Candidate somatic driver mutations 114
5.4.2.3 Discussion on somatic mutations in TETs 118
5.4.3 Germline mutations 127
5.4.3.1 Candidate germline driver mutations 127
5.4.3.2 Discussion on germline mutations in TETs 129
5.5 Conclusions **136**
Chapter 6 ..... **138**
Insights into TETs through SNP genotyping **138**
6.1 The aims of the research described in this chapter were the following: **138**
6.2 Introduction **138**
6.3 Materials & Methods **139**
6.3.1 Sample characteristics and DNA extraction 139
6.3.2 SNP genotyping and data analysis 140

6.4 Results **141**
6.4.1 Patient demographics and sample characteristics 141
6.4.2 Somatic copy number gains 141
6.4.2.1 Somatic copy number gains in aggressive TETs 141
6.4.2.2 Somatic copy number gains in indolent TETs 142
6.4.2.3 Copy number gain discussion 143
6.4.3 Somatic copy number losses 147
6.4.3.1 Copy number losses in aggressive TETs 147
6.4.3.2 Copy number losses in indolent TETs 148
6.4.3.3 Copy number losses discussion 149
6.5 Conclusions **154**

Chapter 7 ..... **156**

Immunohistochemistry for emerging immunotherapy and other targets in TETs **156**
7.1 The aims of the research described in this chapter were the following: **156**
7.2 Cancer immunotherapy targets **157**
7.2.1 Inhibitory checkpoint molecules 157
7.2.1.1 Lymphocyte-activation gene 3 (LAG-3) 157
7.2.1.2 T-cell Immunoglobulin domain and Mucin domain 3 (TIM-3) 158
7.2.1.3 B7-H3 160
7.2.1.4 V-domain Ig- containing suppressor of T-cell activation (VISTA) 160
7.2.2 Stimulatory checkpoint molecules 160
7.2.2.1 T-cell antigen 4-1BB homologue (4-1BB) 160
7.2.3 Other emerging targets 163
7.2.3.1 PDZ And LIM Domain 1 (PDLIM1) 163
7.2.3.2 Hepatocyte Growth Factor/ scatter factor (HGF/ SF) 163
7.3 Methods **165**
7.4 Results **167**
7.4.1 PD-1 ..... **168**
7.4.2 PD-L1 173
7.4.3 PD-L2 183
7.4.4 LAG-3 192
7.4.5 TIM-3 193
7.4.6 B7-H3 197
7.4.7 VISTA 206
7.4.8 4-1BB 208
7.4.9 PDLIM1 210
7.4.10 HGF 217
7.4.11 Staining of normal thymic tissue 222

7.4.12 Inter-rater agreement 225
7.4.13 Co-expression of markers and clustering of TET subtypes according to these 225
7.5 Study limitations and future directions 230
Chapter 8 231
Overall conclusions and future work 231
8.1 The discovery and validation of the GTF2I mutation 232
8.2 WES and CNA of TETs 233
8.3 Genes selected for targeted capture 234
8.4 Immunohistochemistry 238
8.5 Future directions 241
Chapter 9 243
Bibliography 243
Chapter 10 272
Appendices 272
Appendix 1 272
Immunohistochemistry standard operating procedures 272
Appendix 2 312
Baseline clinicopathological data for all TETs including the downstream applications of the extracted DNA 312
Appendix 3 316
Clustering of TET subtypes based upon staining with all antibodies – Raw data for figure 7.29 316
Appendix 4 320
Copy of licence to use two illustrations from Elsevier 320
Appendix 5 326
Presentations and publications 326

List of Figures

Figure 1.1. The thymus gland 21

Figure 1.2. The thymus gland and its relationship to the mediastinal structures. 22

Figure 2.1. Example of a VCF file 75

Figure 2.2. Distribution of QUAL score for variants passing *MPB, msb30* and *tdb0.05* metrics 77

Figure 2.3. Final output format of somatic copy number calls from ASCAT 82

Figure 2.4 Somatic copy number calls in raw format from ASCAT 82

Figure 3.1. TET subtypes 90

Figure 4.1. Frequency of *GTF2I* mutant cases according to TET subtype. 107

Figure 4.2. Sanger sequencing traces demonstrating the T>A mutation in *GTF2I* on chromosome 7, 74146970 108

Figure 5.1. Number of detected somatic mutations in 15 paired TETs 113

Figure 5.2. Somatic CHEK2 and EGFR mutations seen across all tumour types 134

Figure 5.3. Somatic *ERBB4* mutations seen across all tumour types 135

Figure 7.0. Emerging immunotherapy targets in TETs 162

Figure 7.1. Box plot showing the distribution of the lymphocytic PD-1 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive 170

Figure 7.2. Heat map of TET subgroup PD-1 H-score comparison 171

Figure 7.3. Lymphocytic PD-1 staining in TETs 172

Figure 7.4. Box plot showing the distribution of the epithelial PD-L1 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive 177

Figure 7.5. Box plot showing the distribution of the lymphocytic PD-L1 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive 178

Figure 7.6. Heat map of TET subgroup epithelial PD-L1 H-score comparison 179

Figure 7.7. Heat map of TET subgroup lymphocytic PD-L1 H-score comparison 180

Figure 7.8. PD-L1 staining in TETs 181

Figure 7.9. Box plot showing the distribution of epithelial PD-L2 H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive 187

Figure 7.10. Box plot showing the distribution of lymphocytic PD-L2 H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive 188

Figure 7.11. Heat map of TET subgroup epithelial PD-L2 H-score comparison 189

Figure 7.12. Heat map of TET subgroup lymphocytic PD-L2 H-score comparison 190

Figure 7.13. *PD-L2 staining in TETs* 191

Figure 7.14. LAG-3 staining in TETs 192

Figure 7.15. Box plot showing the distribution of TIM-3 epithelial H-scores per TET subtype, including GTF2I mutation status and whether dead or alive 194

Figure 7.16. Heat map of TET subgroup epithelial TIM-3 H-score comparison 195

Figure 7.17. Epithelial TIM-3 staining in TETs 196
Figure 7.18. Box plot showing the distribution of epithelial B7-H3 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive 201
Figure 7.19. Heat map of TET subgroup epithelial B7-H3 H-score comparison 202
Figure 7.20. Epithelial B7-H3 staining in TETs 204
Figure 7.21. VISTA staining in TETs 207
Figure 7.22. 4-1BB staining in TETs 209
Figure 7.23. Box plot showing the distribution of epithelial PDLIM1 H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive 212
Figure 7.24. Heat map of TET subgroup epithelial PDLIM1 H-score comparison 213
Figure 7.25. Epithelial PDLIM1 staining in TETs 215
Figure 7.26. Box plot showing the distribution of epithelial HGF H-scores per TET subtype, including GTF2I mutation status and whether dead or alive 219
Figure 7.27. Heat map of TET subgroup epithelial HGF H-score comparison 220
Figure 7.28. Epithelial HGF staining in TETs 221
Figure 7.29 Clustering of TET subtypes based upon staining with all antibodies 229
Figure 8.1. Epithelial H-scores per TET subtype 239
Figure 8.2. Lymphocytic H-scores per TET subtype 240

List of Tables

Table 1.1. WHO histopathological classification of TETs 28

Table 1.2. Classification of NETTs 29

Table 1.3. Staging of TETs according to Masaoka-Koga and ITMIG definition of details thereof 31

Table 1.4. Ten-year survival in TETs according to Masaoka-Koga Stage 34

Table 1.5. Summary of the largest studies on PD-1 staining in TETs 43

Table 1.6. Summary of the largest studies on PD-L1 staining in TETs 44

Table 1.7. FDA-approved anti-PD-L1 drugs and associated assay systems 45

Table 1.8. Frequently recurrent chromosomal copy number gains in TET subtypes in selected studies 58

Table 1.9. Frequently recurrent chromosomal copy number losses in TET subtypes in selected studies 59

Table 2.1. Kits and sets used 64

Table 2.2. Reagents used 64

Table 2.3. Antibodies used for immunohistochemistry 65

Table 2.4. Clinical data fields collected 67

Table 2.5. Antibody Staining 87

Table 3.1 Baseline clinicopathological data for all TETs (n=156) 91

Table 3.2. MG prevalence in current TET cohort 97

Table 4.1. Prevalence of the *GTF2I* mutation in TETs according to subtype in our cohort Vs Petrini *et al.*, TCGA and Feng *et al*. 103

Table 5.1. Baseline patient and tumour characteristics for TET WES cases 113

Table 5.2. Likely somatic driver mutations in NHLI TET cohort (n=17) 123

Table 5.3. Overview of somatic mutations in TETs from TCGA, Petrini *et al.* and Zehir *et al.* 125

Table 5.4. Likely somatic and germline driver mutations in the NHLI TET cohort (n=17) 132

Table 6.1. Protein coding genes with copy number gains in TETs 146

Table 6.2. Protein coding genes with copy number losses in TETs 152

Table 7.1. Antibody Staining 166

Table 7.2. TET subtype comparison matrix for PD-1 H-scores 171

Table 7.3. PD-L1 staining in the epithelial and lymphocytic component of TETs 176

Table 7.4. TET subtype comparison matrix for epithelial PD-L1 H-scores 179

Table 7.5. TET subtype comparison matrix for lymphocytic PD-L1 H-scores 180

Table 7.6. PD-L2 staining in the epithelial and lymphocytic component of TETs 186

Table 7.7. PD-L1 and PD-L2 co-expression in TET epithelia 186

Table 7.8. TET subtype comparison matrix for epithelial PD-L2 H-scores 189
Table 7.9. TET subtype comparison matrix for lymphocytic PD-L2 H-scores 190
Table 7.10. TET subtype comparison matrix for epithelial TIM-3 H-scores 195
Table 7.11. TET subtype comparison matrix for epithelial B7-H3 H-scores 200
Table 7.12. PD-L1 and B7-H3 co-expression in TET epithelia. 203
Table 7.13. TET subtype comparison matrix for epithelial PDLIM1 H-scores 213
Table 7.14. PD-L1 and PDLIM co-expression in TET epithelia. 214
Table 7.15. TET subtype comparison matrix for epithelial HGF H-scores 216
Table 7.16. Mean percentage of cells from normal thymic tissue staining for PD-L1, PD-L2 and B7-H3. 222
Table 7.17. Inter-rater agreement scores for B7-H3, PD-L1 and PD-L 225
Table 8.1. Prevalence of the *GTF2I* mutation in TETs according to subtype in our cohort Vs Petrini *et al.,* TCGA, and Feng *et al*. 232
Table 8.2. Genes selected for Targeted Capture 236
Table 8.3. Genes with targeted hotspot mutations 237

List of abbreviations

Abbreviation	Full Name
4-1BB	1 T-cell antigen 4-1BB homologue
AChR	Acetyl choline receptor
ACT	Adoptive cellular therapies
AIRE	Autoimmune Regulator
AIRE	Autoimmune regulator gene
ALK	Anaplastic lymphoma kinase
AML	Acute myelogenous leukemia
APNS	Autoimmune paraneoplastic syndromes
ARID1A	AT-Rich Interaction Domain 1A
ASCAT	Allele-Specific Copy Number Analysis of Tumours
ASXL1	Additional Sex Combs
ATM	Ataxia Telangiectasia Mutated
AxAC	Exome Aggregation Consortium
BAP 1	BRCA1 associated protein-1
BRU	Biomedical Research Unit
CDKN2A	Cyclin Dependent Kinase Inhibitor 2A
CGH	Comparative genomic hybridization
CI	Confidence interval
CKE2	Checkpoint Kinase 2
CNA	Copy number alterations
CNV	Copy number variation
CT	Computed tomography
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4
CYLD	CYLD Lysine 63 Deubiquitinase
EGFR	Epidermal growth factor receptor
EMA	European Medicines Agency
EP300	E1A Binding Protein P300
ERBB4	Erb-B2 Receptor Tyrosine Kinase 4
FDA	Food and Drug Administration
FEZ2	Fasciculation And Elongation Protein Zeta 2
FFPE	Formalin-fixed paraffin-embedded
FISH	fFuorescence in situ hybridization
GATK	The Genome Analysis Tool Kit
GNA11	G Protein Subunit Alpha 11
GNAQ	G Protein Subunit Alpha Q
GTF2I	General Transcription Factor IIi
H-score	Histo-score
H&E	Hematoxylin and eosin
HDACs	Histone deacetylases
HER2	Human epidermal growth factor receptor 2
HLA	Human Leukocyte Antigen
HNSCC	Head and neck squamous cell carcinomas
HR	Hazard Ratio
HRAS	V-Ha-Ras Harvey Rat Sarcoma Viral Oncogene Homolog

ICL	Imperial College London
IGF-1R	Insulin-like growth factor 1 receptor
IHC	Immunohistochemistry
ITMIG	International thymoma interest goup
KMT2D	Lysine (K)-specific methyltransferase 2D
KRAS	Kirsten rat sarcoma viral oncogene homolog
LAG-3	*1 Lymphocyte-activation*
LCNEC	Large Cell neuronedocrine tumour
mAb	Monoclonal antibody
MANAs	Mutation-associated neoantigens
MEN1	Multiple endocrine neoplasia 1
MG	Myasthenia Gravis
MMR	Mismatch-repair
MNT LS	Micronodular thymoma with lymphoid stroma
mRNA	Messenger RNA
mRNA	microRNA
mTOR	Mechanistic Target Of Rapamycin
MUTYH	MutY DNA Glycosylase
NETT	Neuroendocrine thymic tumour
NF1	Neurofibromin 1
NGS	Next-generation sequencing
NHLI	National Heart and Lung Institute
NRAS	Neuroblastoma RAS Viral Oncogene Homolog
NSCLC	Non-small cell lung cancer
OCT	Optimal Cutting Temperature Compound
OS	Overall survival
PABPC3	Poly(A) Binding Protein Cytoplasmic 3
PBRM1	Polybromo 1
PCR	Polymrase cahin reaction
PD-1	Programmed Death 1
PD-L1	Programmed Death Ligand 1
PD-L2	Programmed Death Ligand 2
PDLIM1	PDZ And LIM Domain 1
PFS	Progression free survival
RB	Retinoblastoma
RGMB	Repulsive guidance molecule B
SCC	Squamous Cell Carcinoma
SETD2	SET Domain Containing 2
SMAD4	SMAD Family Member 4
SMARCCA4	SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 4
SNP	Single nucleotide polymorphism
SNV	Single-nucleotide variants
TAMG	Thymoma-associated MG
TCGA	The Cancer Genome Atlas
TET	Thymic Epithelial Tumour
TIM-3	T-cell Immunoglobulin domain and Mucin domain 3
TKIs	Tissue Kinase inhibitors

TKR	Tyrosine kinase receptor
TNM	Tumour, Necrosis, Metastasis
TP53	Tumor Protein P53
VAF	Variant allele frequency
VCF	Variant Call Format
VEGFR	Vascular endothelial growth factor receptor
VISTA	4 V-domain Ig- containing suppressor of T-cell activation
WHO	World Health Organization

Chapter 1
Introduction

1.1 The aims of the chapter are:

- To give an overview of the anatomy and function of the thymus gland
- To provide an overview of the current pathological classification and staging of thymic epithelial tumours (TETs)
- To outline the current methods of diagnosis and prognostic factors for TETs
- To give an overview of the link between TETs and autoimmune conditions
- To give an outline of the importance of the immune system in cancer in general and in TETs specifically
- To outline current and developing therapies for TETs
- To provide a comprehensive summary of all the molecular changes known to contribute to the formation of TETs
- To state the aims of this research project

1.2 Background

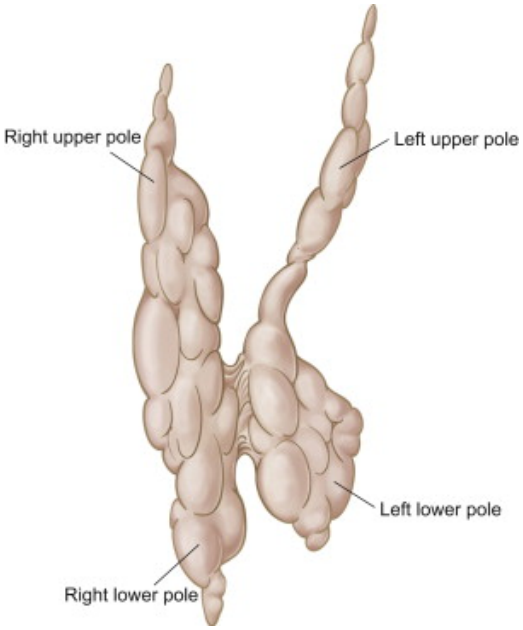
1.2.1 Anatomy and function of the normal thymus gland

The thymus gland arises from the third and possibly fourth branchial pouches of the endoderm as early as in the sixth week of gestation. It assumes its position in the anterior superior mediastinum by the eight week. It is composed of two almost identical lobes, the right usually being a bit larger than the left, which are surrounded by a fibrous capsule. The two lobes are connected by connective tissue at the midline and in some individuals by an intermediate lobe. The capsule septates the gland dividing its lobes into several lobules (1). These lobules consist of a number of follicles. Each one of these has an inner medulla and an outer cortex. The cortex consists mainly of lymphocytes, which are connected with a network of epithelial reticular cells. The thymus is the organ that is primarily responsible for inducing central tolerance to self-antigens. It does this through the control of the differentiation, and positive and negative selection of immature T-lymphocytes. It is in the cortex that T-lymphocytes undergo positive selection. In the medulla the network of epithelial reticula cells is larger and there are fewer lymphocytes. It is in the medulla that T-lymphocytes undergo negative selection. The regulator genes *AIRE* and *FEZ2* are expressed in the medulla (2). Involution of the thymus is a physiological process and occurs from the early teens, by which the thymus parenchyma atrophies and is replaced by fat (1).

1.2.2. Origin of thymic epithelial tumours (TETs)

TETs arise from the thymic epithelium. Although immunohistochemistry suggests they show either a medullary or cortical epithelial phenotype it is likely that they all originate from an epithelial progenitor or a stem cell that co-expresses medullary and cortical molecules (MR6+MR19+) and can give rise to both cortical (MR6+MR19−) and medullary (MR6−MR19+) malignant epithelial subpopulations (3).

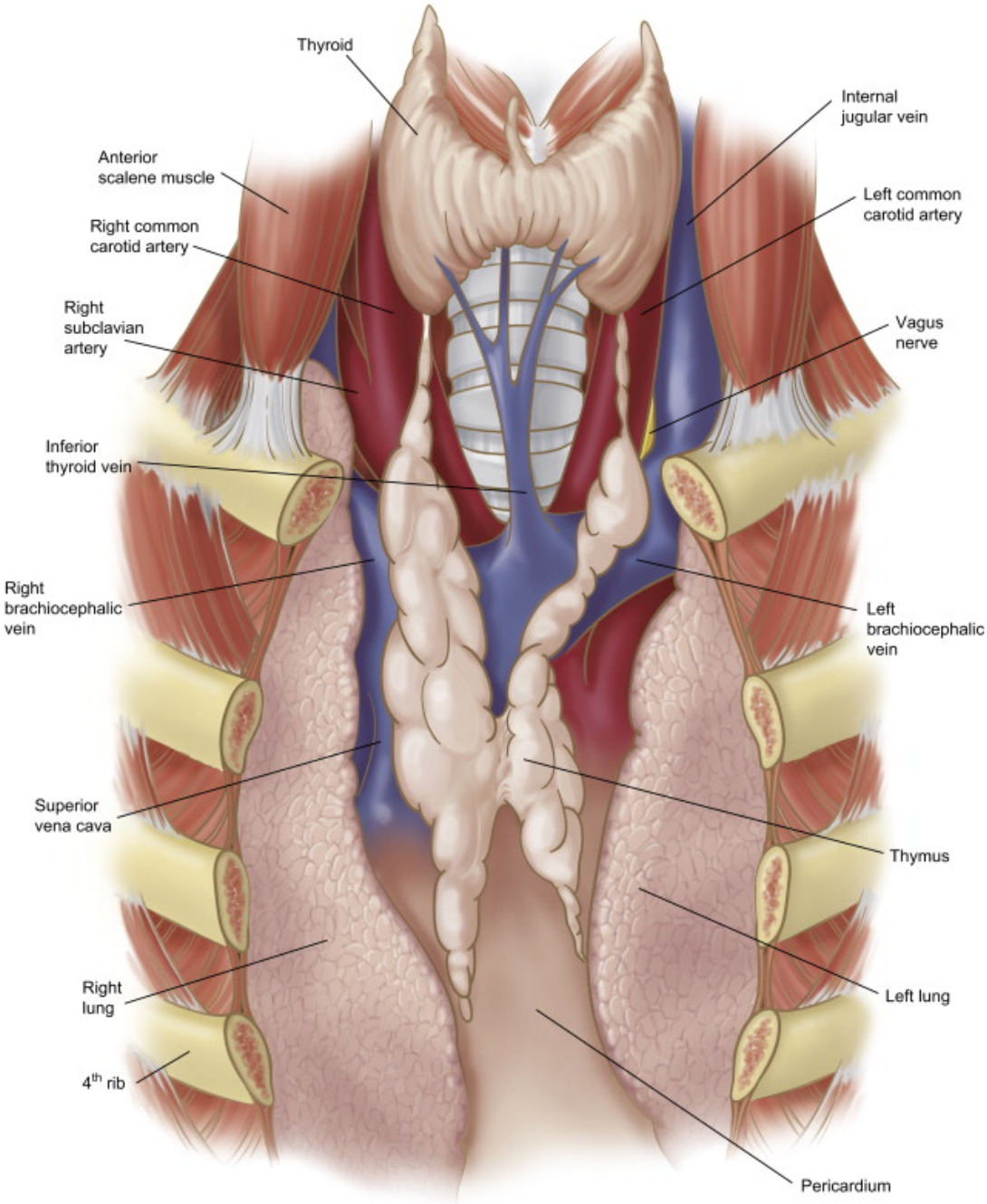
Figure 1.1. The thymus gland



This image has been reused with permission from Elsevier (license number: 4145650684424) from the 'Anatomy of the Thymus Gland' by Najib Safieddine and Shaf Keshavjee (Publication: Thoracic Surgery Clinics, Publisher: Elsevier, Date: May 2011)

Figure 1.2. The thymus gland and its relationship to the mediastinal structures.

Figure 0.1



This image has been reused with permission from Elsevier (license number: 4145650684424) from the 'Anatomy of the Thymus Gland' by Najib Safieddine and Shaf Keshavjee (Publication: Thoracic Surgery Clinics, Publisher: Elsevier, Date: May 2011)

1.2.3 Epidemiology and presentation of TETs

Thymic epithelial tumours (TETs) are rare, epithelial, intrathoracic tumours with an annual incidence ranging from 1.3 to 3.2 per million (4). Although together they represent about 0.2-1.5% of all malignancies (0.13 per 100,000 person-years in the United States (5,6)). They are however the commonest malignancy in the anterior mediastinum, accounting for 47% (7-9). They are broadly divided into thymomas and thymic carcinomas. Thymomas are much more frequent than carcinomas, which have an incidence around 0.2 - 0.5 per million (10). The mean age at diagnosis is 50 – 60 years although individuals present with TETs from early childhood to very old age. Males (49%) and females (51%) are overall almost equally affected with incidence of A, AB and B1 thymomas being slighter higher in females and that of carcinomas being higher in males (10-12). The aetiology of these tumours is largely unknown. No environmental risk factors including infectious causes have so far been recognized. In rare cases, infection with Epstein-Barr virus (EBV) has been linked with thymic carcinomas. Rarely TETs have also been encountered as part of genetic conditions, such as multiple endocrine neoplasia 1 (MEN1) (11-16).

More than 50% of thymomas present with localised disease, and are confined within the capsule of the thymus. The majority of patients with TETs have their tumours diagnosed as incidental findings as they are otherwise completely asymptomatic (17). When symptomatic, signs and symptoms are those of a mediastinal mass and include, cough, chest pain and dyspnoea (17).

TETs, and thymomas in particular, are very commonly linked with a range of autoimmune conditions, myasthenia gravis (MG) being the commonest. More than one autoimmune condition can present at the same time and the timing of those can be prior to, concurrent with, or after the TET diagnosis(18). Overall around one in three patients with a TET will develop an autoimmune condition at some point in their lives (19,20).

As is the case with suspected malignancies, the diagnosis of TETs depends on visualising the extent of disease and establishing a tissue diagnosis. Computed tomography (CT) is the preferred imaging modality to visualize the mediastinum. If the mass is deemed to be a TET that can be fully excised a pre-treatment biopsy is not required (13,21). In all other cases a biopsy is essential. This can be a percutaneous core-needle biopsy or a surgical biopsy either by mediastinotomy or mini-thoracotomy (22).

1.3 Pathology

The 2015 World Health Organisation cellular classification is the accepted classification scheme for TETs (Table 1.1) (23-26). Thymomas are subdivided into five major subtypes: A, AB, B1, B2 and B3. Carcinomas (also referred to as type C thymomas) are further subdivided into several subgroups (Table 1.1). Due to the morphological continuum between some of the thymoma subtypes but also some morphological overlap between thymomas and carcinomas, the International Thymic Malignancy Interest group (ITMIG) recently studied 72 TET cases and achieved a consensus in refining the criteria for each subtype aiming to make them more meaningful and reproducible (10). There are also several rare variants for example micronodular thymoma with lymphoid stroma (MNT LS). Table 1.1 includes all the thymoma and thymic carcinoma subtypes currently included in the WHO classification. Given the rarity of the majority of these subtypes, the ones that are most frequently described are also those that have been included in all major studies on TETs (These are highlighted in blue font in Table 1.1). Description of the major subtypes as well as the reported invasiveness and link with myasthenia gravis are also included to provide a comprehensive overview (11,16,24,26,33-35).

1.3.1 Thymomas

Thymomas are TETs in which there is evidence of differentiation towards the thymus. Varying proportions of immature lymphocytes can be seen intermixed with the epithelial cells. These are not neoplastic and their proportion to epithelial cells differs according to the specific thymoma subtype (27).

Type A

Type A thymomas account for 4-7% of all thymomas. They are comprised of spindle-/oval shaped epithelial cells, with no nuclear atypia. Non-malignant lymphocytes are seen but these are few to none. They are the most indolent subtype and the vast majority are encapsulated although invasive cases have been described (23,24,27,28). Although they have an excellent prognosis and are commonly reported to have 100% survival at 15 years, aggressive cases with metastatic or recurrent disease have been described (29). Atypical variants have also been recently described (27).

Type AB

Type AB thymomas account for 28-34% of all thymomas. These TETs are composed of areas of type A thymoma mixed with areas of a non-malignant lymphocyte-rich B-like component. In the B-like areas, cells have nuclei that are smaller and paler compared to those seen in B1 and B2 thymomas. There are more lymphocytes than are normally seen in Type A thymomas but usually

less than are seen in B1 thymomas. There is a wide spectrum in the variation of these two components (11,23). Prognosis is very good with survival at 15 years reported for around 90% of patients (24,27,28).

Type B1

Type B1 thymomas are TETs whose cells histologically look like the cortex of normal thymus with areas of epithelial cells intermixed with a large number of non-malignant lymphocytes as well as some areas that resemble normal thymic medulla. B1 thymomas account for 9-20% of all thymomas. 20-year survival is reported in up to 90% of patients (27).

Type B2

B2 thymomas account for 20-36% of all thymomas. Cells are large and polygonal and are arranged in a loose network giving an appearance very similar to the normal thymic cortex. There are still a significant number of non-malignant lymphocytes but there are very few areas that resemble normal thymic medulla. The number of lymphocytes still usually is larger than the number of epithelial cells. 20-year survival is significantly less that for the before mentioned subtypes and is around 60% (11,23,24).

Type B3

B3 thymomas account for 10-14% of all thymomas. Cells are medium in size, round or polygonal with mild atypia. There are only very few non-malignant lymphocytes that result in a sheet-like growth of malignant epithelial cells. 20-year survival is as low as 40%.

MNT LS (micronodular thymomas with lymphoid stroma)

These TETs are characterized by multiple, discrete nodules of epithelial cells, which are separated by a stroma of lymphocytes. The epithelial cells resemble those of type A thymomas and are bland, oval-spindle- shaped with few lymphocytes amongst them. They only account for 1-5% of all thymomas (11). Long-term survival is very good, being close to 100%(30). It is not uncommon for components of other, 'conventional' thymomas to be intermixed with MNT LS, which on occasions creates of transition between the 2. A small subset of these rare tumours can transform to micronodular thymic carcinoma with lymphoid hyperplasia or even to low-grade B-cell lymphomas (30,31).

1.3.2 Thymic Carcinomas

Thymic squamous cell carcinoma (SCC)

Thymic SCCs resemble squamous cell carcinomas seen in other solid tumours and completely lack non-malignant lymphocytes. They no longer resemble the normal thymus. Thymic SCCs are very aggressive, often diagnosed at an advanced stage, and have a high rate of recurrence and poor over survival, which is around 28% at 10 years (6,11).

Neuroendocrine thymic tumours (NETTs)

NETTs are TETs made up of epithelial neuroendocrine cells of the thymus. These contain highly atypic cells that behave very aggressively and present commonly at an advanced stage. Neuroendocrine differentiation of thymic carcinomas is routinely demonstrated with the use of immunohistochemistry (positive for chromogranin, synaptophysin, neuron-specific enolase and CD56) (11). In previous classifications, poorly-differentiated thymic neuroendocrine carcinomas were divided into small cell carcinomas (SCCs) and large cell neuroendocrine carcinomas (LCNECs) as both are high-grade and very aggressive (11). However in the most recent, fourth edition of TET classifications this terminology has been abandoned as both LCNECs and SCCs can be highly differentiated in regards to their neuroendocrine features (26). Following the rationale behind the classification of lung neuroendocrine tumours NETTs are also split into typical carcinoids (low grade), atypical carcinoids (intermediate grade) and LCNECs and SCCs (high-grade). The number of mitoses and the presence or absence of tumour necrosis has been used in this classification (Table 2) (26).

1.3.3 Clarifications from the 4th edition of the WHO criteria

The agreed subtypes within the 2004 WHO classification for TETs has led to a decade of clinical application, although shortcomings remained and some of these were addressed by an ITMIG initiative to propose refinements, (10) that subsequently became part of the 2015 WHO classification. A key area was recognition that some thymomas are composed of more than one histological subtype. Diagnoses should now include all the subtypes seen listed as a percentage, with the major component first followed by all minor components in 10% increments. This excludes AB thymomas and any TETs with even a minor carcinoma component. Also, immunohistochemical markers are recommended in cases where simple histology is inadequate to provide a diagnosis. These include cytokeratins and p63 expression for neoplastic epithelial cells, and terminal deoxynucleotidyl transferase expression in immature T cells (typically expressed in AB, B1, B2 and B3 thymomas but not expressed in Type A thymomas or carcinomas). In carcinomas detection of KIT and anti-CD5 antibodies increases the likelihood of a mediastinal carcinoma being of thymic origin. Clinical correlation is always advised especially as these markers are negative in 20% of carcinomas. A final point made is that all thymoma subtypes irrespective of the stage and exact subtype are labeled as malignant given that cases of all subtypes have been reported to behave in

an aggressive manner with associated deaths, excluding microscopic and micronodular thymomas (17,26). In a study of 37 patients at a cancer center in the US 18 of 37 patients with A or AB thymomas (9 of each subtype) had metastatic disease or developed recurrence of their disease. Despite the obvious selection bias as these were patients form an oncology center it becomes evident that there are rare cases of A and AB thymomas that are aggressive (32).

Table 1.1. WHO histopathological classification of TETs (1999, 2015)

TET subtypes	Description/ other names	Invasiveness	Myasthenia Gravis
Thymoma			
A (including atypical)	Spindle cell; medullary	10-40%	26% (0-35%)
AB	Mixed	6-40%	25% (5-42%)
B1	Lymphocyte-rich; lymphocytic; predominantly cortical; organoid	19-50%	35% (5-69%)
B2	Cortical	25-70%	49% (23-73%)
B3	Epithelial; atypical; squamoid; well-differentiated thymic CA	41-90%	40% (25-65%)
MNT LS	Micronodular thymoma with lymphoid B-cell hyperplasia	<10%	5%
Metaplastic	Thymoma with pseudosarcomatous stroma"; "biphasic thymoma, mixed polygonal and spindle cell type	0%	0%
Other rare thymomas			
Microscopic			
Sclerosing			
Lipofibroadenoma			
Thymic carcinoma			
Squamous cell	Epidermoid keratinizing and non-keratinizing carcinoma	90-95%	< 5%
Basaloid			
Mucoepidermoid			
Lymphoepithelioma-like			
Clear cell			
Sarcomatoid			
Adenocarcinomas			
Papillary			
Thymic carcinoma with adenoid cystic carcinoma-like features			
Mucinous			
Adenocarcinoma, NOS			
NUT carcinoma			

Undifferentiated			
Other rare thymic carcinomas			
Adenosquamous			
Hepatoid			
Thymic carcinoma, NOS			
Thymic neuroendocrine tumours			
Carcinoid tumors			
Typical carcinoid			
Atypical carcinoid			
Large cell neuroendocrine carcinoma (LCNEC)			
Combined LCNEC			
Small cell carcinoma			
Combined small cell carcinoma			
Combined thymic carcinomas			

Adapted from Weis et al.(16), Marx *et al.* (26), Travis *et al.*(11), Lamarca *et al.*(33), Girard *et al.* (34), Okumura *et al.*(24) and Kondo *et al.*(35). The commonest subtypes clinically are indicated in blue font.

Table 1.2. Classification of NETTs

Carcinoid tumors	
Typical carcinoid	<2 mitoses/2mm2; no necrosis
Atypical carcinoid	<2 mitoses/2mm2; with necrosis; or 2–10 mitoses/2mm2; + or − necrosis
LCNEC	>10 mitoses/2mm2; no small cell features
Combined LCNEC	LCNEC combined with any thymoma(s) or thymic carcinoma(s)
Small cell carcinoma (SCC)	Typical SCC histology
Combined small cell carcinoma	SCC combined with any thymoma(s) or thymic carcinoma(s)

Classification of NETTs as per the 4th edition of the WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart. Adapted from Marx et al. (26).

1.4 Staging

Up until now the most accepted staging system for TETs has been the Masaoka system with the modifications proposed by Koga *et al*. Despite wide acceptance, the Union Internationale Contre le Cancer and the American Joint Commission on Cancer do not officially recognize this as the standard staging system (36-39) as its prognostic accuracy is questionable (40,41). ITMIG had selected the Masaoka-Koga staging system and had further defined its classifications to achieve consistency amongst clinicians and researchers (Table 1.3) (39). This is the staging system that has been used in most studies so far (39,42) and the one that was used in the current study.

In the last (8^{th}) edition of the American Joint Committee on Cancer's *AJCC Cancer Staging Manual* a new tumour, nodes, metastasis (TNM) staging of TETs was introduced, which has been developed by ITMIG together with the Thymic Domain of the Staging and Prognostic Factors Committee (TD-SPFC) of the International Association for the Study of Lung Cancer (IASLC). This will start being used in January of 2018 (43). It was based on 10808 TET cases for which staging was assessed (44,45). Briefly, T is defined by the level of invasion of the TET. It does not take into account the size of the TET. T1 are localized tumours irrespective of capsular invasion, T2 signifies pericardial invasion and T3 refers to tumours that directly involve structures in the mediastinum, such as the lung, the superior vena cava, the chest wall, etc. Nodal involvement as indicated by N, can be 0 for no involvement, N1 for anterior and N2 for deep intra-thoracic regions. M refers to metastasis with M1a indicating pleural/ pericardial or intra-parenchymal pulmonary nodules and M1b to distant metastases (39).

Table 1.3. Staging of TETs according to Masaoka-Koga and ITMIG definition of details thereof

Stage	Masaoka-Koga	ITMIG Definition of Details of the Masaoka-Koga Staging System
I	Grossly and microscopically completely encapsulated tumour	Grossly and microscopically completely encapsulated tumour This includes tumours with invasion into but not through the capsule, or Tumours in which the capsule is missing but without invasion into surrounding tissues
IIa	Microscopic transcapsular invasion	Microscopic transcapsular invasion Microscopic transcapsular invasion (not grossly appreciated)
IIb	Macroscopic invasion into thymic or surrounding fatty tissue, or grossly adherent to but not breaking through mediastinal pleura or pericardium	Macroscopic invasion into thymic or surrounding fatty tissue, or grossly adherent to but not breaking through mediastinal pleura or pericardium Gross visual tumour extension into normal thymus or perithymic fat surrounding the thymoma (microscopically confirmed), or Adherence to pleura or pericardium making removal of these structures necessary during resection, with microscopic confirmation of perithymic invasion (but without microscopic extension into or through the mediastinal pleura or into the fibrous layer of the pericardium)
III	Macroscopic invasion into neighboring organ (i.e., pericardium, great vessel, or lung)	Macroscopic invasion into neighbouring organ (i.e., pericardium, great vessel, or lung) This includes extension of the primary tumour to any of the following tissues: Microscopic involvement of mediastinal pleura (either partial or penetrating the elastin layer); or Microscopic involvement of the pericardium (either partial in the fibrous layer or penetrating through to the serosal layer); or Microscopically confirmed direct penetration into the outer elastin layer of the visceral pleura or into the lung parenchyma; or Invasion into the phrenic or vagus nerves (microscopically confirmed, adherence alone is not sufficient); or Invasion into or penetration through major vascular structures (microscopically confirmed); Adherence (i.e., fibrous attachment) of lung or adjacent organs only if there is mediastinal pleural or pericardial invasion

		(microscopically confirmed)
IVa	Pleural or pericardial metastases	Pleural or pericardial metastases Microscopically confirmed nodules, separate from the primary tumor, involving the visceral or parietal pleural surfaces, or the pericardial or epicardial surfaces
IVb	Lymphogenous or heamatogenous metastasis	Lymphogenous or haematogenous metastasis Any nodal involvement (e.g., anterior mediastinal, intrathoracic, low or anterior cervical nodes, any other extrathoracic nodes) Distant metastases (i.e., extrathoracic and outside the cervical perithymic region) or pulmonary parenchymal nodules (not a pleural implant)

Adapted from Koga *et al*.(38) and Detterbeck *et al*. (39)

1.5 Prognosis

The prognosis of TETs depends first and foremost on tumour stage at presentation and the tumour resection status.

A large retrospective review into the prognostic factors in TETs in 2011, which included 29 studies reporting a multivariate analysis of prognostic factors for survival and 12 for recurrence; identified stage at presentation as the most consistent factor significantly associated with overall survival (OS) and recurrence. The Masaoka or Masaoka-Koga staging was used and in the majority of studies the difference was seen between Stage I and II vs. Stage II and IV. The other consistent factor predictive of significant longer OS was R0 (complete) resection, which was often the case independent of the tumour stage. Establishing the link between WHO subtype and OS was more difficult given the variability of the studies and the way results had been reported often grouping subtypes together (46).

In 2015 ITMIG compiled a worldwide retrospective database of TETs to establish the clinical and demographic differences between subtypes and determine the effect that the WHO histotypes, stage and resection status have on overall survival of patients with these tumours. The evaluable thymomas were 4221. Overall survival (OS) for patients with thymomas (with an R0/ complete resection) was significantly longer for patients with stage I and II disease vs. stage IV ($p < 0.0001$). R0 resection was also linked with a significant increase in OS vs. R1 and R2 resection ($p < 0.0001$). Thymoma subtype was a prognostic factor only in Stage I and II disease with an R0 resection. On multivariate analysis (age, stage, resection status and TET subtype) age ($p < 0.0001$), stage (I and II vs. III; $p<0.001$, and I and II vs. IV; $p<0.0001$) and resection status (R0 vs. R1 vs. R2) were all associated with OS. Surprisingly R0 vs. R1 and R0 vs. R2 were not significant. WHO subtype was not significantly linked with OS ($p = 0.4$) but was linked with recurrence ($p < 0.04$) (16).

Table 1.4 shows the 10-year survival for all TET subtypes according to the Masaoka-Koga stage at presentation based on the largest review on the topic so far (44).

Table 1.4. Ten-year survival in TETs according to Masaoka-Koga Stage (44)

Masaoka-Koga	10-year survival (n=8994)
Stage I	84% (81%–86%)
Stage II	83% (79%–87%)
Stage III	70% (64%–75%)
Stage IVA	42% (26%–58%)
Stage IVB	53% (32%-73%)

Regarding thymic carcinomas a subset analysis of 1042 thymic carcinomas cases as collected by ITMIG were assessed to determine the factors that influence OS in this cohort. Only 42 (5%) patients presented with Stage I disease, 138 (17%) had stage II, 370 (45%) had stage III, and 274 (33%) had stage IV disease. Of those 22% had induction chemotherapy and 6% had pre-operative radiotherapy. An R0 resection was achieved in a remarkable 61% of cases. Of all the carcinomas, squamous cell carcinomas (SCCs) were 79%. 31% of patients received adjuvant chemotherapy and 60% received adjuvant radiation therapy. Median OS for the entire cohort was 6.6 years (95% CI, 5.8 - 8.3) and the cumulative incidence of recurrence at 5 years was 35% (95% CI, 30% - 40%). On multivariate analysis the only factors associated with a prolonged OS were radiation therapy and male gender (47).

1.6 Autoimmune paraneoplastic syndromes (APNS)

About one-third of patients with TETs also develop APNS with MG being the commonest (48). The most recent data on APNS come from a single institution review at Indiana University, which retrospectively studied 733 patients with TETs (71% thymomas, 26% carcinomas and 3% NETTs). 203 had an APNS (28%). The commonest was myasthenia gravis (MG) with 130 cases (17.7%), followed by red cell aplasia (20 cases, 2.7%), hypogammaglobulinemia (14 cases, 1.9%), systemic lupus erythematosus (12 cases, 1.6%) and other APNS (64 cases, 8.7%). 37 patients has two or more APNS (5%). They were most commonly seen in thymomas (183 of 523, 35%), followed by NETTs (3 of 20, 15%) and carcinomas (16 of 187, 9%). There was a statistically significant protective effect of having an APNS on both recurrence rates (8% vs. 13%) and 5-year mortality (10% vs. 16%) with $p< 0.05$ compared to individuals without APNS. Surprisingly however intrathoracic recurrences were commoner in patients with APNS (89% vs. 77%; $p = 0.016$) (19).

1.6.1 Myasthenia Gravis

15-20% of MG patients will also have a thymoma at some point in their lives and of patients with TETs, 24-40% will develop MG (48). Thymoma-associated MG (TAMG) occurs predominantly in patients aged 40-60 and affects male and female patients alike. In certain cases MG onset or deterioration is a sign of disease relapse or recurrence (48). The risk of developing TAMG has been reported to be between 20% and 60% in thymomas and is more commonly encountered in cases of B2 or B3 thymomas compared to B1, AB and A (49). These patients almost invariably have anti-AChR autoantibodies, which are typically produced from normal thymic tissue adjacent to the thymoma tissue. Also at diagnosis patients demonstrate a unique range of additional autoantibodies including to striational muscle autoantigens such as titin (90% of cases), ryanodine receptors (50% of cases) and type I interferons, IL-12 (in 50%) amongst others (49). Patients with thymoma and MG appear to have a better prognosis than those without MG(42). Some studies suggest there is a significantly higher rate of second malignancies in patients with TETs in non-MG patients, suggesting MG has a protective effect (50). Defective expression of the autoimmune regulator gene (AIRE), a gene important for ectopic expression of peripheral self-antigen and central thymic T-cell education, which encodes a protein involved in nuclear transcription, may be partly responsible for the pathogenesis of autoimmune diseases in TETs. *AIRE* is located in chromosome 21q22.3 with over 100 mutations having been reported so far in the Human Gene Mutation Database. Impaired AIRE expression has been reported in 95% of thymoma cases with MG (51). The molecular basis underlying its defective expression in TETs is unknown but potentially represents an exciting link between autoimmunity and malignancy (52,53).

TAMG is pathologically and genetically different to other cases of MG. MG has been found to be linked with the following HLA alleles: HLA-A1, HLA-B8, HLA-DR3 and HLA-DQ2 (which together form the AH8.1 haplotype) in the Caucasian population (54). IN TAMG however no consistent HLA association has been drawn. A study with exclusively French participants with TETs and MG revealed an increased frequency of HLA-A*25 allele in the entire patient cohort. The same study also showed a negative association between Type B2 thymoma and the HLA-A*02 allele raising the possibility that it confers a potentially protective role in the development of B2 thymoma. In an Asian population HLA-DQA1 and DQB1 have been associated with TAMG. Within that group B2 was the commonest TET subtype (55).

1.7 The immune system as a target for treatment in cancer

Although TETs are closely linked with a large number of autoimmune conditions the importance of the immune system in the treatment of solid tumours is being increasingly understood and exploited.

1.7.1 The role of immunotherapy in the treatment of solid tumours

Immunotherapy is defined as "the approach to treating cancer by generating or augmenting an immune response against it". It has only been recently that Phase II and III clinical trials have proven the efficacy of immunotherapy agents leading to an increased progression-free and overall survival in some tumour types. Immunotherapy drugs currently in clinical practice can be divided into two main subgroups: Immune-cell-targeted monoclonal antibodies (mAb) and adoptive cellular therapies (ACT).

Immunomodulatory mAbs target T- cells rather than tumour cells and therefore do not constitute a form of targeted treatment. The best-described molecule so far is PD-1. This is one of several receptors implicated in immune-checkpoint signaling whose blockade with monoclonal antibodies leads to the T-cells' antitumour activity. Hence blocking its activity blocks the effects of such an inhibitory receptor.

The other means of activating the immune system against cancer is by co-stimulation of T-cells with mAbs that target their stimulatory receptors. Given that the response to checkpoint blockade with checkpoint inhibitors is very variable the idea of using stimulatory mAbs is quite appealing. As with all anticancer drugs identifying markers of response to treatment is now the focus of many groups. The main aim of these studies is to make sure that patients, who are not likely to benefit from a particular therapy, are not exposed to it, thereby avoiding un-necessary toxicity or delaying exposure to a more appropriate treatment that is more likely to be efficacious.

The best-established treatment targets in this field have so far been T-lymphocyte protein 4 (CTLA-4) and PD-1 and these are discussed further below.

1.7.1.1 CTLA-4

CTLA-4 (cytotoxic T-lymphocyte-associated protein 4) has been shown to be key in suppressing T-cell function and consequently dampening down T-cell mediated anticancer activity. Its blockade in

melanoma opened the way for modern immunotherapy (56). Ipilimumab, the first anti-CTLA-4 human mAb has been shown to improve overall survival in patients with metastatic melanoma. It was the first immunotherapy to achieve an increase in overall survival in any tumour type and the first drug to improve overall survival in metastatic melanoma in a phase III trial (56). What is even more interesting is that patients who did not progress at 2 years continued to have very durable responses beyond that. This is very rare, not only in melanoma, but in most solid tumours (57). These findings are also important in demonstrating that while these treatments control tumour growth via activation of the immune system, their effects persist when treatment is discontinued.

However administration of Ipilimumab in melanoma patients gave rise to a whole new class of toxicity, collectively known as immune-related adverse events (58). These are the systemic toxicities that result from disinhibiting the immune system. As with most toxicities some organs are affected more than others (58). Another important lesson from using Ipilimumab in melanoma has been that in certain cases patients progress before eventually responding to therapy, and that the response in these cases can often be durable. Therefore careful assessment of a patient's disease status along with careful clinical judgment is important when it comes to making decisions about treatment discontinuation (59).

1.7.1.2 PD-1

The second inhibitory receptor on T cells that has been utilized therapeutically is PD-1. This is expressed predominantly on activated $CD4^+$ T and $CD8^+$ T cells as well as on peripheral B cells. It regulates late-phase immune responses, such as the memory response and chronic infection amongst others in peripheral tissues (60). It has two known ligands, PD-L1 and PD-L2, which are expressed in antigen presenting cells and tumours. PD-1 And PD-L1 have become major targets of interest in a large number of malignancies. Three anti-PD-1 agents have been given FDA approval in patients with advanced non-small cell lung cancer (NSCLC): Pembrolizumab (Keytruda [Merck & Co., Inc., Kenilworth, NJ]) (61,62), nivolumab (Opdivo [Bristol-Myers Squibb, Lawrenceville, NJ]) (63,64), and atezolizumab (Tecentriq [Genentech/Roche, South San Francisco, CA]) (65). Pembrolizumab and nivolumab have both been approved for the treatment of patients with advanced-stage melanoma, even if they had been previously been treated with Ipilimumab (66). Following these results immune checkpoint inhibitors are currently being investigated in a number of other malignancies. These include pembrolizumab in breast cancer (67), gastric cancer (68) (KEYNOTE-012 study), endometrial cancer (69), carcinoma of the anal canal (70) and mesothelioma (71) (KEYNOTE-028 study). These are Phase 1b trials that have shown promising activity in these tumour types with manageable toxicity.

1.7.1.3 PD-L1 and other biomarkers of response to immune checkpoint inhibitors

As with all treatments, biomarkers of response are of the outmost importance. This is also the case with immune checkpoint inhibitors that so far have shown very encouraging results in a range of solid tumours but have also highlighted cases that are highly responsive that have had prolonged responses that cannot be explained by PD-L1 expression or RNA expression signatures (72,73). A high mutational load (greater than 100 non-synonymous single-nucleotide variants (nsSNV) per exome) has been linked with better clinical outcomes in melanoma patients on anti-CTLA-4 therapy (74). In NSCLC, in two independent patient cohorts, a high non-synonymous mutation load was linked with and improved objective response, progression-free survival and overall survival (75).

It has also been suggested that response to anti-PD-1 inhibitors relies on mutation-associated neoantigens (MANAs) encoded by cancers. Cancers that are mismatch-repair (MMR) deficient are thought to have a high number of MANAs. A Phase II trial (NCT01876511) demonstrated that colorectal cancers that are MMR deficient are indeed highly sensitive to pembrolizumab as were non-colorectal cancers that were MMR-deficient and were also treated with the same drug. Whole exome sequencing (WES) showed a mean of 1782 and 73 somatic mutations per tumour in in mismatch repair–deficient and mismatch repair–proficient tumours respectively ($p = 0.007$). A high somatic mutation load was linked with prolonged PFS ($p = 0.02$). This was the first time that MMR status was proven to predict for response to pembrolizumab (76). The same author (NCT01876511) prospectively looked into the efficacy of pembrolizumab in 86 consecutive patients with 12 different solid malignancies that were MMR deficient as assessed by IHC or PCR. All patients had received at least one line of treatment prior to participating into the trial. The commonest tumour type was colorectal followed by osteosarcoma. There were no cases of TETs included in this study. Thirty-two patients (48%) had tumours with germline mutations consistent with a diagnosis of Lynch syndrome. Another seven had a family history that was consent with the diagnosis of Lynch syndrome but germline testing was not performed. Forty-six of 78 evaluable patients had an objective response (as assessed with the Response Evaluation Criteria In Solid Tumors (RECIST v1.1) guidelines), 18 of whom (21%) had a complete response. 20 patients had stable disease. The overall disease control rate was 77%. Objective response was observed in 52% and 54% of colorectal cancer and non-colorectal cancer patients respectively. Response was lower in patients with Lynch syndrome (46%) and higher in the remaining tumours (59%), which was not significant ($p = 0.27$). Remarkably 11 of the 18 patients had exhibited complete response radiologically. When taken off pembrolizumab two years later no cases of recurrence were reported at a median of 8.3 months off therapy. Median progression free survival (PFS) was not reached (95% CI = 14.8 months to not estimable); PFS at two years was 53%. Median overall

survival (OS) was not reached (95% CI = not estimable) and OS at two years was 64%. The side effects for pembrolizumab were consistent with those previously reported with 21% reported as having any grade of endocrine disorders. Most of these had hypothyroidism that was treated with thyroxine replacement. These results are consistent with the hypothesis that in MMR-deficient tumours, the large proportion of MANAs renders these tumours more sensitive to anti-PD-1 treatment. As testing for MMR deficiency is widely available it is proposed that this is tested routinely in patients that progress on standard treatment options to be offered anti-PD-1 therapy as the following line of therapy (77).

Very few studies have attempted to correlate copy number alterations (CNA) in PD-L1 and PD-L2 and their expression. A study of 654 patients with non-small cell lung cancer using fluorescence in situ hybridization and IHC revealed that PD-L1 genomic gains correlated with higher levels of PD-L1 expression whereas the same was not observe with PD-L2 (78). Another study on with head and neck squamous cell carcinomas (HNSCC) analyzed the TCGA data on patients with this tumour type and found that a significant proportion of HNSCC tumors have genetic alterations in PD-L1 (6.2%). There were no somatic mutations in PD-L1 in any of the 279 samples that underwent whole exome sequencing but amplifications were found in 25 of 530 samples (4.7%) and deletions in 8 of 530 (1.5%). Unfortunately these results could not be correlated with PD-L1 expression (79).

1.7.1.4 The role of PD-L2

Although the role of PD-L1 as a predictive biomarker for response to anti PD-1 treatment has been extensively researched, the significance of PD-L2, which is also a ligand of PD-1, remains elusive. This is particularly important, as there are several cases where patients whose tumour is negative for PD-L1 still respond to anti-PD-1 therapy raising the possibility that PD-L2 expression might also be predictive of a response in certain cases.

Many different antigen presenting cells, including monocytes, macrophages and dendritic cells, express PD-L2. However most studies have found its expression to be lower than that of PD-L1 making its significance questionable. PD-L2 has recently also been implicated in autoimmune pneumonitis, which is a rare but devastating autoimmune toxicity experienced by some patients on anti-PD-1 therapy. The proposed mechanism for this is that PD-L2, which is also present on lung dendritic cells, can bind to the repulsive guidance molecule B (RGMB) found on alveolar cells and interstitial macrophages. Inhibition of PD-1 by PD-1 inhibitors allowed PD-L2 to preferentially bind RGMB, which leads to clonal expansion of lung T-cells to cause inflammation and pneumonitis (80).

A recently published study tried to address this by investigating the prevalence and distribution of PD-L2 in 400 tumour samples from seven different tumour types: renal cell carcinoma, bladder carcinoma, melanoma, NSCLC, triple-negative breast cancer, gastric carcinoma, and head and neck squamous cell carcinoma (HNSCC). The relationship between PD-L2 positivity and clinical response to pembrolizumab was assessed in patients with recurrent or metastatic HNSCC. A novel IHC assay was used to assess PD-L2 expression. PD-L2 was expressed in all tumour types and was present in stromal, tumour, and endothelial cells. Although there was a significant correlation between PD-L2 and PD-L1 expression and distribution ($p = 0.0012 – <0.0001$) PD-L2 was seen in a number of cases when there was no PD-L1 expression. In HNSCC PD-L2 expression was associated with improved response to pembrolizumab independent of PD-L1 expression. Response was however higher in cases that were positive for both (27.5%) compared to those positive for PD-L1 only (11.4%). PD-L2 positivity was also significantly associated with prolonged PFS and OS in this group again independent of PD-L1 status. These findings can perhaps explain why some PD-L1 negative tumours respond to anti-PD-1 treatment and confirm the role of PD-L2 as an independent prognostic biomarker in HNSCC cases treated with pembrolizumab. This will need to be verified in other tumour types but is likely to become key in identifying patients likely to benefit from such treatments and even stratifying the likelihood and extent of response (81).

An outline of what is known for the inhibitory and co-stimulatory receptors relevant to this project can be found in section 7.2.

1.7.2 PD-1/ PD-L1 expression in TETs

A few studies have attempted to examine the expression of PD-L1 and PD-1 in TETs. A summary of the most important studies to date can be found in Tables 1.5 (PD-1) and 1.6 (PD-L1).

One of the first and largest studies to investigate PD-L1 expression in TETs included 101 thymomas and 38 thymic carcinomas. The PD-L1 antibody E1L3N was used to stain tissue microarrays (TMAs) of the above samples. A PD-L1 score was calculated using an H-score (intensity of staining, 0-3, multiplied by the percentage of cells that stain positive, 0-100%). The score used as a cut-off for positivity was 3 (1%). 22 of 102 thymomas (23%) were positive and 26 of 37 carcinomas (70%) were positive for PD-L1. Carcinoma histology vs. types A, AB, and B thymomas was significantly linked with positive PD-L1 expression ($p = 0.006$). PD-L1 positivity was not significantly associated with OS in multivariate analysis in thymomas or thymic carcinomas

(hazard ratio = 0.99, 95% CI: 0.35 - 2.73; *p = 0.987*). WHO stage, Masaoka–Koga stage and pre-operative chemotherapy were significantly linked with OS (82).

In another study TMAs of 69 TETs and 17 thymic controls were stained with an anti-PD-L1 monoclonal antibody (clone 15, Sino Biological). The staining was scored from 0-3, with 0 being no staining and 3 being intermediate to strong. Cases with a score of 1 were defined as PD-L1 low and those with a score of 3 as PD-L1 high. There were no cases with no staining. PD-L1 high scores were seen in 68.1% and 17.6% of TETs and controls respectively. There was significantly higher PD-L1 staining in more aggressive subtypes (overall *p value* for WHO histology = *0.039*) with PD-L1high scores seen as follows: 7/7 B3 (100%) > 10/13 B1 (76.9%) = 15/20 B2 (75.0%;) = 3/4 SCC (75%) > 10/17 AB (58.8%) > 2/8 A (25%). 14.8% of all TETs also had lymphocytic staining in this study. PD-L1high TETs were found to have a significantly worse OS (HR 5.40, 95% CI 1.13-25.89, p = 0.035) (83).

Another study evaluated PD-L1 and PD-1 expression in TETs before and after chemotherapy. FFPE sections were stained with anti-PD-L1 (E1L3N; Cell signaling) and anti-PD-1 antibodies (NAT105; Abcam). PD-L1 expression was scored using an H-score (≥1 being positive) and PD-1 staining was scored on intensity (0-3, with anything over 0 being positive). 67% of thymomas and 41% of carcinomas were positive for PD-L1. The median score was 30 and 0 for thymomas and carcinomas respectively. It is worth pointing out that within the carcinomas there were 2 neuroendocrine carcinomas that both had a PD-L1 score of 0. Surprisingly the objective response rate was 50% in PD-L1 positive and 0% in PD-L1 negative thymomas. In carcinomas there was no such association. Progression free survival (PFS) was not affected by PD-L1 or PD-1 status. PD-L1 and PD-1 scores were found to increase post chemotherapy in 6 cases that were re-biopsied post treatment. This raises the possibility of using anti-PD-1 immunotherapy sequentially post chemotherapy for PD-L1 positive patients (84).

An attempt has been made to assess the relationship between PD-L1 IHC and CNA in the PD-L1 gene via qRt-PCR. PD-L1 and PD-1 expression were determined in 25 cases of thymic carcinomas. The median PD-L1 value was 100 (range 0-300), with 20 of 25 found to be positive (80%). This percentage was much lower for PD-1 where only 5 of 25 samples were positive (20%). qRT-PCR for CNA of PD-L1 was performed in 11 cases where FFPE specimens were available. Only 3 of the 11 carcinomas showed an increase in the PD-L1 copy number. In those cases it was significantly overexpressed. High PD-L1 expression was associated with a higher DFS (85).

A recent retrospective review of 35 TET cases (32 thymomas and 3 carcinomas) by IHC was described (PD-1 clone: NAT105 and PD-L1 clone: 22C3 antibodies). PD-L1 was expressed in 29 of the 35 samples (83%), including all 3 carcinomas. PD-1 expression was seen in 27 of 35 cases

(77%), including only one of the three carcinomas. High levels (as defined by IHC ≥ 3/5, zero being the lowest and five the highest intensity) of PD-L1 and PD-1 were seen in 57% and 31% of cases respectively. There was a trend towards worse survival for PD-L1 expressing cases but this was not significant (*p = 0.097, log-rank test*). There was no association between PD-L1 expression and more aggressive subtypes (B2/ B3/ CA) or more advanced stage at presentation. On the other hand PD-1 expression was significantly linked to more indolent tumours (*p = 0.031*). There was no significant association between expression of either PD-1 or PD-L1 and stage at presentation, smoking status, recurrence or the presence of myasthenia gravis (86).

One of the largest studies in the field to date was performed on whole-tissue sections from 100 TETs (74 thymomas and 26 thymic carcinomas). PD-L1 and PD-1 status was assessed via IHC and how these related to clinical and pathological characteristics was determined. Positivity for each staining was determined as more than 5% of epithelial cells and lymphocytes staining positive. PD-1 was positive in 52/100 cases (52%). This was higher in thymomas (46/74; 62%) and lower in carcinomas (6/26; 23%). PD-L1 was positive in 61/100 cases, with 47/74 thymomas (64%) and 14/26 carcinomas (54%) being positive. There was a significant association between PD-1 positive cases and higher stage at diagnosis for thymic carcinomas (p = 0.01). PD-L1 positive cases were associated with thymoma patients having received primary chemotherapy (*p = 0.0065*) but not with tumour stage, thymic carcinoma subtype, lymph node status or having had primary chemotherapy in carcinomas. PD-1 and PD-L1 status were not found to be independent predictors of OS (87).

Given the conflicting results of previous studies on the expression of PD-L1 and PD-1 in TETs, an attempt has been made to compare PD-L1 and PD-1 expression in TETs using immunohistochemistry (IHC) vs. a molecular assay. TMAs from 68 patients with completely resected TETs were used (63 thymomas and 5 carcinomas). For IHC the BMS PD-L1 (clone 28-8) assay from Dako was used. mRNA expression was evaluated by RT-PCR. PD-L1 mRNA tumour expression was found to be significantly associated with a poorer prognosis in patients whose tumour diameter was over 5cm (HR: 5.40 CI 95% 1.9-78.1, p=0,0083). Patients whose tumour had >1% PD-L1 expression by IHC and PD-L1 mRNA expression had an overall survival of 82 months whereas the median overall survival was not reached for those who were negative in IHC and had no mRNA expression (p=0.0178). PD-L1 positivity was significantly linked with the presence of myasthenia gravis (88)

Table 1.5. Summary of the largest studies on PD-1 staining in TETs (84-87,89,90)

Author	Year	Thymomas		Thymic Carcinomas		Positivity threshold	Antibody clone
Weissferdt et al.	2017	46/74	62%	6/26	23%	≥5% of cells staining	E1L3N (Cell signaling)
Owen et al.	2017	26/32	81%	1/3	33%	≥1/5	NAT105 (Abcam)
Katsuya et al.	2016	4/9	44%	8/17	47%	Intensity score ≥0 (Score range 0-3)	NAT105 (Abcam)
Yokoyama et al.	2016	N/A	N/A	5/25	20%	≥H-score of 20	NAT105 (Abcam)
Marchevsky et al.	2016	24/38	63%	3/8	38%	≥1/3	MRQ22 (Cell Marque)

Table 1.6. Summary of the largest studies on PD-L1 staining in TETs (83-87,89-91)

Author	Year	Thymomas		Thymic Carcinomas		Positivity threshold	Antibody clone
Weissferdt et al.	2017	47/74	64%	14/26	54%	≥5% of tumour cells staining	EPR4877(2) (Abcam)
Owen et al.	2017	26/32	81%	3/3	100%	≥1/5	22C3 (Dako)
Katsuya et al.	2015	22/102	23%	26/37	70%	≥H-score of 3 (≥1% of cells staining)	E1L3N (Cell signaling)
Padda et al.	2015	65/65	100%	4/4	100%	Intensity score ≥0 (Score range 0-3)	Clone 15 (Sino Biological)
Katsuya et al.	2016	6/9	67%	7/17	41%	≥H-score of 1	E1L3N (Cell signaling)
Yokoyama et al.	2016	44/82	54%	20/25	80%	≥H-score of 20	EPR1161[2] (Abcam)
Naido et al.	2015	11/12	94%	4/12	34%	≥ 25% of tumour cells staining	E1L3N (Cell signaling)
Marchevsky et al.	2016	35/38	92%	4/8	50%	≥1/3	SP142 (Spring BioScience)

1.7.3 Antibody clones and kits for PD-L1 staining

So far three anti-PD-1 agents have been given FDA approval in patients with advanced non-small cell lung cancer (NSCLC): Pembrolizumab (Keytruda [Merck & Co., Inc., Kenilworth, NJ]) (61,62), nivolumab (Opdivo [Bristol-Myers Squibb, Lawrenceville, NJ]) (63,64), and atezolizumab (Tecentriq [Genentech/Roche, South San Francisco, CA]) (65). A number of other agents have been granted approval for other use in other solid tumours. Of the three drugs mentioned above, pembrolizumab is the only that has been approved also in the first line setting (62) whereas the rest have only been approved only as second line options. The approval for each drug has come with specific immunohistochemistry assays alongside a specific antibody clone for each. These are summarized in table 1.7.

Table 1.7. FDA-approved anti-PD-L1 drugs and associated assay systems

Drug	**Nivolumab**	**Pembrolizumab**	**Atezolizumab**
Antibody clone	28-8 (Dako)	22C3 (Dako)	SP142 (Ventana)
Detection kit	EnVision Flex on AutostainerLink 48	EnVision Flex on AutostainerLink 48	OptiView detection and amplification on Benchmark ULTRA
Company	Bristol-Myers Squibb	Merck	Genentech

Table adapted from Hirsch *et al*.(92)

As a result of the above approach each drug is recommended only to be used with its 'companion'/ 'complementary' diagnostic. One of the questions raised is how results from these different assay systems correlate. For this reason, the 'Blueprint PD-L1 IHC Assay Comparison Project' was recently established. Phase 1 of their project included 39 patients with NSCLC. These were assessed for how they stained with four PD-L1 IHC assays used in clinical practice: 22C3, 28-8, SP142, and SP263. Of these 22C3, 28-8, and SP263 produced comparable results. SP142 resulted in a smaller proportion of cells staining positive overall. Interestingly in 14 of 38 cases (37%) the PD-L1 classification would change with different assays (92,93).

However a recent study on 29 TETs (26 thymomas and 3 carcinomas) revealed that immunohistochemical staining for PD-L1 using four different antibody clones (E1L3N, Cell Signaling Technology; 28-8, Epitomics; SP142, Spring Bioscience and CAL1, BioCare) produced significantly different results, highlighting the difficulty in exploiting PD-L1 as a predictive biomarker for anti-PD-1 therapy (94).

1.8 Treatment of TETs

1.8.1 Standard of care for patients with TETs

Currently the standard of care for early-stage TETs is surgery aiming for a complete (R0) resection. For completely encapsulated tumours this this will provide a cure in 90% of patients (4) Complete resection rates are 100% in stage I, 88% in stage II, 49% in stage III and 29% in stage IV disease (95,96).Given the rarity of TETs and the multiple subtypes that exist their management should be based on a multidisciplinary effort. The key issue remains the complete resectability of the tumour as assessed by the imaging modalities used (97). When TETs present at stage III or IVA (according to Masaoka-Koga), a biopsy is essential to confirm the suspected diagnosis. If that is confirmed, the aim for patients that are medically fit, is to administer primary chemotherapy to downstage the tumour followed by radiotherapy or surgery (98). Platinum-based chemotherapy is the standard of care, with cisplatin, doxorubicin and cyclophosphamide and cisplatin and etoposide being the most commonly used regimens. Two to four cycles of treatment are usually administered prior to repeating the imaging to assess the response to chemotherapy and the resectability of the tumour. Resection is usually offered to those where complete excision (R0) of the tumour is thought to be achievable (97).

When patients are evaluated as being inoperable either because of comorbidities or because complete resection cannot be realistically achieved chemoradiotherapy is the treatment of choice, with chemotherapy (cisplatin and etoposide) followed by radiotherapy sequentially.

Cisplatin-based chemotherapy alone is given in all cases that are both unresectable and cannot be irradiated, and in all cases that present with metastatic disease (Stage IVB) as well as cases of recurrence that have not previously received any medical treatment. The aim of treatment in these cases is to control and contain the disease and manage associated symptoms. There is lack of randomised trials to compare the most commonly used regiments that include cisplatin, doxorubicin and cyclophosphamide and cisplatin, etoposide and ifosfamide. Carboplatin and paclitaxel are also used in thymic carcinomas. In rare cases following chemotherapy surgery can be offered when a complete resection of the primary tumour is realistic and known metastatic disease can also be excised completely through metastatectomies. In cases of recurrence the same rules apply and when complete excision of metastatic disease is feasible it is the preferred option. When surgery is not an option, chemotherapy is again the mainstay of treatment. If there has been a significant disease free interval re-challenging with a regimen that has previously been successful is an option taking into account the cumulative toxicity of these drugs.

A number of different second-line regimens can be used, including cis/ carboplatin and etoposide, carboplatin and paclitaxel, or capecitabine and gemcitabine. Thymomas that are octreoscan-positive can also be treated with octreotide monotherapy or steroids (97).

1.8.2 Targeted therapies for patients with TETs

There are currently no licensed targeted treatments for TETs.

Although KIT mutations are rare in TETs, KIT inhibitors such as imatinib, sunitinib, or sorafenib have been used in isolated cases sometimes with prolonged clinical benefit. Of those sunitinib was recently shown in a phase II open-label trial (NCT01621568) to induce an overall response rate in thymic carcinomas of 26% and a disease control rate of 91%. This was significantly lower in thymomas with an overall response rate of 6% and a disease control rate of 81%. Patients in this study received sunitinib upon progression following first-line platinum-containing treatment. There was no patient selection based on any biomarker. The commonest grade three or four treatment side effects were lymphocytopenia (20% of patients), fatigue (20%), and oral mucositis (20%). Five (13%) patients were found to have a decrease in their left-ventricular ejection fraction, with 3 presenting with a grade 3 event. Three (8%) patients passed away while on the study and one was likely related to sunitinib. Therefore, sunitinib can be considered as a second line option for the treatment of thymic carcinomas. Molecular tumour profiling that was performed on these tumours, including KIT, FLT3 and PDGFRA, did not reveal any predictive biomarkers for response to sunitinib (99).

Everolimus, an mTOR inhibitor commonly used in other solid tumours, has shown encouraging results in a recent single arm, single-stage, open label, multicentre phase II trial (NCT02049047) where it exhibited a 22% response rate and 93% disease control rate in TETs in the first 35 patients to register. Patients with TETs that had progressed on first line platinum-containing chemotherapy were eligible to participate in the trial. 20% of patients presented with a serious treatment-related serious adverse event with 3 (of the 7) having to discontinue the therapy. Again there was no biomarker of response for the treatment (100).

Trials with anti-PD-1 inhibitors are currently underway with some encouraging results but also severe autoimmune toxicity**.** There are currently no licensed immune checkpoint inhibitors for the treatment of TETs.

1.8.3 PD-1 inhibitors in TETs: The story so far

A recent Phase II trial (NCT02607631) of pembrolizumab (anti-PD-1) in patients with TETs following progression on fist-line platinum-based chemotherapy recruited 26 patients with thymic carcinoma and 7 with thymomas of all subtypes. The median PFS for the thymic carcinoma cases was 6.2 months but was not reached for the thymomas at the time this study last reported its findings. Although patients that had received treatment for any autoimmune condition within the previous year were not included in the study, immune-related adverse events ≥ grade 3 were common. The three most frequent were hepatitis (12.1%), myocarditis (9.1%), and myasthenia gravis (6.1%). 24.2% (8 of 33) of patients had to discontinue their treatment due to an immune related side effect. This included five of seven patients (71%) with a thymoma (101).

Another phase II study (NCT02364076) of pembrolizumab included 40 patients with recurrent thymic carcinoma that had progressed on first-line chemotherapy. In this study, in an attempt to minimize cases of immune-related toxicity, neither patients with thymomas nor anyone with any history of an autoimmune condition were included. The overall response rate was 22.5%, with 8 partial and one complete response. Five of those responses have so far lasted more than a year. Twenty patients had stable disease and 11 progressed. Twenty-nine of the 40 cases were tested for PD-L1 (Dako 22C3) with only 8 (28%) having ≥50% of tumour cells positive. Of the responders 6 of 9 had a high expression of PD-L1 (66.7%). Mutational burden did not predict for response to treatment as determined by targeted next generation sequencing. Six of 40 patients (15%) had multiple ≥ grade 3 immune-related adverse events. Of these myositis and hepatitis were the commonest, and in three cases both were encountered. Two of the patients had myocarditis, both requiring a pacemaker. Five patients developed hypothyroidism. Three patients, who were all responders had to stop treatment due to immune-related toxicity (102,103).

Another phase I study (NCT01772004) investigated the activity of avelumab (MSB0010718C), a fully human, IgG1 anti-PD-L1 antibody in patients with relapsed TETs following at least one line of standard treatment. Again patients with any history of an autoimmune disorder were completely excluded. Of the 7 patients with a thymoma, 2 (29%) had a confirmed partial response (1 at 20 mg/kg, and 1 at 10 mg/kg) and 2 had unconfirmed partial responses. Two patients had stable disease and 1 (14%) had progressive disease. A single case of thymic carcinoma had stable disease. Five patients (63%) had ≥ grade 3 immune-related adverse events. Four of these required steroid treatment and one was given cyclosporine A. All 4 patients, who had had a partial response experienced immune-related adverse events: 3 had myositis and 1 had enteritis. Remarkably all 3 cases of myositis occurred after only the first dose of avelumab. In 1 of the 4 cases a post-treatment biopsy revealed necrotic tissue only and no viable tumour (104).

Immune-related side effects of checkpoint inhibitors remain a huge concern for all solid tumours and particularly for TETs. This is not a surprise as the inflammatory response induced against the tumour is also induced against normal tissue. This is even more pronounced in cases of dual checkpoint inhibition. While it has become the focus of intensive research to establish who the individuals that are more likely to develop such conditions are, it is also essential to discover biomarkers that will determine who is more likely benefit from checkpoint inhibitors so as to avoid unnecessary exposure. As has been the experience with other treatments and their side effects, it has been suggested that individuals with immune-related side effects are also the ones that are more likely to have tumours that respond to that treatment.

Another very intriguing observation in refractory TETs was recently made in a study on sunitinib, a receptor tyrosine kinase inhibitor to all platelet-derived growth factor receptors, vascular endothelial growth factor receptors and c-KIT amongst other kinases. This showed that following sunitinib administration, PD-1 expression in T-reg and CTLA4 expression in CD8+ T-cells both increased in most patients. This raises the possibility that sunitinib activates the immune system. This also triggered the idea of using combination therapies with sunitinib and immune checkpoint inhibitors, probably in a sequential fashion (99).

1.9 Known molecular changes involved in TET formation

Although TETs have been largely under-researched being a rare tumour type, a summary of the current understanding of the following topics will be presented: Gene mutations, copy number alterations/ aberrations, DNA methylation, gene expression profiling and microRNA expression.

1.9.1 Gene mutations

Genomics is becoming increasingly significant in oncology with tumour types such as melanoma and non-small cell lung cancer (NSCLC) being routinely genotyped to identify targetable mutations. Next-generation sequencing (NGS) even allows clinical trials to evaluate the efficacy of targeted therapies (105). The first step however is to identify driver mutations and key related pathways than can be targeted *in-vitro* and *in-vivo*. This section aims to summarize what is currently known about TETs in this context.

The largest study published to date on whole exome sequencing (WES) of TETs was by Petrini et al, who sequenced 28 TETs. Of these seventeen were thymomas (A, B2 and B3 subtypes only)

and eleven were carcinomas (four SCCs, six undifferentiated and one NETT). This study found 722 somatic single nucleotide variations (SNVs) and 68 insertions/deletions (indels). The mutation burden was very low with an average of 28 mutations per sample (range 3 – 94). Carcinomas were found to have a significantly higher number of mutations compared to thymomas (Mann-Whitney; $p = 0.001$). A custom 197–cancer gene panel was used to sequence 26 of the TETs that had already undergone WES and an additional 26. The two methods were found to correlate highly (chi-square $p < 0.001$). The major finding in this study was the presence of a highly recurrent missense mutation (chromosome 7 c.74146970T>A) in *GTF2I*, which was highly prevalent in the more indolent subtypes (A and AB) as determined by expanding to a total of 274 TETs using targeted methods. This is described in depth in section 4.2. From the total of 52 cases that were sequenced with either method several genes, known to be involved in carcinogenesis, were mutated more than once in carcinomas and also in isolated thymoma cases. These include *TP53* (four carcinomas and two B3 thymomas), *CYLD* (three carcinomas), *BAP 1* (two carcinomas and one thymoma), *CDKN2A* (two carcinomas), *RET* (one thymoma and one carcinoma), *PBRM1* (two carcinomas), *ALK* (two thymomas), *PPP2R1A* (two thymomas), and one of each of an *HRAS*, a *KRAS* and an *NRAS* hotspot mutation. It is worth pointing out that no B1 or MNT thymomas were included in the analysis (106).

The Cancer Genome Atlas (TCGA) research network recently presented limited findings from a comprehensive molecular analysis of 117 TETs they recently conducted. They performed whole-exome, transcriptome, methylome and targeted proteome analysis. The median patient age was 60 years (range 17-84); and the male to female ratio was 52:48. Of the 117 cases, 107 were thymomas (10 A; 48 AB; 12 B1; 25 B; 10 B3 and 2 micronodular) and were 10 carcinomas (one of which was a large cell neuroendocrine carcinoma). The breakdown of the Masaoka Stage at presentation was the following: 36 with Stage I, 39 with Stage IIA, 19 with Stage IIB, 15 with Stage III, 1 with Stage IVA and 5 with IVB. Thirty-two of 117 patients had MG (27%). They confirmed that TETs have a very low tumour burden and was in fact the lowest of all adult tumour types they have examined. They also confirmed the findings of a highly recurrent mutation in GTF2I (overall frequency across all subtypes was 39%) predominantly in A and AB thymomas, with the mutations being present only in TETs and no other solid tumour in the entire database. There were no other frequently recurrent mutations across any TET subtypes. HRAS, NRAS and TP53 mutations were amongst the other mutations found at low frequency. MG was found to correlate with increased aneuploidy and overexpression of muscle auto-antigens. It was also commoner in the B subtypes compared to A, AB, MNT and SCCs. All samples were found to belong to one of four clusters based on all the above molecular tests (107).

A very recent study by Zehir et al. recently published some of their results from a ‘large-scale, prospective clinical sequencing initiative’ from the Memorial Sloan Kettering Cancer Center. They

used a hybridization capture–based NGS panel (MSK-IMPACT) that can detect 'all protein-coding mutations, CNAs, and selected promoter mutations and structural rearrangements in 341 cancer-associated genes'. They included 12670 tumours from 11369 patients that had 62 different tumour types (over 300 subtypes), all presenting at an advanced stage. Matched normal tissue and detailed pathological and clinical data were available for the vast majority. Of these there were only eighteen patients with TETs, ten of whom had thymic carcinomas and the other eight, thymomas (subtype information was not available). Although sequencing results for this cohort were not presented they are publically available at: http://cbioportal.org/msk-impact. As this was a pre-selected gene panel of known cancer-associated genes, *GTF2I* was not included. Also although the tumours studied were all at an advanced stage, which is key in potentially revealing genes that are linked with more aggressive tumour behaviour, this meant that approximately half of these were derived from biopsies, rather than surgical removal of the thymus. Also, about 50% of the samples were not from the primary tumour but were obtained from metastatic deposits, which is likely to influence their genetic make-up. Of the eight thymomas, somatic mutations were identified in only one gene (*CREBBP*, frameshift deletion) in one sample and fusions were identified in *ERCC3, PAX8, CRIM1* and *DRC1* in the same sample. Of the ten carcinomas, only one did not have any mutations. In terms of recurrence, *KMT2D* was mutated in three samples (30%), followed by *TP53* and *CDKN2A* that were mutated in two samples each (20%). In one of these cases a *KMT2D* and *CDKN2a* mutation co-existed. Of the remaining four carcinomas, one had a *KIT* W557S hotspot mutation, another had an *NRAS* Q61K hotspot mutation and a *PIK3CA* E542K hotpot mutation, another had an *EZH2* mutation only and one had a *CDK12* mutation only (105).

A recent study focused on more aggressive TETs, investigating eighteen type A (for comparison), nineteen B3 thymomas and 35 carcinomas (34 SCCs and one lymphoepithelioma-like carcinoma) and six cases of normal, non-neoplastic thymus all derived from FFPEs. Hotspot mutations of 50 genes (oncogenes and tumour suppressor genes) were part of the targeted panel used. The most commonly mutated gene was TP53, which was mutated in nine of 35 carcinomas (26%). Four of 35 (11%) carcinomas harboured mutations in CDKN2A, and two (6%) in each of the fibroblast growth factor receptor 3 (FGFR3) and KIT. ALK, ATM, ERBB4 and NRAS each were mutated in one carcinoma (3%). B3 thymomas were characterized by an almost complete absence of any mutation with only two mutations in non-coding regions of the SMARCB1 and STK11 genes discovered. Three of the eighteen (17%) type A thymomas harboured an HRAS hotspot mutation, one of which was an atypical A thymoma with a lung metastasis (108).

A number of genes that have been targeted in other tumour types have also been the focus of other, smaller studies on TETs. Although a small number of patients respond to certain such therapies the lack of a systematic, biomarker-based approach has overall resulted in overall disappointing results. Some of the most salient findings are summarized

KIT

KIT is a trans-membrane growth factor with tyrosine kinase activity. KIT overexpression has been observed in 73-86% of thymic carcinomas but in only 0-2% of patients with thymomas. The rate of KIT gene mutations in carcinomas is very low at around 9-12.5 % (34,109-112). It has been suggested that KIT is used as a diagnostic marker of carcinoma of thymic origin vs. other mediastinal tumours and vs. thymomas (110). The proportion of patients with KIT mutations is very small, expression of KIT does not directly correlate with mutation rate and trials of imatinib on patients with KIT-mutant carcinomas have given mixed results without an obvious biomarker for patient selection (113-116) Two reports of single patient cases with heavily pre-treated thymic carcinomas and different KIT mutations have demonstrated an impressive, prolonged clinical and radiological responses. In a case report of a single patient with a c-KIT mutation a very good partial response to sorafenib treatment was achieved (117).

Epidermal growth factor receptor (EGFR)

EGFR is a tyrosine kinase receptor (TKR) that is involved in the genesis of many epithelial tumors. Tyrosine kinase inhibitors (TKIs) have been used successfully against activating mutations of the EGFR in non-small cell lung cancer (118-120). EGFR expression is observed in 20-100% of TET cases investigated by immunohistochemistry (121-128). High EGFR expression is associated with advanced tumour stage and aggressive phenotype (123,127). EGFR gene amplification has been identified in 0-30% of TETs (125,129,130). Somatic activating EGFR mutations are extremely rare in TETs (130). In a study of 26 patients with TETs, gefitinib achieved 1 partial response, while 14 patients had stable disease (131-133). Other studies have shown no response (134). Erlotinib has been evaluated in combination with bevacizumab with a minority achieving stable disease only (131-133,135). A handful of responses to cetuximab have been reported (136,137) An ongoing phase II clinical trial is evaluating the role of neoadjuvant cetuximab in combination with chemotherapy (ClinicalTrials.gov Identifier: NCT01025089). EGFR TKIs have not currently got an indication for administration to patients with TETs.

Human epidermal growth factor receptor 2 (HER2)

HER2 belongs to the EGFR family and is also a proto-oncogene that encodes a TKR, whose significance is already established in breast cancer (138). In TETs its role is considerably less clear and it has only been explored in three studies where a low percentage of HER2 positivity was seen, with only a single case of amplification. (123),(125). Based on these studies anti HER2 treatment is not indicated in TETs.

Kirsten rat sarcoma viral oncogene homolog (KRAS)

KRAS is a member of the RAS family of proto-oncogenes that are downstream of the EGFR pathway. Mutations of KRAS have been described in other solid tumours, including lung cancer

(139) and these are know to confer resistance to EGFR inhibitors (140) Based on two studies activating KRAS mutations have been observed in only three of 63 TET cases (122).

Vascular endothelial growth factor receptor (VEGFR)

VEGF is an important angiogenic factor that is involved in normal endothelial cell proliferation and migration as well as in carcinogenesis. Protein expression has been demonstrated in up to 62.5% - 75% of patients with TETs (141,142) although this is much less in other studies (117,143). Anti-angiogenesis treatments, such as sorafenib, SU14813 (TKIs) and aflibercept have been tested in the phase 1 setting and have shown a very modest benefit although some responses have been prolonged (117,143-145). Sunitinib that targets VEGFR amongst other receptors has yielded very encouraging results despite an absence of a biomarker for its action and is discussed in more detail in section 1.8.2. A recent study on B3 thymomas and carcinomas revealed that VEGF expression in carcinomas was higher than in B3 thymomas and higher expression in either was linked to poor overall survival (146).

Mammalian target of rapamycin (mTOR)

MTOR is a serine/threonine kinase and belongs to phosphatidylinositol-3 kinase (PI3K)-related kinases (PIKKs) family. It regulates cellular metabolism, growth, and proliferation. In a single Phase I study, 10 of 21 patients with a TET received treatment with an mTOR inhibitor. Median time to treatment failure (TTF) for these patients was 11.6 months vs. 2.3 months on their last conventional treatment (147). Everolimus is also currently used in clinical trials with some encouraging results despite the absence of a biomarker for response. This study is discussed in section1.8.2.

Histone deacetylases (HDACs)

HDACs play an important role in regulating gene expression through chromosome remodeling (131). Belinostat, an HDAC pan-inhibitor, has been investigated in phase I and II trials of patients with advanced TETs showing very rare but prolonged responses despite the absence of a clear biomarker (148,149). Belinostat has shown very promising results in combination with chemotherapy and a Phase II trial is currently ongoing including molecular profiling of tumours (38,39,150).

Insulin-like growth factor 1 receptor (IGF-1R)

Overexpression of the TKR IGF-1R has been described in many solid tumors, including TETs and is associated with tumor development and progression (124,151). IGF-1R protein overexpression has been demonstrated in up to 86% of thymic carcinomas and 43% of thymomas (124). Heat shock protein 90 (Hsp90) inhibitors significantly reduce viability of TETs, by acting upon a number

of targets, including IGF-1R (152). IGF-1R protein levels have been shown to be a negative prognostic indicator (153). Figitumumab and cixutumumab, monoclonal antibodies against IGF-1R, have shown activity in isolated thymoma cases but not in carcinomas (154),(155).

Bcl-2

Bcl-2 is a proto-oncogene known to regulate apoptosis. Array comparative genomic hybridization in 59 TETs has revealed recurrent copy number gain of Bcl-2 underlying its likely involvement in carcinogenesis and potential value in diagnosis and prognosis of TETs (156). This finding was validated in a further study from the same group as described in section 1.9.2.

Tropomyosin receptor kinase A (TrkA)

TrkA activation promotes tumor growth, probably through activation of the MAPK pathway. In a study of 60 patients with TETs by Kelly et al., a copy number gain of TrkA was demonstrated in 45% and 72% of thymomas and carcinomas respectively. PHA-848125-AC, an oral, potent inhibitor of the CDK2/cyclin A complex and TrkA has shown very promising results in Phase 1 trials and is now being investigated further (132,157).

Src family kinase

Saracatinib (AZD0530), a small-molecule inhibitor of Src was given in a Phase II trial in 21 patients with previously treated advanced TETs. It showed no clinical activity (158).

p53

p53 protein overexpression is common in TETs. Several studies have detected p53 immunoreactivity in 33.3 - 100% of cases (157,159-164). Mutational studies have been few and contradictory, some having detected no mutations while others have detected a high prevalence of p53 mutations (157),(161). p53 expression has been associated with higher malignant potential in some studies (157,159-161).

RB and p16INK4

The expression of p16INK4 and RB has been investigated in TETs in a single study so far. p16INK4 is a protein that inhibits phosphorylation of the RB protein leading to blockade of the cell cycle at G1 to S. In one study p16INK4 expression was found in 21 of 36 thymomas and 1 of 3 carcinomas. RB was identified in 28 of 36 thymomas and 3 of 3 carcinomas (157).

1.9.2 Copy number alterations/aberrations (CNA)

The most systematic CNA determination in TETs to date has been performed by TCGA through WES (107). Although results have not been published by TCGA yet, the data are publically available. These have also been analysed by Lee et al. and a summary of their findings can be found in chapter 1.9.4.

Comparative genomic hybridization (CGH) is a method commonly used to identify areas of chromosomal imbalances, including deletions and amplifications.

Petrini et al used array CGH to look for CNAs in 65 TETs. CNAs affecting entire chromosome arms were found in more aggressive TET subtypes (CA, B3 and B2) and were very infrequent in A and AB subtypes, hence they divided TETs into two main clusters. The most frequent chromosomal arm deletions amongst all TETs affected chromosome 6 (6p, 26%; 6q, 29%), 3p (22%) and 13q (18%). The most frequent arm-level gains were in chromosomes 1q (55%), 7 (7p, 20%; 7q, 15%) and 20p (17%). Using GISTIC, which allows the identification of 'peaks' of focal copy number gains or losses, seven peak regions of copy number gains (two of which were in chromosome 1) and ten of copy number losses (two of which were in chromosome 6 and two in 9) were identified. One of the amplifications was of the BCL2 locus (chromosome 18q). This correlated with increased BCL2 expression as seen in RNA sequencing data from the same group (106). This was an expanded cohort from what had been previously reported by the same group in 59 cases. What had been reported for this smaller number of cases was that the most frequent copy number gains were of PBX1 (chr 1q) and GLI4 (8q) and most frequent losses were of FOXC1 (6p) and IGF2R (6q) ($p < 0.05$) (165). A summary of these results can be found in Tables 1.7 (CN gains) and 1.8 (CN losses).

Girard et al. also used CGH to analyze 45 TETs (eight Type A, 22 B2, 8 B3 thymomas and seven carcinomas). This study confirmed that increasing number and extent of aberrations are seen from indolent to more aggressive TETs. B2, B3 thymomas and carcinomas had gains of chromosome 1q and losses of 1q, 3p and 17p. Gains of 4p, 17q and 18 were seen in B3 thymomas and carcinomas only as were losses in 6q, 13q and 16q (122).

Lee et al. used CGH to investigate CNAs in 39 thymomas (six A, eleven AB, seven B1, seven B2 and eight B3). Type A thymomas had the fewest CNAs overall. Results for all subtypes are summarized in Tables 1.8 and 1.9 (166).

Inoue et al. investigated 57 TETs, again with a notable exception of B1 thymomas. Overall, CNAs were rare in Type A thymomas, with many more were seen in B2, B3 thymomas and carcinomas.

AB thymomas were somewhere in the middle. The commonest deletions across all subtypes were seen in chromosomes 6p and 6q followed by 13q. Amplifications were overall a lot fewer with 3 seen in chromosome 16 (AB, B2 and B3 thymomas) (167).

Penzel et al. used CGH to investigate CNAs in 28 TETs (Type A, B2, B3 thymomas and carcinomas). In Type A thymomas there were recurrent gains in 17p and 20q (3 of cases for both). Gains were also seen in 16q in 2 cases and all other changes were isolated. In type B thymomas six of ten cases had gains in 1q and two of ten in each of 16p, 20 and Xq. Recurrent losses were seen in three of ten cases in 6p and 3p and in two cases in 12p and 16q. The eight carcinomas studied were of five different subtypes. Most frequent gains were seen in 1q and Xq (five of eight) and 4q (three of eight). Recurrent losses were seen in five of eight samples in 16q, in four cases in 6q, in three cases in 17p and two cases in 9p (168).

Zettl et al. looked into 37 TET cases (twelve Type A, sixteen B3 and nine SCCs) with CGH. No recurrent gains or losses were identified in Type A thymomas. In B3 thymomas recurrent gains were seen in 1q (eleven of sixteen) and recurrent losses in chromosome 6 (six cases) and 13q (five cases). In SCCs recurrent gains were seen in 1q (five of nine cases) and 18 (three cases). Recurrent losses were observed in 16q (six of nine cases), 6 (four cases) and 3p and 17p (three cases) (169).

Nakanishi et al. used OncoScan (affymetrix) to determine CNAs in 10 thymic carcinomas. They recently reported their findings that confirmed gains in 1q and losses in 6p, 6q and 16q (170).

In a small, recent study on B3 thymomas and thymic carcinomas, fluorescence in situ hybridization (FISH) detected CDKN2A deletions in 38% of carcinomas studied, in TP53 in 32% of carcinomas and in ATM in 8%. Only one B3 thymoma also had a CDKN2A deletion and none were found in type A thymomas that had been used for comparison (108).
Moreira et al performed exome-capture of 275 cancer-related genes in fifteen thymic carcinomas and six B3 thymomas. The median number of mutations per tumour was one (range zero to 26). Recurrent mutations were seen in carcinomas in TP53 (4), KDM6A (3), SMAD4 (2), CYLD (2) SETD2 (2), MLL3 (2), and MLL2 (2). There were also mutations in isolated cases of the following genes: ARID1A, EP300, BAP1 and NF1 amongst others. BCOR was mutated in three B3 thymomas (171).

Overall it is striking that there have hardly been any B1 thymomas investigated in any of these studies although this may simply reflect this is a less common TET subtype. It is also apparent that there are overall fewer CNAs in A and AB thymomas compared to B2 and B3 thymomas and carcinomas. All studies also reveal a lack of recurrent chromosomal gains in TETs. The only

exception is chromosome 1q, where gains have been reported in B2 thymomas in three of five studies, and in B3 thymomas and carcinomas in five of six studies. Losses in 6q have been reported across all TETs. Across B2 and B3 thymomas and carcinomas recurrent losses have been reported in 3p, 6p and 6q. It is striking that deletions in the whole of chromosome 6 or either arm are reported in all studies that have looked into B3 thymomas and carcinomas. Deletions in 13q have been observed in four of six studies in B3 thymomas but only once in carcinomas. 16q and 17p deletions are also frequently observed in B3s and carcinomas.

The clinical significance of these findings remains to be determined.

Table 1.8. Frequently recurrent chromosomal copy number gains in TET subtypes in selected studies (106,122,165-170)

TET subtype	**Petrini et al**. (n=65)	**Girard et al.** (n=48)	**Lee et al.** (n=39)	**Inoue et al.** (n=57)	**Penzel et al.** (n=28)	**Zettl** (n=37)	**Nakanishi et al.** (n=10)
A	-	-	-	16p	16q, 17p, 20q	--	*N/A*
AB	-	-	-	16q	-	*N/A*	*N/A*
B1	*N/A*	*N/A*	9q	*N/A*	-	*N/A*	*N/A*
B2	1q, 7p, 7q, 20p	1q, 4p, 17q, 18	-	16p, 16q	1q, 16p, 20, Xq	*N/A*	*N/A*
B3	1q, 7p, 7q, 20p	1q, 4p, 17q, 18	1q	16q	1q, 16p, 20, Xq	1q	*N/A*
CA	1q, 7p, 7q, 20p	1q, 4p, 17q, 18	*N/A*	-	1q, 4q, Xq	1q, 18	1q

Recurrent 1q gains are highlighted in red font.

N/A signifies that these subtypes were not included in the study. The symbol '-' signifies the lack of any findings despite the subtype having been included in the study.

Table 1.9. Frequently recurrent chromosomal copy number losses in TET subtypes in selected studies (106,122,165-170)

TET subtype	**Petrini et al.** (n=65)	**Girard et al.** (n=48)	**Lee et al.** (n=39)	**Inoue et al.** (n=57)	**Penzel et al.** (n=28)	**Zettl et al.** (n=37)	**Nakanishi et al.** (n=10)	**Lee et al. (TCGA)** (n=120)*
A	-	-	2, 4, 6q, 13	6p, 6q, 13q, 17p	-	-	*N/A*	
AB	-	-	2, 4, 5, 6q, 8, 13, 18	6p, 6q, 7p, 8p, 13q, 17p	-	*N/A*	*N/A*	
B1	*N/A*	*N/A*	1p, 2q, 3q, 4, 5, 6q, 8, 13	N/A	-	*N/A*	*N/A*	
B2	3p, 6p, 6q, 20p	1q, 3p, 17p	1p, 2q, 3q, 4, 5, 6q, 8, 13, 18	5q, 6p, 6q, 7p, 8p, 16q	6p, 3p, 12p, 16q	*N/A*	*N/A*	2q, 3p, 6p, 6q, 7q, 9p, 9q, 10q, 11q, 22q
B3	3p, 6p, 6q, 20p	1q, 3p, 6q, 13q, 16q 17p	2q, 4, 5, 6, 8, 12q, 13, 18	5q, 6p, 6q, 7p, 8p, 13q, 16q, 17p	6p, 3p, 12p, 16q	6, 13q	*N/A*	2q, 3p, 6p, 6q, 7q, 9p, 9q, 10q, 11q, 22q
CA	3p, 6p, 6q, 20p	1q, 3p, 6q, 13q, 16q 17p	*N/A*	6p, 6q, 16q, 17p	6q, 9p, 16q, 17p	3p, 6, 16q, 17p	6p, 6q, 16q	

Recurrent chromosomal (arm) deletions across studies highlighted in different colours.

N/A signifies that these subtypes were not included in the study. The symbol '-' signifies the lack of any findings despite the subtype having been included in the study.

* Although the entire cohort of TET samples consisted of 120 tumours, these results are based on a total of 25 B2 and B3 thymomas

1.9.3 DNA methylation

Silencing of tumor suppressor genes, including *FHIT*, *MLH1* and *CDH1* by hypermethylation of their promoter is known to occur in TETs (172). These have been correlated with Masaoka stage and WHO subtype, as has DNA hypomethylation. Aberrant methylation affects other genes such as DAPK1, *CDKN2A/p16*, *MGMT* and *HPP1* (173). Methylation of the MGMT promoter and loss of its protein expression has been found to be significantly more frequent in thymic carcinomas vs. thymomas (74% vs. 29%), and in TETs diagnosed in stages III/IV vs. I/II (174). Methylation of the of the CDKN2 promoter has been described in up to 14% of thymomas vs. 25% of carcinomas (157). The significance of DNA methylation and its potential therapeutic implications remain to be elucidated (132).

1.9.4 Gene expression profiling

Badve et al. were the first group to conduct a retrospective RNA microarray analysis in 34 TET with AB and B thymomas. Unsupervised cluster analysis of 8260 genes allowed four clusters to be identified. Although these contained samples of certain subtypes more than others, AB thymomas were present in all four and B1-B2 in three of the four. Ingenuity pathway analysis revealed that tumours from metastatic cases had an increased expression of COL1A1 (thirteen-fold), AKR1B10 (six-fold) and JPH1 (three-fold) compared to non-metastatic cases (175).

A retrospective study by Gökmen-Polar in 111 TETs was able to determine a nine gene signature that can predict the likelihood of metastasis. This divided all cases to a low-risk and a high-risk group with more accuracy than the current staging system. The genes whose expression was linked with a higher likelihood of metastasis were AKR1B10, JPH1, and NGB and those linked with a lower metastatic potential were DACT3, SLC9A2, PDGFRL, FCGBP, PRRX1, and SERPINF1. In multivariate analysis these genes were the only independent predictor of the risk of developing metastasis (HR 5.26, 95% CI = 1.12–24.8; $p = 0.036$) (176).

Following the discovery of the *GTF2I* mutation in TETs, Petrini et al. performed RNA-sequencing (RNA-seq) analysis and demonstrated the expression of *GTF2I* in all five type A and two AB thymomas that had been whole exome sequenced. Mutated and wild-type alleles were both expressed. The cDNA underwent Sanger sequencing, which confirmed the presence of *GTF2I* mutant alleles in the transcripts of all samples. Transcriptome sequencing proved that the β and δ isoforms of *GTF2I* are the main isoforms expressed in TETs compared to the other three isoforms. qRT-PCR for the specific isoforms confirmed their expression and the T>A mutation was detected

in both isoforms. This results in p.Leu404His in the β isoform and p.Leu383His in the δ isoform of TFII-I (106).

The same group also used RNA-seq data to look for fusions in TETs. Fusions were detected in seven of the 25 tumours that were evaluated. Every case had an average of one fusion transcript (range, 0–16). There was a single B2 thymoma that expressed 16 fusions, which was considerably higher than any of the other TETs. *GTF2I* was not involved in any of these (106)

Nakanishi et al recently investigated the transcriptome of ten thymic SCCs using the GeneChip Human Transcriptome Array 2.0 (affymetrix). Three Type A thymomas were used for comparison. Gene expression analysis showed down-regulation of BRD2, HSP90AB1, FOXO3, and MARCKS in chromosome 6, and MTSS1L in chromosome 16q in the thymic SCCs (170). The full publication is still pending.

TCGA have been the first group to conduct transcriptome analysis of TETs. A recent study by Lee et al. sub-grouped TETs into four molecular subtypes, analyzing 120 TET cases from TCGA, incorporating analyses from WES, mRNA expression, and somatic CNA. These results were then validated in two independent TET cohorts. The first molecular subtype included all cases that were *GTF2I* mutant. GTF2I was confirmed to be the most frequently mutated gene in TETs (38% across all subtypes in this cohort) and its mutation status was well-reflected in mRNA expression patterns as all the cases with mutations clustered together in hierarchical clustering analysis. As already known, the GTF2I group consisted primarily of A and AB thymomas and was associated with favorable OS. The second molecular subtype was identified by investigating GTF2I wild-type tumours only and using unsupervised mRNA clustering. This resulted in the identification of a group of TETs that demonstrated significant enrichment in the expression of genes associated with T cell signaling (named TS group, accounting for 33% of TETs). TETs in this group were of B1 and B2 subtypes. The remaining two subtypes were defined using data from somatic CNA. These were the chromosomal stability (CS; 8% of all TETs) and instability subtypes (CIN; 21%). The CS group included almost exclusively B2 thymomas whereas the CIN group included B2, B3 thymomas and SCCs. Significant deletions were seen in 9p21.3 (which contains tumour suppressors such as CDKN2A and CDKN2B), 22q12.1, 6p25.2, 3p22.2, 2q37.1, 6q21, 10q26.3, 9p13.11, 22q13.32, 7q36.3, 11q22.1, and 9q22.33. In the CIN group, CDKN2A at 9p21.3 was significantly deleted in seven of 25 cases (28%). Differences in the genomic stability of these two groups were also expressed at the mRNA level. Amongst the four established groups the CIN was linked with least favorable OS. Another interesting finding was high mRNA PD1 expression and CD8A in the TS group (177).

1.9.5 MicroRNA (miRNA) expression

MiRNAs are non-coding RNAs involved in post-transcriptional regulation. Altered miRNA expression is correlated with the initiation and progression of several tumors (178). Ganci et al. investigated miRNA expression by microarray analysis of formalin-fixed paraffin-embedded (FFPE) tissue from 54 TETs and matched normal tissue from 12 of these patients. In comparison to normal tissue, in TETs, 56 miRNAs were up-regulated and 31 were down-regulated. Differential miRNA expression was observed between thymomas and thymic carcinomas and between arbitrary TET subtype groups. Of those, miR-21-5p, miR-34b-5p and miR-141-3p were the most up-regulated and miR-486-5p and miR-145-5p were the most down-regulated in 54 TETs vs. twelve normal samples using an array method and in 20 TETs vs. five normal samples by RT-qPCR. miR-21 has been found to be up-regulated in the majority of carcinomas, including lung and breast carcinomas and melanomas, amongst others. It has also been shown to target PTEN, an important tumour suppressor gene. The downregulation of miR-145-5p has been linked with overexpression of the EGFR pathway. Bioinformatics analysis suggested possible molecular pathways targeted by differentially expressed miRNAs that include important proteins such as PDGFR, IGF-1R, c-kit, Notch, Hsp90 and p53, which are likely to be involved in thymic carcinogenesis, progression and metastasis (179).

A recent study by Radovich et al used next-generation RNA sequencing on a total of 48 TETs and three cases of normal thymus gland. Hierarchical clustering proved there was a 100% agreement between gene expression clusters and TET subtype. The most significant finding was a large microRNA cluster on chr19q13.42, which was overexpressed in A and AB thymomas whereas there was no expression in any of the other subtypes. As this cluster activates the PI3K/AKT/mTOR pathway an AB cell line was treated with eight PI3K or dual PI3K/mTOR inhibitors, with two of the dual inhibitors causing the most significant reduction in cell viability (180).

1.10 Conclusions

It is clear that despite recent advances and an increased interest in TETs there is still a significant need to identify novel therapies for thymic malignancies. The identification of genetic lesions is a first step in this process. Identification of predictive and prognostic biomarkers is also lacking. The above data also indicate that TETs are comprised of a number of different subsets, which sometimes coincide with their histological subtype that can be defined at the molecular level. *GTF2I* is mutated in 40% of TETs, the majority being A and AB thymomas. For Type B thymomas and carcinomas (SCCs and NETTs), all known genomic changes only account for a very small

proportion of patients and for these the key drivers are yet to be discovered underscoring the importance of this research project.

1.11 Hypothesis

1. Different TET subsets can be defined and differentiated at the molecular level.
2. The key drivers leading to TETs are yet to be discovered and such information could reveal novel therapeutic targets.

1.12 Aims

The primary aim of this project is to identify genetic alterations leading to thymic epithelial malignancies and to identify biomarkers that differentiate thymomas versus thymic epithelial carcinomas. Access to clinical information will also allow me to look for correlations between these DNA changes and other conditions, such as autoimmune conditions and prior, synchronous or metachronous malignancies, as well as survival and tumour recurrence.

The specific aims are:

1. To identify a cohort of TETs and corresponding clinical information.
2. To use state of the art molecular approaches to identify genetic alterations that drive the development of thymic epithelial carcinomas versus thymomas.

Chapter 2

Materials & Methods

2.1 Materials

Table 2.1. Kits and sets used

Name	Supplier	Catalogue Number
AllPrep DNA/RNA FFPE Kit (50)	Qiagen, Crawley, UK	80234
AllPrep DNA/RNA/miRNA Universal Kit (50)	Qiagen, Crawley, UK	80224
HotStarTaq Plus PCR Master Mix Kit (250)	Qiagen, Crawley, UK	203646
QIAamp DNA FFPE Tissue Kit (50)	Qiagen, Crawley, UK	56404
Quant-iT™ dsDNA Assay Kit, broad range	ThermoFisher Scientific, Waltham, MA, USA	Q-33130
Quant-iT™ PicoGreen® dsDNA Assay Kit	Life Technologies, California, USA	P11496
RNase-Free DNase Set	Qiagen, Crawley, UK	79254
TissueRuptor (230 V, 50/60 Hz, EU/CH)	Qiagen, Crawley, UK	9001272
TissueRuptor Disposable Probes	Qiagen, Crawley, UK	990890

Table 2.2. Reagents used

Reagent	Supplier
100-bp DNA ladder	New England Biolabs, Herts, UK
1X Tris-Borate-EDTA buffer	Sigma-Aldrich, Dorset, UK
Agarose	Sigma-Aldrich, Dorset, UK
Deionised distilled water	MilliQ, Hertfordshire, UK
Ethanol	Fisher Scientitfic, Loughborough, UK
GelRed	Biotium, California, USA
Histoclear	Agar Scientific, Essex UK
Hyclone™ Water	Fisher Scientific, Loughborough, UK

Table 2.3. Antibodies used for immunohistochemistry

Antibody Description	**Supplier**	**Product Code**
Rabbit polyclonal anti-CD137 (**4-1BB**) antibody	Abcam, Cambridge, UK	Ab197942
Mouse monoclonal [6A1] anti-CD276 (**B7-H3**) (clone 6A1) antibody	Abcam, Cambridge, UK	Ab105922
Mouse monoclonal anti-human **CD45,** Leucocyte Common Antigen, Clone 2B11 + PD7/26	Dako, Agilent Technologies, California, USA	IS751
Cytokeratin MNF116	Dako, Agilent Technologies, California, USA	M082101-2
Rabbit polyclonal anti-**HGF** antibody	Sigma-Aldrich, Dorset, UK	HPA040360
Mouse monoclonal anti-**LAG-3** (17B4) antibody	Novus Bio, Abingdon, UK	NBP1-97657
Mouse monoclonal anti-**PD-1** (NAT105) antibody	Ventana, Roche, Basel, SWI	07099029001
Rabbit monoclonal anti-**PD-L1** antibody [28-8]	Abcam, Cambridge, UK	ab205921
Rabbit polyclonal anti-CD273/ **PD-L2** antibody	ThermoFisher Scientific, Waltham, MA, USA	PA5-20344
Rabbit polyclonal anti-**PDLIM1** antibody	Abcam, Cambridge, UK	Ab64971
Rabbit polyclonal anti-**TIM-3** antibody	Abcam, Cambridge, UK	Ab185703
Rabbit monoclonal anti-**VISTA** antibody	Cell Signalling Technology	64953

2.2 Methods

2.2.1 Thymic epithelial tissue/ DNA ascertainment:

1. The Royal Brompton Hospital (RBH) and Harefield NHS Trust Biomedical Research Unit (BRU) BioBank and Diagnostic Bank harbours an extensive range of specimens from highly unique patient populations, including thymomas and thymic carcinomas. It prospectively collects, processes and stores specimens, together with clinical data, consented for donation to research by patients undergoing treatment at the Trust. It provides a standardised, equitable, accessible and transparent resource for researchers. Standardized operating procedures for sample collection and storage are in place and have been developed with rigorous quality assurance (181). The BRU Cancer consortium has approved the use of this material and ethics approval has been granted (Ethics Ref No: 10/H0504/9). The following were obtained:

    a) Formalin-fixed, paraffin-embedded (FFPE) tissue blocks from patients diagnosed with a TET at RBHT that had been removed during a thymectomy or metastatectomy and were surplus to diagnostic requirement.
    b) Snap frozen tissues from patients with a TET obtained in a similar fashion that were surplus to diagnostic requirement. At least one tissue sample from each of TET and normal tissue for each patient that has been collected since 2010 was stored immediately in RNA*later.*

    In 76% of all cases matched normal (non-malignant) tissue was also collected.

2. A collaboration between Imperial College London (ICL) and the King's Health Partners (KHP) Cancer BioBank gave access to DNA from FFPE and frozen TET tissue from patients that had been treated at Guy's and St Thomas's NHS Foundation Trust. DNA extraction was performed exclusively by KHP and anonymized clinical data were also received.

2.2.2 Patient data

Patient demographic and clinical data were collected by myself in a link-coded manner on a password-protected excel spreadsheet. The data fields completed for all patients where available are summarized in Table 2.4.

Table 2.4. Clinical data fields collected

Patient Data	TET Data
Date of birth	Stage at diagnosis
Date of diagnosis	Histological subtype
Gender	Chemotherapy Y (Type)/N
Smoking History	Best response to chemotherapy
Prior/ synchronous/ metachronous malignancies	Radiotherapy Y/N
Autoimmune conditions	Response to radiotherapy
On/ off treatment for AC	Date of surgery
Date of death or last follow up	Complete resection Y/N
Family history of any solid malignancy	Date of disease relapse

2.2.3 Histologic assessment of tumour samples

Snap-frozen and FFPE samples were subject to histological assessment by either of two certified, expert histopathologists (Professor Nicholson or Dr Monterro-Fernandez). Frozen tissue was weighed and cut to approximately 10mg pieces when the starting amount was in excess of that. Each sample was embedded in Optimal Cutting Temperature Compound (OCT) and two 5-micron sections were cut on a microtome-cryostat. Histologic assessment involved application of haematoxylin and eosin (H&E) staining. The oxidation product of haematoxylin is haematin, the active ingredient that stains the nuclei blue. Eosin will stain the other structures red and is highly soluble in water. Over-staining was washed out in running water before fixation. FFPE samples already had matched H&E histological slides (one per sample) in all cases. This was performed to confirm the diagnosis (with cross-reference to the original histopathological report), to provide information on the percentage of viable-appearing tumour cells (vs. necrotic) on each sample and to choose the ones with the highest tumour content to be used for DNA/ RNA extraction. The inclusion threshold for samples in this study was at least 60% viable-appearing tumour cells, which is in accordance with the International Cancer Genome Consortium (ICGC) guidelines (182) while taking into account the rarity of this tumour type.

2.2.4 Polymerase chain reaction (PCR)

As described in Chapter 1, in June 2014, Petrini and co-workers reported in Nature Genetics that a single frequent recurrent missense mutation (T>A, at the same position on chromosome 7, 74146970) in the *GTF2I* gene was strikingly prevalent in A and AB thymomas (106). We

examined the frequency of this alteration in our patient cohort using PCR to amplify the particular region followed by Sanger sequencing.

PCR was performed with semi-nested primers (same reverse primer was used) to increase specificity of DNA amplification from the FFPE specimens. The second forward primer was specific for the *GTF2I* gene whereas the reverse anneals to both *GTF2I* and two pseudogene sequences. See details regarding *GTF2I* pseudogenes in Chapter 1.

The primers for PCR were the following:

Forward primer 1 (P_{F1})	AGAGCACTACAGATTCAATGA
Reverse primer (P_R)	TCAACAGGAATCCAAGAGTCTTA
Forward primer 2 (P_{F2})	ATCCCGTACCCTCTTTTCC

To design the primers the genomic location of the mutation was identified in the UCSC Genome Browser. A sequence of 50 bases upstream and 50 downstream was selected:

TTTTAGAAGGCCATCTACTTACGGAATTCCTCGCCTGGAGAGGATATTACTTGCAAAGGA AAGGATTCGTTTTGTGATTAAGAAGTAAGACTCTTGGATTC

This was copied to primer3web, where forward (P_{F1}) and reverse (P_R) primers were picked. Sequence specificity was checked using 'In-Silico PCR' in UCSC. These primers annealed to the sequence of interest as well as 2 pseudogenes. The second forward primer (P_{F2}) was the same as in the original publication by Petrini *et al* (106) and again sequence specificity was checked as above together with the reverse (P_R) primer. These primers were specific for the *GTF2I* sequence of interest. Each of the three primers was made into a 10mM solution by adding the appropriate amount of Hyclone ™ molecular biology grade water (Fisher Scientific, Loughborough, UK).

For PCR the HotStarTaq *Plus* Master Mix Kit (Qiagen, Crawley, UK) was used.
For the first PCR run each 20µL reaction consisted of 10µL HotStarTaq *Plus* Master Mix (1 unit HotStarTaq Plus DNA polymerase, 1 x PCR Buffer, containing 1.5mM $MgCl_2$ and 200µM of each dNTP), 1µL of each of the forward (P_{F1}) and reverse (P_R) primers at 10mM, 6µL of RNase-free water and 2µl of sample DNA.
For the semi-nested PCR run the 20µL reaction there were 2 changes:
The forward primer used was P_{F2} (semi-nesting).
The DNA sample used was 2µL of the PCR product from the first PCR for each case.

The Peltier Thermal Cycler–PTC225 DNA Engine Tetrad (MJ Research) was used and the conditions were as per the HotStarTaq Plus PCR Master Mix Kit (Qiagen) instructions, following tedious optimization for the above machine and in my hands:

Step 1: Heat activation for 5 minutes at 95^0C
Step 2: Denaturation for 30 seconds at 94^0C
Step 3: Annealing for 30 seconds at 54^0C
Step 4: Extension for 1 minute at 72^0C
Steps 2-4 were repeated 34 times (35 cycles)
Step 5: Final extension for 10 minutes at 72^0C
Step 6: Hold at 4^0C
Exactly the same PCR protocol was used for both runs.
Positive and negative controls were used every time a PCR was run.

The PCR products were examined by visual inspection following separation by agarose gel electrophoresis.

Each agarose gel was prepared as follows:
100ml of 1 x Tris-Borate-EDTA buffer (TE) were added into a conical flask. 1.8g of 2% agarose were weighed and added to the same flask. The solution was microwaved for a minute and shaken to ensure the agarose had dissolved completely. Once the solution was back to room temperature 3µL of GelRed™ were added and mixed with the solution by rotation to achieve homogenisation. The solution was poured into a gel moulding apparatus, set up with the insertion of a plastic comb to form wells within the agarose gel. Once solidified the comb was removed. 2µL of a 100-bp DNA ladder was loaded into the first well to serve as a molecular size marker. 7µL of each PCR product were mixed by pipetting with 3µL of loading dye and the 10uL of dye and product were loaded into each well.

Electrophoresis was carried out in 1 x TE at 100 Volts for 40 minutes. The gel and amplicons were then visualised under ultraviolet light using the FirstLight UV illuminator (Ultra Violet Products Inc, BioDoc-It ™ imaging system).

2.2.5 Sanger sequencing

Sanger sequencing is the sequencing method by which DNA polymerases add 2',3'dideoxynucleotides (ddNTP's) at the 3' end of a primer. Termination of the chain elongation takes place selectively at A, C, G, or T with the incorporation of dideoxynucleotides, which causes the chain to be short of a 3'-hydroxyl group rendering further elongation of the chain impossible (183).

Sanger sequencing was performed commercially by Eurofins Genomics. Both the second forward (P_{F2}) and the reverse (P_R) primers were used generating two .seq files for each sample. The analysis was performed with the use of SEQUENCHER, a DNA Sequence Analysis Software by Gene Codes Corporation. This software is designed to assist to easily assemble, trim and edit sequences to identify mutations of interest. The use of both forward and reverse primers increased our confidence in the identification of the mutation even in those cases where there was a strikingly smaller proportion of mutant alleles present.

Sanger sequencing remains the gold standard for sequencing single genes at the lowest cost and to confirm NGS findings (184,185). Its specificity relies on the specificity of the primer target sequences, which can be verified for sequence specificity, using online tools such as Basic Local Alignment Search Tools (BLAST) or using 'In-Silico PCR' in UCSC (186). This allows specify of 100% to be reached. Sanger sequencing is still used as the benchmark to which NGS technologies are compared to when it comes to an assay's sensitivity (185) and is the method of choice to exclude false-positive NGS variant calls (184).

DNA and RNA extraction from FFPEs and snap frozen tissue were performed exclusively by me.

2.2.6 DNA extraction from FFPEs

DNA from tumour (with at least 60% tumour content) and matched normal tissue samples is extracted using Quiagen's QIAamp DNA FFPE Tissue Kit (Catalogue No. 56404). The steps, as well as diversions from the original protocol, are as follows:

1. 3 - 5 10µm sections were cut from each sample block. Using a scalpel excess paraffin was trimmed at the same time. The sections were placed in a 2 ml microcentrifuge tube and 1ml of Histo-clear (instead of xylene) was added to the sample. This was vortexed vigorously for 10 seconds and centrifuge at 14,000 rpm for 2 minutes at room temperature. The supernatant was removed by pipetting.
2. 1ml of 99% ethanol was added to the pellet, mixed by vortexing and centrifuged at 14,000rpm for 2 minutes at room temperature. The supernatant was removed by pipetting and residual ethanol was allowed to evaporate by incubating the open tube at room temperature (15–37°C) fir approximately 10 minutes.
3. The pellet was re-suspended in 180 µl Buffer ATL and 20µl of Proteinase K, and mixed by vortexing. For multiple samples a master-mix of Buffer ATL and proteinase K was prepared.

4. This was incubated at 56°C for 1 hour. Due to the nature of the tissue that was very fatty and cellular this step was subsequently modified and the sample was left to incubate under the same conditions overnight.
5. The following day the sample was further incubated at 90°C for 1 hour (to partially reverse the formaldehyde modification of nucleic acids).
6. The sample was centrifuged at room temperature, 2 µl RNase A (1mg/ml) was added and it was left to incubate for 2 minutes.
7. 200 µl Buffer AL was added to the sample, and mixed thoroughly by vortexing. 200µl 99% ethanol were added and mixed again thoroughly by vortexing followed by centrifuging. For multiple samples a master-mix of buffer AL and 99% ethanol was prepared.
8. The entire lysate was transferred to the QIAamp MinElute column (in a 2 ml collection tube) and centrifuged at 8000rpm for 1 minute. The column was placed in a clean 2ml collection tube. The previous collection tube was discarded. Sometimes re-centrifugation at a higher speed was required until the column was completely empty.
9. 500 µl Buffer AW1 was added and the sample was centrifuged at 8,000 rpm for 1 minute. The column was placed in a clean 2 ml collection tube.
10. 500µl Buffer AW2 was added and the sample was centrifuged at 8,000rpm for 1 minute. The column was placed in a clean 2 ml collection tube.
11. The sample was centrifuged at 14,000 rpm for 3 minutes to dry the membrane completely.
12. The column was transferred to a clean 1.5 ml microcentrifuge tube and 50-200 µl Buffer ATE was added. This was incubated at room temperature (15–25°C) for 5 minutes and centrifuged at 14,000 rpm for 1 minute.
13. The column was discarded and the tube with the DNA-containing elution was stored at -80°C.

2.2.7 Simultaneous purification of genomic DNA and total RNA, including miRNA, from frozen tissues

DNA and RNA from snap-frozen tumour and matched normal tissue samples was extracted with Quiagen's AllPrep DNA/RNA/miRNA Universal Kit (Catalogue No. 80224). 10 - 20mg of frozen tissue were used. The steps, as well as diversions from the original protocol, are as follows:

1. Each tissue was ruptured and homogenised using a TissueRuptor® disposable probe and the Qiagen TissueRuptor® for approximately 45-60 seconds (to prevent RNA degradation due to heat) in 350-600 µl Buffer RLT (containing guanidine thiocyanate to prevent activity of RNase enzymes, and beta mercaptoethanol) in a fume-hood. The actual amount of RLT varied according to the size of the tissue (10 mg in 350 µl and 10–30 mg in 600 µl).

2. The homogenized lysate was transferred to an AllPrep DNA Mini spin column placed in a 2 ml collection tube and centrifuged for 30 seconds at 14,000 rpm. Sometimes re-centrifugation until all liquid had passed through the membrane was required. The column was placed in a new 2 ml collection tube and stored at room temperature (15–25°C) for DNA purification later in this protocol (See step 14). The flow-through was transferred into a new 2 ml microcentrifuge tube.

Total RNA (including miRNA) purification

The amount of reagents and buffers from this stage were added in two different amounts according to the sample weight: 10mg Vs 10-30mg

3. 50 µl or 80 µl Proteinase K was added to the flow-through and mixed by pipetting. 200 µl or 350 µl of 99% ethanol was added and mixed by pipetting followed by a 10-minute incubation at room temperature.
4. 400 µl or 750 µl of 99% ethanol was added and mixed by pipetting.
5. 700 µl of the sample, including any precipitate were transferred to an RNeasy Mini spin column placed in a 2 ml collection tube and centrifuged for 15 seconds at 14,000 rpm. The flowthrough was discarded. This step was repeated until the entire sample has passed through the spin column.
6. 500 µl Buffer RPE was added to the spin column and centrifuged for 15 seconds at 14,000 rpm. The flowthrough was discarded.
7. 10 µl DNase I stock solution was added to 70 µl Buffer RDD and mixed gently by inverting the tube and centrifuging briefly to collect residual liquid from the sides of the tube. The DNase I incubation mix (80 µl) was directly added onto the spin column membrane and place on the benchtop (20–30°C) for 15 minutes.
8. 500 µl Buffer FRN was added to the spin column and centrifuged for 15 seconds at 14,000 rpm.
9. The flow-through (that contains small RNAs) was applied to the spin column in a new 2 ml collection tube and centrifuged for 15 seconds at 14,000 rpm. The flowthrough was discarded.
10. 500 µl Buffer RPE was added to the spin column and centrifuged for 15 seconds at 14,000 rpm. The flowthrough was discarded.
11. 500 µl of 99% ethanol was added to spin column and centrifuged for 2 minutes at 14,000 rpm to wash the spin column membrane. The long centrifugation dries the spin column membrane, ensuring that no ethanol is carried over during RNA elution as this may interfere with downstream reactions. The flow-through was discarded and the spin column was centrifuged for a further 2 minutes at 14,00 rpm to eliminate any possible carryover of ethanol

12. The spin column was transferred to a new 1.5 ml collection tube and 50 µl of RNase-free water was added directly to the spin column membrane. This was incubated at room temperature for 2 minutes and centrifuged for 1 min at 10,000 rpm to elute the RNA. As the RNA yields expected were high this step was repeated twice giving a final eluate of 100 µl.
13. The column was discarded and the tube with the RNA-containing elution was stored at -80°C.

Genomic DNA purification

14. Using the AllPrep DNA Mini spin column from step 2, 350 µl Buffer AW1 was added and centrifuged for 15 seconds at 14,000 rpm. The flow-through was discarded.
14. 20 µl Proteinase K was added to 60 µl Buffer AW1, mixed gently, and the mixture was applied to the spin column membrane. For multiple samples a master-mix of Proteinase K and Buffer AW1 was prepared. The sample was left to incubate for 5 min at room temperature.
15. 350 µl Buffer AW1 was added to the spin column and was centrifuged for 15 seconds at 14,000 rpm to wash the spin column membrane. The flow-through was discarded.
16. 500 µl Buffer AW2 was added to the spin column and was centrifuged for 2 minutes at 14,000 rpm to wash the spin column membrane. The flow-through was discarded.
17. The spin column was placed in a new 1.5 ml collection tube. 100 µl Buffer EB was added directly to the spin column membrane and incubated at room temperature (15–25°C) for 1 minute and then centrifuged for 1 minute at 10,000 rpm to elute the DNA. Due to reasonably high concentrations this step was repeated with further 50 µl Buffer EB, using the same collection tube, giving a single eluate of 150 µl.

2.2.8 DNA quantification and quality control

DNA quality and yield were assessed using the Thermo Scientific NanoDropTM 1000 Spectrophotometer. DNA concentration is determined by the absorbance at 260nm, the peak for DNA and the modified Beer Lambert equation (a reading of 1.0 at 260 nm corresponds to 50µg/ml of double-stranded DNA). DNA integrity is determined by the ratio of absorbance at 260 and 280 nm. Good quality DNA has a ratio of 1.8. A ratio lower than 1.8 or higher than 2.1 usually indicates protein or RNA contamination respectively**.** DNA quantification results were then validated by fluorometry using the Quant-iT PicoGreen dsDNA Assay Kit. This was done on approximately 20% of randomly selected to obtain more sensitive and accurate readings, as this method is more specific for double-stranded DNA. The DNA concentration for these samples was compared with the concentration given from the above method.

2.2.9 Whole exome sequencing (WES)

WES was performed by the Centro Nacional de Análisis Genómico (CNAG, Barcelona, Spain). Paired-end multiplex libraries were prepared according to manufacturer's instructions with Illumina TruSeq DNA Sample Prep kit and enriched with the NimbleGen SeqCap EZ Human Exome Library v3.0 (187). The NimbleGen kit targets ~64Mb, which corresponds to ~30,000 genes. Libraries were loaded to Illumina flowcells for cluster generation prior to producing 100 base read pairs on a HiSeq2000 instrument following the Illumina protocol. Base calling and quality control was done with the Illumina RTA sequence analysis pipeline according to manufacturer's instructions.

2.2.9.1 Data analysis

This part of the data analysis was performed by CNAG.

Reads were hard-trimmed from the end of the read up to the first base with a quality of at least 10. Reads with at least 40nt of length were mapped to Human genome build hg19 (Human reference genome (Broad): ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequence/hs37d5.fa.gz) using the GEM toolkit (188) allowing up to 4 mismatches. Alignment (.bam) files containing only properly paired, uniquely mapping reads without duplicates were processed using picard tools version 1.110 (http://broadinstitute.github.io/picard/) to add read groups and to remove duplicates. The Genome Analysis Tool Kit (GATK) (189,190) version 3.1 was used for local realignment. Processed .bam files were submitted to variant calling for single nucleotide variants and small insertions and deletions using samtools version 0.1.19 (191). Functional annotations were added using snpEff (192) with the GRCh37.75 database. Human dbSNP version 137, population frequencies from 1000 Genomes and Exome Variant Server, as well as conservation and deleteriousness predictions from dbNSFP (193,194) were annotated using snpSift (195).

2.2.9.2 Variant filtering and identification of potential driver mutations

This part of the analysis was performed with the help of Amit Mandal, bioinformatician at NHLI.

The Variant Call Format (VCF) files that were created were sent to NHLI and were the starting point for inspection of various quality metrics per variant call in each sample pair. The aim was to

minimize potential false positives while maximizing sensitivity when differentiating somatic mutations from germline variations.

A VCF file is the standard file format used by mutation calling algorithms to return variant identification records. The first few lines of the file are comment lines (indicated by '#' at the beginning) that document the definition of the metrics, the reference genome sequence used and optionally the command-line parameters used by the respective algorithm to arrive at this set of variant records. For each variant in the file there are multiple metrics specified that measure the confidence in the identification and the probability that values for the call are an artefact or false positive. After the comment lines, the actual variant records follow in tab-delimited format. There are nine mandatory columns (CHROM, POS, ID, REF, ALT, QUAL, FILTER, INFO, FORMAT) followed by one or more sample-specific column(s) at the end. If a variant identified is already known then its name/ ID is present in the third column. Columns four and five hold the reference and the alternate/ variant allele identified. Column six is specific to the algorithm used and is a combined score of different confidence measures. Column seven enlists one or more metrics that could be used to filter the variant call. In cases where no confounding metric is found, the algorithm writes 'PASS'. Column eight holds descriptive information about the variant, like its frequency, heterozygosity and overall depth in the sample cohort examined. When the VCF file has been annotated for genes and phenotypic impact, those values are present there. Column nine shows the order in which different values appear for each variant call in the sample-specific columns. An example VCF file can be seen in Figure 2.1.

Figure 2.1. Example of a VCF file

```
##fileformat=VCFv4.2
##fileDate=20090805
##source=myImputationProgramV3.1
##reference=file:///seq/references/1000GenomesPilot-NCBI36.fasta
##contig=<ID=20,length=62435964,assembly=B36,md5=f126cdf8a6e0c7f379d618ff66beb2da,species="Homo sapiens",taxonomy=x>
##phasing=partial
##INFO=<ID=NS,Number=1,Type=Integer,Description="Number of Samples With Data">
##INFO=<ID=DP,Number=1,Type=Integer,Description="Total Depth">
##INFO=<ID=AF,Number=A,Type=Float,Description="Allele Frequency">
##INFO=<ID=AA,Number=1,Type=String,Description="Ancestral Allele">
##INFO=<ID=DB,Number=0,Type=Flag,Description="dbSNP membership, build 129">
##INFO=<ID=H2,Number=0,Type=Flag,Description="HapMap2 membership">
##FILTER=<ID=q10,Description="Quality below 10">
##FILTER=<ID=s50,Description="Less than 50% of samples have data">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FORMAT=<ID=GQ,Number=1,Type=Integer,Description="Genotype Quality">
##FORMAT=<ID=DP,Number=1,Type=Integer,Description="Read Depth">
##FORMAT=<ID=HQ,Number=2,Type=Integer,Description="Haplotype Quality">
#CHROM POS     ID        REF    ALT     QUAL FILTER INFO                              FORMAT      NA00001        NA00002        NA00003
20     14370   rs6054257 G      A       29   PASS   NS=3;DP=14;AF=0.5;DB;H2           GT:GQ:DP:HQ 0|0:48:1:51,51 1|0:48:8:51,51 1/1:43:5:.,.
20     17330   .         T      A       3    q10    NS=3;DP=11;AF=0.017               GT:GQ:DP:HQ 0|0:49:3:58,50 0|1:3:5:65,3   0/0:41:3
20     1110696 rs6040355 A      G,T     67   PASS   NS=2;DP=10;AF=0.333,0.667;AA=T;DB GT:GQ:DP:HQ 1|2:21:6:23,27 2|1:2:0:18,2   2/2:35:4
20     1230237 .         T      .       47   PASS   NS=3;DP=13;AA=T                   GT:GQ:DP:HQ 0|0:54:7:56,60 0|0:48:4:51,51 0/0:61:2
20     1234567 microsat1 GTC    G,GTCT  50   PASS   NS=3;DP=9;AA=G                    GT:GQ:DP    0/1:35:4       0/2:17:2       1/1:40:3
```

Overall the various quality metrics are present in the QUAL, FILTER and INFO columns. Also, sample-wise read-depth support for the variants identified is present in the last *n* columns (*n* being the number of samples analysed).

Of the seventeen samples (fifteen with matched normal tissue and two unpaired), seven were from *samtools* (*ver. 0.1.19*) and remaining ten from GATK (*ver. 2.4.9*) *HaplotypeCaller*. As the VCFs differed in their origin, the two groups had non-overlapping set of quality metrics. Nonetheless, effort was made to achieve similar filtering stringency in both groups of VCF files. For each sample pair, CNAG had performed joint variation calling and hence the variants in the VCFs were a mixture of germline polymorphisms and potential somatic mutations. To identify and distinguish somatic mutations from germline, *per* variant call, the genotype status provided in the VCF, along with the other VCF *field* values, were utilised. The process was as follows:

For samtools-generated VCF files the following metrics were utilised:

1. *MPB* – This indicated uniqueness (or absence thereof) of the mappability of each variant position on the genome. Repetitive regions or low-complexity stretches had no value assigned. Value 'U' (signifying unique) if present, was used as a positive filter.
2. *msb30* – This metric ('Maximum strand bias above Phred score 30') indicated the *p-value* significance (in phred scale) of strand-bias at the variant position. '30' signified a *p-value* of 10^{-3} or less. This was used as a negative filter.
3. *tdb0.05* – This metric ('tail distance bias p-value 0.05') indicated the *p-value* significance of over-representation of fastq read ends at the variant position. This means that the variant could be an artefact of bad quality bases at the fastq read-ends. This was used as a negative filter.

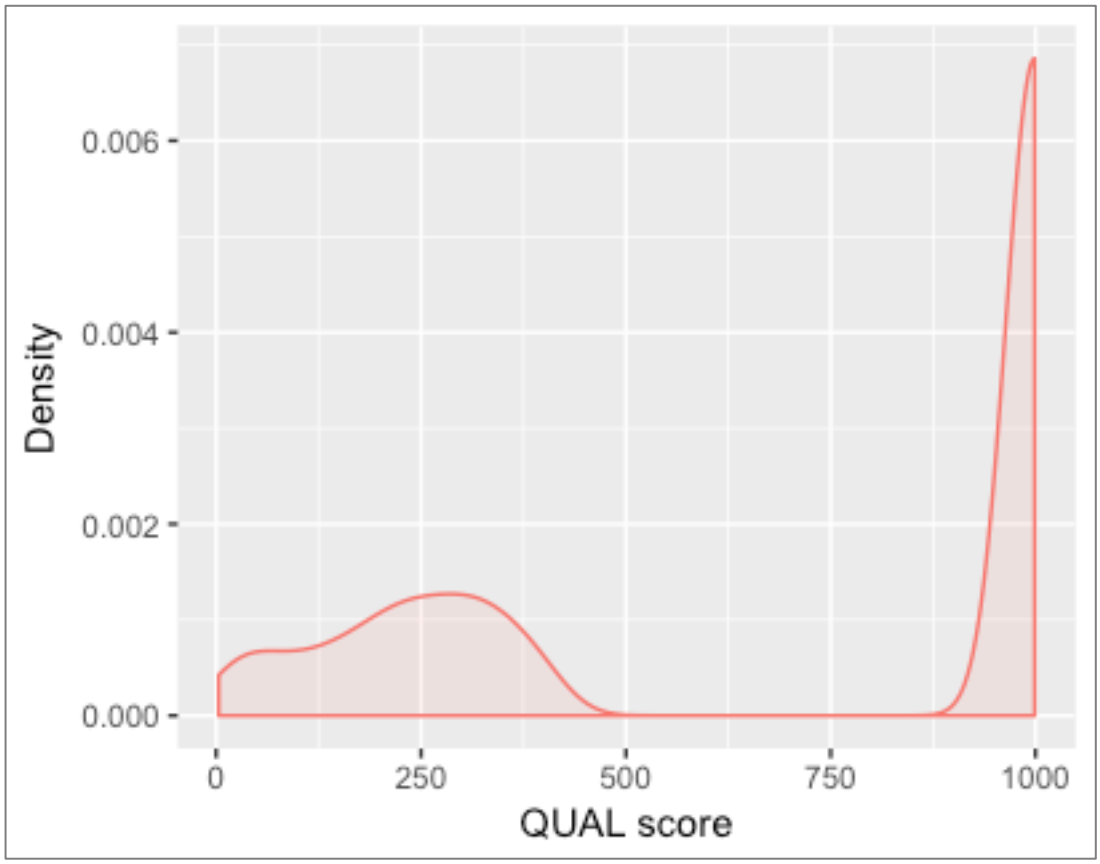
In addition to the above the following filters were used to discriminate between somatic and germline calls:

- **Somatic calls** – Fisher's test comparing the read-depth numbers between the reference and the variant allele in the tumour and normal sample respectively were performed, with a Fisher's p-value ≤0.001 being considered significant. The genotype call had to be 0/1 or 1/1 in the tumour sample and 0/0 in the germline. The total read-depth support for the reference allele, both in the germline and the tumour, had to be ≥10. Read-depth support for the variant allele had to be ≥5 in the tumour sample and the same allele *if seen* in the germline would only be included if its frequency was <3%.

- **Germline calls** – One of the central tenets of assessing confidence in somatic calls was the allele-wise read-depth in both samples. This was not possible for germline calls in the samtools VCF files as the allele-wise read depth was *not* present for germline calls. Another problem was that using Fisher's p-values, with the logic that it should be non-significant for germline calls, was returning too few candidates (<1000). In the GATK VCF files, which have allele-wise read depths for all variant records, filtered germline calls were around 30-50,000. Hence a separate filtering parameter was required in addition to the above filters (*MPB, msb30* and *tdb0.05*). For this reason column 6th (QUAL) was explored.

For all calls passing the above three general filters, the scores of column 6 were extracted and a density plot was created. From the plot (Figure 2.2), it could be seen that there is a flat bump around 200-400 and then a sharp peak around 1000 (max. value of QUAL score was 999). A filtering threshold of >500 for QUAL was therefore used to filter germline calls.

Figure 2.2. Distribution of QUAL score for variants passing *MPB, msb30* and *tdb0.05* metrics



In addition to passing *MPB, msb30* and *tdb0.05* metrics and QUAL score threshold, in order for the variant to be confidently called germline it also had to satisfy the following conditions:

1. Total read-depth in each sample had to be ≥10 (N.B. allele-wise read depth was not available in these cases)
2. Presence of '1' in genotype field ('GT'; i.e. in FORMAT column) in both samples was essential
3. 'GT' of tumour had to equal 'GT' of germline. (This was to exclude cases of LOH and cases that went from 1/1 to 0/1, probably due to copy number effects)
4. Neither sample could have a value greater than '1' in the genotype field. (Genotypes like 1/2 or 1/3 likely indicate multi-allelic sites according to the algorithm but could also potentially represent 'noise'.)

The reasons more filters and metrics were used for germline calls vs. somatic were the need for stricter criteria in calling these due to the very high number of germline variants prior to filtering and the potential confounding cases of LOH when looking for germline variants

For GATK-generated VCF files the following metrics were utilised:

1. *map* – This is the same metric as MPB in samtools-generated VCF files and was also used as a positive filter.
2. *msb30* and *FS* – Both metrics assess strand-bias at the variant position. 'msb30' was used as before as a negative filter. 'FS' (Fisher exact test's *p-value* in phred scale) had to be <30.
3. *tdb0.05* and *ReadPosRankSum* – Both metrics assess the bias in the variant's location among the supporting fastq reads. As before, tdb0.05 was used as a negative filter. The ReadPosRankSum values were represented as z-scores. A very moderate range of > -2 or < 2 z-score was required by variants to pass this metric.
4. *ClippingRankSum* – This metric assesses the over-representation, if any, of hard-clipping of fastq reads at the variant location. Hard-clipping can indicate low complexity sequence around, mapping to a pseudogene and/ or un-ambiguous mapping. The scores were present as z-score and variants needed to be in the range of > -2 or < 2 z-score to pass this metric.

In addition to the above the following filters were used to discriminate between somatic and germline calls:

- **Somatic calls** – In GATK VCF files, Fisher p-values were not available. Hence only read-depth and genotype assignment value-based selection was performed. Briefly, the genotype call had to be 0/0 in the germline and 0/1 or 1/1 in the tumour. Read-depth-based support was used as in the samtools VCF files: ≥10 for the reference allele in both samples, ≥5 for the alternate allele in the tumour sample and *if seen* in the germline would only be included if its frequency was <3%.

- **Germline calls** – Unlike samtools VCF files, read-depth information was available for reference as well as alternate/ variant allele in both samples. Hence no additional filtering parameters were required other than those for the somatic calls. Briefly –
    i. The presence of '1' in genotype field ('GT'; i.e. in FORMAT column) in both samples was required
    ii. 'GT' of tumour had to equal 'GT' of germline. (This was to exclude cases of LOH and cases like 1/1 to 0/1, probably due to copy number effects)

iii. Neither sample could have a value greater than '1' in the genotype field. (Genotypes like 1/2 or 1/3 indicate multi-allelic site *as per the algorithm's understanding* but could potentially be 'noise')

iv. The reference allele supporting reads had to be ≥10 in both samples. The frequency of the alternate allele had to be >5% in both samples.

After arriving at the somatic mutation or germline variant calls, phenotypic impact-based filtering was applied to prioritise the list of variants. Two main strategies were applied:

1. *Gene-level based impact* – Here the variant was assessed for how critically it could affect the gene's function. SnpEff had been run previously and annotation data were available in the VCF files. SnpEff, a variant annotation and prioritisation software, can categorise variants into four impact categories, in descending order: *High*, *Moderate*, *Low* and *Modifier*. These categories are based on whether the variant is within the CDS of the gene and whether on translation the variant would lead to an amino-acid change. Briefly, all changes that disrupt the ORF frame (stop-codon affecting, InDels, nonsense mutations and splice-donor/ acceptor site affecting) get classified as *High* impact. Missense mutations get classified as *Moderate* impact.

   As a first step in the prioritisation process, only variants with a *High* or *Moderate* predicted impact were filtered in. Further gene-level based assessment was carried out using CADD (Combined Annotation Dependent Depletion), which is a scoring system used to rank the *predicted* deleteriousness of variants. It aggregates evidence from variant position (within the protein sequence), conservation at the site, observed variability in sequenced human population and interaction between different evidences. For this, the latest release of CADD (*ver. 1.3*) was used to re-assign scores. There was no specific score cut-off but scores were documented in the result as a valuable indicator of predicted importance.

2. *Human population-level based fitness* – Exome Aggregation Consortium (ExAC) is a recent endeavour by a large group of collaborating laboratories to aggregate over 60,000 WES datasets from presumably 'healthy' population. As of *version 0.3.1* of ExAC release, observed WES variants in the population are available for 60,706 individuals. These individuals belong to diverse ethnicities of African American (8.6%), Latino (9.5%), East Asian (7.1%), Finnish (5.5%), Non-Finnish European (~55%) and South Asian (13.6%). Hence this resource is useful to test the fitness of a potentially deleterious variation, the hypothesis being that variations seen at an appreciable frequency in ExAC are very less likely to have a deleterious effect. A moderate threshold of <0.001 for filtering of any variant was chosen in the ExAC non-TCGA cohort (a subset of the total that excludes 7,601

individuals from the TCGA study, which are all cancer patients and which therefore could harbour important cancer causing mutations).

Following these steps genes were prioritised based on their recurrence in the entire cohort of samples or in each TET subtype, the presence of mutations in specific gene domains (mutation hotspots) and on whether they were known to be implicated in tumourigenesis.

2.2.10 SNP genotyping of TETs

SNP genotyping of TETs was performed by Eurofins Genomics (https://www.eurofinsgenomics.eu/en/genotyping-gene-expression/genotyping-services/snp-genotyping.aspx.). A total of 144 patient tissue samples (97 paired and 47 unpaired) were used using the Illumina Infinium® OmniExpressExome-8 v1.3 and 1.4 arrays. The OmniExpressExome comprise >950k markers (ver. 1.3 = 958,497 markers; ver. 1.4 = 960,919 markers). These Single Nucleotide Polymorphism (SNP) markers are spaced across the genome at median and average distances of 1.36kb and 3.03kb respectively. More than half (~58%) of the markers fall within RefSeq genes, with over 274,000 of them targeting specifically functional exonic regions, which cover common variants from over 12,000 individual exome and whole-genome sequences as assessed by the 1000 Genomes Project (1kGP).

2.2.10.1 Preparation of LRR and BAF files

This part of the analysis was performed with the help of Amit Mandal, bioinformatician at NHLI.

After processing raw data (array scan images and configuration files) using Illumina GenomeStudio software (ver. 1.9.4), LogR ratio (LRR; log transformed signal intensities) and B-allele frequencies (BAF) were exported out. As the samples were run on two different version arrays the content had to be made uniform. This was done as follows:

a) SNPs/ markers needed to be present in *both* array versions.
b) For some markers, genomic location information was not present in the annotation filesthat were available from the Illumina website. Only markers with precise genomic location provided were retained.
c) SNPs belonging to pseudo-autosomal regions (PAR) were excluded.
d) Mitochondrial SNPs were also excluded to ensure compatibility of the final dataset with both hg19 and GRCh37 genome versions. Though in effect hg19 and GRCh37 are the same release version of the human genome, mitochondrial content and unplaced contig IDs differ.

The above resulted in a final set of 929,133 SNPs/ markers that belong only to the primary chromosomes and were present in both versions of the Infinium arrays.

2.2.10.2 Segmentation and Somatic Copy number calling

Allele-Specific Copy Number Analysis of Tumours (ASCAT; ver. 2.4.4) was then used for normalization, segmentation and copy number calling using the R statistical language platform (ver. 3.4.0). ASCAT uses, along with LogR ratio, allele frequencies to derive copy number information. Since ASCAT uses BAF, it allows the ascertainment the allele-specific copy number of solid tumours, taking ploidy and non-aberrant cell admixture into account. This allows "ASCAT profiles" to be determined (genome-wide allele-specific copy-number profiles), which allow identification of copy number gains and losses and loss of heterozygosity (196).

- Custom GC Content file creation and Normalization – Before segmenting the data, the software performs a normalization process that rectifies for 'genomic wave artefacts' i.e. a genome-wide spatial autocorrelation of signal intensities that gives the impression of wave-like pattern when plotted (197). These artefacts interfere with accurate CNV detection and are in part linked to GC content of the local genomic region. Hence ASCAT employs a GC wave correction as a normalization process (command ascat.GCcorrect). For it to do that it needs a *GC Content file* which has for every SNP/ marker, the local GC content in genomic windows of 25, 50, 100, 200, 500, 1k, 2k, 5k, 10k, 20k, 50k, 100k, 200k, 500k, 1M, 2M, 5M and finally 10M base pairs. This file is available, from the ASCAT website [https://goo.gl/widYLS], for some of the major SNP genotyping array platforms. Since at the time of this study such a file was not available for the Infinium OmniExpressExome array, it was created using the custom script provided in the software package. Again only the primary chromosomes were chosen (excluding mitochondrial) and any SNP/ marker *not* present on both versions of the arrays were excluded.

- Predicted Germline genotypes for unpaired samples – Using the tumour data available (LRR and BAF) and the information about the array platform used, ASCAT can predict germline genotype. This function (command ascat.predictGermlineGenotypes) was utilised for the 47 unpaired samples.

- Allele-specific segmentation and copy number calling – Once all the paired samples had been normalized (for 'genomic wave' artefacts) and predicted germline genotypes were ready for the unpaired samples, the ASPCF (Allele-Specific Piecewise Constant Fitting) algorithm (command ascat.aspcf) was applied *separately* on the paired and unpaired sample subsets of the total cohort. It is worth noting that the normalization procedure is unavailable for unpaired samples. Nonetheless, the predicted germline genotype is provided as an additional argument while running the ASPCF algorithm.

- Output of ASCAT – The software on completion returns the following:

    - Somatic Copy number segments – For each copy number segment identified per sample, copy number states in whole number form are provided for the major as well as the minor allele (Figure 2.3).

Figure 2.3. Final output format of somatic copy number calls from ASCAT

	A	B	C	D	E	F	G
1		**sample**	**chr**	**startpos**	**endpos**	**nMajor**	**nMinor**
2	1	THRBF25324	1	82154	4269491	4	1
3	2	THRBF25324	1	4271526	4335893	80	18
4	3	THRBF25324	1	4341843	16260392	4	1
5	4	THRBF25324	1	16260513	16277647	1	0
6	5	THRBF25324	1	16285342	37962756	4	1

    - Somatic Copy number segments (raw format) – This is the same as described above with two additional columns for the raw copy number call for each allele (Figure 2.4).

Figure 2.4 Somatic copy number calls in raw format from ASCAT

	A	B	C	D	E	F	G	H	I
1		sample	chr	startpos	endpos	nMajor	nMinor	nAraw	nBraw
2	1	THRBF25324	1	82154	4269491	4	1	3.90375523	0.86116462
3	2	THRBF25324	1	4271526	4335893	80	18	79.6351724	17.5674419
4	3	THRBF25324	1	4341843	16260392	4	1	4.34163075	0.95775954
5	4	THRBF25324	1	16260513	16277647	1	0	1.2095067	0.26681601
6	5	THRBF25324	1	16285342	37962756	4	1	4.39537435	0.96961532

    - Failed arrays – These are the tumour samples where the ASCAT software's ASPCF algorithm was unable to resolve the ploidy status and aberrant cell fraction. For these samples no somatic copy number information could be determined using ASCAT.

- Aberrant cell fraction – This is the estimated content of tumour (or aberrant cells) for each named tumour sample successfully processed.
- Non-aberrant arrays – For samples where ASCAT determined nil content for aberrant cells.
- Ploidy – For all tumour samples successfully processed by ASCAT, a ploidy estimation is returned.

2.2.10.3 Custom thresholding on ASCAT copy number output

The copy number segments received from ASCAT were subject to custom thresholding to increase the stringency and minimize false positives. To implement the thresholding the allele-wise copy number status was used. This was performed as follows:

- For segments with amplifications – Based upon the distribution of the copy number amplification calls across all the samples, the following filtering was implemented:
  - Copy number status of *either* allele >3
  - AND, both major and minor allele has copy number status >1

- For segments with deletions – A strategy similar to the above could not be implemented for deletions as for these there were only two values for the major allele, either 0 or 1 and only a single value, 0, for the minor allele. Stratification of values was therefore impossible. For this reason raw copy number calls were used. The distribution of the raw calls was studied for both the alleles and the filter implemented was: Raw copy number call of major allele < median value of all major allele raw calls.

The same filtering process was applied to the calls from the paired and unpaired sample sets. The final segments were then pooled for the amplification and deletion groups, prior to proceeding to the annotation of the segments.

2.2.10.4 Gene annotation of Copy number segments and recurrent aberrations

Ensembl GTF (Gene Transfer Format) (rel.87) was downloaded from the FTP site and using a custom script converted to BED (Browser Extensible Data) format. This was done so that the copy number segment data, already amenable to BED format (i.e. tab-delimited text file with mandatory 3 columns in order of: chromosome, coordinate start and coordinate end), could be overlapped with the Ensembl annotation data using Bedtools (ver. 2-2.26). Further to annotation

using Ensembl genes, another custom script was used to select out only those genes, which satisfied *both* of the following criteria:

a. Were protein-coding
b. Did not belong to a cDNA clone ID

This annotation set became the starting point for the creation of recurrence matrix, which allowed the determination of the most frequent somatic copy number aberration across the samples in the cohort. Using the package reshape2 (ver. 1.4.2) in R statistical platform (ver. 3.4.0) a recurrence matrix (rows are genes and samples are columns) was created such that for any gene affected in any sample there was a value of 0 or 1 in its respective *cell* for each sample-wise column. This matrix could be easily imported in excel to sum up the value for each row (i.e. gene) and thus sort the matrix by recurrence.

2.2.11 Immunohistochemistry (IHC)

All IHC procedures took place at the Histopathology Core Facility at the Royal Brompton Hospital by expertly trained histopathology staff.

2.2.11.1 Construction of Tissue Microarray Blocks (TMAs)

TMAs are increasingly used in histopathology as a means of obtaining high throughput of results from multiple FFPE samples. The FFPEs from which cores were obtained all came from the RBH as described above. H&E slides from each of the available blocks were assessed to blindly confirm the diagnosis, including the TET subtype. For each case the slide with the highest tumour content was marked in three areas. This assessment was carried out by a consultant histopathologist who was an expert in TETs. The matched blocks were inspected and if the tissue was of sufficient depth for 3 x 4mm cores, these were taken. A tissue Microarrayer (Alphelys Minicore® Microarrayer) was used to embed cores into the recipient paraffin blocks. The reason three cores per sample were taken was to ensure that the tissue cores were representative of each block and to address tumour heterogeneity. Each sample was therefore represented with 3 x 1mm in diameter x 4mm in depth cores. A TMA map was created for each TMA block indicating the exact position of each core. 145 cases were distributed across 6 TMAs. The details of the above procedure are outlined in the relevant Standard Operating Procedure (SOP), which can be found in Appendix 1. Clinical data for the cases were obtained via the RBH EPR as described in section 2.2.2.

2.2.11.2 Microtomy

From each of the 6 TMA blocks sections were cut at a thickness of 3μm using a standard microtome (Shandon Finesse 325, Thermo Scientific). Each section was then floated on a water bath (Section Bath, Thermo Scientific) at 45°C and mounted onto extra-adhesive slides (Superfrost Plus, Thermo Scientific). These were left to air dry for 30 minutes and were finally placed in a 60°C oven for 60 minutes. The relevant SOP can be found in Appendix 1.

2.2.11.3 IHC staining

Fully automated staining was performed with the Ventana Benchmark Ultra platform. Each 3μ section was deparaffinised for 8 minutes at 60°C (EZ-Prep, Ventana), followed by antigen retrieval and incubation of the primary antibody (see Table 2.5). Endogenous peroxidase activity was blocked with the Ventana Optiview Peroxidase Inhibitor (3%) for 8 minutes. The Ventana Optiview DAB Detection Kit was used for amplification and visualization as per the manufacturer's guidelines followed by counterstaining with Ventana Hematoxylin II for 8 and Ventana Bluing Reagent for 12 minutes. In between steps rinses were performed with the Ventana Reaction Buffer. Each slide was then removed from the immunostainer and manually dehydrated through IMS followed by xylene, and cover-slipped with DePeX in standard fashion.

The antibodies used all underwent standard manufacturer's quality control and were all validated by the manufacturer to be used for laboratory research although none of these are used yet in routine clinical practice. Positive and negative controls were used for IHC with all antibodies. Isotype controls and knockout validation using a lung adenocarcinoma cell line were also used by the manufacturer to confirm antibody specificity in the case of PD-L1.
All details, including the controls used, can be found in the following antibody datasheets:
PD-L1: http://www.abcam.com/pd-l1-antibody-28-8-ab205921.html
PD-1: https://www.e-labeling.eu/CMC48950050/60800/EN#remarkPopup
PD-L2:https://www.thermofisher.com/order/genome-database/generatePdf?productName=CD273%20(B7-DC)&assayType=PRANT&detailed=true&productId=PA5-20344
B7-H3: http://www.abcam.com/cd276-antibody-6a1-ab105922.html
PDLIM1: http://www.abcam.com/pdlim1-antibody-ab64971.html
TIM-3: http://www.abcam.com/tim-3-antibody-ab185703.html
4-1BB: http://www.abcam.com/cd137-antibody-ab197942.html
HGF: https://www.sigmaaldrich.com/catalog/product/sigma/hpa040360?lang=en&region=GB
LAG-3: https://www.novusbio.com/products/lag-3-antibody-17b4_nbp1-97657#datasheet
VISTA: https://media.cellsignal.com/pdf/64953.pdf

Following the manufacturer's recommendations on the dilution of each antibody to be used this was optimized by expertly trained histopathology staff at RBH. Further validation took place at RBH. Positive controls were used for all antibodies as per the manufacturer's instructions and are detailed in Table 2.5. In-house negative, isotype controls were for all antibodies to ensure the level of background, non-specific staining was established and kept to a minimum. The isotype controls were done by running them under the same conditions as used for the antibodies. Pipetting the primary antibody onto the slides is a manual step in the protocol, whereby the machine draw beeps and slides out so that a human can interact with it. At this stage, instead of adding the actual antibody, the isotype control primary antibody is added instead. The results of antibody labeling were all assessed by two expert Consultant Histopathologists, who were in complete agreement regarding the findings of both positive and negative controls.

Table 2.5. Antibody Staining

Antibody*	Antigen Retrieval Reagent Used	Length/Temperature of Antigen Retrieval	Antibody Dilution Factor	Primary Antibody Incubation	Positive controls
4-1BB	Ventana Cell Conditioning 1 pH8.5	72 mins at 100°C	1/40	60 mins at 37°C	Tonsil
B7-H3	Ventana Cell Conditioning 1 (pH8)	48 mins at 100°C	1/800	32 mins at 37°C	Tonsil and stomach
CD45	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	1/100	32 mins at 36°C	Tonsil and placenta
Cytokeratin MNF116	Ventana Protease 1	12 mins	1/200	32 mins at 36°C	Tonsil and placenta
HGF	Ventana Protease 1	8 mins	1/200	32 mins at 37°C	Colon and placenta
LAG-3	Ventana Cell Conditioning 1 (pH8)	40 mins at 100°C	1/100	32 mins at 37°C	Tonsil and lymph node
PD-1	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	Predilute	16 mins at 36°C	Tonsil and placenta
PD-L1	Ventana Cell Conditioning 1 (pH8)	72 mins at 100°C	1/200	32 mins at 37°C	Tonsil and placenta
PD-L2	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	3µg/ml	32 mins at 36°C	Kidney
PDLIM1	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	1/200	32 mins at 37°C	Colon and placenta
VISTA	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	1/150	60 mins at 37°C	Tonsil and Kidney

*Antibody details, including description, supplier and product code can be found in Table 2.3.

Antibody dilution factor = Concentration. Positive controls were all healthy, 'normal', human tissue.

2.2.11.4 Scoring

All IHC-stained sections were scored by two expert pathologists blinded to the TET subtype and to each other's scoring. The score for each case was calculated as the mean of each of the three scored cores per case. For all cases I was present throughout the scoring procedure. CD45 and cytokeratin MNF116 stains were used to differentiate epithelial cells and lymphocytes. A Histo-score (H-score) was assigned to each sample, and the epithelial and lymphocytic component were scored separately. This is a semi-quantitative approach that assigns an intensity score (0-3) to each cell fixed in a field together with the percentage of cells staining at that intensity. A final H-score is assigned following this formula:

[1 × (% cells 1+) + 2 × (% cells 2+) + 3 × (% cells 3+)]

The final score ranged between 0 and 300. In our case there was no specific threshold set for positivity and all scores were recorded and compared(198,199).

2.2.11.5 Statistical analyses

A comparison of the H-scores of each subtype was performed with the non-parametric Mann-Whitney test in R. The p-values were then extracted and were used to create comparison matrices and heat-maps for a graphical representation of the results. For all statistical tests significance was set at $p\text{-value} \leq 0.05$, which when appropriate was adjusted for multiple-testing with the Benjamini-Hochberg method.

Chapter 3
TET samples and clinical data

3.1 The aims of this chapter were the following:

- To describe the major clinicopathological characteristics of the TETs collected for this research project
- To specifically assess the prevalence of MG in all TET subtypes in this research project
- To give an overview of the nature of the samples used in this research project

3.2 Clinicopathological characteristics

Tumour samples of 144 patients as well as hyperplastic thymic tissue from a further twelve patients was collected from the two institutions described in Chapter 2, RBH and KHP. Figure 3.1 shows the breakdown of the TET subtypes. Of the 126 thymomas, 22 were Type A (17%), 35 were AB (28%), 22 B1 (17%), 34 B2 (27%), nine B3 (7%) and four MNT (3%). There were also 18 carcinomas, twelve of which were SCCs (67%) and six were NETTs (33%). Normal, non-malignant tissue was available for 119 cases (76%). The mean age of the 144 patients with TETs was 59. The male: female ratio was 1: 1.136. The patient's baseline clinical characteristics and basic sample characteristics are summarized in Table 3.1. An extended version of this table, including the downstream applications of the extracted DNA for these cases can be found in Appendix 2. The mean tumour percentage of all tumours was 91%. In only eight of 127 (6%) TETs was the tumour content below 70%. These were all the patients who had received primary chemotherapy prior to surgery. In terms of survival only two of 22 (9%) type A thymomas are dead. Unfortunately the cause of death is not known for any of the cases so establishing a link with their TET was not possible. The two patients with a type A thymoma that passed away presented with stage I and Stage IIA disease. All four MNTs were alive. Of the AB thymomas six of 35 (17%) patients were dead, with only one of these presenting with stage IV disease. Half of these harboured *GTF2I* mutations, which comes in contrast with the reports that this mutation is linked to good prognosis. Two of 22 patients (9%) with B1 thymoma (the one with stage I disease and the *GTF2I* mutation, the other with stage IV disease), five of the 22 patients (23%) with B2 (four of whom had stage IV disease an none with *GTF2I* mutations) and only one of nine (11%) patients with B3 thymoma were dead (Stage IV disease). Of the carcinomas, four of the twelve (33%) SCC patients and all six with NETTs had died.

Figure 3.1. TET subtypes

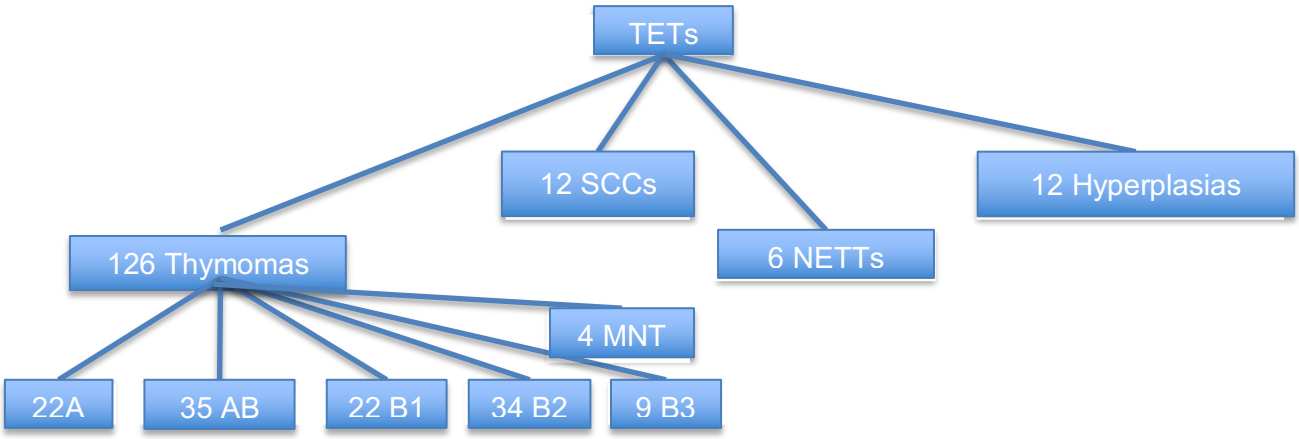


Table 3.1 Baseline clinicopathological data for all TETs (n=156)

#	Sample collection	Tumour Samples (n=156)	Sample Type	Tumour content (%)	Matched Normal Tissue (n=119)	Sample Type	WHO histotype	Sex	Stage	Age at Diagnosis	Resection	Survival (months / or alive)
1	>2010	150098PT	FFPE				A	M	I	81	R0	A
2	>2010	150113PT	FFPE				A	F	IIB	43	R0	A
3	2000-9	THRBF04707T1	FFPE	100	✓	FFPE	A	M	I	77	R0	18
4	2000-9	THRBF05579T1	FFPE	100			A	M	I	63	R0	A
5	2000-9	THRBZ07321T2	Frozen	90			A	F	IIA	69	R0	A
6	>2010	THRBF11200T2	FFPE	80	✓	FFPE	A	F	I	55	R0	A
7	2000-9	THRBF13203T1	FFPE	80	✓	FFPE	A	F	I	69	R0	A
8	2000-9	THRBF16927T1	FFPE	100	✓	FFPE	A	M	I	57	R0	A
9	2000-9	THRBF17788T1	FFPE	75	✓	FFPE	A	F	II	74	R0	A
10	>2010	THRBF20533T1	FFPE	100	✓	FFPE	A	M	II	65	R1	A
11	2000-9	THRBF20683T1	FFPE	100	✓	FFPE	A	M	IIA	62	R0	
12	2000-9	THRBF21048T1	FFPE	100	✓	FFPE	A	M	IIA	47	R0	A
13	<2000	THRBZ23369T2	Frozen	100	✓	FFPE	A	F	I	55	R0	A
14	>2010	THRBZ23415T1	Frozen	100	✓	FFPE	A	M	IIA	75	R0	A
15	>2010	THRBF26625T1	FFPE	90	✓	FFPE	A	M	I	72	R0	
16	2000-9	THRBF2887T1	FFPE	90	✓	FFPE	A	M	I	62	R0	A
17	>2010	THRBF30253T1	FFPE	90	✓	FFPE	A	F	IIA	53	R0	A
18	2000-9	THRBZ57149T2	Frozen	100	✓	FFPE	A	F	I	85	R0	A
19	2000-9	THRBZ85801T2	Frozen	100	✓	Frozen	A	F	I	63	R0	A
20	<2000	THRBF94124T1	FFPE	100	✓	FFPE	A	M	I	57	R0	0
21	>2010	THRBZ61001T1	Frozen	100	✓	FFPE	A	M	I	77	R0	A
22	>2010	THRBZ80427T1	Frozen	100	✓	Frozen	A	M	IIA	75		11
23	2000-9	THRBZ03917T2	Frozen	100	✓	FFPE	A (MNT LS)	M	IIA	72	R0	A
24	<2000	THRBF15012T1	FFPE	95	✓	FFPE	A (MNT LS)	M	IIA	66	R0	A

25	>2010	THRBF20233T2	FFPE	100	✓	FFPE	A (MNT LS)	M	IIA	75	R1	A
26	<2000	THRBF86319T1	FFPE	90	✓	FFPE	A (MNT LS)	M	IIB	60	R0	A
27	>2010	150077PT	FFPE				AB	F	I	74	R0	A
28	>2010	150101PT	FFPE				AB	F	II	56	R0	A
29	<2000	THRBF00959T1	FFPE	100	✓	FFPE	AB	F	I	61	R0	A
30	2000-9	THRBZ02130T2	Frozen	100	✓	Frozen	AB	M		60		A
31	2000-9	THRBF05377T1	FFPE	100	✓	FFPE	AB	F	I	62	R0	88
32	>2010	THRBZ00852T2	Frozen	100	✓	FFPE	AB					
33	>2010	THRBF09409T1	FFPE	100			AB	M	I	67	R0	A
34	>2010	THRBF13694T1	FFPE	80	✓	FFPE	AB	M	IIA	60		A
35	>2010	THRBZ14059T1	Frozen	100	✓	Frozen	AB	F	I	53	R0	A
36	>2010	THRBF14859T1	FFPE	80	✓	FFPE	AB	F	IIA	64	R0	A
37	>2010	THRBF15330T1	FFPE	100	✓	FFPE	AB	F	I	61	R0	A
38	>2010	THRBZ16097T3	Frozen	100	✓	Frozen	AB	M	I	72	R0	A
39	2000-9	THRBF19872T1	FFPE	100	✓	FFPE	AB	F	I	72	R0	44
40	<2000	THRBZ21732T2	Frozen	100	✓	FFPE	AB	M	I	57	R0	A
41	2000-9	THRBF24855T2	FFPE	80	✓	FFPE	AB	M	I	70	R0	A
42	2000-9	THRBF25392T1	FFPE	100	✓	FFPE	AB	F		31		A
43	>2010	THRBZ25604T1	Frozen	100	✓	FFPE	AB	F	I	66	R0	A
44	>2010	THRBF25774T1	FFPE	100	✓	FFPE	AB	F	II	67	R1	A
45	2000-9	THRBF28521T1	FFPE	100	✓	FFPE	AB	F	I	44	R0	A
46	<2000	THRBZ32960T2	Frozen	100	✓	FFPE	AB	F	IIA	66	R0	A
47	<2000	THRBZ33408T2	Frozen	100			AB	M	I	74	R0	A
48	<2000	THRBZ42901T2	Frozen	100			AB	M	I	68	R0	134
49	2000-9	THRBF54962T1	FFPE	80	✓	FFPE	AB	F	I	50	R0	A
50	2000-9	THRBF55773T1	FFPE	100	✓	FFPE	AB	F	IIA	41	R0	A
51	<2000	THRBF73550T1	FFPE	100	✓	FFPE	AB	F	I	60	R0	A
52	2000-9	THRBZ73761T2	Frozen	100	✓	FFPE	AB	F	IIA	45	R0	A
53	2000-9	THRBF79419T1	FFPE	80	✓	FFPE	AB	F	I	75	R0	42
54	<2000	THRBZ8209T2	Frozen	100	✓	FFPE	AB	M	IIA	72	R0	A
55	2000-9	THRBF84054T1	FFPE	80			AB	F	IVA	49	R2	98

56	2000-9	THRBZ95591T2	Frozen	95	✓	Frozen	AB	M	I	53		R0	61
57	2000-9	THRBZ95929T2	Frozen	90	✓	Frozen	AB	F	2A	73		R0	A
58	2000-9	THRBZ96635T2	Frozen	90	✓	Frozen	AB	F	I	56		R0	A
59	>2010	THRBZ00546T2	Frozen	100	✓	Frozen	AB	M	I	70		R0	A
60	>2010	THRBZ08002T1	Frozen	100	✓	Frozen	AB	F	IIA	77		R0	A
61	2000-9	THRBZ67094T1	Frozen				AB	F	I	54		R0	A
62	>2010	THRBF28180T1	FFPE	60	✓	FFPE	B1	F	IV	65		R0	A
63	>2010	150092PT	FFPE		✓	FFPE	B1	F		60		R1	A
64	>2010	150099PT	FFPE				B1	F	II	60		R0	A
65	<2000	THBRF20178T1	FFPE	90	✓	FFPE	B1	M	IIA	49		R0	A
66	>2010	THRBF00036T2	FFPE	90	✓	FFPE	B1	F	IVA	43			
67	2000-9	THRBF10445T1	FFPE	100	✓	FFPE	B1	F	I	46		R0	40
68	>2010	THRBF12588T1	FFPE	100	✓	FFPE	B1	F	II	70		R1	A
69	>2010	THRBF12695T1	FFPE	95	✓	FFPE	B1	F	IIA	54		R0	A
70	2000-9	THRBF13228T1	FFPE	80			B1	F	III	58		R0	A
71	2000-9	THRBF15530T1	FFPE	75	✓	FFPE	B1	F	I	53		R0	A
72	2000-9	THRBF16249T1	FFPE	100	✓	FFPE	B1	F	IVA	48		R0	A
73	>2010	THRBF19516T2	FFPE	100	✓	FFPE	B1	F	IIA	64		R0	A
74	>2010	THRBF23148T1	FFPE	100	✓	FFPE	B1	F	IIA	57			A
75	2000-9	THRBF25622T1	FFPE	100			B1	M		26			A
76	2000-9	THRBF25781T1	FFPE	100			B1	F	III	53		R0	A
77	2000-9	THRBZ56585T2	Frozen	100	✓	FFPE	B1	F	I	41		R0	A
78	<2000	THRBF87472T1	FFPE	100	✓	FFPE	B1	M	IVB	37			114
79	>2010	THRBZ02570T1	Frozen	100	✓	Frozen	B1	M	IIA	73		R0	A
80	>2010	THRBZ03029T1	Frozen	100	✓	Frozen	B1	M	II	46		R1	
81	2000-9	THRBZ84946T1	Frozen	90	✓	Frozen	B1	F	III	69		R0	
82	<2000	THRBF59802T1	FFPE	100	✓	FFPE	B1	M	IVA	44		R0	A
83	>2010	THRBF17723T1	FFPE		✓	FFPE	B1	M	II	60		R0	A
84	2000-9	THRBZ74105T1	Frozen	100	✓	FFPE	B2	M	IVA	54		R0	A
85	>2010	10LW00453TIS1	Frozen	95			B2	F	IVB	56			28
86	>2010	150095PT	FFPE				B2	F		55		R0	A

87	<2000	THRBF00151T1	FFPE	100	✓	FFPE	B2	F	III	48	R0	A
88	>2010	THRBF02631T1	FFPE	100	✓	FFPE	B2	M	II	59	R0	A
89	>2010	THRBF05018T1	FFPE	90	✓	FFPE	B2	M	III	53	R0	A
90	2000-9	THRBF05109T2	FFPE	30	✓	FFPE	B2	M	IVA	68	R1	
91	2000-9	THRBF06854T1	FFPE	100			B2	M	IIB	64	R0	A
92	2000-9	THRBF08261T1	FFPE		✓	FFPE	B2	M	III	66	R0	A
93	>2010	THRBF08297T1	FFPE	100	✓	FFPE	B2	M	III	76	R0	A
94	2000-9	THRBF08609T1	FFPE	80	✓	FFPE	B2	F	IVA	42	R0	
95	2000-9	THRBF13057T1	FFPE		✓	FFPE	B2	M	IIB	58	R0	A
96	>2010	THRBF15321T1	FFPE	100	✓	FFPE	B2	F	II	67		A
97	>2010	THRBF19801T2	FFPE	50	✓	FFPE	B2	M	IIB	67		A
98	2000-9	THRBF20036T1	FFPE	95			B2	F	IVA	40	R1	65
99	>2010	THRBZ21401T1	Frozen	100	✓	Frozen	B2	F	IIA	60	R1	A
100	<2000	THRBF2161T1	FFPE	80			B2	F	IVA	65	N/A	
101	>2010	THRBF21890T1	FFPE	80	✓	FFPE	B2	M	II	60	R0	A
102	>2010	THRBF23542T1	FFPE	100	✓	FFPE	B2	M	III	63	R2	
103	2000-9	THRBF25074T1B	FFPE	90	✓	FFPE	B2	M	IIB	40	R0	
104	<2000	THRBF25324T1	FFPE	100	✓	FFPE	B2	M	IIA	61	R0	A
105	>2010	THRBF27768T2	FFPE	100	✓	FFPE	B2	F	IVB	41	R0	
106	<2000	THRBF34401T1	FFPE	100	✓	FFPE	B2	M	IVA	70	R2	69
107	<2000	THRBZ36836T2	Frozen		✓	FFPE	B2	M	IVA	30		
108	2000-9	THRBF47489T2	FFPE	90	✓	FFPE	B2	F	IVA	21	R0	
109	2000-9	THRBZ48563	Frozen	100	✓	FFPE	B2	F	IVA	68	R1	
110	2000-9	THRBF49184T1	FFPE	95			B2	F		57		A
111	2000-9	THRBF60143T1	FFPE	90	✓	FFPE	B2	F	IVA	66	R0	A
112	2000-9	THRBF62382T1	FFPE	50	✓	FFPE	B2	F	IIA	38	R0	A
113	2000-9	THRBF89748T2	FFPE	20	✓	FFPE	B2	F	I	18	R0	A
114	<2000	THRBF99775T1	FFPE	100	✓	FFPE	B2	F	III	56	R1	A
115	<2000	THRBF99782T1	FFPE	100	✓	FFPE	B2	M	IIB	78	R0	A
116	>2010	THRBZ11462T2	Frozen	100	✓	FFPE	B2	F	IIA	71	R0	3
117	<2000	TCRBZ16327T2	Frozen	100	✓	FFPE	B2 (!)	F	IVA	62	R0	15

118	>2010	150108PT	FFPE				B2&B3	M	IV (R)	47	R0	A
119	>2010	150089PT	FFPE				B3	M	III	73		A
120	>2010	150096PT	FFPE				B3	M		77		A
121	>2010	THRBF10175T1	FFPE	100	✓	FFPE	B3	F	II	48	R0	A
122	2000-9	THRBF13799T1	FFPE		✓	FFPE	B3	F	IIA	63	R0	A
123	>2010	THRBF23949T2	FFPE	80	✓	FFPE	B3	M	II	74	R0	A
124	2000-9	THRBF70621T1	FFPE	95	✓	FFPE	B3	F	IVB	30	R0	25
125	>2010	THRBZ16503T2	Frozen	100			B3	M	IIA	58	R0	A
126	>2010	TCRBF12844T1	FFPE	100	✓	FFPE	B3 (!)	M	IVB	59	R1	A
127	<2000	TCRBZ03870T1	Frozen	50	✓	FFPE	CA	M	IVB	48	R0	160
128	>2010	TCRBF06753T1	FFPE	100			CA	M	III	64	R1	
129	>2010	TCRBF08757T1	FFPE	80	✓	FFPE	CA	F		63		
130	2000-9	TCRBF15505T3	FFPE	70	✓	FFPE	CA	F	IVB	58	R0	46
131	2000-9	TCRBF15586T1B	FFPE	50	✓	FFPE	CA	F	IIB	50	R0	A
132	<2000	TCRBF16264T1	FFPE	60	✓	FFPE	CA	M	III	60	R0	75
133	>2010	TCRBF19400T1	FFPE	100	✓	FFPE	CA	M		73	R0	A
134	2000-9	TCRBF21432T3	FFPE	80	✓	FFPE	CA	M	III	64	R2	
135	>2010	TCRBF25163T3	FFPE	30	✓	FFPE	CA		III		R0	
136	>2010	TCRBF26901T1	FFPE	100			CA	M	III	69	R1	A
137	2000-9	TCRBF31499T1	FFPE	100	✓	FFPE	CA					
138	<2000	TCRBF91244T4	FFPE	90			CA	M	IIB	51	R0	108
139	>2010	150086PB	FFPE	N/A			Hyperplasia	F	N/A	29	R0	A
140	2000-9	TPRBF97974P1	FFPE	N/A			Hyperplasia	F	N/A	23	R0	A
141	2000-9	TPRBF26458	FFPE	N/A	✓	FFPE	Hyperplasia	F	N/A	40		A
142	2000-9	TPRBF018	FFPE	N/A	✓	FFPE	Hyperplasia		N/A			A
143	2000-9	TPRBF1519	FFPE	N/A			Hyperplasia	F	N/A	28		A
144	2000-9	TPRBF04051	FFPE	N/A			Hyperplasia		N/A			
145	2000-9	TPRBF02155	FFPE	N/A			Hyperplasia		N/A			A
146	2000-9	TPRBF26464	Frozen	N/A			Hyperplasia	F	N/A	23		A
147	>2010	150075FB	Frozen	N/A	✓	FFPE	Hyperplasia	F	N/A	39	R0	A
148	>2010	150087PB	FFPE	N/A			Hyperplasia	M	N/A	27	R0	A

149	>2010	150104PB	FFPE	N/A			Hyperplasia	F	N/A	44	R0	A
150	>2010	150114PB	FFPE	N/A			Hyperplasia	F	N/A	34	R0	A
151	2000-9	TNRBZ63876T2	Frozen	100	✓	FFPE	NETT	M	IVB	52	R0	D
152	2000-9	TCRBZ64925T1	Frozen	100	✓	FFPE	NETT	F				108
153	2000-9	TNRBF73531T1	FFPE	90	✓	FFPE	NETT	M	IIA	58	R0	81
154	<2000	TNRBF76261T1	FFPE		✓	FFPE	NETT	F	IVB	49	R0	46
155	2000-9	TNRBF92784T1	FFPE	100	✓	FFPE	NETT	F	IVB	57	R0	D
156	2000-9	TNRBF95414T1	FFPE	70	✓	FFPE	NETT	F	IVB	67	R0	7

3.3 Prevalence of myasthenia gravis (MG) in this patient cohort

In this study MG was reported in 27 of 144 TET cases (19%) and two of twelve hyperplasia cases (17%). The prevalence of MG was highest in B2 thymomas (12/33, 36%), followed by B3 (2/9, 22%) and B1 thymomas (4/21, 19%). The prevalence dropped in Type A thymomas (4/22, 18%) and AB thymomas (4/37, 11%). In MNTs there were no cases of MG (0/4, 0%). Only one of twelve cases of SCC had MG (8%) and there were no cases in the NETTs. The overall MG prevalence in TETs has been reported to be around 24-40% (48). It is therefore very likely that the figures in the current study underestimate the true prevalence of MG in TETs. This could be due to incomplete or inaccurate history taking or documentation but also because they are unlikely to include any case that developed MG following their TET diagnosis and surgery to remove it. These findings are however in agreement with those reported by TCGA that found MG to be commoner in B thymomas compared to A, AB, MNT and carcinomas (107).

Table 3.2. MG prevalence in current TET cohort

Subtype	MG (Y/N)	MG %
A	4/22	18
AB	4/37	11
B1	4/21	19
B2	12/33	36
B3	2/9	22
CA	1/12	8
NETT	0/6	0
MNT LS	0/4	0
HP	2/12	17
TOTAL	29/156	19

MG prevalence in all TET subtypes and thymic hyperplasias in the current study. All MG cases include diagnoses before or at time of TET diagnosis.

3.4 Nature of collected samples

3.4.1. Suitability of FFPE for NGS

Of the 156 TET samples, 42 (27%) had been either snap frozen or were put in *RNAlater* and snap frozen. The remaining tumour samples were all FFPEs. Of the matched normal tissue only 14 of 119 samples had been snap-frozen with or without *RNAlater* (12%).

The quality of FFPE tissue for any downstream application, including DNA extraction for NGS has been questioned. Isolated nucleic acids are often fragmented and possibly modified compared to those from frozen tissue from the same tumour(200). Cross-linking between nucleic acids and proteins as well as addition of methylol groups can also occur(201). Some of these issues have been effectively dealt with. For example cross-linking can be reversed by digestion with proteinase K(202). Other studies have shown that as little as 5ng of DNA from FFPEs can generate good quality NGS data and provide CNA data similar to these from DNA extracted from frozen tissue (203). Another study recently looked at 23 frozen samples 38 FFPEs from three different cancer types (bladder, prostate an colon carcinoma) and liver and colon normal tissue as well as reactive tonsil tissue to establish the use of FFPEs in NGS-based studies, including RNA sequencing. DNA and RNA were isolated with two different methods and samples were used for exome and transcriptome analysis. Interestingly RNA libraries could be prepared from RNA extracted from FFPEs from 20 years earlier. Exome sequencing data in this study showed a concordance of 70-80% between frozen and FFPE samples that had been collected within three years. Unexpectedly perhaps RNA sequencing data revealed high correlation of expression profiles between the two that was irrespective of the duration of storage (Pearson Correlations of 0.90 +/- 0.05), overall showing promising results in terms of the utility of FFPEs in prospective and retrospective studies for which frozen tissue is unavailable (200).

3.4.2 Normal tissue for germ line DNA sequence analysis

Availability of 'normal' DNA from patients alongside DNA from tumour tissues is key to determine whether mutations or CNAs are somatic or germline. Blood is more commonly used for this purpose than matched normal tissue. One of the reasons for this is that it is often very hard to obtain such normal tissue when performing a biopsy or surgery to remove a tumour due to its proximity to the actual malignancy. In prospective studies it is much easier to collect blood samples for germline DNA sequence analysis. There is also the worry that tissue assessed to be normal by IHC is not truly normal and may carry genomic and translational aberrations that are shared with the nearby tumour (204). The majority of studies in cancer genomics have therefore based their

analyses on malignant tissue only when it comes to identifying driver mutations that can subgroup cancer types and/ or predict the natural history of the disease or response to certain therapies (205). It has been suggested however in certain tumour types, such a breast and colorectal carcinomas, gene expression of tissue outside the tumour area may predict clinical outcomes (206,207). There are a number of theories to explain this that include the tumour cell 'contamination' theory, by which cells that have escaped from the original tumour are influencing genomic signals in surrounding tissues (208). This is quite similar to the tumour 'field' theory (also called field cancerization), which proposes that tissue nearby tumours is in a state between that of normal cells and malignant ones (209). The third theory is that non-malignant tissues in close proximity to tumours themselves contain signals that influence the tumour microenvironment to the extent that these contribute to tumour development and growth (210). In fact a recent study tried to answer the above questions by analyzing data available on tumours and matched normal tissue from TCGA. Using gene expression profiling it was established that tumour and normal tissues of breast cancer patients cluster very distantly from one another. Surprisingly the transcriptional profiles of the adjacent normal tissues were more predictive of patient survival compared to those of tumours. This could mean that a 'field effect' and the tumour microenvironment are working together. It certainly highlights the importance of the tumour microenvironment as it can provide invaluable information for the tumour and the patient. It is however worth considering that the definition of adjacent normal tissue can be problematic as the distance from the actual tumour needs to be defined and establishing whether that distance makes any difference to the above findings is key (205).

Another very interesting study in patients with hepatocellular carcinoma examined tumour tissue, adjacent normal tissue and blood with whole genome sequencing. Surprisingly they found that somatic mutations in tumour samples were significantly fewer compared to those in adjacent 'normal' tissue. Certain other gene mutations that were absent in white blood cells were detected in adjacent tissues and at even higher frequency in tumour tissues. There were however other mutations that were seen more frequent in normal tissue compared to tumour tissue. A total of 506 genes with non-silent mutations were seen in tumours only. This study, that performed a three-way analysis gives credit to the theory that adjacent 'normal' tissue is not in fact normal and is perhaps in a state between normal and malignant at least in this tumour type (211).

It's worth pointing out that the nature of TETs is such that macroscopically (and histologically) there is, in the majority of cases, a very clear distinction between tumour and normal tissue. Also quite often normal tissue was collected from nearby lymph nodes often as an explorative procedure to look for disease spread from the primary TET. Therefore drawing conclusions about the suitability of such tissue for germline DNA is hard although given the above literature it is considerably more likely that this 'adjacent' matched normal tissue was indeed normal.

Chapter 4
GTF2I mutation in TETs

Abstract

Sanger sequencing was used to assess the frequency of the c.74146970T>A *GTF2I* mutation in 141 TET and thymic hyperplasia cases. It was seen in 48 of 141 evaluable TETs (34%), its prevalence being 90% and 69% in type A and AB thymomas respectively. This decreased to 16%, 6% and 13% in B1, B2 and B3 thymomas respectively. None of the twelve SCCs or eight NETTs had the mutation. It was not detected in any of the eight thymic hyperplasias (mutation in 40/52 A and AB thymomas vs. 0/8 hyperplasias, p = 0.020). No link between myasthenia gravis or overall survival in our TET cohort and the *GTF2I* mutation status was seen. We have also shown that AB thymomas have a much lower percentage of mutant alleles compared to the majority of the A thymomas. These findings could be of value as a diagnostic test in difficult to diagnose mediastinal masses when there is a diagnostic dilemma between thymic hyperplasia and type or AB thymomas.

4.1 The aims of the research described in this chapter were the following:

- To provide an overview of the importance of GTF2I mutations in TETs
- To assess the frequency of GTF2I mutations in our patient cohort
- To assess the frequency of GTF2I mutations in the TCGA cohort
- To compare these findings with the existing literature on GTF2I and correlate them with key clinical characteristics

4.2 Background

In June 2014 Petrini *et al* identified with next generation sequencing (NGG) a novel, recurrent missense mutation (T>A, chr7:74146970 (hg19)) in the *General Transcription Factor II-I (GTF2I)* gene. This mutation changes a leucine-to-histidine in the δ and β isoforms of TFII-I (p.Leu383His and p.Leu404His respectively). TFII-I's function is to regulate genes involved in cell cycle, cell proliferation and developmental processes. This mutation was found in *GTF2I* but not its pseudogenes *GTF2IP1* and *LOC10093631* and Sanger sequencing of tumour and normal DNA of all cases used for NGS confirmed it as a somatic mutation. This alteration is predicted to be deleterious by PolyPhen-2 and SIFT. The frequency of the *GTF2I* mutation was assessed in 274 TETs (270 tumours and four cell lines) using Sanger sequencing and a deep sequencing approach and shown to be 43.4%. It was strikingly prevalent in the A (82%) and AB (74%) histotypes, which

are the most indolent subtypes. The majority of cases with the mutation presented with Stage I-II disease (57%) (106).

For the 204 patients with available survival data those with the mutation had a better 10-year survival than wild-type cases (96% Vs 70% respectively; log-rank $p < 0.001$). The three carcinomas with the mutation were all alive (median follow-up time of 27.6 months, 95% CI 0–70 months). These cases also had a lower Ki67 staining than wild-type cases (5% Vs 14%). The favorable effect of having the *GTF2I* mutation was seen amongst thymomas, with a 10-year survival of 96% in mutant Vs 88% in wild-type cases (log-rank $p = 0.057$) although there was no difference in Ki67 staining. RNA sequencing showed expression of the *GTF2I* mutation in all five A and two AB thymomas that were sequenced. It also showed that these tumours cluster together in an A and AB-rich group. Although not the case at the mRNa level, at the protein level *GTF2I* mutant tumours exhibited a higher TFII-I expression. As this *GTF2I* mutation was highly prevalent in the A and AB subtypes it was concluded that it is likely a marker of good prognosis (106).

Recently another group analyzed whole exome sequencing data from 120 TET cases that have become available from the Cancer Genome Atlas (TCGA). They found the prevalence of the mutation to be 38% in this cohort (46/120). They also found the presence of the mutation to be linked with a more indolent WHO subtype (A, AB Vs B1, B2, B3 and CA), with a less advanced stage and the absence of MG. Consistent with the previous findings, 7 of 54 (13%) patients with B2/ B3/ CA and GTF2I mutant tumours did not have a recurrence of their TET. This was also the case for the 5 of 21 patients (23.8%) with Stage III or IV disease at presentation and were GTF2I mutant, who did not have a single recurrence (177).

4.3 Methods

Tumour samples were collected from 144 patients as described in chapter 2. These were subject to histological assessment by expert Consultant Histopathologists to confirm the diagnosis and determine tumour abundance (See section 2.2.3). Of the 141 evaluable TETs, 20 were type A, 32 AB, 19 B1, 32 B2 and eight B3 thymomas. There were four cases of micronodular thymoma with lymphoid stroma (MNT LS), and eighteen carcinomas, twelve of which were squamous cell carcinomas (SCCs) and six were neuroendocrine carcinomas (NETTs). There were also eight case of thymic hyperplasia. Only eight of 141 (6%) patients had had primary chemotherapy prior to resection of their tumour and none had had radiotherapy. Stage was recorded as per the Masaoka-Koga classification. DNA from FFPEs was extracted with Quiagen's QIAamp DNA FFPE Tissue Kit (Catalogue No. 56404) as described in 2.2.6 and from frozen tissue with Quiagen's AllPrep DNA/RNA/miRNA Universal Kit (Catalogue No. 80224) as described in 2.2.7. PCR and

Sanger sequencing were performed with semi-nested primers as described in 2.2.4 and 2.2.5. DNA extraction from FFPEs (2.2.6) and snap frozen tissue (2.2.7) as well as DNA quantification (2.2.8) were performed exclusively by me. Sanger sequencing was performed commercially by Eurofins Genomics. Data analysis was performed exclusively by me with the use of SEQUENCHER (2.2.5).

A two-tailed Mann Whitney Test was used to compare the mean age of the mutant and wild-type groups and a chi-square test to compare the male: female ratio. Fisher's exact test was used to compare the percentage of mutant to wild-type cases across the four groups in the different TET subtypes.

4.4 Results

The frequency of the GTF2I mutation was assessed in a total of 144 TETs. Excluding the eight cases of thymic hyperplasia the mean age for all patients with malignancy was 59. In the GTF2I mutant group the mean age was 63 whereas it was 56 in the wild-type group (Two-tailed Mann Whitney Test Z score = -3.109, p = 0.002). For the entire cohort the male: female ratio was 1:1.30. In the GTF2I mutant group this was 1:1.05 (23M: 24F) whereas in wild-type group it was 1:1.46 (37M: 54F), which is not significant (The chi-square statistic is 0.864, p=0.35).

In this cohort the GTF2I mutation was seen in 48 of 141 evaluable TETs (34%) and was present more commonly in type A (90%) and AB (69%) thymomas. The frequency decreased to 16%, 6% and 13% in B1, B2 and B3 thymomas respectively. The mutation was not detected in any carcinomas, neuroendocrine tumours or in any thymic hyperplasias. Table 1 shows the number and percentage of GTF2I mutations per TET subtype compared with the findings of Petrini et al, TCGA and Feng et al. The average tumour percentage for all 141 TETs was 91%. For GTF2I mutant and wild type cases it was 94% and 89% respectively. Forty-seven of the 48 patients (98%) with the mutation had Stage I – II disease, according to the Masaoka-Koga staging system, at presentation and 45 of the 48 (94%) had complete resection (R0) of their TET. The prevalence of MG in the GTF2I mutant and wild type groups was 21% (6/28) and 20% (19/93) respectively.

Establishing a reliable correlation between the GTF2I status and overall survival was not possible given that only five of the 48 patients with a GTF2I mutation had died (10.4%). Seventeen of the 93 wild-type patients had died (18.3%). Eight of the seventeen had either a squamous cell or neuroendocrine carcinoma. Of the GTF2I mutant cases, one of 18 patients with a type A thymoma and one of 22 with an AB had a recurrence of their thymoma. Of the five patients that were GTF2I

mutant and died, one had a Type A, three an AB and one a B1 thymoma. Whether the cause of death was linked to their TET is not known.

Table 4.1. Prevalence of the *GTF2I* mutation in TETs according to subtype in our cohort Vs Petrini *et al.* (106), TCGA (107) and Feng *et al*. (212)

TET subtype	**NHLI (n=141)**	**Petrini (n=274)**	**TCGA (n=117)**	**Feng (n=296)**	**ALL (n=828)**
CA	0/12 (0%)	3/36 (8%)	0/9 (0%)	3/36 (8%)	6/93 (6%)
NETT	0/6 (0%)	0/4 (0%)	0/1 (0%)	N/A	0/11 (0%)
A	18/20 (90%)	46/56 (82%)	10/10 (100%)	20/23 (87%)	94/109 (86%)
MNT LS	2/4 (50%)	1/2 (50%)	1/2 (50%)	N/A	4/8 (50%)
AB	22/32 (69%)	40/54 (74%)	34/48 (71%)	70/89 (79%)	166/223 (74%)
B1	3/19 (16%)	9/28 (32%)	0/12 (0%)	17/58 (29%)	29/117 (25%)
B2	2/32 (6%)	7/32 (22%)	0/25 (0%)	8/40 (20%)	17/129 (13%)
B3	1/8 (13%)	13/62 (21%)	1/10 (10%)	3/30 (10%)	18/110 (16%)
Mixed B	N/A	N/A	N/A	3/17 (18%)	3/17 (18%)
Hyper	0/8 (0%)	N/A	N/A	N/A	0/8 (0%)
TOTAL	48/141 (34%)	119/274 (43%)	46/117 (39%)	124/296 (42%)	337/828 (41%)

NHLI: National Heart and Lung Institute; **TCGA:** The Cancer Genome Atlas; **CA:** Carcinoma; **NETT:** Neuroendocrine thymic tumour; **MNT LS:** Micronodular thymoma with lymphoid stroma; **Mixed B:** Cases of more than one B subtype; **Hyper:** Hyperplasia

4.5 Discussion

The *GTF2I* (chromosome 7 c.74146970T > A) is the commonest mutation in TETs to date, with no other recurrent mutations seen across any subtype (106,107). The results of this study confirm the presence of the *GTF2I* mutation at a high frequency in type A and AB thymomas in a separate patient cohort. Between the four groups, (NHLI, Petrini *et al*., TCGA and Feng et al) a total of 828 TETs were evaluated for the presence of the mutation in *GTF2I* (chromosome 7 c.74146970T > A). The overall frequency across all subtypes was 41% (Figure 4.1). Overall the proportion of mutant A and AB thymomas is very similar across the three groups (86% and 74% respectively). In the TCGA cohort all ten Type A thymomas were mutant. For B1 thymomas the overall frequency was 25% although in TCGA there were no mutant cases out of a total of twelve. In the case of B2 thymomas NHLI, Petrini *et al,* TCGA and Feng et al had a mutation frequency of 6% (2/32), 22% (7/32), 0% (0/25) and 20% (8/40) respectively. This difference was significant (*p=0.020*). For B3 thymomas NHLI, TCGA and Feng et al. show a mutation frequency of 13%, 10% and 10% respectively with Petrini et al having found a much higher percentage of mutant cases (21%) although this difference is not statistically significant (*p=0.630*). For carcinomas, the overall frequency of the mutation is very low at 6% with neither NHLI nor TCGA having any mutant cases. There were no mutations in any of the eleven neuroendocrine carcinomas examined.

The most significant observation from this study is that there were no mutations in the eight cases of thymic hyperplasia. Having established the GTF2I mutation as one that is exceedingly more prevalent in more indolent TETs and with theories questioning whether there might progression from thymic hyperplasia to thymomas the absence of the mutation in hyperplasias makes that theory less plausible. This novel data supports GTF21 mutations being part of the neoplastic process. This may also have clinical relevance in cases when there is a diagnostic dilemma between a type A thymoma and thymic hyperplasia, a situation that not infrequently occurs in biopsy material (22,213). Also, thymic neuroendocrine tumours (NETTs) and spindle cell B3 thymomas can also morphologically resemble type A thymomas (10,213). Immunohistochemistry is currently the mainstay for helping differentiate between thymomas, NETTs and other mediastinal neoplasms such as lymphomas and germ cell tumours (214), although is less useful in the distinction of subtypes of thymoma and distinguishing thymic hyperplasia from thymoma. In the entire cohort of tumours sequenced by TCGA this mutation was not seen in any other tumour type (107). As this mutation in GTF2I is present in 90% of Type A thymomas, it could be used as an adjunct to existing tests to aid with their diagnosis and/ or exclude other conditions, including thymic hyperplasia.

Although Petrini et al. reported GTF2I as a good prognostic factor, when it was examined together with stage and WHO classification, the only independent prognostic factor was stage at diagnosis. In that study two of 83 GTF2I mutant evaluable patients died of tumour progression; one had a B2 and the other a B3 thymoma. 26 of the 121 evaluable GTF2I wild type cases died of tumour progression (21%) The difference with the proportion of deaths due to progression in the mutant cases was found to be statistically significant. However 111 of the 155 cases (72%) that were GTF2I wild type (overall number of cases), were B2, B3 thymomas or carcinomas. In other words the cases of more aggressive subtypes were over-represented in this group (106). This was also the case in this study as eight of 17 patients that were GTF2I wild type and passed away had a squamous cell or neuroendocrine carcinoma. One of each of the A and AB thymomas with a GTF2I mutation had a recurrence of their disease. Although this goes against this being a marker of good prognosis these are very few cases.

An intriguing observation in the NHLI cohort is that 20 of the 22 AB thymomas (91%) with the mutation had a much lower percentage of mutant alleles compared to the majority of the A thymomas that had a high percentage (15 of 18, 83%). An example of each can be seen in Figure 2. This was observed with both the forward and reverse primers. The explanation for this is not clear. A hypothesis that the percentage of mutant alleles is linked to the A component of TETs is less likely to be true given that 3 TETs of the 6 B subtype (B1, B2 and B3) had a high percentage of *GTF2I* mutant alleles. Looking at the TCGA data, where we have detailed information on the variant allele frequency (VAF) a similar effect is observed. The average VAF is low at 25.2, ranging from 6.9% to 53.5%. All the cases that were reported as mutant but not were not validated by TCGA and hence excluded from this analysis had a VAF of less than 6.7%. The mean VAF for the 14 Type A mutant cases was 37.5% whereas the mean VAF for the 28 mutant AB thymomas was 20.4%. For the 2 B2, 3 B3 and 1 carcinoma it was 8.5%, 17.7% and 44% respectively.

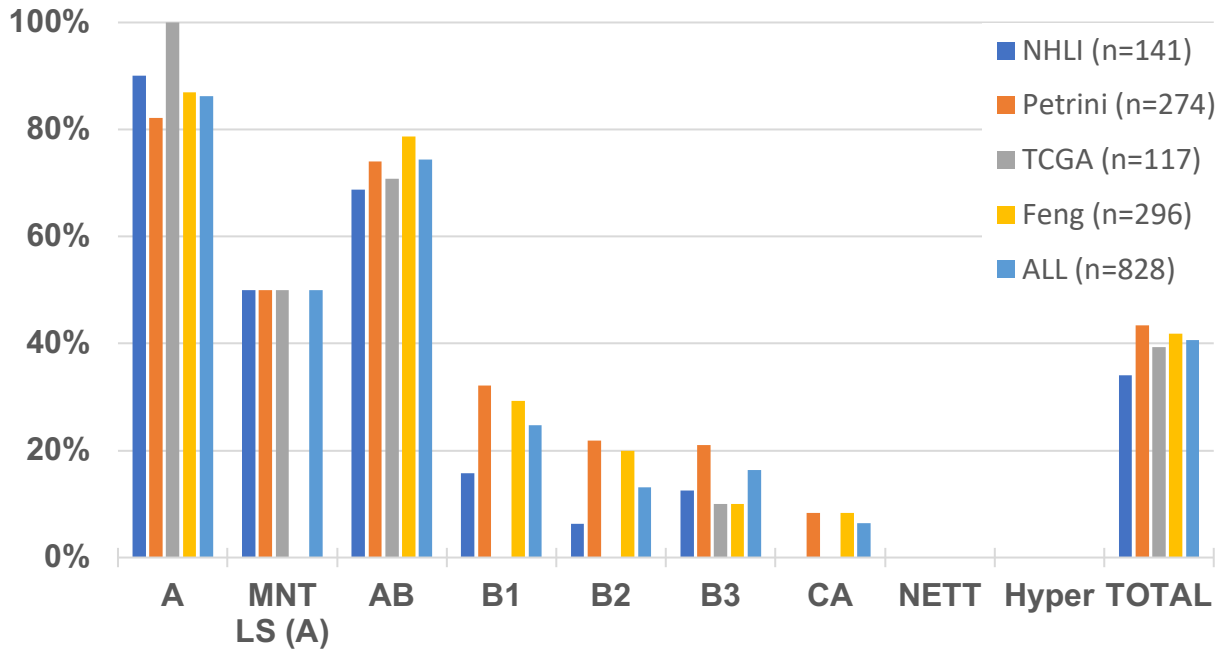
Regarding the MG status of the GTF2I mutant cases in this cohort there was no difference compared to wild type. This is in contrast with the findings by Lee et al. who found GTF2I mutant cases were associated with absence of MG (177). There is no biological reason why this should be the case. In fact the only link between GTF2I and any autoimmune condition so far has been that its rs117026326 polymorphism (a noncoding change in an intron of GTF2I) is associated with anti-SSA-positive primary Sjögren's syndrome. STAT4 and TNFAIP3 were the other 2 non-HLA susceptibility loci(215).

The significant difference observed in mean age of GTF2I mutant and wild type patients, which was 63 and 56 respectively, can be explained by the fact that mean age for patients with Type A thymomas has been found to be 60.5 years whereas it is 48 for patients with B1, B2 and B3 thymomas. The finding that in the wild type group female patients were over-represented, although

not significant, is contradictory to the same study that found the proportion of men with Type A thymoma was only 33%, vs. 49% for AB, 38% for B1, 58% for B2 and 54% for B3 (24).

In conclusion, this study has confirmed the presence of the GTF2I mutation at a high frequency in indolent thymomas (type A and AB). It is not present in thymic hyperplasias or NETTs. This opens the road for the presence of this mutation to be used as a diagnostic test in cases where there is a diagnostic difficulty between thymomas, particularly type A, and other conditions especially if tissue diagnosis is being sought using a biopsy.

Figure 4.1. Frequency of *GTF2I* mutant cases according to TET subtype.



Results from all four groups: NHLI, Petrini et al, TCGA and Feng et al according to TET subtype are shown

Figure 4.2. Sanger sequencing traces demonstrating the T>A mutation in *GTF2I* on chromosome 7, 74146970

Figure 4.2.1. Example of a T>A mutation in *GTF2I* in a type AB thymoma (low % of mutant alleles)

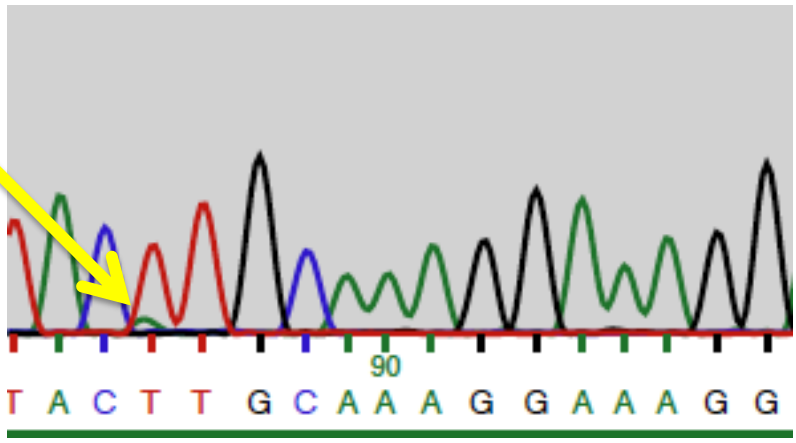
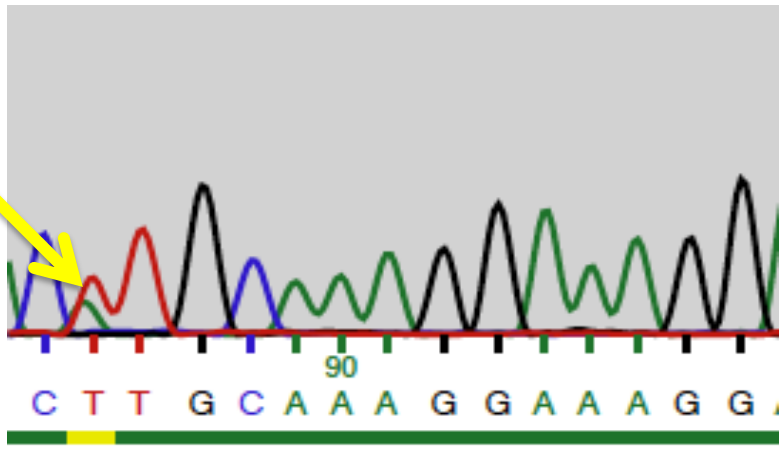


Figure 4.2.2. Example of a T>A mutation in *GTF2I* in a type A thymoma (high % of mutant alleles)



Chapter 5
Whole exome sequencing of TETs

Abstract

WES was performed on DNA extracted from FFPE material from a total of seventeen TETs. Data was analysed together with the WES mutation data from TCGA that have not been yet published. A very low mutation burden in TETs was seen, as was a striking absence of putative somatic driver mutations in B subtype thymomas. On *GTF2I* mutant type A and AB thymomas, approximately 25% of tumours also had an oncogenic mutation affecting the MAP kinase pathway, *HRAS* being the most commonly mutated gene. Including published data on TETs, in thymic carcinomas (SCCs and NETTs) TP53 is the most commonly mutated gene, with mutations found in 30% of cases. Somatic mutations of known cancer genes in our cohort were seen in BRCA2, KMT2D, MSH2 and MSH6 amongst others. We also identified a recurrent germline insertion in *HLA-A* was seen across the TET subtypes at a frequency of 27%. Other salient, isolated cases harboured putative driver, missense, germline mutations in *EGFR* (p.V1044L), *ERBB4* (p.R1259W), *CHEK2* (p.Arg361Leu) and *GNAQ* (p.R104G). A case of an individual with a *BAP1* deletion and a clinical history highly suggestive of BAP1 predisposition syndrome is likely the first reported case of thymic SCC presenting as part of this syndrome. Targeted re-sequencing of selected genes in an additional cohort of 125 matched TETs will help validate and expand these findings.

5.1 The aims of the research described in this chapter were the following:

- To identify the likely driver somatic mutations in TETs through WES
- To identify likely driver germline mutations in TETs through WES
- To identify pathways that are of importance for each TET subtype

5.2 Introduction

Next-generation sequencing (NGS), also referred to as massively parallel, deep sequencing or second-generation DNA sequencing refers to a non-Sanger-based DNA sequencing technology that allows billions of strands of DNA to be sequenced in parallel (216,217). As all cancers accumulate genomic alterations, analysis of their genomes provides the opportunity to investigate the cause of each cancer that can in turn help in their diagnosis and treatment. NGS technologies encompass whole-genome, whole-exome and transcriptome sequencing approaches. As the exome is only approximately 3% of the genome current platforms can achieve a very high coverage (216).

As the amount of data generated from such methods is vast, algorithms are used to identify mutations based on the number of reads in the tumours and their absence from normal DNA. As somatic mutations are quite rare it is really challenging to identify those amongst 'background noise', false positive changes/ mutations that often arise for many reasons, such as sequencing or alignment errors. Other issues arise from the fact that tumours are often not 100% made of neoplastic cells but instead are intermixed with normal cells (218). Another increasingly recognized issue is intra-tumour heterogeneity that means that even within one tumour there may be areas that are genetically very different from one another, harbouring entirely different allelic-imbalance profiles. This renders identification of driver mutations a lot more complex and challenging (219). Increasing the number of reads per base (sequencing depth) can in part improve mutation calling. Even then the recognition of driver vs. passenger mutations remains problematic. It relies on bioinformatics analysis that can take background mutation rates in the genome, variations within tumours of the same type and intra-tumour heterogeneity to identify mutations that are more frequent than 'random expectation' (218).

Despite the fact that other tumour types have been subject to WES multiple times in numerous studies only very few have focused on TETs. Petrini et al performed WES of 28 TETs, which formed the basis for the discovery of the *GTF2I* mutation. This is the commonest somatic mutation in TETs (overall frequency 40%) and is discussed in detail in chapter 4. Otherwise the main study findings from WES and targeted capture that was subsequently performed in an expanded cohort of 52 TETs, were an overall low mutation burden and very infrequently recurring mutations in genes such as *TP53* (four carcinomas and two B3 thymomas), *CYLD* (three carcinomas), *BAP 1* (two carcinomas and one thymoma), *CDKN2A* (two carcinomas), *RET* (one thymoma and one carcinoma), *PBRM1* (two carcinomas), *ALK* (two thymomas), *PPP2R1A* (two thymomas), and one of each of an *HRAS*, a *KRAS* and an *NRAS* hotspot mutation (106). TCGA confirmed TETs have a very low mutation burden and that the *GTF2I* mutation is the most highly recurrent mutation in TETs. HRAS, NRAS and TP53 mutations were amongst the other mutations found at low frequency (107). Other studies, using more targeted methods have also found TR53 and CDKN2A to be most commonly mutated in thymic carcinomas as well as KIT and NRAS at a lower frequency (105,108).

5.3 Materials & Methods

5.3.1 Samples and DNA extraction

TET genomic DNA was extracted from seventeen TETs from RBH that had been removed during a thymectomy or metastatectomy and were surplus to diagnostic requirement. Fifteen of these had matched normal tissue from which DNA was also extracted. All samples were subject to histological assessment by expert Consultant Histopathologists to confirm the diagnosis, using the 2015 WHO criteria, and determine tumour abundance as described in chapter 2.2.3 and DNA was extracted with Quiagen's QIAamp DNA FFPE Tissue Kit (Catalogue No. 56404) as described in chapter 2.2.6. The histological distribution of samples was as follows: Six SCCs, three NETTs, two B3 thymomas, three B2 thymomas, one of which was a rare case of an anaplastic B2 thymoma, one B1 and two AB thymomas. The focus was on more aggressive subtypes, for which there are fewer known drivers and are the ones, which are more likely to benefit for discovery of new, targeted therapies.

5.3.2 Whole exome sequencing

WES was performed by the Centro Nacional de Análisis Genómico (CNAG, Barcelona, Spain). Paired-end multiplex libraries were prepared according to manufacturer's instructions with Illumina TruSeq DNA Sample Prep kit and enriched with the NimbleGen SeqCap EZ Human Exome Library v3.0 (187) (chapter 2.2.9).

5.3.3 Data analysis and variant calling

CNAG provided VCF files (2.2.9.1). Variant filtering and identification of potential driver mutations was performed by me with the help of Amit Mandal, bioinformatician at NHLI (2.2.9.2)

5.4 Results

5.4.1 Patient demographics and sample characteristics

Seventeen TETs (6 SCCs, 3 NETTs, 2 B3, 3 B2, 1 B1, 2 AB) were subject to WES. The baseline patient and tumour characteristics are described in Table 5.1. There were nine male (53%) and eight female (47%) patients. The median age was 58 (range 30 - 70). Three patients were known

to have MG. Six presented with Stage IV disease (Masaoka-Koga classification). Nine of seventeen are known to be dead (53%). The mean tumour content for all cases was 79% (range 50 - 100).

5.4.2 Somatic mutations

5.4.2.1 Overview of somatic mutations

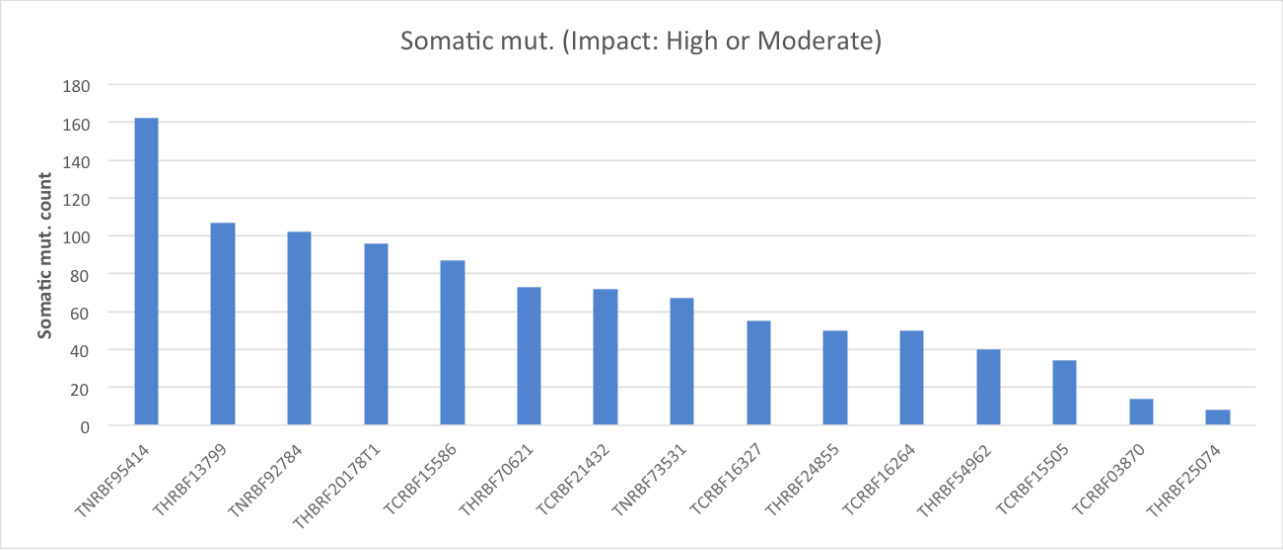
The exomes of all seventeen tumours and fifteen matched normal samples were sequenced successfully. Following filtering of the variants, for the fifteen matched cases, 1017 somatic mutations were identified in 793 genes. These included variants that were predicted to have a moderate or high impact only. The median number of mutations was 67 (average was 67.8; range eight – 162) (Figure 5.1). The low number of mutations observed is consistent with the other two studies that have performed WES on TETs (107,220). In fact Petrini et al. (106) only found an average of 28 mutations per sample (range 3 – 94) in the 28 samples studied and TCGA recently reported that of the 21 tumour types they have studied, TETs have the lowest mutation burden of all adult cancers (107). Although in our samples aggressive subtypes were over-represented with nine of seventeen (53%) being carcinomas and five (29%) B3 or B2 thymomas, which render such comparisons more difficult, we did not find higher number of mutations in carcinomas vs. thymomas as previously reported (107,220).

Table 5.1. Baseline patient and tumour characteristics for TET WES cases

Sample	Histology	Sex	Age	Tumour %	Stage	MG	S	Other	TMA code
TCRBF03870	SCC	M	48	50	IVB	Y	D		Pat 110
TCRBF15505	SCC	F	58	70	IVB	N	D		Pat 111
TCRBF15586	SCC	F	50	50	IIB	N	A		Pat 109
TCRBF16264	SCC	M	60	60	III	N	D		Pat 108
TCRBF21432	SCC	M	64	80	III	N	A		N/A
TCRBF91244	SCC	M	51	70	IIB	N	D	No Normal	Pat 112
TNRBF73531	NETT	M	58	90	IIA	N	D		Pat 120
TNRBF92784	NETT	F	57	100	IVB	N	D		Pat 124
TNRBF95414	NETT	F	67	70	IVB	N	D		Pat 125
THRBF13799	B3	F	63	70	IIA	N	A		N/A
THRBF70621	B3	F	30	95	IVB	N	D		Pat 104
TCRBF16327	B2 An	F	62	100	IVA	Y	D		Pat 95
THRBF06854	B2	M	64	100	IIB	N	A	No Normal	Pat 103
THRBF25074	B2	M	40	90	IIB	N	A		Pat 81
THRBF20178	B1	M	49	90	IIA	N	A		Pat 54
THRBF24855	AB	M	70	80	I	N	A	GTF2I mut.	Pat 25
THRBF54962	AB	F	50	80	I	Y	A	GTF2I mut.	Pat 24

Histology: WHO histotypes; **Age**: Age at diagnosis; **S**: Status (**D**=Dead, **A**=Alive), **TMA code**: Patient TMA code in chapter 7 (IHC); **Other**: Indicates two cases with no matched normal tissue and two cases known to be GTF2I mutant from Sanger sequencing; **B2 An**: Anaplastic B2 thymoma

Figure 5.1. Number of detected somatic mutations in 15 paired TETs



Mutations predicted to have a moderate or high impact only were included.

5.4.2.2 Candidate somatic driver mutations

5.4.2.2.1 Gene mutation recurrence in the NHLI cohort

As expected from the low overall number of somatic mutations in TETs in this and other studies, genes that were recurrently mutated were very few. In the fifteen, paired samples there were 131 genes with mutations in two or more samples (17% of the total number of genes). Of those only 47 were mutated in three or more samples (6% of all mutated genes) and 22 were mutated in four or more samples (3%). The eleven most commonly mutated genes (mutated in five samples or more) were the following: HRNR, NBPF1, CDC27, FRG1, HLA-DRB1, MUC4, PABPC1, PABPC3, MUC12, NBPF3 and ZNF91. None of these are known to be implicated in oncogenesis according to the Cancer Gene Census by COSMIC (http://cancer.sanger.ac.uk/census) (221). MUC4 and MUC12 are both members of the mucin family of genes. These encode for very large proteins and commonly thought to be a false positive in this context. A few studies have associated some members of the mucin family with poor prognosis in some malignancies, including lymphomas and breast cancer, but there has been no evidence that they constitute driver mutations for any of these (222,223). Many other mucins have been found to be less frequently mutated in these samples.

Given the nature of the samples in this study, with a high proportion of carcinomas, including SCCs but also NETTs that have not been included in most thymic studies so far, assessment of recurring gene mutations was performed separately in these subtypes despite the small numbers. Any mutations that were more frequent than 10^{-3} in ExAC_non-TCGA_adj_AF or had a CADD_phred score lower then 15 were filtered out.

In SCCs PABPC3 was mutated in three of six cases (50%). It was also overall the most commonly recurrent gene mutation across all subtypes following this filtering. It was mutated in one NETT, both B3 thymomas and one B2 thymoma with an overall mutation frequency of 7/17 (41%). A majority of these were missense and two were stop-gain protein truncation mutations. Missense and truncating mutations in this gene have been seen by TCGA in 3% of lung cancer studies but never in TETs (224,225).

Cell division cycle 27 (CDC27) was mutated in five of seventeen samples (29%), across all the subtypes that were sequenced, two SCCs (33%), the anaplastic B2, one B3 and one AB. This mutation has been seen as a putative passenger in one thymic carcinoma by TCGA (Accessed via http://www.cbioportal.org). However its role in cancer has not been established. A study has shown downregulation of CDC27 to be associated with decreased proliferation in colorectal cancer cells via the accumulation of p21Cip1/Waf1 (226).

SLC9B1 and HHLA3 were mutated in three cases each. Mutations in these genes have been seen in TCGA in lung cancer and are considered to be passenger mutations (Accessed via http://www.cbioportal.org). They have not been seen in TETs before.

5.4.2.2.2 Gene mutation recurrence in the TCGA cohort

As TCGA have sequenced the largest number of TETs so far, these results have been used as a reference and WES data were analysed alongside data from Petrini et al, who performed WES in 28 TETs and targeted sequencing of these and a further 26 samples. Of the 124 TETs sequenced by TCGA 99 were found to have mutations. These are unpublished data that have been accessed via the cBioPortal (http://www.cbioportal.org) and The National Cancer Institute GDC Data Portal (https://portal.gdc.cancer.gov). The GTF2I mutation was the commonest with 39% of all cases being mutated. This was followed by TTN that was mutated in ten cases (8%) and HRAS mutated in nine (7%) (Table 5.3).

Overall in TCGA, of the 50 genes that were most often mutated, the ones that are known cancer genes in order of frequency were: HRAS (9), TP53 (4), BCOR (3), SMARCA4 (3), NF1 (3), NRAS (3) and SPEN (2). This is a very small number of genes and a very low recurrence rate. These genes were predicted to be driver mutations in the majority of cases. MutSig ("Mutation Significance") analysis (227), which allows identification of genes mutated more frequently than expected by chance, taking mutational heterogeneity into account, was significant for GTF2I (Q-value = 1.84e-11), HRAS (Q = 7.88e-11), TP53 (Q = 4.730e-4) and NRAS (Q = 6.070e-3). Of the genes not known to be involved in tumourigenesis, which were also mutated in the TET TCGA cohort, MutSig was significant for UNC93B1 (five mutations; Q = 3.74e-7), CAPNS1 (three mutations; Q = 2.670e-5), MUC4 (six mutations; Q = 1.290e-4), BACH1 (nine mutations; Q = 0.0187), ATRN (three mutations; Q = 0.0252) and FOXD1 (two mutations; Q = 0.0478) (Accessed via http://www.cbioportal.org) (Table 5.3). Of the remaining genes that were not in the 50 most mutated, another 90 known ‘cancer genes’ were mutated.

It is striking how in certain subtypes, like B1, from the fourteen samples in TCGA, there were no putative drivers identified, no GTF2I mutations and not a single gene was mutated in more than one sample. Similar findings were seen in the 32 B2 cases, where there were only two putative drivers TP53 and CEBPA each in one sample, with hardly any recurrence even of putative passenger mutations. In the fifteen B3 thymomas there were no recurrent mutations and the only putative drivers identified were TP53 in one sample, AKT1 and STAG2 both in one sample and ASXL1 also in one sample.

5.4.2.2.3 Mutations in known cancer genes in TETs

In our cohort the moderate and high impact genes mutated across all seventeen samples were compared to a list of known cancer genes including the 976 genes known to be implicated in tumourigenesis according to COSMIC (http://cancer.sanger.ac.uk/census) (221), Leiserson et al. (228), Vogelstein et al. (229) and Tamborero et al. (230). Given the low mutation recurrence rate and in order to maximize the chances of finding putative driver mutations all four databases were queried even if they were only found in one sample. Filtering based on frequency in the EXAC database and CADDphred scores as described in chapter 2.2.9. Sixty known 'cancer genes' were found to have at least one mutation based on these criteria. With the aim of increasing the likelihood of identifying real driver mutations, these 60 genes were compared to the WES TCGA findings on TETs, breast and lung cancer (Accessed via http://www.cbioportal.org), to the WES findings by Petrini et al in TETs (106) and to Zehir et al (105) (looking specifically at TETs and then at the entire tumour cohort). Results are summarized in table 5.2.

Sanger sequencing for the *GTF2I* mutation (chromosome 7 c.74146970T>A) had already been performed as described in chapter 4 and both AB thymomas in the WES cohort had been found to harbour the mutation. This was missed in both cases by the samtools pipeline used for variant calling of the WES data for the current study.

A very interesting observation is that in TCGA of the ten *HRAS* mutant cases, nine also had a *GTF2I* mutation (Table 5.3). *NRAS* was mutated overall in three tumours and was mutated once in a case that was also a *GTF2I* mutant. *KRAS* was mutated only in one sample and that hotspot mutation was in a type A thymoma that also had a *GTF2I* mutation. There were no *BRAF, GNAQ* nor *GNA11* mutations in this cohort. Overall in eleven of the 46 cases (24%) that had the *GTF2I* mutation there was also a mutation in *HRAS, NRAS or KRAS*. In the current study both AB thymomas had the *GTF2I* mutation, and one had a hotspot *NRAS* mutation.

Another interesting observation was that the NETT, in the NHLI cohort, with the highest mutation burden (TNRBF95414) and a pathogenic, hotspot *TP53* mutation (chromosome 17 c.527G>A) shared a very similar mutation profile with one of the SCCs in the TCGA cohort. The eleven genes mutated in both samples (NHLI NETT and TCGA SCC) were the following: *ARHGEF26, CDC27, DCN, DOCK4, EEF1A1, KCND2, LGALS9C, MSH2, PARP4, PKD1* and *TP53*. In the TCGA SCC there was also a likely driver mutation in *KMT2C* and in the NHLI NETT there was a high impact *KMT2D* mutation. The same NETT also had mutations in the following known cancer genes: A missense *ETV1* mutation with a CADD-phred score of 28.4, a stop-gain mutation in *JAK1* with a CADD-phred score of 41 and a missense *PTPRT* mutation with a CADD-phred score of 22.8 (Table 5.2). Although these are known cancer genes they haven't been seen in TETs before.

Another of the three NETTs (TNRBF73531) had a somatic missense BRCA2 mutation with a CADD-phred score of 24.9 and a somatic KRAS hotspot mutation (chromosome 12 c.25380276T>C) with a CADD-phred score of 23.2. BRCA2 has been found to be a putative driver in an AB thymoma from TCGA but has not been implicated in NETTs before. Petrini et al identified damaging missense mutations in *BRCA1* in two B3 thymomas by targeted capture**.**

No likely driver mutations were identified for the third NETT.

Of the six SCCs sequenced, putative driver mutations were identified in four of them.

One of the SCCs (TCRBF91244) had a missense KMT2C mutation (no CADD-phred score) (Table 5.2). This has been seen as a putative driver before in a thymic SCC in TCGA (TP53 mutant sample). The same individual had a missense CHEK2 mutation with a CADD-phred score of 21.3. Unfortunately there was no matched normal tissue for this case to establish that these were definitely somatic mutations. The same individual also had a BAP1 mutation that given the clinical history was very likely a germline mutation and this is discussed in chapter 5.4.3.

There was a high impact frameshift (stop-gained) MSH6 somatic mutation in one of the SCCs (TCRBF21432). Although this has been found to be a driver mutation in other tumour types it has so far not been found to be a driver in TETs in previous studies. An MSH2 mutation was also seen in one of the NETTs (TNRBF95414) already discussed.

However, some of the other TETs harboured mutations in several other known cancer genes that have never been described in TETs before. In one of the SCCs there was a missense AR mutation with a CADD-phred score of 33 and an inframe-deletion in ATRX. In another SCC (TCRBF15505) there was a missense mutation in SMARCB1 with a CADD-phred score of 33 and a missense CYLD mutation with a CADD-phred score of 29.4. In two of the SCCs no likely somatic driver mutations were identified although one of these (TCRBF16264) had a germline EGFR mutation, which is discussed in section 5.4.3.

One of the B3 thymomas (THRBF13799) had an inframe deletion in MED12 and the anaplastic B2 thymoma (TCRBF16327) had a missense MAP2K4 mutation with a CADD-phred score of 22.4. No putative somatic driver mutations were seen in the other B3 thymoma, two B2 and the one B1 thymoma in this cohort.

5.4.2.3 Discussion on somatic mutations in TETs

Results from the current study, as well as analysis of the WES data from TCGA, demonstrate some very interesting albeit infrequent mutations in TETs.

Petrini *et al*, sequenced 28 samples (eleven carcinomas, seven type A thymomas, three B2 and seven B3 thymomas. There were no AB or B1 thymomas in this cohort. Genes with deleterious mutations in more than three samples (11%) were the following: GTF2I (six cases), PRAMEF10 (five), TAS2R30 (four), ANKRD36C (four), TP53 (three) and EP300 (three). TP53 and EP300 are both known 'cancer genes' according to COSMIC (http://cancer.sanger.ac.uk/census) (221). Mutations were detected in another 34 known cancer genes: CDKN2A was mutated in two samples and the remaining 33 in one sample each. These were the following: ATR, ATRX, BRD4, BRAF, CHD4, CYLD, EML4, FGFR2, FOXO3, GNA11, GRIN2A, HOXA13, HRAS, IDH1, KMT2C, KMT2D, KRAS, MAML2, MLLT4, MSH6, NFIB, PAX7, PBRM1, PHF6, RBM10, RET, ROS1, RSPO3, SLC34A2, TET2, TRRAP, USP6 and ZNF521. Putative driver mutations were only seen in the carcinomas and there were none in the A, B2 or B3 thymomas, including *GTF2I* that was not thought to be a driver in the samples in which it was mutated (106). Known cancer genes recurrently mutated in this cohort can be found in (Table 5.3).

Sequencing of eighteen TETs (ten carcinomas and eight thymomas) by Zehir *et al* with a hybridization capture–based NGS panel (MSK-IMPACT) showed similar results with only one of eight thymomas having only one mutation detected in *CREBBP.* Of the ten carcinomas, nine had mutations detected. *KMT2D* was mutated in three samples (30%), followed by *TP53* and *CDKN2A* that were mutated in two samples each (20%). One carcinoma had a *KIT* W557S hotspot mutation, another had an *NRAS* Q61K hotspot mutation and a *PIK3CA* E542K hotpot mutation, another had an *EZH2* mutation only and one had a *CDK12* mutation only (105).

Although with the exception of the TCGA cohort, other studies, including the current, have included a small number of samples, it is remarkable that apart from very few genes found to be mutated across all cohorts there is little overlap in terms of the mutated genes and the frequency with which they are mutated.

The *GTF2I* mutation is present at a high frequency in more indolent thymomas (Type A and AB) and has been found to be an independent prognostic factor in TETs of all subtypes (106,212). Petrini et al showed that it has a deleterious effect on the protein structure of TFII-I and/or its function by affecting the second conserved TFII-I repeat protein domain, which is near its DNA binding site (106). TFII-I is involved in a number of biological processes relating to transcription and growth factor signaling. In vitro the *GTF2I* mutation was found to increase TFII-I expression

post-transcriptionally, which in turn likely accelerates cell proliferation through up-regulation of cell-cycle control proteins such as FOS. However cell transformation was not seen raising the possibility that *GTF2I* acts as a cell growth promotor only and has no oncogenic function (106). Hence the exact and full impact of this mutation is not fully understood.

Although Petrini *et al* found one mutation in each of *KRAS, HRAS* and *NRAS*, none of those also had the *GTF2I* mutation (106) An approximately 25% co-mutation rate between the *RAS/RAF/MEK* pathway and the *GTF2I* mutation have been seen in TCGA and our cohort. This observation means that it is likely that both mutations are required for tumourigenesis but also probably indicates that *GTF2I* mutations do not have an affect on proliferation since the *KRAS*, *HRAS* or *NRAS* mutations are responsible for that aspect of tumourigenesis. Alternatively, mutations in both *GTF2I* and RAS family members work synergistically to upregulate proliferation. *HRAS* mutations in TETs were also commonly seen in a targeted-re-sequencing study, which however did not include GTF2I so that co-mutation status could not be assessed (231).

The RAS/RAF/MEK pathway is one of the most-studied signalling pathways due to its significance in human cancer. Numerous cell surface molecules result in rat sarcoma (*RAS*) activation, which includes Kirsten rat sarcoma viral oncogene homolog (KRAS), neuroblastoma RAS viral oncogene homolog (*NRAS*), and Harvey rat sarcoma viral oncogene homolog (*HRAS*), which are GTPases responsible for turning on the RAF family of kinases, which include BRAF, CRAF, and ARAF. Downstream of RAF are the MAPK/ERK kinases, MEK1, and MEK2, which have a single main substrate, which is ERK. Downstream of ERK are multiple substrates whose activation affects cell survival, cell migration, cell invasion, cell cycle progression and differentiation, protein synthesis, metabolism, and senescence but also pro-apoptotic molecules (232). As RAS and RAF are important oncogenes, targeting this pathway has been the focus of extensive research. RAS mutations are overall the commonest oncogenic mutations in human cancer accounting for approximately 30% of all human cancers (233). RAS mutations are most commonly found in pancreatic cancer (90%), followed by colorectal cancer (40%), NSCLC (30%), bladder cancer (30%), cholangiocarcinoma (25%), and melanoma (15%) according to the COSMIC database (http://cancer.sanger.ac.uk/cosmic). So far attempts to target RAS directly through the use of small molecule inhibitors has been unsuccessful given the affinity of RAS for GTP, which exists intracellularly at very high concentrations (234). MEK inhibitors have been shown to be effective in the treatment of RAS-mutated cancers such as lung cancer and melanoma (235,236).

TP53 was only one of the two known cancer genes (the other being *NRAS*) to be mutated in the NHLI cohort, in TCGA, Petrini et al (106) and Zehir et al (105). In fact it has been found to be mutated in a total of thirteen TETs (out of 213 TETs from all four cohorts), all being hotspot, putative driver mutations with eleven of those having been found in SCCs, making this the most

commonly mutated gene in thymic SCCs and second most commonly mutated known cancer gene in all TETs after *HRAS* and third overall when including *GTF2I*. Another study that performed targeted sequencing of TETs also found TP53 to be the most commonly mutated gene in TETs (17% overall and 26% in thymic carcinomas) (231).

TP53 is a tumour suppressor gene that encodes for a transcription factor that in response to various cellular stresses regulates gene expression by exerting numerous anti-proliferative functions (237). It is very frequently mutated in human cancers with more than 75% of mutations being missense substitutions leading to loss of the tumour suppressor function of any remaining wild-type *TP53*. However it can also acquire oncogenic functions (238). Somatic mutations in *TP53* vary in frequency depending on the tumour type. Germline mutations in *TP53* are linked with the development of hereditary cancers such as Li-Fraumeni and Li-Fraumeni-like syndrome (237). Overall over 50% of human cancers have *TP53* mutations highlighting its significant tumour-suppressing role. Treatments aiming to restore normal TP53 function or change the ratio of mutant to wild-type TP53 have had minor success and there are currently no licensed treatments for this indication (239)

Lysine (K)-specific methyltransferase 2D (*KMT2D;* also known as MLL2) that was also mutated in this NETT is a histone methyltransferase key in regulating gene transcription. As is the case with other chromatin remodeling and histone modulators it acts as a tumour suppressor gene whose loss of function can lead to different cancers, including medulloblastoma, lymphoma and gastric cancer amongst others (240). It is very intriguing that Zehir et al (105) found mutations in this gene in three of ten thymic SCCs investigated (30%) making this the most frequently mutated gene in thymic SCCs in that cohort. It also established that the frequency of *KMT2D* mutations is 7.9% (867 / 10945 cases) across all tumour types and was most commonly mutated in penile cancer, non-Hodgkin lymphoma, bladder, vaginal and small cell lung cancer (105). There were no putative driver mutations in *KMT2D* in TETs in TCGA or Petrini et al (106). *KMT2D* has also been recently found to be very rarely mutated in pancreatic neuroendocrine tumours (1/98, 1%) (241).

Somatic *BRCA2* mutations have not been described in thymic neuroendocrine tumours before and following a thorough search on cBioportal putative driver somatic mutations in *BRCA2* have only been reported in two cases of small cell lung cancer (http://www.cbioportal.org/index.do). Recently a germline BRCA2 mutation was described in a case of pancreatic neuroendocrine cancer (241). BRCA2 was identified in chromosome 13q12.36 in 1995. BRCA1 and BRCA2 are tumour suppressor genes key for DNA repair through homologous recombination. Cells lacking wild-type BRCA2, usually by mutation and then loss of the wild-type allele, which can be identified through loss of heterozygosity studies, are more sensitive to ionizing radiation, which is a sign of defective double stranded base repair (242). Germline mutations in BRCA1 and BRCA2 are known to confer

an increased risk of developing breast and ovarian cancer. For women with germline BRCA1 mutations the average cumulative rick of developing breast cancer by the age of 70 is 65% (95% CI 44%–78%) and of developing ovarian cancer 39% (CI 18%–54%). For BRCA2 germline mutants this is 45% (CI 31%–56%) for breast cancer and 11% (CI 2.4%–19%) for ovarian cancer (243). Damaging BRCA2 mutations also increase the risk of ocular melanoma, prostate, pancreatic, and colon cancers as well as breast cancer in men (244). Poly(ADP-ribose) polymerase (PARP) inhibitors, such as Olaparib (maintenance) and Rucaparib (third-line monotherapy) have been licensed by the US Food and Drug Administration (FDA) and European Medicines Agency (EMA) for the treatment of BRCA1 and BRCA2 mutated ovarian cancer. Indications include both germline and somatic mutations in either gene (245). PARPs are activated very early following DNA damage to repair DNA using the base excision repair and single strand repair pathways. The rationale behind utilizing these agents against tumours with BRCA1 or 2 mutations lies in that these tumours already have an inadequate activation of homologous repair pathways as a result of these mutations, hence inactivating PARP leads to their death, a phenomenon described as “synthetic lethality” (246).

One of the limitations of the current study was that the number of samples subjected to WES was only seventeen, in what is a tumour type that consists of many subtypes. Therefore reliably establishing mutation spectra is very challenging. What makes this particularly difficult is also the very low mutation burden of TETs. In the NHLI cohort the median number of somatic mutations was very low at 67 (average was 67.8; range eight – 162). These were mostly missense mutations. The biggest published study on TETs so far by Petrini et al identified a total of 722 SNVs and 68 indels in the coding regions of a total of 28 TETs that underwent WES an overall very low mutation burden (106). This is in keeping with results presented from TCGA although not yet published demonstrating that TETs have the lowest mutation burden amongst all adult solid tumours (107). All groups have also failed to reveal any genes that are recurrently mutated across all subtypes with *GTF2I* being a notable exception of unknown significance (106,107). In fact of the likely somatic driver mutations in NHLI TET cohort there is no other gene mutated in more than sample (Table 5.2). The most commonly mutated genes across all cohorts, as already discussed are HRAS and NRAS, which are oncogenes and key components of the *RAS/RAF/MEK* pathway.

Of the 66 most commonly mutated genes in TETs (Table 5.3) the only gene related to DNA repair is *TP53*. Of the mismatch excision repair genes, *MSH2* was found to have one missense mutation in one NETT the NHLI cohort and has a frameshift deletion in a thymic SCC from TCGA that also harboured a hotspot missense *TRP53* mutation and an *MLH1* nonsense mutation. There is also a single high impact frame=shift/ stop-gained MSH6 somatic mutation in one of the NHLI SCCs. No BRCA1 mutations were seen but two missense mutations in BRCA2 were seen. One was a NETT from the NHLI cohort and the other was in an AB thymoma from TCGA that also harboured a hotspot SF3B1 mutation. CHEK2 had a nonsense mutation in a thymic SCC from TCGA and a

missense mutation in an NHLI SCC. FANCM was mutated in only one NHLI SCC. Although the numbers are small it is worth pointing out that the majority of these genes are mutated in the more aggressive NETTs and SCCs and not in thymomas.

Table 5.2. Likely somatic driver mutations in NHLI TET cohort (n=17)

	TCRBF03870	TCRBF15505	TCRBF15586	TCRBF16264	TCRBF21432	TCRBF91244	TNRBF73531	TNRBF92784	TNRBF95414	THRBF13799	THRBF70621	TCRBF16327	THRBF06854	THRBF25074	THRBF20178	THRBF24855	THRBF54962
Subtype	SCC	SCC	SCC	SCC	SCC	SCC	NET	NET	NET	B3	B3	B2*	B2	B2	B1	AB	AB
GTF2I																Ss	Ss
AR			33														
ATRX			S														
BAP1						G											
BRCA2							25										
CHEK2						21											
CYLD		29															
ETV1									29								
JAK1									41								
KMT2C						S											
KMT2D									23								
KRAS							23										
MAP2K4												22					
MED12										S							
MSH2									23								
MSH6					S												
NRAS																30	
PTPRT									23								
SMARCB1		32															
TP53									S								

B2*: anaplastic B2 thymoma

Ss: Sanger Sequencing

Numbers indicate the CADD_phred(scaled) score for likely driver mutations

S indicates there was no CADD_phred(scaled) score

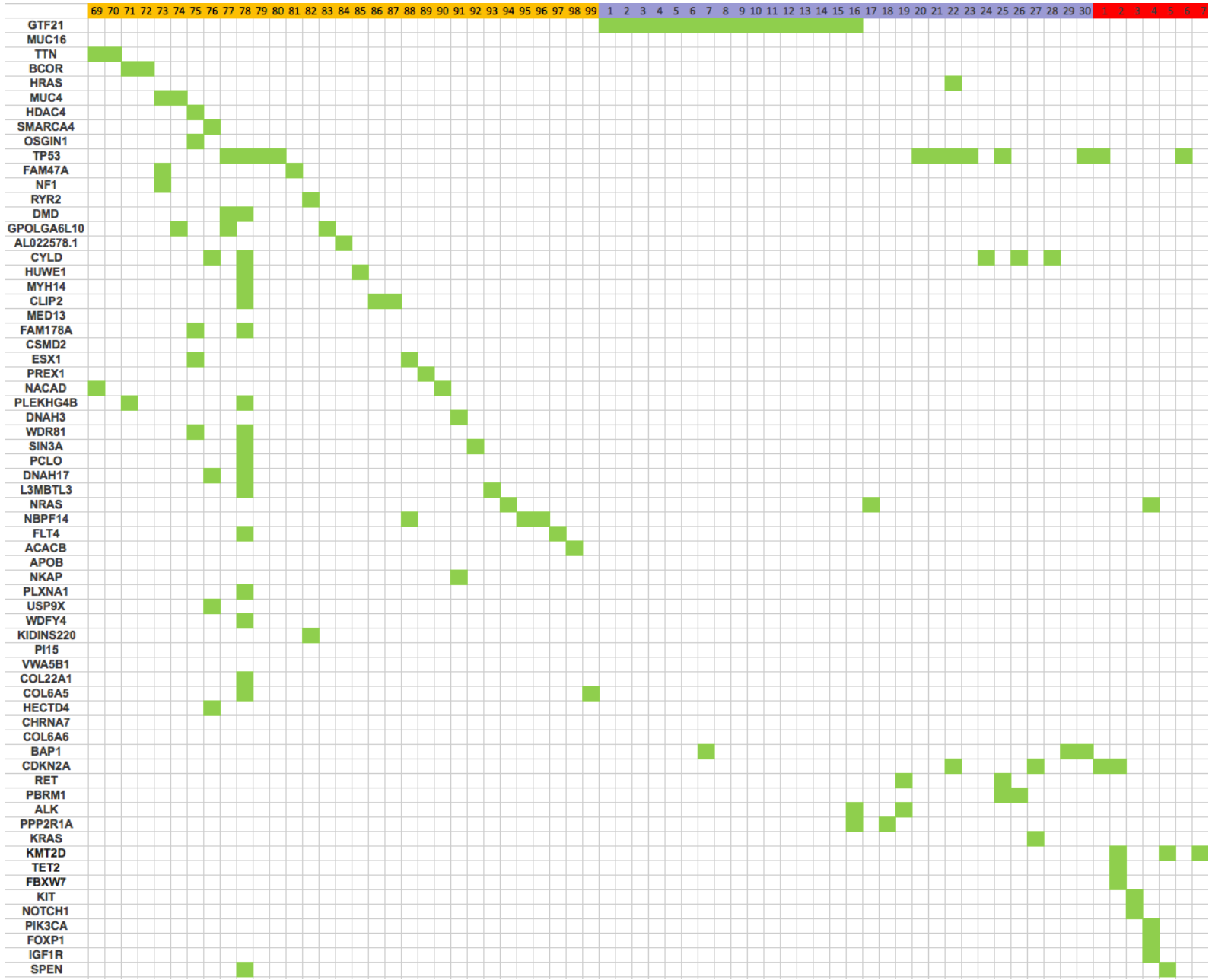
Genes (and associated scores) in red font indicate these genes have been found by other groups to have putative driver mutations in TETs. The remaining have been putative drivers in other tumour types.

The two samples in light grey background did not have matched normal tumour

Table 5.3. Overview of somatic mutations in TETs from TCGA, Petrini *et al.* and Zehir *et al.*



Table 5.3. Overview of somatic mutations in TETs from TCGA, Petrini *et al.* and Zehir *et al.*



This table has been put together using published data from Petrini et al (106), unpublished TCGA that have been accessed via the cBioPortal (http://www.cbioportal.org) and The National Cancer Institute GDC Data Portal (https://portal.gdc.cancer.gov) and by unpublished data, which are based on the publication by Zehir et al (105) that have also been accessed via the cBioPortal. In TCGA 124 TETs were whole exome sequenced and mutations were seen in 99 of the tumours. The top genes in terms of frequency of mutations are included in this table. Petrini et la performed WES of 28 TETs and targeted sequencing of 52 TETs (26 of which had also undergone WES) with a custom 197–cancer gene panel. The top twelve known mutated cancer genes are included in this table in the 30 samples in which these were seen. Zehir et al used targeted sequencing of ten thymic SCCs, and identified mutations in nine of these. Mutations that were thought to be putative drivers have been included in the table only hence the number of samples included from this study is seven. KIT and SPEN were also found to have putative driver mutations each in one sample in the TCGA cohort and were added to the table for completion, as initially these were not part of the top 50 genes for that cohort.

5.4.3 Germline mutations

5.4.3.1 Candidate germline driver mutations

There are currently no published sequencing data on germline mutations in TETs. Of the fifteen samples in this study with matched normal tumour a screen for germline mutations was carried out. All variants that were not predicted to have a moderate nor high impact or were present in EXAC with a frequency of more than 10^{-3} were filtered out. Of these, known cancer genes as per COSMIC (http://cancer.sanger.ac.uk/census) (221) were prioritised as were variants that were recurrently mutated. Based on the somatic mutation profile of TETs, selected genes were specifically screened for germline mutations on the other allele. These were the following: *ATM, ATRX, BAP1, BRCA1, BRCA2, CHEK2, TSC1, CDKN2A, RET, MEN1, MEN2, OMIM, RAD51L, BMP2, ALK, FLT3, MSH2, SETD2, TET1, EGFR, TET2, NOTCH2, ROS1, HLA-A, TP53, KIT* and *MET*. Table 5.4 includes all candidate predisposing germline variants identified in this screen as well as likely somatic driver mutations from Table 5.2. Five of the six SCCs (one unmatched) had germline mutations in one of these genes (83%), four of the six B subtype thymomas (one unmatched) (67%), one of the two AB thymomas (50%) and only one of three NETTs (33%).

A recurrent germline, high impact loss of function mutation (chromosome 6 c29912395insGG) (p.Gly341fs) was seen in *HLA-A* in four samples (one of each of an SCC, a B3, a B2 and a B1 thymoma) (Table 5.4). These were all heterozygous mutations. This is a known polymorphism albeit with a 0.8% frequency in ExAC but with a significantly higher frequency in this cohort at 27% (x^2 *p value* < *0.001*).

The individual with the unmatched SCC (TCRBF91244) had a deleterious mutation in *BAP1* (chromosome 3 c.52437204delC) (Table 5.4). Although there was no matched normal tissue to confirm this as a germline mutation, the clinical history for this individual is highly suggestive of BAP1 predisposition syndrome. This was a male patient that developed a thymic SCC at the age of 51, who had also previously been treated for a renal cell carcinoma, and also had a metachronous mesothelioma and a melanoma at the ages of 50, 58 and 59 respectively. Germline mutations in BAP1 have been shown to predispose to all of these tumour types (247-252).

Missense germline mutations were seen in isolated cases in a number of genes. A missense germline mutation in *GNAQ* (chromosome 9 c.80336403G>C) (p.R104G) with a CADD_phred score of 20.2 was seen in one of the two AB thymomas (THRBF54962), which also has a somatic *GTF2I* mutation (Table 5.4). This increases the number of TET cases with a *GTF2I* mutation and an additional mutation in a member of the MAP kinase pathway. In our cohort, both AB thymomas

had the *GTF2I* mutation. One of those (THRBF54962) had the germline *GNAQ* mutation described here and the other had a somatic NRAS mutation as discussed in section *5.4.2.2.5*.

One of the two B3 thymomas (THRBF13799) had a missense mutation in *ASXL1* (chromosome 20; c.31025008C>T) (p.Thr1498Met) with a CADD-phred score of 24.8 (Table 5.4). This mutation, which is in the DNA-binding domain (PHD domain) of the gene, has not been seen before according to cBioPortal (http://www.cbioportal.org/index.do) (Figure 5.3a). This was one of the cases that also had an *HLA-A* germline mutation. The other B3 thymoma (THRBF70621) in this cohort had no likely somatic nor germline mutation. ASXL1 somatic mutations have been seen before in WES by TCGA (107) and with targeted capture in another study on TETs (231).

Of the three B2 thymomas, one (THRBF06854) had the HLA-A germline mutation, another (THRBF25074) had a germline missense mutation in *ERBB4* (chromosome 2 c.212248444G>A) (p.R1259W) with a CADD_phred score of 33 and the third, the anaplastic B2 (TCRBF16327), had none (Table 5.4). Somatic, putative driver ERBB4 mutations have been seen in a minority of non-melanoma skin cancers, melanoma, oesophageal, bladder and large and small bowel carcinomas amongst others but not in TETs (105). These mutations have been mostly missense and distributed across the entire gene. A query of cBioPortal (http://www.cbioportal.org/index.do) across all tumour types generated a map of known mutations, which can be seen in Figure 5.3a. The specific mutation identified in the current study (registered as R1275W in cBioPortal) has been found once before in a highly mutated bladder carcinoma (105), where it was one of many putative driver mutations, which also included TSC1, FDFR3, ATM, CDKN2A, KMT2A and KMT2D amongst others. This can be seen in Figure 5.3b.

One of the NETTs (TNRBF73531) had a germline missense mutation with a high CADD phred score of 23.6 in *CHEK2* (chromosome 22 c.29095881C>A) (p.Arg361Leu). It is worth pointing out that another SCC (TCRBF91244; the unmatched case) also had a somatic missense *CHEK2* mutation in the same region with a CADD_phred score of 21.3. CHEK2 deletions were found in four cases in the TCGA cohort and in the two samples where it was mutated it was deemed to be a putative passenger. This mutation, which lies in the gene's protein kinase domain, has not been described before in any tumour type according to a search on cBioPortal (http://www.cbioportal.org/index.do). Putative driver mutations of CHEK2 in TETs have not been described before (Figure 5.2b).

A missense germline *EGFR* mutation (chromosome 7 c.55272966G>C) (p.V1044L) was seen in one of the SCCs (TCRBF16264) with a CADD_phred score of 10.06. Somatic mutations have not been reported in that position in cBioPortal (http://www.cbioportal.org/index.do) so far (Figure 5.2c).

Another SCC (TCRBF21432) had a germline missense mutation in *TSC1* (p.Arg190Cys), which had a high CADD-phred score of 29.1. The matched normal tissue was heterozygous for the mutation whereas the tumour tissue was homozygous or hemizygous for the mutation. This SCC also had the germline *HLA-A* mutation.

One of the SCCs (TCRBF15586) had a germline splice acceptor *FANCA* mutation.

5.4.3.2 Discussion on germline mutations in TETs

A number of interesting germline variants were seen.

A high impact insertion (chromosome 6 c29912395insGG) (p.Gly341fs) seen in *HLA-A* was seen in four different TET subtypes (an SCC, a B3, a B2 and a B1 thymoma) making this the most frequently mutated gene in this cohort. Putative driver somatic mutations in HLA-A have been seen in a variety of solid tumours, including NSCLC, as well as lymphomas but not in TETs (105). In humans major histocompatibility complex (MHC) class I alleles are encoded by three genes (HLA-A, HLA-B and HLA-C) mapping to chromosome 6. Identifying mutations in HLA genes through WES is particularly difficult given the high polymorphism rate of the HLA loci, which render alignment of sequencing reads to the reference genome very challenging (253). A study of HLA mutations in 7,930 samples in 20 different tumor types revealed that somatic mutations in class I HLA genes are commonly seen in colon adenocarcinomas, head and neck tumours, and SCCs of the lung and stomach cancer, whereas they are very rare in other tumour types such as glioblastoma, ovarian cancer and CLL (253). So far somatic class I HLA mutations have been shown to be loss-of-function mutations, which are preferably found at the N-terminal end and are likely to render the protein completely non-functional (253). Hence, loss of function of HLA-A appears to be a predisposing factor in TETs and is a finding, which hasn't been described before in this tumour type.

One of the most salient cases in this cohort was an individual with an SCC with a germline BAP1 mutation and a history highly suggestive of BAP1 predisposition syndrome. The *BAP1* gene is located at chromosome 3p21.124. It encodes for a deubiquitinating enzyme that helps regulate key cellular pathways and by associating with different transcription factors may result in suppression or promotion of proliferation in a cell-type specific manner (254). One of its functions is proposed to be to bind to BRCA1 via its RING finger domain to act as a tumour suppressor. It is also involved in the regulation of transcription and cell cycle, DNA damage response and cell death (254). *BAP1* tumour predisposition syndrome (*BAP1*-TPDS) is inherited in autosomal dominant fashion and is

linked with an increased risk of developing atypical Spitz tumours and a number of solid tumours such as uveal melanoma (commonest), mesothelioma, cutaneous melanoma, clear cell renal cell carcinoma, and basal cell carcinoma amongst others. TETs have so far not been included in the list of malignancies involved making this the first likely such case (255). Carriers of deleterious germline BAP1 mutations can develop many of these primary tumours at a median age of onset that is younger than the rest of the population, who develop these malignancies and that lack such a germline variant. This was the case with the affected individual **(**TCRBF91244) described here. It is interesting that somatic *BAP1* mutations are also found in the cancer types that are part of the *BAP1*-TPDS. This points towards this pathway being of particular importance for these tumours. Somatic BAP1 mutations have also been described in TETs albeit in small frequency, which is consistent with our finding (220,231).

In *Drosophila* the Additional Sex Combs (*ASX)* gene encodes a protein that belongs to the enhancers of the Trithorax and Polycomb (ETP) group. It is in a complex with the *Drosophila* homolog of human BAP1, called *Calypso*. This is an ubiquitin carboxy-terminal hydrolase that deubiquitylates H2AK119ub1. Additional Sex Combs Like 1 (*ASXL1*) is one of the mammalian homologs of *ASX* and has been found to be mutated in a numbers of human cancers, which include hematological cancers as well as breast and prostate cancer. Activation of NRAS or loss of TET together with loss of ASXL1 have been associated with increased severity in hematological cancers (256). Human ASXL1 interacts with BAP1 to form a polycomb repressive deubiquitinase complex, which regulates histone H2A ubiquitination (256) in a similar manner to its activity in *D. Mel*. Loss of function mutations in either gene leads to loss of their tumour suppressor activity. ASXL1 is often mutated together with other genes involved in myeloid diseases, including those encoding proteins that are involved in gene signaling (CBL, JAK2, NF1, RAS) and mRNA splicing (SF3B1, SRSF2, U2AF35) (257). Somatic ASXL1 and BAP1 mutations are mutually exclusive in TETs and in thymic carcinomas are reported to account for 17% of mutations of all samples (231). Somatic, putative driver mutations were also seen in one thymic SCC and one B3 thymoma in TCGA tumours (107). Germline ASXL1 mutations have been found in patients with Bohring-Opitz syndrome (258), although the individual in the current study had no history of this condition.

The epidermal growth factor receptor (EGFR or ErbB1 or HER1) is a member of the family of receptor tyrosine kinases that are responsible for activating a large range of signalling pathways that lead to cell growth, proliferation and survival, such as the MAP kinase and the PI3K-AKT-mTOR pathways (259). Ten to seventeen per cent of NSCLC cases in North America and Europe and 30–50% in Asia harbour EFGFR mutations (260) with the vast majority of those being in exons 18–21 (90% of these are exon 19 deletions or exon 21 L858R point mutations) which encode a portion of the gene's kinase domain. They are usually heterozygous mutations, with amplifications of the mutant allele also being common. These mutations lead to aberrant EGFR kinase activity

and subsequent hyperactivation of downstream pathways (259). Tissue kinase inhibitors (TKIs) of EGFR have become the mainstay of treatment for EGFR mutant patients. These include gefitinib (Iressa®, AstraZeneca, London, UK), erlotinib (Tarceva®, F. Hoffmann-La Roche, Basel, Switzerland), and afatinib (Giotrif®, Boehringer Ingelheim, Ingelheim, Germany). They have all shown improvements in response rate, progression-free survival and quality of life compared to standard chemotherapy in patients with these mutations although overall survival advantages have been more difficult to establish given treatment crossover in the majority of trials (260). As resistance to the above agents quickly emerge, with one of the commonest (approximately 50% of cases) being the development of the T790M EGFR mutation, a number of drugs alone or in combination have been used to overcome resistance. Of those osimertinib (AZD9291), a third-generation TKI has been the first to receive FDA and EMA approval for metastatic *EGFR* T790 mutant NSCLC, following progression on *EGFR* TKI therapy (261). The following germline EGFR variants have also been described in lung cancer: the T790M and R776X in exon 20, and the V843I and P848L in exon 21 (262). Somatic T790M mutations have been established as a common resistance mechanism to EGFR TKis. Germline EGFR T790M mutations considerably rarer (0.54%) but have recently been increasingly recognized in families with early-onset, hereditary NSCLC (263,264). To date germline EGFR mutations in TETs have not been described.

A study by Wang et al, which employed targeted re-sequencing of 197 cancer-associated genes in 78 paired TETs (31 thymomas and 47 carcinomas) identified recurrent mutations in the following genes: ASXL1, BAP1, BRCA2, CDKN2A, CYLD, DCC, DNMT3A, HRAS, KIT, SDHA, SETD2, SMARCA4, TET2, TP53 and WT1 (231). Combining this with the list of somatic and germline mutations in our NHLI cohort (Table 5.3) there was some overlap and five of these fifteen genes (33%) were mutated in both cohorts: ASXL1, BAP1, BRCA2, CYLD and TP53.

Table 5.4. Likely somatic and germline driver mutations in the NHLI TET cohort (n=17)

	TCRBF03870	TCRBF15505	TCRBF15586	TCRBF16264	TCRBF21432	TCRBF91244	TNRBF73531	TNRBF92784	TNRBF95414	THRBF13799	THRBF70621	TCRBF16327	THRBF06854	THRBF25074	THRBF20178	THRBF24855	THRBF54962
Subtype	SCC	SCC	SCC	SCC	SCC	SCC	NET	NET	NET	B3	B3	B2*	B2	B2	B1	AB	AB
GTF2I																Ss	Ss
AR			33														
ASXL1										25							
ATRX			S														
BAP1						G											
BRCA2							25										
CHEK2						21	24										
CYLD		29															
EGFR				10													
ERBB4														33			
ETV1									28								
FANCA			G														
GNAQ																	20
HLA-A					G					G			G		G		
JAK1									41								
KMT2C						S											
KMT2D									23								
KRAS							23										
MAP2K4												22					
MED12										S							
MSH2									23								
MSH6					S												
NRAS																30	
PTPRT									23								
SMARCB1		32															
TP53									S								
TSC1					29												

B2*: anaplastic B2 thymoma

Ss: Sanger Sequencing

Numbers indicate the CADD_phred(scaled) score for the mutation in that sample.

S indicates a somatic mutation with no CADD_phred(scaled) score

G indicates a germline mutations with no CADD_phred(scaled) score

Genes (and associated scores) in red font indicate these genes have been found to have putative driver mutations in TETs. The remaining have been putative drivers in other tumour types.

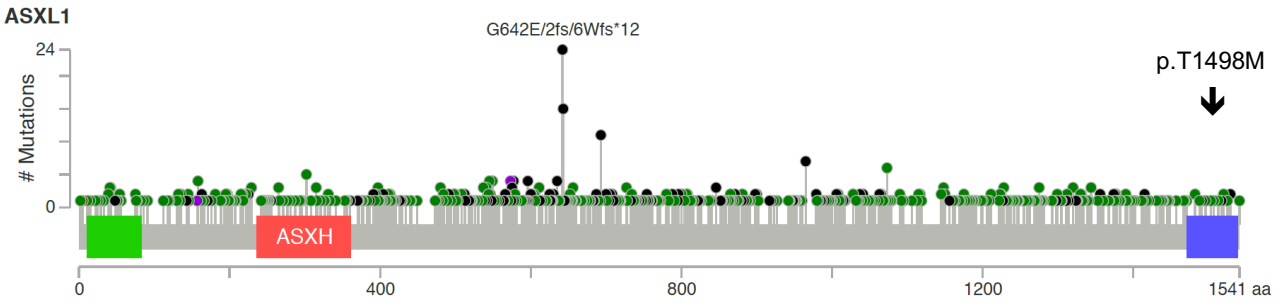
The two samples in light grey background did not have matched normal tumour

A light green background indicates a somatic mutation

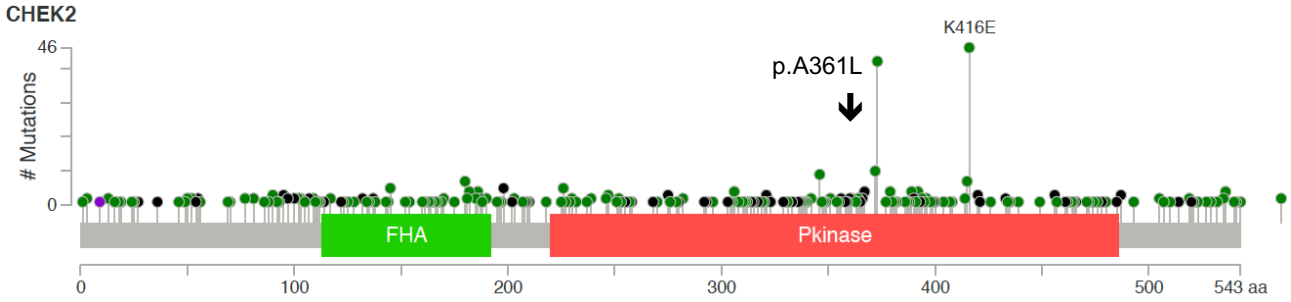
A light blue background indicates a germline mutation

Figure 5.2. Somatic CHEK2 and EGFR mutations seen across all tumour types

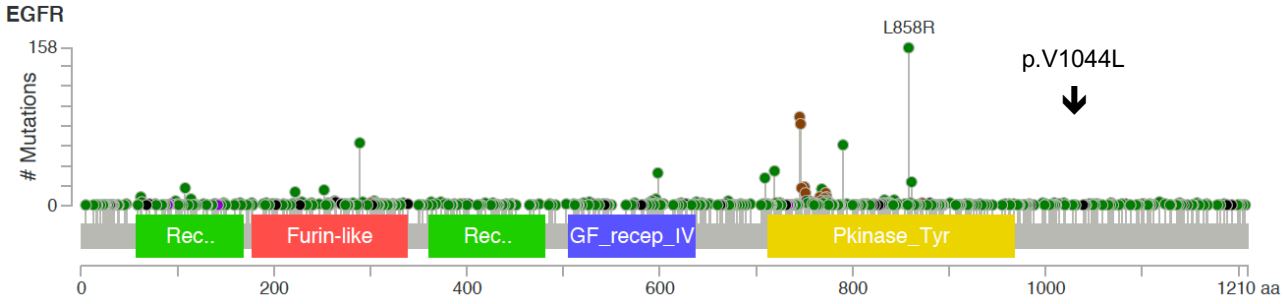
5.2a. *ASXL1* mutations



5.2b. *CHEK2* mutations



5.2c. *EGFR* mutations

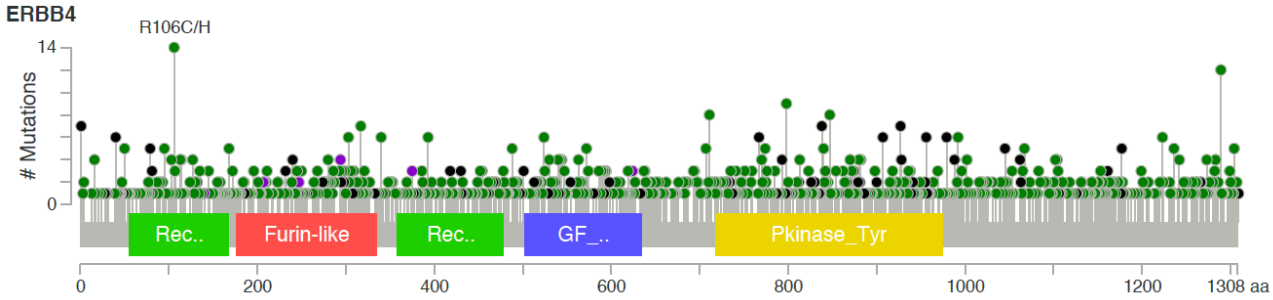


Somatic *ASXL1*, *CHEK2* and *EGFR* mutations as visualised on cBioPortal (http://www.cbioportal.org/index.do) are shown. Figures 5.2a, 5.2b and 5.32c show all the *ASXL1*, *CHEK2* and *EGFR* mutations respectively that have so far been identified in all tumour types from all studies that have data in cBioPortal. The position of the germline mutation from the current study is shown with a black arrow. Mutation types are denoted with the following colours: **Missense Mutations; Truncating Mutations**: Nonsense, Nonstop, Frameshift deletion, Frameshift insertion, Splice site; **Inframe Mutations**: Inframe deletion, Inframe insertion; **Other Mutations**: All other types of mutations

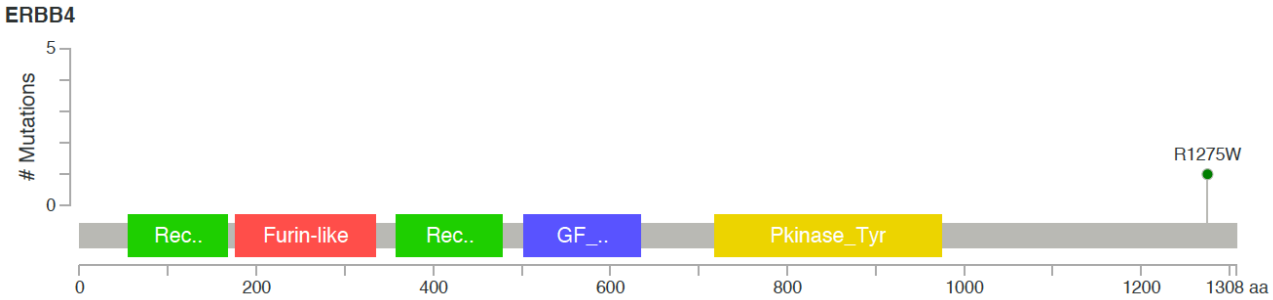
The Y axis shows the number of mutations at each position

Figure 5.3. Somatic *ERBB4* mutations seen across all tumour types

5.3a. All mutations



5.3b. R1275W mutation



Somatic *ERRB4* mutations as visualised on cBioPortal (http://www.cbioportal.org/index.do).

Figure 5.3a shows all mutations and figure 5.3b shows the mutation identified in the current study (as germline) that has been previously been identified as somatic, in a single case of a bladder carcinoma.

Mutation types are denoted with the following colours: **Missense Mutations; Truncating Mutations**: Nonsense, Nonstop, Frameshift deletion, Frameshift insertion, Splice site; **Inframe Mutations**: Inframe deletion, Inframe insertion; **Other Mutations**: All other types of mutations; The Y axis shows the number of mutations at each position

5.5 Conclusions

Although the suitability of FFPE material for WES has been questioned (see chapter 3.4.1) a satisfactory read-depth of sequencing across the exome of the samples was achieved in the current study. It also confirmed a very low somatic mutation burden in TETs. In the A and AB thymomas, apart from the frequent *GTF2I* mutation, approximately a quarter of the tumours appeared to have oncogenic mutations affecting the MAP kinase pathway, with *HRAS* being most commonly mutated gene. Somatic *NRAS* and *KRAS* mutations have also been seen in *GTF2I* mutant cases and in the NHLI cohort a germline *GNAQ* mutation was identified in one AB TET. In thymic carcinomas, including SCCs and NETTs, the most commonly mutated gene was TP53, which was mutated in 14 of 46 cases (30%). This figure includes the published data from Petrini et al (106). Otherwise very few genes appear to be recurrently mutated in any TET subtype. In the current and previous studies B subtypes had very few putative somatic driver mutations (105-107). Results from TCGA which is the largest cohort screened so far, are particularly notable with putative driver mutations identified in only five of 61 (8%) B1, B2 and B3 thymomas (220). In the study by Zehir et al, of the eight thymomas sequenced with a targeted gene panel, mutations were seen in only one case and were all were predicted to be passenger mutations (105). It is worth considering a significant contribution of germline mutations for these tumours. In the current cohort were only six B subtype thymomas and candidate somatic mutations were only seen in one gene in each of two of those (MAP2K4 and MED12). Neither of these has been described in TETs before (Table 5.2). Germline mutations of note in ASXL, ARBB4 and HLA-A (x3), were seen in four of the six tumours (Table 5.4).

In the entire TET cohort in the current study, in terms of germline mutations the current study identified a number of interesting candidate driver mutations. A recurrent insertion in HLA-A was seen across the TET subtypes at a frequency of 27%. This was much higher than its 0.8% frequency reported in the ExAC database. In regards to the different populations included in ExAC, Africans have the highest frequency of this mutation in the 'normal population with a frequency of 1% and East Asians have the lowest with a frequency of 0.2%. Europeans (Non-Finnish) that are overrepresented in our cohort, have a reported frequency of 0.9%. HLA-A has been found to have putative somatic driver mutations in other tumour types (253) although germline loss of function of this gene in TETs has not been described before and warrants validation and further investigation**.** A high number of loss-of-function events in HLA genes, including frameshift indels, splice-site and nonsense mutations have been identified in a subset of tumours in other studies and it is hypothesized that these can decrease HLA class I expression on tumour cells, which in turn would decrease the host immune response to tumours (253). From this it could also be hypothesized that PD-L1 expression would decrease in such tumours, which upon confirmation of the HLA-A mutation would also warrant investigation.

Other salient, isolated cases harboured putative driver, missense, germline mutations in *EGFR* (p.V1044L), *ERBB4* (p.R1259W), *CHEK2* (p.Arg361Leu) and *GNAQ* (p.R104G). A single case of a thymic SCC with a high impact deletion in *BAP1* (chromosome 3 c.52437204delC) was also seen. Despite the absence of matched normal tissue this was likely to be a germline mutation in an individual whose clinical history is highly suggestive of *BAP1*-TPDS, making this the first reported case of thymic SCC presenting as part of this syndrome.

As has been the case with more extensively researched commoner tumour types (265), the few recent genetic screens of TETs have focused on exome sequencing (106). The remaining regions of the genome, including intronic and intergenic regions have not been studied. Given the apparent lack of driver mutation in the exome from both WES and targeted capture studies it is possible that key answers regarding the molecular pathogenesis of this tumour type lie within these areas. One other possibility is that gene fusions are pathogenic. Studying gene fusions and possible driver mutations within non-coding regions, such as in the TERT gene promoter via targeted re-sequencing or genome sequencing, should therefore be the next step.

Chapter 6
Insights into TETs through SNP genotyping

Abstract

SNP genotyping was performed on DNA extracted from frozen and FFPE tissue from 143 TETs of all subtypes and 96 matched normal tissues. ASCAT gave copy number gain results for 55 cases and copy number loss results for 78. We have shown an increasing number of CNAs from indolent to more aggressive TETs. Across all TET subtypes the current study has confirmed previous findings showing multiple genes with copy number gains in chromosome 1q. The single gene with the most recurrent copy number gains across all TETs in our study was *MKI67* (chromosome 10q26.2) (30.91%). This finding warrants further follow-up. Very few known cancer genes were shown to have copy number gains, the most frequent being *PDE4DIP* mapping to chromosome 1q in13% of all TETs and in 38% of SCCs. *MYC* was amplified in 25% of NETTs. Very few genes were found to have recurrent copy number losses across all TETs. CDKN2A has been found to have copy number losses in the current as well as previous studies with varying frequency. An intriguing finding from our study is recurrent, homozygous copy number losses in *MUTYH* mapping to chromosome 1p34 in 18% of all TETs and in 25% of SCCs. Loss of function of this gene with germline mutations is known to predispose individuals to colorectal cancer and its role in TETs warrants further investigation.

6.1 The aims of the research described in this chapter were the following:

- To discover important chromosomal copy number alterations in TETs
- To integrate these results with the copy number alteration data in TETs from TCGA

6.2 Introduction

Copy number alterations (CNAs) in cancer can lead to amplification of oncogenes via chromosomal gain or loss of tumour suppressor genes via chromosomal loss. Interrogation of genome-wide SNP arrays is a way of identifying somatic or germline CNAs in cancer genomes. These aberrations can range from very small to very large structural changes. A single commercially available assay can be used to genotype over one million SNPs with over 99.5% accuracy (266).

The importance of studying focal or large-scale (whole chromosome or chromosome arm) CNAs is being increasingly recognized in cancer. Large-scale gains or deletions affect approximately 25% of a cancer's genome and focal events affect around 10% of tumours. Focal deletions or

amplifications are now routinely localized to 'peaks' with around six to seven genes, although this number can vary considerably. Despite that, identifying driver mutations within those still remains a challenge (218).

The main difficulty with SNP array data, as with WES data, lies in their analysis and interpretation. There are several reasons for this, including tumour aneuploidy, intratumour heterogeneity and the presence of non-malignant cells within the samples genotyped. A number of different tools have been developed to attempt to analyze data taking these factors into account in an attempt to increase the sensitivity of the analysis, increase true positive results and minimize false positive findings (196). One of these is ASCAT (allele-specific copy number analysis of tumours), which has been used in the current study and is described in more detail in chapter 2.2.10.2.

In TETs the most systematic CNA determination has been performed by TCGA, whose data have been analysed by Lee et al (177). The significant findings from that study were few. Significant deletions in CDKN2A at 9p21.3 were seen in a sub-cohort of B2 and B3 TETs (seven of 25, 28%). Apart from 9p21.3, significant deletions were seen in the following chromosomal regions: 22q12.1, 6p25.2, 3p22.2, 2q37.1, 6q21, 10q26.3, 9p13.11, 22q13.32, 7q36.3, 11q22.1, and 9q22.33. Analysis was performed using SNP6 Copy number analysis (GISTIC2). A Q value of less than 0.20 was used for significance. In terms of copy number gains there were no results reaching significance at a Q value of 0.20. There was however a number of gains with a Q value marginally below 0.25 (Q = 0.24554 for all), which were the following: 1q, 10q, 12p, 17q, 21q, Xq and Xp (177). Determination of CNAs in TETs has also been previously performed using CGH (106,122,165-170) and results are described in chapter 1 and summarized in Table 1.8 for gains and Table 1.9 for losses.

6.3 Materials & Methods

6.3.1 Sample characteristics and DNA extraction

TET genomic DNA was extracted from 143 TETs that had been removed during a thymectomy or metastatectomy and were surplus to diagnostic requirement. 96 of these also had matched normal tissue from which DNA was also extracted. All samples were subject to histological assessment by expert Consultant Histopathologists to confirm the diagnosis, using the 2015 WHO criteria, and determine tumour abundance as described in chapter 2.2.3. Stage was recorded as per the Masaoka-Koga classification. DNA from FFPEs was extracted with Quiagen's QIAamp DNA FFPE Tissue Kit (Catalogue No. 56404) as described in 2.2.6 and from frozen tissue with Quiagen's AllPrep DNA/RNA/miRNA Universal Kit (Catalogue No. 80224) as described in 2.2.7. Of the 143

TETs, 22 were type A, 36 AB, 19 B1, 33 B2 and nine B3 thymomas. There were four cases of MNT LS, and eighteen carcinomas, twelve of which were SCCs and six were NETTs. There were also two cases of thymic hyperplasia.

6.3.2 SNP genotyping and data analysis

SNP genotyping was performed by Eurofins Genomics (https://www.eurofinsgenomics.eu/en/genotyping-gene-expression/genotyping-services/snp-genotyping.aspx) (2.2.10). The samples were genotyped using the Illumina Infinium® OmniExpressExome-8 v1.3 and 1.4 arrays. This is described in detail in chapter 2.2.10 as is the analysis that was performed performed with the help of Amit Mandal, bioinformatician at NHLI (2.2.10.1-4). Briefly, ASCAT (ver. 2.4.4) was used for normalization, segmentation and copy number calling using the R statistical language platform (ver. 3.4.0). Copy number segments from paired and unpaired samples from ASCAT underwent custom thresholding to increase stringency and minimize false positives. The final segment data were then pooled separately for the amplification and deletion groups. These were then overlapped with the Ensembl annotation data using Bedtools (ver. 2-2.26). A further custom script was used to select genes that were protein coding and did have a cDNA clone ID. Following this annotation, a recurrence matrix was created in R (ver. 3.4.0) to determine the most frequent somatic copy number aberrations in the entire cohort. The resulting two tables, one each for copy number gains and deletions, were then studied as a whole and each was subsequently split into two separate tables, to undergo analysis as two groups, a) A, AB, B1 and B2 thymomas and b) B3 thymomas and thymic carcinomas (SCCs and NETTs). This was done to increase the likelihood of identifying specific CNAs for each TET subtype or for overall more aggressive or indolent subtypes.

Of these, known cancer genes (n=616) as per COSMIC (http://cancer.sanger.ac.uk/census) (221) were prioritised as candidate genes within regions harbouring both amplifications and deletions. In the case of deletions, homozygous deletions were also specifically screened for and prioritized over heterozygous deletions.

6.4 Results

6.4.1 Patient demographics and sample characteristics

Of the 143 patients with TETs that underwent SNP genotyping, 64 were male (45%) and 79 were female (55%). The median age was 60 (range 21 – 85). 25 of the 143 (17%) patients were known to have MG. The mean tumour content for all tumours was 91% (range 30 - 100). 26 (18%) were known to be dead although the cause of death is uncertain. 134 had had their GFT2I status determined with Sanger Sequencing (chapter 4) and 48 (36%) were found to be mutant.

ASCAT that was used for normalization, segmentation and copy number calling was unable to confidently call a large number of cases for either gains or losses. Specifically 55 cases (38%) gave copy number gain results and 78 (55%) copy number loss results.

6.4.2 Somatic copy number gains

9175 genes were found to have copy number gains in this study. The gene that was found to be most commonly amplified across all 55 samples was *MKI67* (chromosome 10q26.2) (17/55, 30.91%).

There were only another four genes that were found to have copy number gains in over 15% of TETs. These were the following: *MUC17*, *FLG*, *HRNR* and *AHNAK*, none of which are known cancer genes (Table 6.1). Interestingly there were four SCCs, which all had gains in all four of these genes (TCRBF06753T1, TCRBF19400T1, TCRBF21432T3 and TCRBF31499T1).

Copy number gains were analyzed in two major groups as described in 6.3.2.

6.4.2.1 Somatic copy number gains in aggressive TETs

The first group consisted of 16 TETs, eight SCCs, four NETTs, three B3 thymomas and the single case of an anaplastic B2 thymoma. Of the 2168 genes that were amplified in two or more cases (12.5%), only 51 were in the COSMIC database as known cancer genes. 462 genes were amplified in three or more SCCs. These amplifications were seen in the following chromosomal arms: **1q,** 2q, **7q,** 10q, 11p, **11q,** 12q, 14q, 16q, 17p, 19p and 20q. Of those only ten harboured known cancer genes, which were the following: *ASXL1* (chromosome 20q11.21), *ELK4* (chromosome 1q32.1), *LMNA* (chromosome 1q22), *MUC*1 (chromosome 1q22), *NTRK1*

(chromosome 1q23.1), *PDE4DIP* (chromosome 1q21.1), *PRCC* (chromosome 1q32.1, *SDC4* (chromosome 20q13.12), *SLC45A3* (chromosome 1q32.1) and *TPM3* (chromosome 1q21.3 (Table 6.1). Of those only *ASXL1* had a germline mutation in a B3 thymomas, which is important as tumour suppressor mutations commonly occur due to mutation on one allele and loss of the other allele through CNA. *PDE4DIP* had two somatic mutations that were common polymorphisms as per ExAc. *PDE4DIP* was also the gene from the list of known cancer genes that was amplified the most across all 55 TETs (n=7, 12.73%). This was followed by *ELK4*, *LMNA*, *MUC1*, *NTRK1*, *PRCC*, *SLC45A3*, and *ERBB3* that were all found to have gains in four TETs (7%). Apart from *ERBB3* (chromosome 12q13) all of these genes are on chromosome 1q and were in three of the four samples with copy number gains that all happened to be SCCs (TCRBF06753T1, TCRBF19400T1 and TCRBF21432T3) (Table 6.1).

TREML4 (chromosome 6p21.1) was the only gene amplified in two of three B3 thymomas, which is not a known cancer gene. There were no genes amplified in more than one of four NETTs. *MYC* (chromosome 8q24.21) was amplified in one of the four (25%) NETTs (TNRBF95414T1). The same NETT had *AKT1* amplified (chromosome 14q32.32) and *POU2AF*1 (chromosome 11q23.1). These were the only known cancer genes with copy number gains in this cohort of NETTs.

6.4.2.2 Somatic copy number gains in indolent TETs

The second group of TETs analyzed was composed of a total of 39 thymomas (eleven type A, four MNT, eight AB, two B1 and fourteen B2). Only twenty genes were amplified in more than four cases (10%) and of those *PDE4DIP* (chromosome 1q21.1) being the only known cancer gene. This is a significantly lower number of amplified genes compared to the group of more aggressive TETs. The remaining genes were in the following chromosomes: 1p, 1q, 6p, 6q, 7q, 10q, 11p, 11q, 17p, 18q and 19p. Of the eight AB thymomas only two had amplifications, which were in the following genes: *OR4A5*, *OR4C46* and *APOB*, and none of these is a known cancer gene. Two type A, one B2 and one MNT thymomas did not have gains in any genes.

Table 6.1 summarizes the most salient findings from this study in terms of genes that were found to have copy number gains. The genes that were found to have the most copy number gains across all samples are listed first. These are followed by the ones that are **known cancer** genes and were found to have the most copy number gains across all samples and the last four are known cancer genes that were recurrent in B3 thymomas or found in NETTs.

6.4.2.3 Copy number gain discussion

The gene with the most recurrent copy number gains in this cohort was *MKI67* (chromosome 10q26.2) (30.91%). Marker of Proliferation Ki-67 (*MKI67*) encodes for the proliferation marker Ki-67, whose function is to prevent chromosomes from collapsing into a chromatin mass and instead move independently to interact with the mitotic spindle and ultimately to daughter cells during mitosis. It effectively acts as a surfactant during chromosome separation to facilitate intracellular compartmentalization (267). Clinically Ki-63 4xpression is used routinely as a marker of proliferation in human solid and haematological tumours as it is associated with cell proliferation and growth (268).

MKI67 was not mutated in our cohort of samples that underwent WES and has not been found to be mutated (or amplified) in the TET WES by TCGA (107) or Petrini et al (106) or any other of the studies looking at CNAs in TETs (106,122,165-170). WES of 1144 lung cancers by TCGA found MKI67 to be mutated in 6.6% of cases, all of which were described as putative passenger mutations and with 84% of those being different missense mutations with hardly any recurrence. Amplification has been described in 0.2% of tumours and deletions have been described in 0.4% of tumours (269). The significance of this finding is therefore unclear but warrants further investigation.

In Petrini *et al.* the most frequent arm-level gains were in chromosomes 1q (55%), 7 (7p, 20%; 7q, 15%) and 20p (17%) (106). Apart from the notable exception of 20p these results are in agreement with the findings of the current study.
Using GISTIC, the following peaks of focal copy number gains were identified:
chr1:207,973,966-207,984,254
chr1:145,340,062-147,824,207
chr7:130,558,381-130,691,879
chr8:48,724,203-48,810,613
chr12:116,439,775-116,672,665
chr15:62,167,567-62,341,841
chr18:60,783,213-60,793,676 (BCL2 was the locus within this peak region)
BCl2 amplifications were a particularly salient finding as it correlated with increased BCL2 expression from RNA sequencing data (106). There were no cases of copy number gains in BCL2 in our NHLI cohort.

The *BCL2* (18q21.33) gene encodes for the Bcl-2 (B-cell lymphoma 2) protein. It was the first apoptotic regulator to be discovered and a chromosomal translocation in follicular lymphomas was the first mechanism identified by which the oncoprotein is activated. The Bcl-2 family is one of the

two major apoptotic pathways in human cells, and is called the 'intrinsic or 'mitochondrial' or 'stress' pathway. The mechanism by which the Bcl-2 family acts to switch on the apoptosis pathway has been described as an on-off switch, which integrates signals between the anti-apoptotic sub-family (including (Bcl-2, Bcl-xL, Bcl-w, Mcl-1, A1 and Bcl-B) and the pro-apoptotic sub- family (including (Bad, Bak, Bax, Bid, Bik, Bim, Hrk, Noxa and Puma), which determine the cell's fate (270). *BCL2* is overexpressed in a small subset of leukemias and the t(14;18) chromosomal translocation causing over-expression of the protein is found in over 90% of follicular lymphomas (271). *BCL2* is also overexpressed in solid tumours, such as breast cancer and has also been linked to resistance to chemotherapy as well as targeted therapies such as the anti-*HER2* transtuzumab (272).

Other studies that have grouped B2 and B3 thymomas have all shown gains in chromosome 1q in both subtypes as they have in carcinomas but there has otherwise been very little overlap in their findings (see Table 1.8) (106,122,165-170). This was the case in this study for all TET subtypes but particularly for the group of aggressive TETs. The largest study so far to look at CNAs in TETs was TCGA which showed gains in 10q, which were also seen in the current study, as well as in 12p, 17q, 21q, Xq and Xp that were not (177).

Chromosome 1 is the largest human chromosome. Many important oncogenes are present on chromosome 1 such as *NRAS*, *JUN*, *MYCL* and *TAL1*. However these are not amongst those showing gains in this study. Gains in chromosome 1q have also been seen in other tumour types. For example in breast cancer 1q gains are seen in 50-60% of tumours. As is the case with TETs, they have been found across the spectrum of low to high-grade breast cancers raising the possibility that any oncogenic process relating to that region is an early event. The exact significance of these findings is yet to be established (273). However as such copy number gains have been seen in other tumour types and are the only consistent finding in TETs further attempts need to be made to pin down the gene or genes that are more likely to be responsible and establish their likely oncogenic function in TETs and, if not already established, in other cancers.

Regarding NETTs, despite the small number of samples, our results are in keeping with the largest study to date, which used CGH to determine CNAs in 73 NETTs (13 typical carcinoids, 40 atypical carcinoids and 20 high-grade neuroendocrine carcinomas). The most frequent gains were of chromosome 8q24 (*MYC* gene locus) across all NETT subtypes (16%). 1q gains were also seen in 12% of all NETTs. No other chromosomal arm showed gains in over 10% of cases (274).

The results for the indolent thymomas are in keeping with the findings from other groups that have also shown an increasing number and extent of CNAs from indolent to more aggressive TETs (106,122). Very few groups have reported gains in A and AB thymomas and there are hardly any

data on B1 thymomas at all. In the current study it was likely that the quality of the DNA from FFPE samples hindered the analysis of some of the results. It is however interesting that for B1 thymomas in particular ASCAT gave copy number gain results for only two of nineteen (11%), which is by far the lowest proportion of samples between all subtypes.

Table 6.1. Protein coding genes with copy number gains in TETs

Gene	Chromosome	Recurrence in all TETs (n=55)	Recurrence in TET subtypes	Known CA gene
MKI67	10q26.2	17 (31%)	4/8 (50%) SCCs 6/14 (43%) B2 4/11 (36%) A	Tamborero *et al*
MUC17	7q22.1	12 (22%)	4/8 (50%) SCCs 3/14 (21%) B2	NO
FLG	1q21.3	9 (17%)	4/8 (50%) SCCs	NO
HRNR	1q21.3	9 (17%)	4/8 (50%) SCCs	NO
AHNAK	11q12.3	9 (17%)	4/8 (50%) SCCs	NO
PDE4DIP	1q21.1	7 (13%)	3/8 (38%) SCCs 2/14 (14%) B2	COSMIC
ELK4	1q32.1	4 (7%)	3/8 (38%) SCCs	COSMIC
LMNA	1q22	4 (7%)	3/8 (38%) SCCs	COSMIC
MUC1	1q22	4 (7%)	3/8 (38%) SCCs	COSMIC
NTRK1	1q23.1	4 (7%)	3/8 (38%) SCCs	COSMIC
PRCC	1q32.1	4 (7%)	3/8 (38%) SCCs	COSMIC
SLC45A3	1q32.1	4 (7%)	3/8 (38%) SCCs	COSMIC
ERBB3	12q13	4 (7%)	2/8 (25%) SCCs 2/14 (14%) B2	COSMIC
ASXL1	20q11.21	3 (5%)	3/8 (38%) SCCs	COSMIC
SDC4	20q13.12	3 (5%)	3/8 (38%) SCCs	COSMIC
TPM3	1q21.3	3 (5%)	3/8 (38%) SCCs	COSMIC
TREML4	6p21.1	2 (4%)	2/3 (67%) B3	NO
MYC	8q24.21	2 (4%)	1/3 (33%) NETTs 1/8 (13%) SCCs	COSMIC
AKT1	14q32.32	2 (4%)	1/3 (33%) NETTs 1/8 (13%) SCCs	COSMIC
POU2AF1	11q23.1	1 (2%)	1/3 (33%) NETTs	COSMIC

Important protein coding genes with copy number losses in TETs based on recurrence and on whether these are known to be implicated in cancer. COSMIC refers to genes included in the cancer gene census by the COSMIC database. Recurrence is included for the entire cohort and subtypes with high recurrence are also mentioned separately.

6.4.3 Somatic copy number losses

In the 78 evaluable TETs, copy number losses were seen for a total of 10216 protein-coding genes. *TDRD6, TTN, CHPF2, NLRP12, RP1L1, NDOR1, ANKK1, HELZ2, C1orf116, TAS1R1, COL4A4, PLEC, BZRAP1* and *EMILIN3* were deleted in over 25% of all TETs, and none are known cancer genes (Table 6.2). Only 107 genes had copy number losses in more than 10% of all 78 TETs. Of those the only known cancer gene was MUTYH (chromosome 1p34.1) with copy number losses seen in fourteen (18%) TETs. These were all homozygous deletions. There were only seven known cancer genes that had copy number losses in more than 5% of samples. Apart rom *MUTYH*, these were *BCL9L* (chromosome 11q23.3), *MET* (chromosome 7q31), *NUMA1* (chromosome 11q13), *FIP1L1* (chromosome 4q12), *ARID1B* (chromosome 6q25.1) and *GNAS* (chromosome 20q13.2) (Table 6.2).

As with copy number gains, copy number losses were analyzed in two groups according to the TET subtypes.

6.4.3.1 Copy number losses in aggressive TETs

This group consisted of four B3 thymomas, four NETTs and eight SCCs. 270 genes had copy number losses in two or more cases (13%). Of these the only known cancer gene with copy number losses in three cases (two SCCs and one B3 thymoma) was MUTYH (chromosome 1p34.1). Only 83 genes were lost in three or more cases (17%) and these were found on many different chromosomes: 1p, 1q, 2p, 2q, 3p, 4p, 4q, 5q, 6p, 6q, 7q, 8p, 8q, 9q, 10q, 11q,12q, 14q, 15q, 17p,17q, 19p, 19q, 20p, 20q, 21q and 22q. Copy number losses were seen in two cases (2/16, 13%) in the following known cancer genes: *ARID1B* (6q25.3) (one SCC and one B3), *ECT2L* (chromosome 6q24.1) (one SCC and one B3), *MAML2* (11q21) (one SCC and one NETT), *MET* (7q31.2) (one SCC and one B3), and *MLLT4* (6q27) (one SCC and one B3) (Table 6.2). It is worth pointing out that three of these six genes are located on chromosome 6q although they were not seen in the same cases.

Of the three subtypes, NETTs had the fewest copy number losses, with a median of 38 in the four cases (range 4 – 63), followed by SCCs with a median of 95 (and a vast range of 1-1844) and the B3 thymomas with a median of 161 (range 76 – 185).

BZARP1 (17q22), *ACAN* (15q26.1) and *FAM65C* (20q13.13) were the only genes with copy number losses in two of four NETTs and these were all homozygous. No genes had copy number losses in three or four NETTs. The only previous study on 73 NETTs revealed loss of 13q in 16%

of cases as well as loss of 3p and 3q both in 10% of cases, with no other chromosomal losses in over 10% of cases (274).

29 genes were found to have copy number losses in two or more B3 thymomas, with none of those being known cancer genes and with hardly any being present on the same chromosome apart from three in 1q.

134 genes had copy number losses in two or more SCCs; only thirteen had copy number losses in three (38%) SCCs and two (TTN and COL12A1) in four (50%). Of these 134 the only known cancer gene was *MUTYH* (Table 6.2).

A homozygous deletion of *CDKN2A* (9q21.3) was seen in only one TET (1%), which was one of the eight SCCs (13%). No TP53 deletions were seen in the aggressive TETs.

6.4.3.2 Copy number losses in indolent TETs

This group consisted of a total of 62 TETs, eleven type A thymomas, four MNT, seventeen AB, ten B1, nineteen B2 and one hyperplasia. As was the case with the aggressive TETs, there were no recurrent deletions and only 121 had copy number losses in six or more cases (10%). Of those the only known cancer gene was again *MUTYH* (chromosome 1p34.1), which was seen in eleven cases across all the TET subtypes (Table 6.2). The top fourteen genes had copy number losses in fifteen or more cases (24%) (Table 6.2). Like with the aggressive TETs, the deletions were in many chromosomes and for only these top genes, these were: 1p, 1q, 2q, 3q, 6p, 7q, 8p, 8q, 9q, 11q, 19q and 20q.

The median number of genes with copy number loss was lowest in B1 thymomas at 15 (range 0-2804), followed by AB with a median of 39 (range 9-174), B2 with 49 (range 1-282), A with 68 (range 4-6509) and MNT with 103.5 (range 13-198). It is very interesting that there were outliers in almost all subtypes/. Two B2 and two B1 thymomas had only one gene with copy number losses and there was one B1 with none. At the other end of the spectrum one B1 thymoma had copy number losses in 2804 genes and an A thymoma had copy number losses in 6509 genes.

A deletion in *TP53* (17p13.1) was seen in only one (1%) TET, which was a type A thymoma.

6.4.3.3 Copy number losses discussion

Despite a large number of genes having copy number losses in this TETs cohort (n=78), there were very few genes with recurrent losses across all TETs and exceedingly few known cancer genes amongst those. This is consistent across all groups that have investigated CNAs in TETs before (Table 1.9). From the TCGA CNA data significant deletions can be found only for *CDKN2A* at 9p21.3 in a sub-cohort of B2 and B3 thymomas (seven of 25, 28%). Other significant deletions were seen in chromosomes 22q12.1, 6p25.2, 3p22.2, 2q37.1, 6q21, 10q26.3, 9p13.11, 22q13.32, 7q36.3, 11q22.1, and 9q22.33. In the entire TCGA cohort there were no genes with copy number alterations (gains or losses) in over 4% of all TETs. In a recent study on 28 thymic carcinomas, FISH detected *CDKN2A* (chromosome 9p21.3) deletions in twelve cases (38%), deletions of TP53 (chromosome 17p13.1) in nine (32%) cases and deletions of *ATM* (chromosome 11q22.3) in 8% cases. Of the *CDKN2A* deletions, three were heterozygous, four were homozygous and five were mixed. Of the nine cases with *TP53* deletions, a mutation in the remaining allele was identified in four (44%) of them (heterozygous deletions). Only one of seventeen (6%) B3 thymomas also had a mixed heterozygous/ homozygous *CDKN2A* deletion and none were found in type A thymomas (108). In the NHLI cohort there was only one case of *CDKN2A* copy number loss in an SCC (1% of all TETs and 13% of SCCs).

Cyclin Dependent Kinase Inhibitor 2A (*CDKN2A*) is a tumour suppressor gene, which is commonly lost in a number of solid tumours. *CDKN2A* encodes the p16INK4a protein, which is involved in cell cycle control and senescence. Therefore its loss or silencing leads to altered cell cycle regulation that promotes tumourigenesis and metastasis (275). The majority of deletions are homozygous as was the case with the one SCC in this study. In mesothelioma *CDKN2A* losses were the commonest CNA seen and in a study where homozygous deletions of *CDKN2A* were seen in 70% of those cases there was also a genetic alteration in NF2 (240,276). Deletions are also seen in 53% of NSCLC cell lines and 20% of oesophageal SCCs (275). Another way *CDKN2A* is inactivated in cancer is by promoter hypermethylation, with some studies suggesting this as poor prognostic marker for cancers such as liver carcinomas, colorectal carcinomas, melanomas, head and neck SCCs, pancreatic adenocarcinoma, epithelial ovarian cancer, prostate cancer and mesothelioma amongst others (275-277).

In Petrini *et al* (220), in 65 TETs, the most frequent chromosomal arm deletions amongst all TETs affected chromosome 6 (6p, 26%; 6q, 29%), 3p (22%) and 13q (18%). Analysis with GISTIC revealed copy number losses within peaks in ten regions, two of which are in chromosome 9 and two in chromosome 6:
chr1:72,765,091-72,812,744
chr3:54,321,095-54,545,875

chr4:2,681,584-2,842,143
chr6:932,265-1,376,210
chr6:167,905,600-168,332,222
chr9:21,894,008-21,984,463 (where CDKN2A was the gene in the peak region of loss)
chr9:137,424,346-137,505,498
chr11:131,943,005-132,771,743
chr13:113,039,306-113,063,878
chr16:86,004,137-86,526,617.

In agreement with these findings, most previous studies have found the commonest deletions across all subtypes to be in chromosomes 6p and 6q and 3p in all TET subtypes in addition to those in 13q, 16q and 17p in B3 thymomas and SCCs (122,165-170) (Table 1.8). In the current study recurrent deletions have been seen in genes on 6p and 6q

Tudor Domain Containing 6 (*TDRD6*) (6p12.3) was the most commonly deleted gene in all TETs (32/78, 40%) and was particularly commonly deleted in indolent TETs (26/62, 42%). *TDRD6* is a central component of the chromatoid body (CB) in male germ cells (278). Although it is not a known cancer gene, one putative passenger mutation was seen in this gene in the TET TCGA cohort as well multiple putative passenger mutations in the lung cancer cohort of TCGA. In the same cohort amplifications were seen but deletions were not (279).

Cyclin Dependent Kinase Inhibitor 2A (*CDKN2A*) acts as a tumour suppressor gene, which is commonly lost in a number of solid tumours. The majority of cases are homozygous deletions as was the case with the one SCC in this study. Another way it is inactivated in cancer is by promoter hypermethylation, with some studies suggesting this as poor prognostic marker for cancers such as liver carcinomas, colorectal carcinomas and mesothelioma (276,277). In mesothelioma *CDKN2A* losses were the commonest CNA seen and in a study where homozygous deletions of *CDKN2A* were seen in 70% of those cases there was also a genetic alteration in NF2 (276).

Of the known cancer genes *MUTYH* (chromosome 1p34.1) was the gene with the greatest number of copy number losses in TET samples (14/78,18%). *MUTYH* has not been found to be mutated in the NHLI, TCGA or Petrini et al cohorts and CNAs have not been seen in TETs in TCGA or Petrini et al. WES of 1144 lung cancers by TCGA found putative passenger *MUTYH* mutations in 1% of cases and amplifications in 0.9% (269). Somatic mutations in this gene as well amplifications have also been seen in other tumour types such as breast and pancreatic cancers (105), albeit rarely. MutY DNA Glycosylase, also known as MutY Homolog (*MUTYH* or *MYH*) is a gene that encodes a DNA Glycosylase and it is involved in oxidative DNA damage repair as part of the Base Excision Repair (BER) system (280). *MUTYH* is best known for its role in *MUTYH*-associated polyposis

(MAP), which is an autosomal recessive polyposis syndrome (281). Germline *MUTYH* mutations predispose to colorectal adenomas and increase the risk of developing colorectal cancer. The exact mechanism by which this is achieved is not entirely clear although what has been see in cases of biallelic *MUTYH* germline mutations is a marked increase in spontaneous somatic C:G > T:A transversions in *APC* and in *K-RAS*. Germline *MUTYH* mutations are thought to alter the interactions between *MSH6* and *MUTYH*. A case of a germline *MUTYH* germline and missense *MSH6* mutations has been described although other studies indicated that these mutations are mutually exclusive (281). Few cases of somatic *MUTYH* mutations have been described in cases of sporadic colorectal cancer in a study of Tunisian patients (282).

In the current study it is intriguing that one of the SCCs with a homozygous *MUTYH* deletion (TCRBF03870) was one of the samples that was whole exome sequenced and no putative somatic or germline driver mutations were identified. Another sample that had a *MUTYH* homozygous deletion was THRBF13799, a B3 thymoma that was also sequenced and was found to have germline *ASXL1* and germline *HLA-A* mutations. Of the other samples with a homozygous *MUTYH* deletion, another SCC (TCRBF21432) that was sequenced carried a somatic *MSH6* mutation and a NETT with multiple mutations and a hotspot *TP53* mutation also carried a somatic missense *MSH2* mutation. The potential role of *MUTYH* as a tumour suppressor in TETs warrants further investigation.

Table 6.2. Protein coding genes with copy number losses in TETs

Gene	Chromosome	Recurrence in all TETs (n=78)	Recurrence in TET subtypes	Known CA gene
TDRD6	6p12.3	32 (40%)	26/62 (42%) of indolent TETs	NO
TTN	*2q31.2*	29 (37%)	22/62 (35%) of indolent TETs	NO
CHPF2	7q36.1	25 (32%)	20/62 (32%) of indolent TETs	NO
NLRP12	19q13.42	22 (28%)	17/62 (27%) of indolent TETs	NO
RP1L1	8p23.1	21 (27%)	18/62 (29%) of indolent TETs 3/4 (75%) B3	NO
NDOR1	9q34.3	21 (27%)	18/62 (29%) of indolent TETs	NO
ANKK1	11q23.2	21 (27%)	19/62 (31%) of indolent TETs	NO
HELZ2	20q13.33	21 (27%)	15/62 (24%) of indolent TETs	NO
C1orf116	1q32.2	20 (26%)	15/62 (24%) of indolent TETs 3/4 (75%) B3	NO
TAS1R1	1p36.31	20 (26%)	18/62 (29%) of indolent TETs	NO
COL4A4	2q36.3	20 (26%)	16/62 (26%) of indolent TETs	NO
PLEC	8q24.3	20 (26%)	15/62 (24%) of indolent TETs	NO
BZRAP1	17q22	20 (26%)	14/62 (23%) of indolent TETs 2/4 (50%) NETTs	NO
EMILIN3	20q12	20 (26%)	18/62 (29%) of indolent TETs	NO
COL8A1	3q12.1	17 (22%)	15/62 (24%) of indolent TETs	NO
MUTYH	1p34.1	14 (18%)	2/8 (25%) SCCs >10% of all thymoma	COSMIC

			subtypes	
ACAN	15q26.1	12 (15%)	2/4 (50%) NETTs	NO
FAM65C	20q13.13	11 (14%)	2/4 (50%) NETTs	NO
BCL9L	11q23.3	6 (8%)	3/11 (27%) A	COSMIC
MET	7q31	5 (6%)	1/8 (13%) SCC 1/4 (25%) B3	COSMIC
NUMA1	11q13	5 (6%)	2/11 (18%) A, 210 (20%) B1, 1/8 (13%) SCC	COSMIC
FIP1L1	4q12	4 (5%)	2/17 (12%) AB	COSMIC
ARID1B	6q25.1	4 (5%)	1/8 (13%) SCC 1/4 (25%) B3	COSMIC
GNAS	20q13.2	4 (5%)	2/19 (11%) B2	COSMIC
ECT2L	6q24.1	3 (4%)	1/8 (13%) SCC, 1/4 (25%) B3, 1/17 (7%) AB	COSMIC
MAML2	11q21	3 (4%)	1/8 (13%) SCC, 1/4 (25%) B3, 1/19 (5%) B2	COSMIC
MLLT4	6q27	2 (3%)	1/8 (13%) SCC, 1/4 (25%) B3	COSMIC
CDKN2A	9q21.3	1 (1%)	1/8 (13%) SCC	COSMIC
TP53	17p13.1	1 (1%)	1/11 (9%) A	COSMIC

Important protein coding genes with copy number losses in TETs based on recurrence and on whether these are known to be implicated in cancer. COSMIC refers to genes included in the cancer gene census by the COSMIC database. Recurrence is included for the entire cohort and subtypes with high recurrence are also mentioned separately.

6.5 Conclusions

In a tumour type that has the lowest mutation burden of all solid tumours (107) it is not surprising that there are not many significant, recurrent CNAs. This study has also confirmed that there are overall fewer CNAs in A and AB thymomas compared to B2 and B3 thymomas and carcinomas.

All studies also revealed a lack of recurrent chromosomal gains in TETs with the exception of chromosome 1q where amplification of a number of genes was seen. The gene with the most recurrent copy number gains in the current cohort was *MKI67* (chromosome 10q26.2) (30.91%), which is an intriguing result that needs to be followed up. It has not been described in any previous studies on TETs. Copy number gains in genes such as MYC appeared to account for a very small minority of TETs.

Despite a large number of genes having copy number losses in this cohort (n=10216) there were few with copy number losses in more than 10% of samples (n=107). These were across different chromosomes. In terms of recurrence, copy number losses in chromosomes 6p and 6q and 3p have been seen in all TET subtypes (122,165-170). It is striking that deletions in the whole of chromosome 6 or either arm are reported in all studies that have looked into B3 thymomas and carcinomas. Deletions in 13q have been observed in four of six studies in B3 thymomas but only once in carcinomas. 16q and 17p deletions are also frequently observed in B3s and carcinomas (122,165-170). CDKN2A, known to be lost in other tumour types, has been found to have copy number losses in the current as well as previous studies on TETs although the frequency of those varies considerably (108,177).

One of the most interesting findings in this study was copy number losses in *MUTYH* (chromosome 1p34.1) in 18% of all TETs (25% of thymic SCCs). This gene is involved in Base Excision Repair and germline mutations in this gene have been extensively described in *MUTYH*-associated polyposis (269). Somatic mutations have also been rarely reported (105) and are known to be cancer casing in sporadic colorectal cancer (282). This finding warrants follow up in TETs.

One of the study's major shortcomings was the use of DNA extracted FFPEs, which was of suboptimal quality for SNP genotyping and for ASCAT to perform the required analysis. Lack of expression data also makes it harder to correlate these findings with the expression of genes that are likely to be key.

A pan-cancer analysis of somatic copy number variations identified 70 recurrently amplified regions and 70 recurrently deleted regions. 'Peak' regions that are more likely to include important

oncogenes or tumour suppressor genes were also identified. These accounted only for 21% of all amplifications and 23% of deletions (283). In the case of amplifications 24 peaks contained known oncogenes, which are known to be activated with amplification. Regarding deletions, 22 of the 70 regions contained known tumour suppressor genes. Further analysis in the 37 peak regions of amplifications that have fewer than 25 genes and no known targets revealed that genes on these peaks are mostly associated with epigenetic and mitochondrial regulation and thirteen consisted of chromatin and histone modifier genes. These findings indicate the likely importance of epigenetic alterations and chromatin remodelling in cancer (283). It is also clear that for the majority of CNAs, driver genes are yet to be discovered (218,283).

Chapter 7

Immunohistochemistry for emerging immunotherapy and other targets in TETs

Abstract

Antibody staining on the emerging immunotherapy targets PD-1, PD-L1, PD-L2, LAG-3, TIM-3 B7-H3, VISTA and 4-BB as well another two emerging treatment targets, PDLIM1 and HGF was used on TMAs from 143 TET cases. Lymphocytic PD-1 expression was highest in thymic hyperplasias, followed by AB and B1 thymomas and was lowest in SCCs and NETTs. Epithelial PD-L1 and PD-L2 expression exhibited a pattern of increasing staining from A, AB and B1 thymomas to B2 and B3 being highest in SCCs and NETTs. Lymphocytic PD-L1 and PD-L2 staining were also seen following the opposite pattern with more indolent TETs exhibiting higher expression likely reflecting the lymphocytic component of these tumours. Higher epithelial PD-L1 and PD-L2 in aggressive TETs is encouraging as these are currently targeted with anti-PD1 therapies. B7-H3 staining was overall very high and highest in type A, AB and B3 thymomas and SCCs. Clinical trials with drugs targeting B7-H3 should consider inclusion of appropriate TET cases. TIM-3 staining was overall negative in thymomas and positive in subset of SCCs (4/10) and NETTs (6/8) an interesting finding given the emerging drugs targeting TIM-3. HGF staining was overall negative in thymomas with the surprising exception of MNT thymomas and was highest in NETTs also providing a potential further treatment target for these TET subtypes. Clustering all TETs according to their co-expression of these markers revealed there are cases in all subtypes that behave very differently in terms of their expression and have clustered more with other TET subtypes. This suggests that any TET that behaves more aggressively should be screened using appropriate antibody tests as these might benefit from immunomodulatory therapy, independent of its histology.

7.1 The aims of the research described in this chapter were the following:

- To provide an overview of the importance of immunotherapy targets for the treatment of solid tumours
- To provide an insight into key IHC targets currently under investigation in solid tumours
- To assess the quantity and distribution of a range of immunotherapy targets in all major TET subtypes and correlate these findings with clinicopathological variables
- To cluster TET samples and subtypes according to their expression of these proteins

The proteins investigated were the following:

PD-1	B7-H3
PD-L1	VISTA
PD-L2	4-1BB
LAG-3	PDLIM1
TIM-3	HGF

7.2 Cancer immunotherapy targets

The concept of immune surveillance has been around for many years. Its importance was first questioned in an attempt to explain rare cases of spontaneous tumour regression in cancers such as melanoma, NSCLC and renal cell carcinoma. Today cancer immunotherapy is an exciting, rapidly growing field with clinical trials showing impressive results in a range of solid tumours with PFS and OS improvements already established in melanoma and NSCLC. CTLA4 was the first immune checkpoint protein whose inhibition became a landmark in the treatment of metastatic melanoma and other solid tumours. Other examples of immune checkpoint blockade that have been very successfully applied across many tumour types are the PD-1 inhibitors. PD-1/ PD-L1/ PD-L2 are discussed in detail in sections 1.7.1.2-4. Still it is clear that not all patients respond to these treatments and of those who will, despite some prolonged clinical responses, the majority will ultimately progress. At the same time interactions between tumours and the immune system are being found to be multiple and increasingly complex. Also from experience with these drugs to date, it is likely that combinations of immunotherapy drugs or their combinations with other anticancer agents will likely achieve better outcomes for patients than monotherapies.

For these reasons, in this study expression of a number of protein markers were investigated by antibody labelling of TET TMAs. The aim was to establish which protein markers are expressed in TETs, find differences in expression between TET subtypes and to try to ascertain whether there is a link between their expression and OS, *GTF2I* status and treatment response. A better understanding of some of these molecules in TETs will also provide insights into differences between subtypes and even within subtypes, which might also explain their response or lack thereof to current therapeutic agents.

7.2.1 Inhibitory checkpoint molecules

7.2.1.1 Lymphocyte-activation gene 3 (LAG-3)

LAG-3 (CD223) is a cell surface molecule expressed on activated conventional T-cells(284), T_{REG} cells, B cells and plasmacytoid dendritic cells (Figure 7.0) (285). When it binds to MHC class II molecules on antigen presenting cells, LAG-3 transmits an inhibitory signal in conventional T-cells. At the same time this signaling event boosts the suppressive function of T_{REG} cells. PD-1 and LAG-3 are often co-expressed on T-cells following chronic infections, which renders them dysfunctional. Dual inhibition of these receptors has been shown to confer an additive therapeutic effect in preclinical models. Clinical trials are underway to establish whether this effect is also translated

into cancer patients(285). Specifically a phase I/ II trial is investigating the safety of Anti-LAG-3 with/ without anti-PD-1 in a range of solid tumours (NCT01968109). This does not include TETs.

7.2.1.2 T-cell Immunoglobulin domain and Mucin domain 3 (TIM-3)

TIM-3 (also known as hepatitis A virus cellular receptor 2; HAVCR2) is also an exhaustion-associated inhibitory receptor that serves to dampen down T-cell-effector function and induce T-cell apoptosis (Figure 7.0) (285). It is found on macrophages, monocytes, dendritic and mast cells(286). The four known ligands to TIM-3 are the following: galectin-9, HMGB1, phosphatidyl serine and CEACAM-1. Inhibition of TIM-3 has shown to be effective in preclinical(287) and mouse models against adenocarcinoma of the colon, melanoma, and sarcoma especially in conjunction with PD-1 blockade.

Dual TIM-3 and PD-1 inhibition in mice with acute myelogenous leukemia (AML) has been shown to be very effective and prolong overall survival more than PD-1 blockade does alone(287). TIM-3 inhibition has also been found to enhance $CD8^+$ T cell proliferation and cytokine production of melanoma patients(285).

An intriguing study also found that upon progression to anti-PD1 treatment in two fully immunocompetent mouse models of lung adenocarcinoma, tumours had an increase in TIM-3 staining. With the addition of an inhibitor to TIM-3 a survival advantage was noted. This was later observed in 2 patients who became resistant to anti-PD-1 therapy and also had TIM-3 up-regulation. TIM-3 was up-regulated following anti-PD-1 treatment in T-cells from lungs that had had tumour and not from lymph nodes, peripheral blood or the spleen. In these cases TIM-3 upregulation was in CD4 and CD8 cells that were anti-PD1 antibody bound. Significantly higher TIM-3 positivity was associated with a longer duration of anti-PD1 treatment(288).

TIM-3 has been investigated recently in the context of NSCLC as, in unselected populations, the efficacy of anti-PD1 inhibitors remains at around 20%. Other immune checkpoint inhibitors are also being sought to work alone or in conjunction with anti-PD1 inhibitors to benefit a larger number of patients.

A further study investigated expression of TIM-3 and its correlation with PD-1, PD-L1 and tumour-infiltrating lymphocytes (TILs). 132 NSCLC cancer specimens were evaluated by IHC. The antibodies used were TIM-3 (D5D5R, Cell signaling), PD-1 (NAT 105, Cell marque) and PD-L1 (22C3, Dako). TIL abundance was also scored. On evaluation of TILs, TIM-3 was expressed in 18 of 132 samples (13.6%) and PD-1 was seen in 58 of 132 samples (43.9%). Only 23 tumours

(17.4%) were positive for PD-L1. As expected there was no PD-1 or TIM-3 expression on tumour cells. TIM-3 was significantly overexpressed in TILs from adenocarcinomas vs. non-adenocarcinomas ($p = 0.011$). There was significant correlation between PD-1 and TIM-3 expression on all samples ($p = 0.011$). TIM-3 positivity did not have an effect on overall survival in this cohort(289).

Another study looked into TIM3 mRNA expression in resected NSCLC specimens using microarray analysis. The Agilent Whole Human Genome 4x44K 2-color platform was used to analyze 319 NSCL specimens as well as 15 normal lung resection samples. An equal mixture of 258 of the NSCLC samples was used as reference. Ninety of 319 samples (28%) had high TIM3 expression, 150 (47%) had moderate and 79 (25%) had low expression. What is perhaps even more interesting is that 7 of 15 normal lung samples (47%) also exhibited high TIM3 expression with none having low expression. Consistent with the previous study, 34% of adenocarcinomas showed a high expression vs. only 17% in those with squamous cell histology ($p = 0.03$). TIM3 expression was not prognostic for survival and was not significantly related to gender or tumour stage. There was a non-significant trend towards high TIM3 expression in smokers vs. non-smokers ($p = 0.10$)(290).

TIM-3 has also been investigated in pleural mesothelioma. A recent study used 40 samples of untreated patients and 14 of patients that had been treated with chemotherapy and looked into the expression of PD-1/ PD-L1, TIM-3 and LAG-3. Untreated and pre-treated patients had PD-1 positive TILs and bizarrely 10% of the untreated patients also had only PD-1 positive tumour cells. More than 50% of cases had epithelial or lymphocytic cells positive for PD-L1, which was significantly different between the different mesothelioma subtypes ($p < 0.001$). Sarcomatoid mesotheliomas have the most PD-1 expression. Again unexpectedly TIM-3 was expressed in stromal lymphocytes, plasma cell and tumour cells. This was seen more frequently in pre-treated cases. TIM-3 expression in lymphoid aggregates was a good prognostic factor ($p = 0.001$). LAG-3 was negative in all cases(291).

A small study of 24 TETs has also looked at TIM-3 staining. Here twelve cases were B2/3 thymomas and twelve were squamous cell carcinomas. High TIM-3- expression (>25% of cells staining positive) was seen in 18 of 24 cases (75%). PD-L1 positive tumour cells were highly correlated with a high Tim-3 expression ($p = 0.03$) raising the possibility of trying dual inhibition of both molecules in these patients(91).

A Phase 1 multicentre, open-label, first-in-human study of TSR-022 (NCT02817633), an anti-TIM-3 monoclonal antibody, in patients with advanced solid tumours is currently recruiting. It is being investigated alone and in combination with anti-PD-1 treatment.

7.2.1.3 B7-H3

B7-H3 (also known as CD276) is another important immune checkpoint molecule (Figure 7.0). It is a type I transmembrane protein that belongs to the B7 immunoregulatory family that includes PD-L1 (B7-H1). It is expressed on an array of tissues and cell types. It is constitutively expressed on non-immune resting fibroblasts, endothelial cells, osteoblasts, and amniotic fluid stem cells. Its expression can also be induced on antigen presenting cells, including dendritic cells and monocytes. It can also be found in natural killer cells, B cells, and a small population of T cells. It is crucial in inhibiting T-cell function. It is also present on many solid tumours and is linked with poor overall survival (292). These tumours include hepatocellular/ pancreatic/ prostate/ colorectal/ ovarian/ oral squamous cell/ clear renal cell carcinomas, endometrial, cervical, bladder cancer, NSCLC and melanoma amongst others. The receptor(s) for B7-H3 still remain elusive, as is the mechanism by which it is involved in tumor growth and allows them to evade the immune system (292,293).

7.2.1.4 V-domain Ig- containing suppressor of T-cell activation (VISTA)

VISTA (also known as C10orf54, differentiation of ESC-1 (Dies1), platelet receptor Gi24 precursor, or PD1 homolog (PD1H)) is one of the emerging immunotherapy targets in cancer (Figure 7.0). VISTA–Fc fusion protein and overexpression of VISTA within cells suppress T-cell activation, proliferation, and cytokine production In humans VISTA is expressed predominantly on hematopoietic cells(294). In mice, a combination of PD-1 and VISTA inhibition has achieved impressive synergistic anti-tumour activity providing a rationale for targeting both in humans too(295).

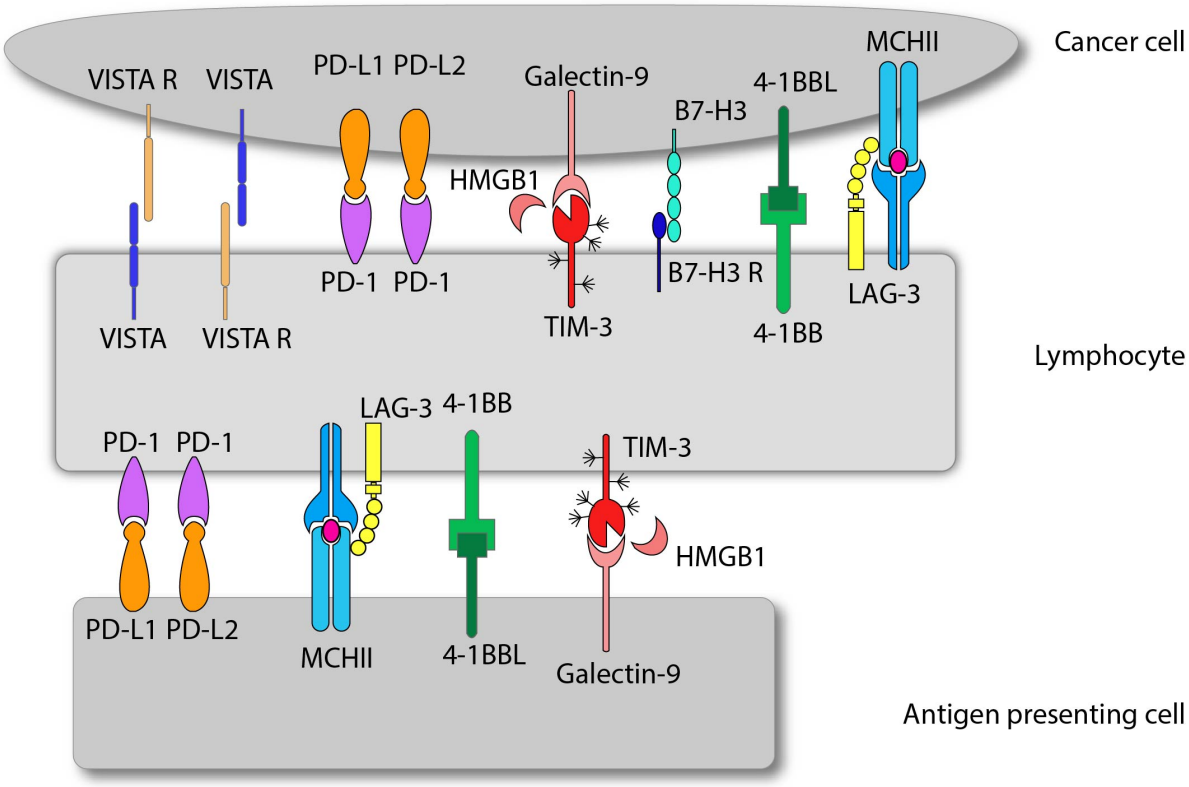
7.2.2 Stimulatory checkpoint molecules

7.2.2.1 T-cell antigen 4-1BB homologue (4-1BB)

4-1BB (also known as CD137 or TNFR superfamily member 9) is a co-stimulatory molecule that is currently being targeted with therapeutic mAbs in early clinical trials (Figure 7.0). It is expressed on T-cells, natural killer cells and monocytes. Stimulation of T-cells by 4-1BBL (its ligand on dendritic cells) leads to T-cell proliferation and the upregulation of the antiapoptotic proteins Bcl-2-like protein 1 (Bcl-xL), Bcl-2-related protein A1 (Bfl-1) and CASP8 and FADD-like apoptosis regulator (c-FLIP) whose function are to protect T cells from activation-induced cell death(285).

Preclinical studies have shown that mAbs to 4-1BB are active against solid tumours alone and in combination with CTLA-4 inhibitors, CD40 activation or radiotherapy. In patients with melanoma, urelumab, a human anti-4-1BB antibody has been shown to be clinically active. Unfortunately a subsequent clinical trial was terminated early to due to severe hepatotoxicity. Other clinical trials with the same agents are still in progress(285). These include combination of urelumab and cetuximab in colorectal and head and neck cancer. Given the different mode of action of such drugs, combining them with anti-PD-1 therapies is also compelling and clinical trials are currently underway in B-Cell Non-Hodgkin lymphoma (NCT02253992) and in patients with solid tumors (PF-05082566 and pembrolizumab; NCT02179918)(296).

Figure 7.0. Emerging immunotherapy targets in TETs



Schematic representation of checkpoint molecules targeted in the current study and their ligands: PD-1, PD-L1, PD-L2, LAG-3, TIM-3, B7-H3, VISTA and 4-1BB.

7.2.3 Other emerging targets

Apart from a number of immune-related markers that are currently under investigation in all cancers there are also other receptors that are under investigation whether because they are part of a known or emerging pathway involved in tumourigenesis or because preclinical studies have found those to be over- or under-expressed in various malignancies. Two of these have been investigated in this study.

7.2.3.1 PDZ And LIM Domain 1 (PDLIM1)

PDLIM1 (also known as CLP36, Elfin, or CLIM1) is a cytoskeletal protein, which works with actin filaments and stress fibres. Its function is to recruit signal proteins to actin and organize signalling complexes at the cytoplasmic or cellular membranes. PDLIM1 is expressed in the heart, in muscles, and in the gastrointestinal tract where it serves as a regulator of cell–cell adhesion and focal adhesion. Its function in cancer is only now beginning to emerge(297).

In colorectal cancer it has been found to be down-regulated, especially in liver metastasis. In these cases it correlates with reduced e-cadherin and membrane β-catenin levels, inducing epithelial to mesenchymal transition. This enhances invasion and increases the metastatic potential of colorectal cancer leading to reduced OS. Hence loss of PDLIM1 expression is likely to be a marker of poorer prognosis(297).

In breast cancer results are inconsistent with this finding, where suppression of expression has been found to impair motility and invasiveness and diminish metastatic potential *in-vivo*. It has also been found to interact with α-actinin-1 and -4 in breast cancer cells and inhibition of either in vitro inhibits migration of breast cancer cells(298).

7.2.3.2 Hepatocyte Growth Factor/ scatter factor (HGF/ SF)

HGF (also known as F-TCF, SF, DFNB39, HPTA and HGFB) is a multi-domain protein that binds the tyrosine kinase protein *MET*, leading it to dimerize and become activated. MET activation leads activation of various effector proteins including PLC, GAB1, GRB2 and SRC. In normal tissue, HGF / *MET* activation is essential for liver and skin regeneration and EMT in embryogenesis. In cancer however it is activated in many solid tumours leading to uncontrolled cell survival, motility, angiogenesis and metastasis(299). Aberrant *MET* activation in cancer is achieved through multiple mechanisms. It can be achieved by overexpression of *MET* in tumours and stromal HGF

overexpression or by point mutation in *MET*. The downstream effectors of this pathway in cancer include the PI3K-AKT pathway, the WNT-beta-catenin pathway and RAS–MAPK pathways. Cross-signaling between MET and he EGFR and VGFR pathways has also been described.

So far MET inhibitors have been developed and have shown some activity against NSCLC, breast, prostate, liver and renal cancer(299). In NSCLC c-MET mutations are relatively uncommon, and abnormal amplification is seen in less than 5% of NSCLC, predominantly in adenocarcinomas. A meta-analysis on the efficacy of c-Met inhibitors in NSCLC found that the currently available agents targeting c-MET (including onartuzumab, tivantinib and crizotinib) significantly prolong PFS but not OS even in cases with c-Met amplification and even when sub-group analyses were contacted(300).

There are currently a large number of trials with numerous drugs targeting HGF and Met in many solid tumours alone or in combination with other drugs such as VEGF inhibitors, EGFR inhibitors and chemotherapy. A full list can be found at:
https://ccrod.cancer.gov/confluence/display/CCRHGF/Home(301).

Preliminary results from some of these are encouraging. Recently results from a phase I trial in head and neck squamous cell carcinoma (NCT02277197) treated with ficlatuzumab, a mAb against HGF and cetuximab in cetuximab-resistant, recurrent or metastatic cases showed a response rate of 17% with a clinical benefit rate, as defined by partial responses and stable disease, of 67%. A Phase II trial is underway(302).

Another study on pancreatic cancer has shown that HGF inhibition *in-vitro* and *in-vivo* was as effective as chemotherapy in inducing local tumour growth arrest and more effective than gemcitabine in reducing metastasis and angiogenesis. A combination therapy with gemcitabine however did not have the same effect(303).

Another area of interest in HGF is its role in cases of EGFR resistance following treatment with EGFR tissue kinase inhibitors. Multiple mechanisms of acquire resistance have been described so far. One of these, although not the commonest, is MET amplification(304). Hence targeting this the MET/ HGF pathway effectively can provide a further line of treatment for these patients.

7.3 Methods

TMAs were constructed from FFPE tumour tissue following thymic resections at the Histopathology Core Facility at the Royal Brompton Hospital by expertly trained histopathology staff. (2.2.11.1-3). From each block 3 x 4mm cores were taken and with the help of a Tissue Microarrayer (Alphelys Minicore® Microarrayer) were embedded into the recipient paraffin blocks. A TMA map was created for each TMA block indicating the exact position of each core. 143 cases were distributed across 6 TMA blocks. Further details can be found in chapter 2.2.12 with detailed SOPs in Appendix 1. From each of the 6 TMA blocks sections were cut at a thickness of 3µm using a standard microtome (Shandon Finesse 325, Thermo Scientific). Each slide was stained using the Ventana Benchmark Ultra platform. All available cases were included and there were no exclusions on any grounds. Antibody staining was performed as detailed in Table 7.1, which for convenience is also included here as Table 8.1. CD45 and cytokeratin MNF116 stains were used to differentiate between epithelial cells and lymphocytes. All IHC-stained sections were scored using a Histo-score (H-score) that was assigned to each sample, scoring the epithelial and lymphocytic component separately, giving a final score range of 0 – 300 scored, by two expert pathologists blinded to the TET subtype and to each other's scoring with me present throughout (2.2.11.4. All statistical analyses were performed by me (2.2.11.5).

To be in line with current published literature on IHC a minimum of 5% of cells had to be positive to be assigned an intensity score. Therefore the minimum positive score included in this study was 1 x 5 = 5.

Table 7.1. Antibody Staining

Antibody*	Antigen Retrieval Reagent Used	Length/Temperature of Antigen Retrieval	Antibody Dilution Factor/Concentration	Primary Antibody Incubation
4-1BB	Ventana Cell Conditioning 1 pH8.5	72 mins at 100°C	1/40	60 mins at 37°C
B7-H3	Ventana Cell Conditioning 1 (pH8)	48 mins at 100°C	1/800	32 mins at 37°C
CD45	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	1/100	32 mins at 36°C
Cytokeratin MNF116	Ventana Protease 1	12 mins	1/200	32 mins at 36°C
HGF	Ventana Protease 1	8 mins	1/200	32 mins at 37°C
LAG-3	Ventana Cell Conditioning 1 (pH8)	40 mins at 100°C	1/100	32 mins at 37°C
PD-1	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	Predilute	16 mins at 36°C
PD-L1	Ventana Cell Conditioning 1 (pH8)	72 mins at 100°C	1/200	32 mins at 37°C
PD-L2	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	3µg/ml	32 mins at 36°C
PDLIM1	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	1/200	32 mins at 37°C
VISTA	Ventana Cell Conditioning 1 (pH8)	32 mins at 100°C	1/150	60 mins at 37°C

*Antibody details, including description, supplier and product code can be found in Table 2.3.

7.4 Results

The histological breakdown of the 143 analyzed samples was: 17 A, 4 MNT LS, 30 AB, 25 B1, 26 B2, 5 B3, 10 CA, 8 NETT and 18 hyperplasias. The mean age of the patients was 59 and the male: female ratio was 1: 1.136. Regarding survival each patient's status was recorded as dead or alive and, given the nature of the cases, months to death from date of surgery were not used for any analysis. Time to last follow up was also missing in the majority of cases since most had had a complete excision of a TET and had not had any regular follow up thereafter.

Each of the following sections deals with each one of these markers separately. The final section integrates the results from all samples and markers. This includes analysis of the co-expression of all the markers that were expressed in TETs as well as cluster analysis of all the samples according to the intensity of the staining for all proteins.

7.4.1 PD-1

PD-1 expression was detected only in the lymphocytic component of 88 of 125 TETs (70%) and in 17 of 18 hyperplasias (94%). Thymomas were positive in 70 of 107 (65%) cases and carcinomas (SCCs and NETTs) only in six of 18 cases (33%). These findings are consistent with the majority of previous studies on PD-1 in TETs, which have also demonstrated an exclusively lymphocytic staining and a higher expression in thymomas compared to carcinomas (84-87,89).

Detailed H-scores for each sample and subgroup comparisons of the H-scores for each subtype as well as the *GTF2I* mutant/ wild-type cases and dead vs. alive status are summarized in Figure 7.1. When scoring the samples it was noted that the exclusively lymphocytic staining exhibited certain distinct patterns. These were therefore noted and classified into the following three groups: Scattered (S), aggregates (A) and scattered aggregates (AS). When assessing the TET subtypes, Type A thymomas were positive in 5/17 cases (29%) with the staining pattern being S in 80% of these. AB thymomas were positive in 29/30 cases (97%) with 45%, 28% and 27% of cases exhibiting an S, A and AS pattern respectively. Of the B1 thymomas 21/24 (88%) were positive with 83% exhibiting an A distribution. Twenty of 26 B2 thymomas were positive with 71% showing an A distribution and 4/5 B3 thymomas were positive with 75% showing an S distribution. Of the carcinomas, only 4/10 (40%) SCCs and 2/8 (25%) NETTs were positive, and neither exhibited a predominant staining pattern. Hyperplasias expressed PD-1 with an impressive 17/18 of cases (94%) being positive and exhibiting the highest mean H-score (56, 95% CI 38-74) with a predominantly A pattern (88% of cases). This could simply reflect the distribution of lymphocytes within the tumour. There are no previous reports on PD-1 staining in thymic hyperplasia. It is worth pointing out that in a single published study (90) that looked into PD-1 staining in normal/ non-neoplastic human thymic tissue, focal PD-1 expression was seen in 55% of cases. This makes the findings of the current study in hyperplasias surprising given the high percentage of cases that were PD-1 positive. It also intriguing that PD-1 expression is high in hyperplasias as its function in normal tissue is to negatively regulates immune responses to prevent autoimmunity, which is however extremely common in thymic hyperplasias that are very commonly associated with MG (55,60). For the same reason however, based on these findings it is logical that thymomas are getting more severe autoimmune toxicities with the administration of anti-PD-1 agents, as the function of PD-1, which is predominantly expressed by these subtypes and whose role is to prevent autoimmunity, is blocked.

The mean H-score value for each subtype was compared and the results are summarized in Table 7.2 and shown schematically in Figure 7.2. Notably, Type A thymomas had a significantly lower mean H-score compared to AB ($p = 0.00057$), B1 ($p = 0.00648$) and B2 ($p = 0.0249$) thymomas and hyperplasias ($p = 0.00079$). Similarly SCCs and NETTs both also had lower mean scores

compares to AB and B1 thymomas and hyperplasias). In the case of SCCs it is likely that the outliers that had very high scores limited the significance of this finding. Although these findings are indeed similar to previous studies (as summarized in table 1.4) it is likely that, as most studies group thymomas together, the exact percentage of positive thymomas depends predominantly on the ratio of type A thymomas to the other subtypes. PD-1 expression has been linked with more indolent TETs in other studies (86), which again with the exception of Type A thymomas, is true in this study. Figure 7.3 demonstrates the low overall staining for PD-1 across TETs but also highlights the difference in distribution of positive cells, showing scattered staining in the A and B1 thymomas but staining of aggregates in the hyperplasias case.

No link between OS and *GTF2I* mutation status (as established in Chapter 4) was found. This is not surprising given that in these 125 TETs both variables are very closely linked to specific TET subtypes, with the majority of A and AB thymomas being *GTF2I* mutant and alive whereas all B3 thymomas, SCCs and NETTs are *GTF2I* wild type and dead.

PD-1 and PD-L1 (epithelial) were co-expressed in 74 of 125 TETs (59%) and five of 18 (28%) hyperplasias. This is predominantly due to less frequent expression of PD-1 overall in these samples. Exceptions are four AB and four B1 thymomas that did not express PD-L1. Only two of these eight did not show any PD-1 staining. In hyperplasias PD-L1 staining was seen only in five of 18 cases hence co-expression was observed only in 28% of cases. PD-1 and PD-L2 (epithelial) were co-expressed in 52 of 125 TETs (41%) and 1 of 18 (6%) hyperplasias. From the carcinomas, only three of ten SCCs (30%) and two of the eight NETTs (25%) co-expressed PD-1 and PD-L1. This was almost identical for PD-1 and PD-L2 co-expression that was seen in four of ten SCCs (40%) and two of the eight NETTs (25%). In both cases this had to do with the lack of PD-1 expression in carcinomas, with a single SCC having no PD-L1 expression but staining strongly for PD-L2.

Figure 7.1. Box plot showing the distribution of the lymphocytic PD-1 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive

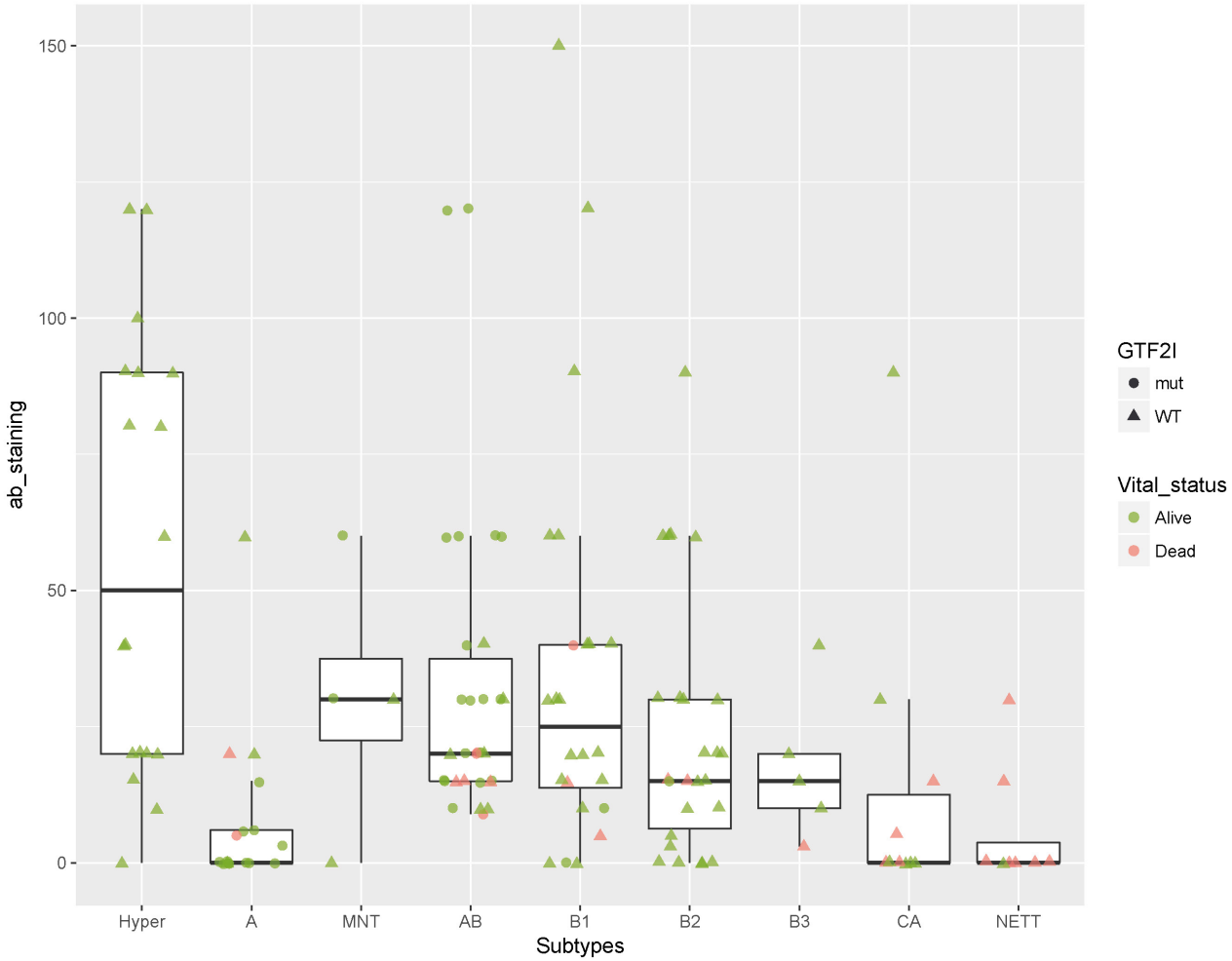


Table 7.2. TET subtype comparison matrix for PD-1 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.00057	-	-	-	-	-	-	-
B1	0.00648	1	-	-	-	-	-	-
B2	0.0249	0.19623	0.31344	-	-	-	-	-
B3	0.11472	0.2955	0.45014	0.94999	-	-	-	-
CA	1	0.0148	0.04449	0.12959	0.21056	-	-	-
Hyper	0.00079	0.11156	0.15625	0.0171	0.11156	0.01467	-	-
MNT	0.14983	0.90833	1	0.61647	0.64631	0.2955	0.44981	-
NETT	0.60311	0.00648	0.01705	0.04802	0.11156	0.65025	0.00648	0.13902

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.2. Heat map of TET subgroup PD-1 H-score comparison

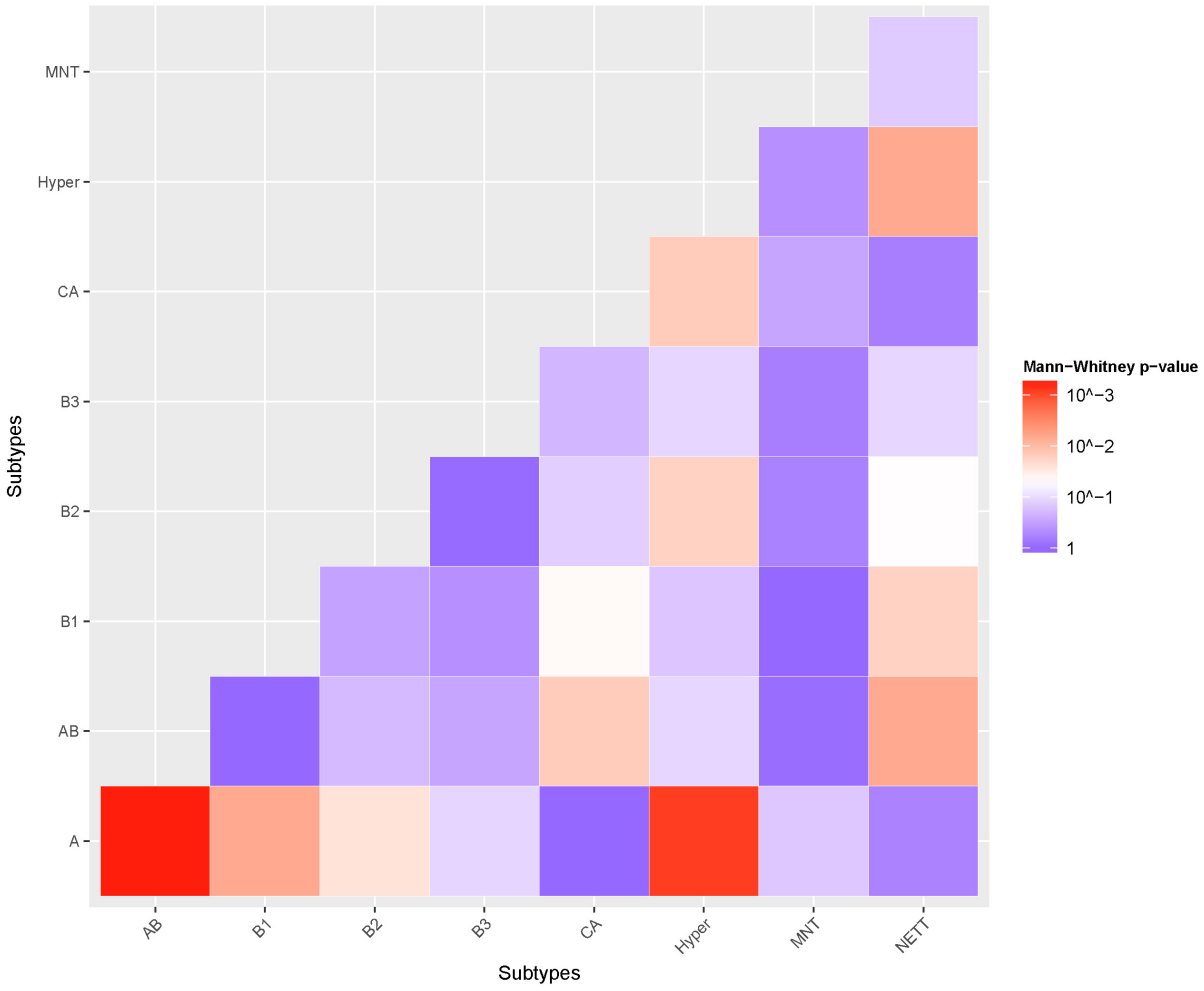
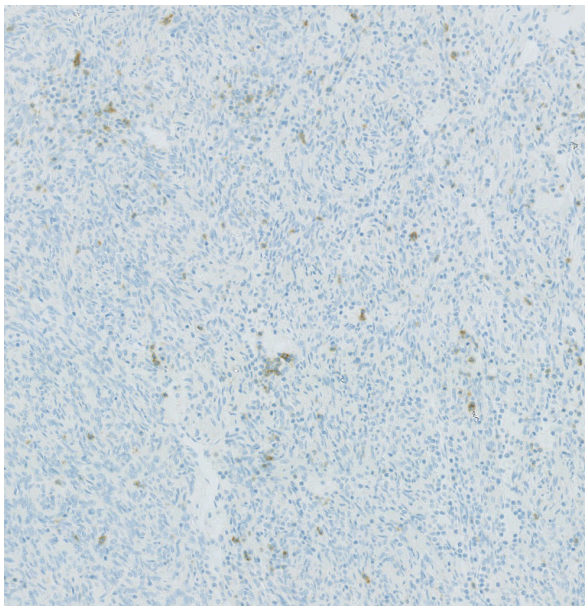
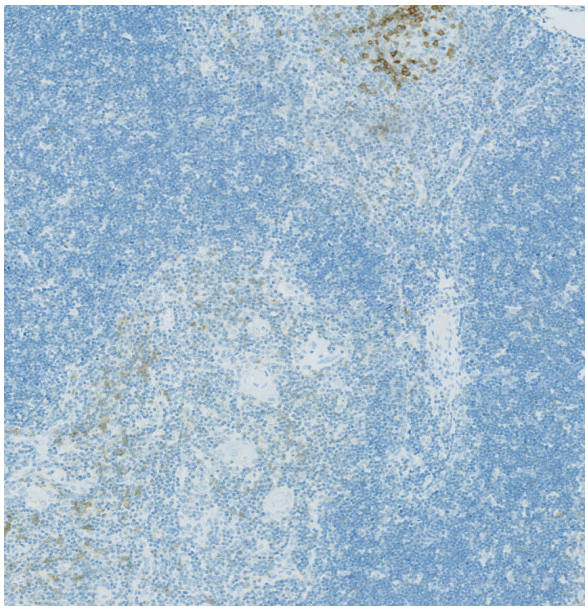


Figure 7.3. Lymphocytic PD-1 staining in TETs

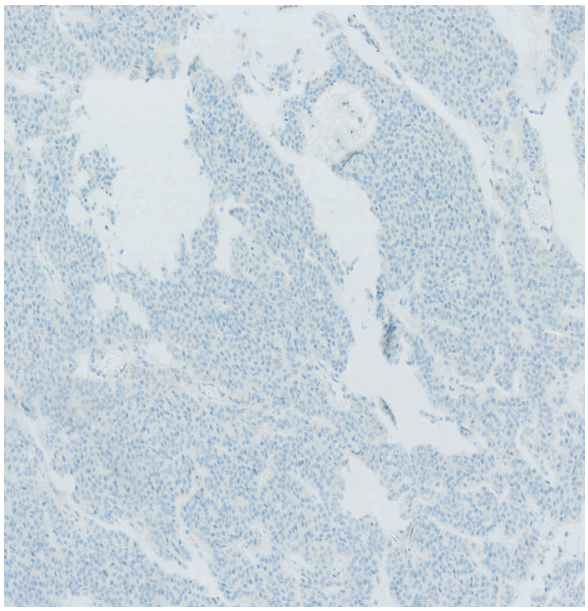
a)



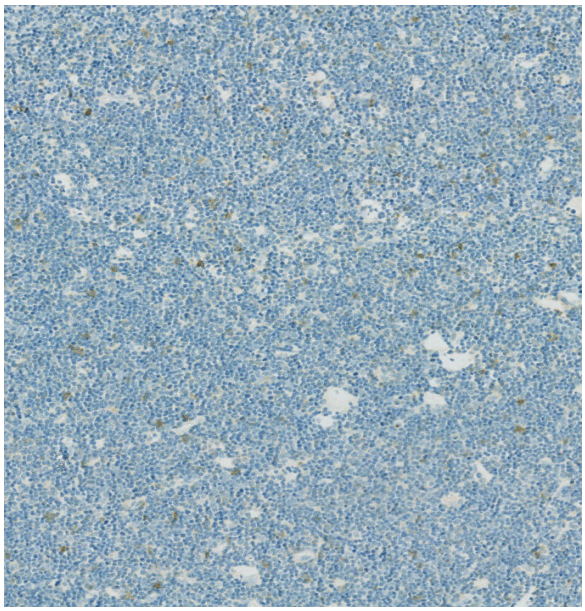
b)



c)



d)



a) A: L H-score: 5 x 3 = 15 S

b) Hyperplasia: L H-score: 20 x 3 = 60 A

c) NETT: H-score 0

d) B1: H score: 50 x 3 = 150 S

7.4.2 PD-L1

PD-L1 is expressed in both lymphoid tissues, including antigen-presenting cells, like dendritic cells, macrophages/monocytes, and B cells and non-lymphoid such as parenchymal cells and vascular endothelial cells in the peripheral tissues. The PD-1 / PD-L1 pathway is crucial in preventing autoimmunity in normal tissues. Tumours expressing PD-L1 can also exploit this pathway by dampening down the immune response against them (60).

In the current study PD-L1 staining was detected in epithelial TET cells and in lymphocytic cells within/ around the tumours. Therefore analysis was performed separately for the 2 components (epithelial and lymphocytic).

PD-L1 was detected in the epithelial component of 104 of 125 TETs (83%) and in five of 18 hyperplasias (28%). Epithelial PD-L1 expression, which is what has been studied as a marker of response to anti-PD-1 therapies was highest in NETTs and lowest in hyperplasias, exhibiting a pattern of increased staining in NETT>B3>B2>CA>B1>AB>A>Hyperplasias. Given the smaller number of samples in some subtypes, NETTs were found to have a significantly higher mean H-score compared to B1, MNT and hyperplasias only and hyperplasias had a significantly lower score compared to all subtypes apart from Type A thymomas and MNT, where significance was not reached.

PD-L1 expression also detected in the lymphocytic component of 73 of 125 TETs (58%), which is higher than previously reported, with other studies having reported lymphocytic staining in TETs in as few as 14.8% of cases (83). Lymphocytic staining was also observed in 16 of 18 hyperplasias (89%). Other studies have not looked at PD-L1 staining of thymic hyperplasias. The significance of the extensive lymphocytic expression of PD-L1 in TETs is unknown. In other cancers, such as head and neck squamous cell carcinoma, PD-L1 staining, in contrast to TETs in the current study, was exclusively either on the tumour cells or the tumour infiltrating lymphocytes. High expression in tumour infiltrating lymphocytes was an independent prognostic factor for increased OS (305).

A breakdown of the percentage of TETs that were positive for PD-L1 in the epithelial and lymphocytic component as well as the proportion of co-expression of epithelial and lymphocytic PD-L1 expression can be found in Table 7.3. Overall more samples exhibited PD-L1 expression in the epithelial component with PD-L1 expression in the lymphocytic component ranging from 100% in MNTs, to 58% in B2s, with, B1s, CAs and ABs having intermediate expression (76%, 70% and 63% respectively). Co-expression in these components was lowest in Type A thymomas (18%) and NETTs (15%). What also stands out was the very small proportion of B3 thymomas and NETTs that were positive for PD-L1 in the lymphocytic component whereas exactly the opposite

was observed in hyperplasias, which were mostly negative for epithelial PD-L1 staining but positive for lymphocytic PD-L1 staining. These are likely to represent the known nature of each TET subtype with some consisting of many more lymphocytes than others. Staining heterogeneity was evident in all TETs, particularly in AB thymomas where even adjacent cores often looked and scored quite differently.

Detailed scores for each sample and subgroup comparisons of the H-scores for each subtype as well as the *GTF2I* mutant/ wild-type cases and dead vs. alive status are summarized in Figure 7.4 for the epithelial component and Figure 7.5 for the lymphocytic component.

The mean H-score value for each subtype was compared and the results are summarized in Table 7.4 and shown schematically in Figure 7.6 for the epithelial component and in Table 7.5 and Figure 7.7 for the lymphocytic. In the epithelial component NETTs had a significantly higher mean H-score compared to A ($p = 0.0224$) and AB ($p = 0.0418$) thymomas and hyperplasias ($p = 0.0029$). B2 thymomas also had a significantly higher shore compared to A (p = 0.0114), AB ($p = 0.0404$) and hyperplasias ($p = 0.0115$). The reason B3 thymomas did not exhibit the same pattern was that, although they had high scores, there were only 5 samples in this cohort. Hyperplasias were completely negative for any epithelial staining in 13 of 18 cases and had a significantly lower H-score compared to all TET subtypes. In terms of the lymphocytic PD-L1 staining the scoring pattern was completely different. Hyperplasias had the highest mean H-score (as well as MNT thymomas), which reached significance only vs. AB ($p = 0.0018$) and B3 thymomas ($p = 0.0396$) and NETTs ($p = 0.0176$). Seven of eight NETTs had no lymphocytic PD-L1 staining at all.

What is perhaps surprising in this study is that all NETTs were positive for epithelial PD-L1 and in fact had the highest mean H-score of all subtypes (mean 190, range 90-300). Results from other studies have been very few and inconsistent with certain studies reporting no staining at all in NETTs(84). One of the reasons for this could be that thymic NETTs encompass a large number of different neuroendocrine tumours and the exact nature of those in other studies is not known. In the current study all NETTs were atypical carcinoids apart from one large cell neuroendocrine tumour of the thymus. Another important point in scoring NETTs is that these demonstrate intense nuclear staining, as opposed to cytoplasmic, which is what is seen in all other TETs. The significance of such staining in unknown and it is possible that inconsistencies between studies have to do with some scoring epithelial cytoplasmic staining only.

Issues complicating comparisons between different studies include not only the antibody clones and kits used for PD-L1 detection (Discussed in detail in Chapter 1.7.3) but also the positivity threshold set by each group when it comes to assessing positive vs. negative PD-L1 expression (Refer to Table 1.2).

Figure 7.8 demonstrates the intense epithelial staining of the more aggressive subtypes, which gets progressively less so in the more indolent TETs. As current trials with anti-PD1 agents are currently recruiting patients with carcinomas only (102,103), or those with no history of autoimmune conditions (101,104), which are much more likely to be B1 and B2 thymomas, the results of the current study are supportive of the this strategy as these are also the cases with the highest PD-L1 expression. Immune-related toxicity remains a major issue and ascertainment of markers that may predict autoimmune toxicity is as important as establishing biomarkers of treatment response.

In the current study there is no link between OS and *GTF2I* mutation status and H-scores in multivariate analysis. This is not surprising given how closely these are linked with specific TET subtypes. This is also consistent with the findings of previous studies on PD-L1 status and OS (82,87).

Table 7.3. Percentage of PD-L1 staining (≥5% of cells) in the epithelial and lymphocytic component of TETs

TET subtypes	PD-L1 E positive	PD-L1 L positive	PD-L1 E & L +ve	TOTAL
A	11 (65%)	7 (41%)	3 (18%)	17
MNT	2 (50%)	4 (100%)	2 (50%)	4
AB	24 (80%)	19 (63%)	17 (57%)	30
B1	19 (76%)	19 (76%)	15 (60%)	25
B2	26 (100%)	15 (58%)	15 (58%)	26
B3	5 (100%)	1 (20%)	1 (20%)	5
CA	9 (90%)	7 (70%)	6 (60%)	10
NETT	8 (100%)	1 (13%)	1 (13%)	8
TET TOTAL	***104* (83%)**	***73 (58%)***	***60 (48%)***	***125***
Hyperplasia	5 (28%)	16 (89%)	3 (17%)	18

E = Epithelial, L = Lymphocytic. Percentage of PD-L1 positive TETs was identical when using 1% or 5% as the positivity threshold as there were no samples that were positive in over 1% but fewer than 5% of cells.

Figure 7.4. Box plot showing the distribution of the epithelial PD-L1 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive

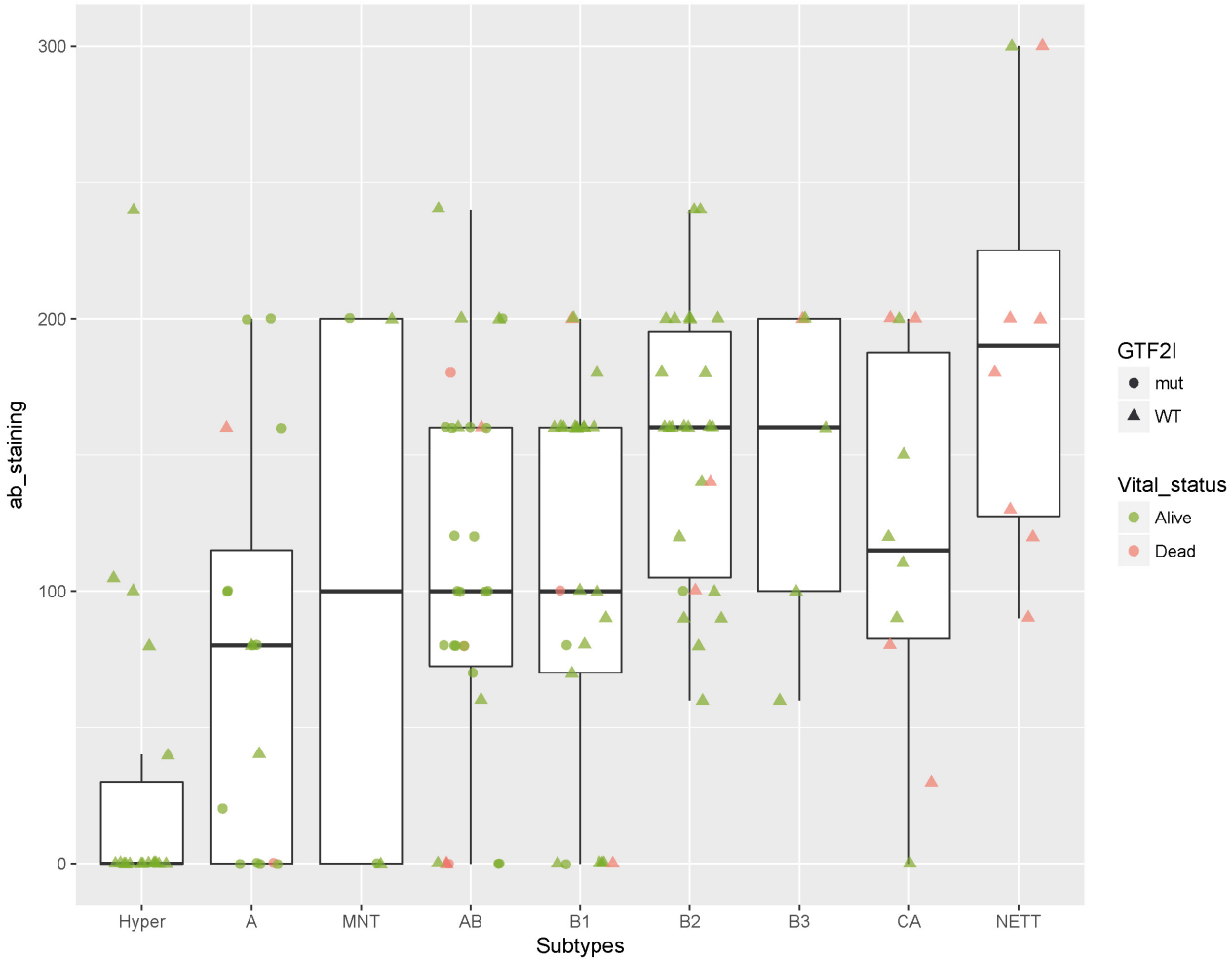


Figure 7.5. Box plot showing the distribution of the lymphocytic PD-L1 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive

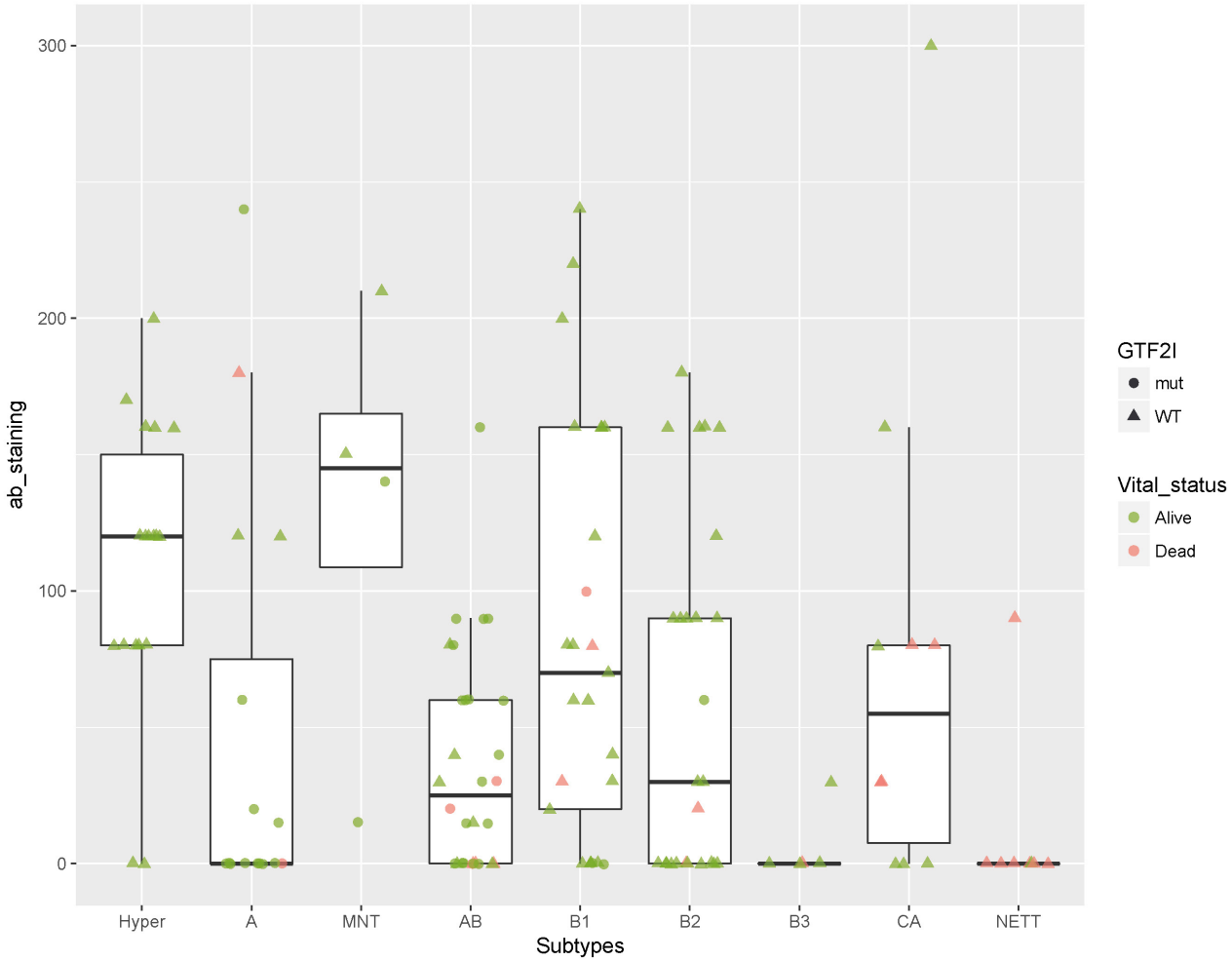


Table 7.4. TET subtype comparison matrix for epithelial PD-L1 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.4375	-	-	-	-	-	-	-
B1	0.4375	0.9392	-	-	-	-	-	-
B2	0.0114	0.0404	0.0686	-	-	-	-	-
B3	0.2214	0.4375	0.4375	0.9392	-	-	-	-
CA	0.3208	0.7662	0.8151	0.4216	0.7662	-	-	-
Hyper	0.091	0.0093	0.0114	8.50E-05	0.0203	0.0115	-	-
MNT	0.9027	0.9784	0.9027	0.7703	0.8398	0.9027	0.4375	-
NETT	0.0224	0.0418	0.0589	0.4375	0.5956	0.2308	0.0029	0.5098

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.6. Heat map of TET subgroup epithelial PD-L1 H-score comparison

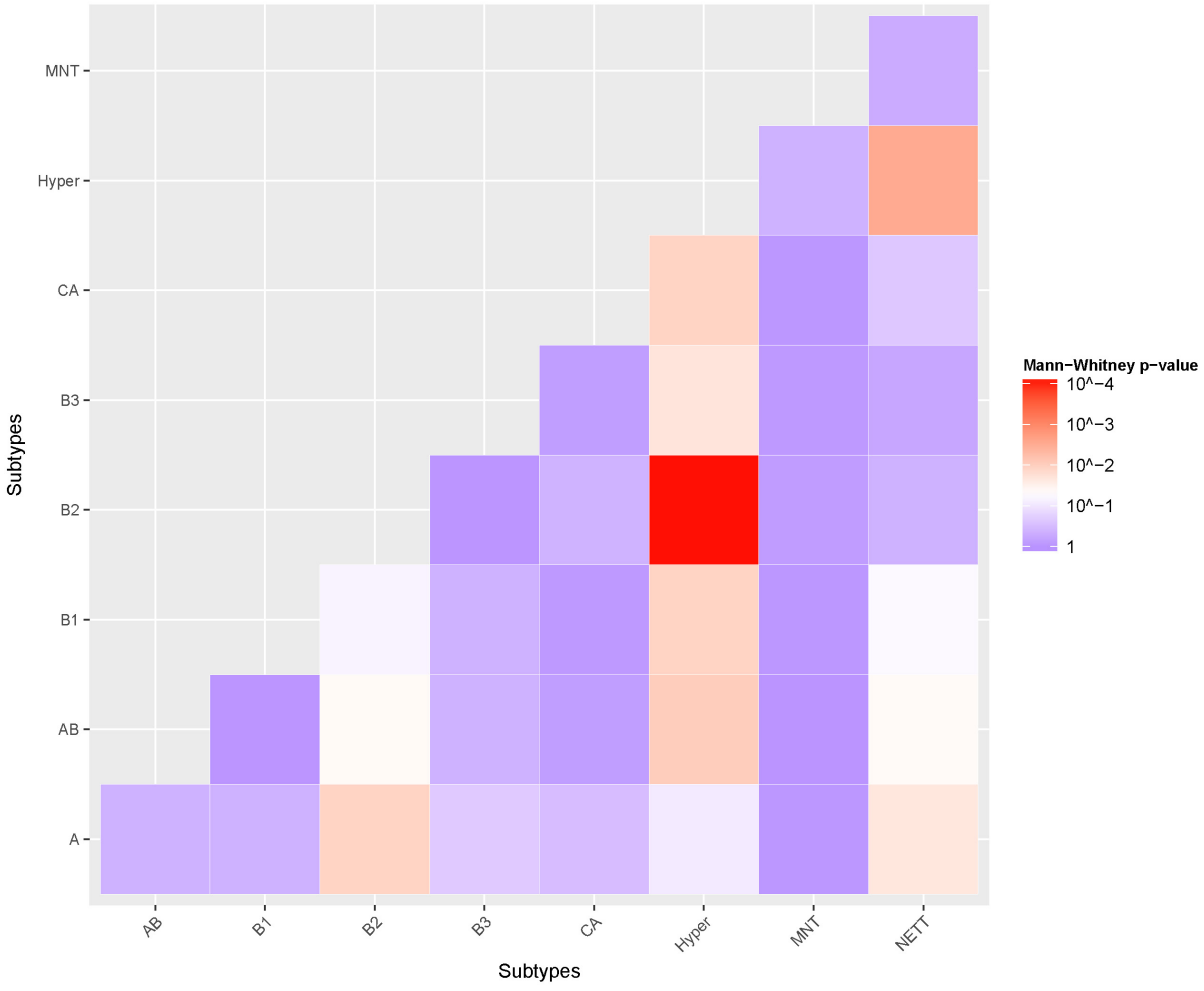


Table 7.5. TET subtype comparison matrix for lymphocytic PD-L1 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.6319	-	-	-	-	-	-	-
B1	0.1356	0.0889	-	-	-	-	-	-
B2	0.4880	0.4305	0.3743	-	-	-	-	-
B3	0.4126	0.1448	0.0875	0.1643	-	-	-	-
CA	0.3651	0.3787	0.7242	0.8501	0.1331	-	-	-
Hyper	0.0556	0.0018	0.1961	0.0889	0.0396	0.1448	-	-
MNT	0.1137	0.0971	0.4305	0.1648	0.0971	0.3787	0.4769	-
NETT	0.2259	0.109	0.048	0.1077	0.9074	0.109	0.0176	0.048

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.7. Heat map of TET subgroup lymphocytic PD-L1 H-score comparison

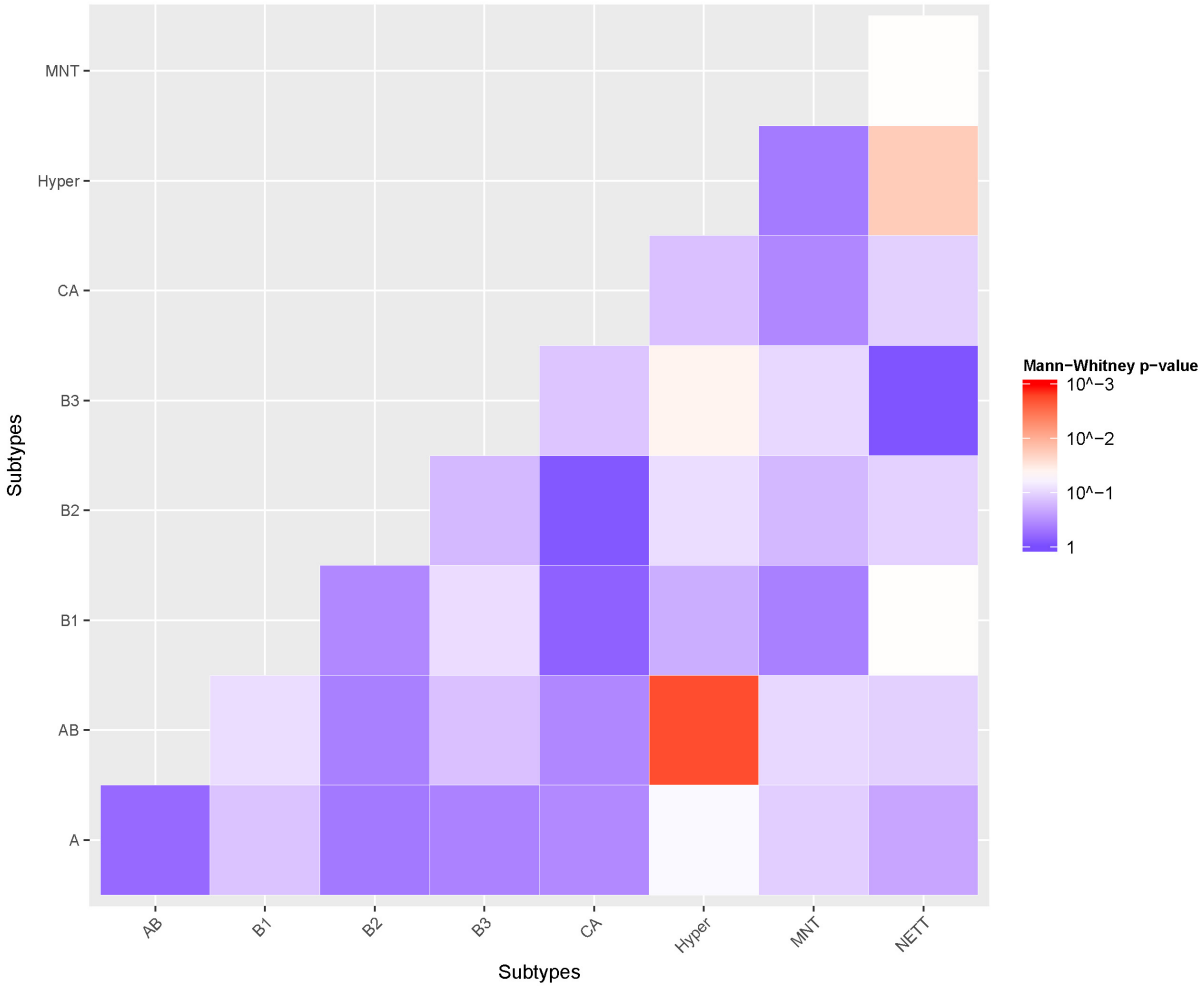
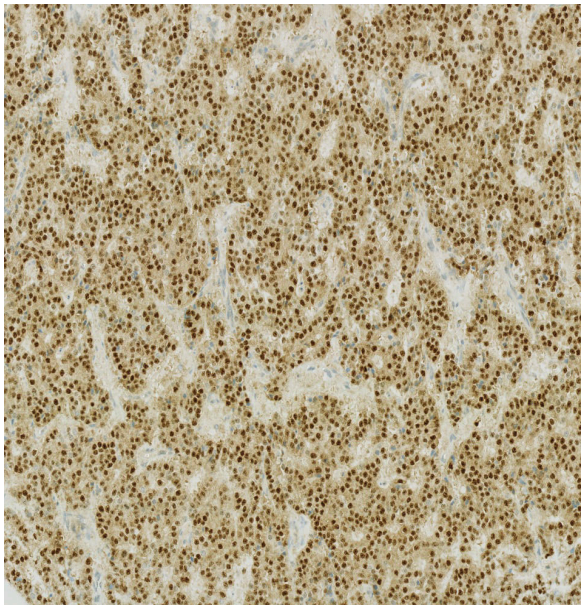
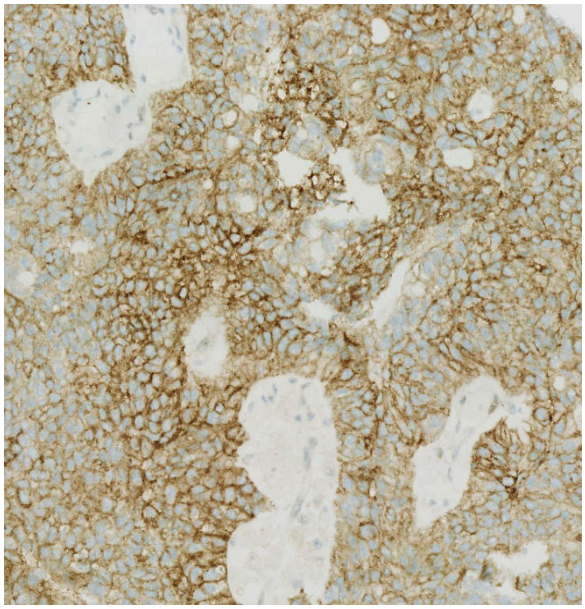


Figure 7.8. PD-L1 staining in TETs

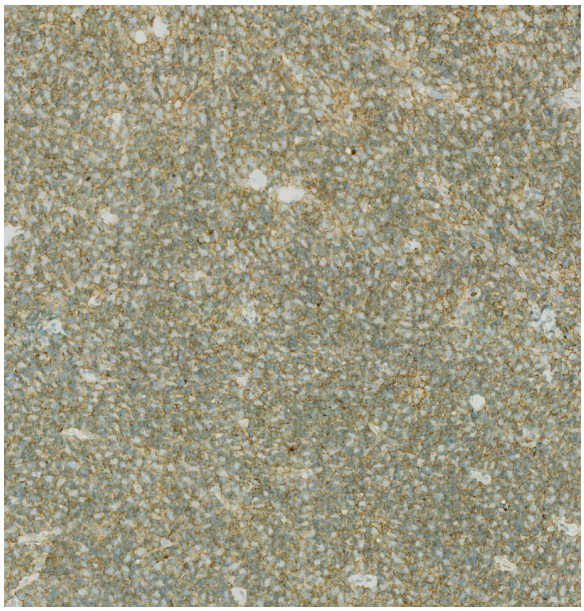
a)



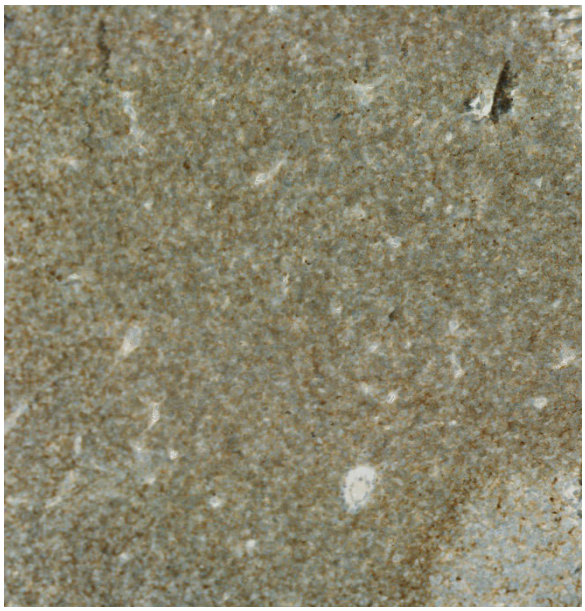
b)



c)



d)



a) NETT: E H-score: 100 x 3 = 300, L H-score: 0

b) CA: E H-score: 100 x 2 = 200, L H-score: 0

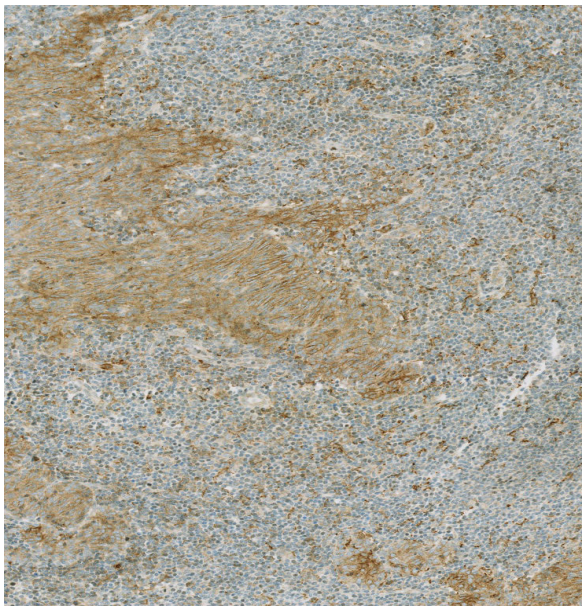
c) B2: E H-score: 100 x 2 = 200, L H-score: 0

d) **B1:** E H-score: 80 x 2 = 160, L H-score: 80 x 2 = 160

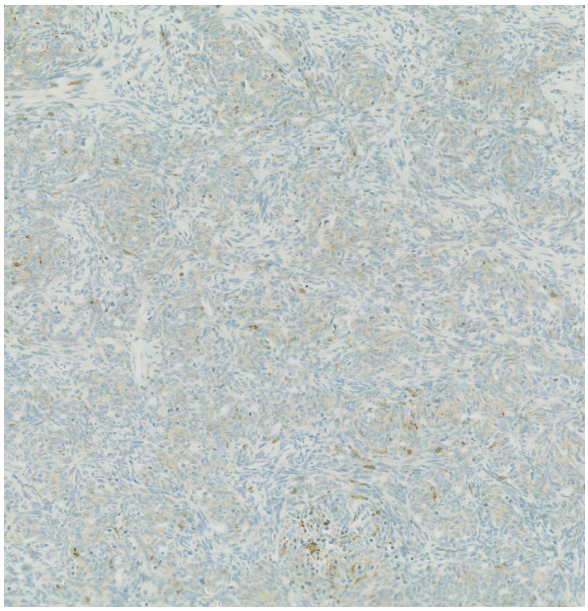
e)



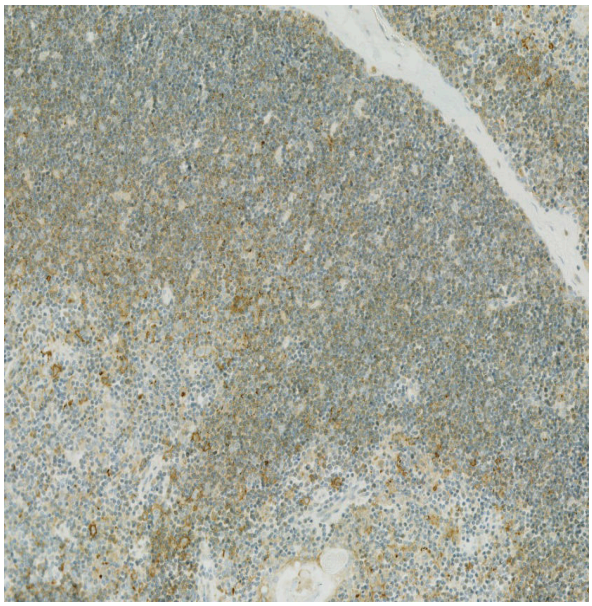
f)



g)



h)



e) AB: E H-score: 40 x 2 = 80, L H-score: 10 x 3 = 30

f) MNT: E H-score: 100 x 2 = 200, L H-score: (30 x 3) + (30 x 2) = 150

g) A: E H-score: 100 x 1 = 100, L H-score: 0

h) Hyperplasia: E H-score: 0, L H-score: 80 x 2 = 160

E = Epithelial

L = Lymphocytic

7.4.3 PD-L2

In this study PD-L2, like PD-L1, was detected in epithelial TET cells and in lymphocytic cells within/ around the tumours. Therefore analysis was performed separately for the 2 components (epithelial and lymphocytic). PD-L2 was detected in the epithelia of 83 of 125 TETs (66%) and in 1 of 18 hyperplasias (6%) and the lymphocytic component of 89 of 125 TETs (71%) and in 18 of 18 hyperplasias (100%).

A breakdown of the percentage of TETs that were positive for PD-L2 in the epithelial and lymphocytic component as well as the proportion of tumours that co-expressed PD-L2 in epithelial and lymphocytic components can be found in Table 7.6. Detailed scores for each sample and subgroup comparisons of the H-scores for each subtype as well as the *GTF2I* mutant/ wild-type cases and dead vs. alive status are summarized in Figure 7.9 for the epithelial component and Figure 7.10 for the lymphocytic component. Overall PD-L2 expression was higher in the lymphocytic component vs. the epithelial component in all TET subtypes apart from B3, SCCs and NETTs where no lymphocytic staining was detected. Although 59% of Type A thymomas exhibited PD-L2 staining by epithelial cells and 66% by lymphocytes, only 29% of cases were positive for both. Similar results were observed in type AB thymomas where 40% and 87% of cases were positive for PD-L2 staining in the epithelial and lymphocytic components respectively but expression in both components was only observed in 27% of cases. In B1 and B2 thymomas both components were positive in 48% and 77% of cases respectively, which is very close to the percentage these are positive in the epithelial component (52% and 81% respectively). In the hyperplasias only 1 case was positive for PD-L2 in the epithelium and all 18 were positive in the lymphocytes. Although as discussed for PD-L1 certain thymoma subtypes are characterized by their lymphocytic component and it is not clear why PD-L2 would be more abundant in this component than in the epithelial component. It is also not clear why in the case of A and AB thymomas, there were many tumours that expressed PD-L2 in only one of the components. This observation has not been described before.

The mean H-score value for each subtype was compared and the results are summarized in Table 7.8 and shown schematically in Figure 7.11 for the epithelial component and in Table 7.9 and Figure 7.12 for the lymphocytic component. In the epithelial component NETTs and SCCs had significantly higher scores compared to all other subtypes, excluding B3s and in the case of SCCs, with the exception of B2 thymomas. This was the same pattern exhibited by the PD-L1 staining. As PD-L2 expression in NETTs is not well documented it is also very important to note how in this study, as was the case for PD-L1, PD-L2 had the highest epithelial mean H-score (172.5) compared to all other subtypes. At the other end of the spectrum, AB thymomas and hyperplasias had significantly lower scores compared to all TET subtypes (MNT and B1 did not reach

significance vs. AB). There was only a single outlier in the case of the hyperplasias with none of the others staining for PD-L2 in the epithelial component. The pattern of PD-L2 staining of the lymphocytic component was entirely different resembling that of lymphocytic PD-L1 expression with the exception of B3 thymomas, SCCs and NETTs where it was not expressed.

Although in most studies in other tumour types it has been shown that PD-L2 expression is overall lower than that of PD-L1(80) this was not the case with TETs where the H-scores did not differ significantly when those of PD-L1 and PD-L2 were compared. Co-expression of PD-1 PD-L1 and PD-L2 is summarized in Table 7.7. The epithelial PD-L1 and PD-L2 staining of TETs was used for this comparison only.

In regards to PD-L1 and PD-L2 all TETs showed co-expression in 57% of cases. It increased from the less aggressive to the more aggressive subtypes, being as low as 47% in Type A thymomas, 37% in ABs, and 36% in B1s to 81% in B2s, 80% in B3s, 90% in SCCs and 100% in NETTs. Despite a large number of samples co-expressing both, the finding that certain TETs do not is consistent with findings from other tumour types where PD-L2 was seen in a number of cases when there was no PD-L1 expression (81). This could explain why in other tumour types certain cases have responded to anti-PD-1 therapies in the absence of PD-L1 staining. Response to pembrolizumab has however been shown to be higher in cases that were positive for both PD-L1 and PD-L2 compared to either alone. This is very important when it comes to determining who are the patients that are more likely to benefit from anti-PD1 treatment.

Although trial information on the significance of PD-1, PD-L1 and PD-L2 co-expression is not available, examining this as a possible determinant of the efficacy of anti-PD-1 therapies is not unreasonable. In the current study co-expression of all three was observed in only 49 TETs (39%). Only the epithelial staining of PD-L1 and PD-L2 was included in this analysis. Co-expression was lowest in NETTs and MNTs (25% for each), followed by A thymomas (29%), SCCs (30%), B1 (32%), AB (37%), B2 (58%) and finally B3 thymomas (80%).

As PD-L2 has been implicated in the development of autoimmune pneumonitis (80), which is one of the many immune related toxicities experienced by patients on anti-PD-1 therapies, it is certainly worth investigating its role in TETs further. What is known so far from on-going clinical trials in TET patients with anti-PD-1 agents is that those with thymomas get more severe autoimmune toxicities compared to carcinomas. It would be very interesting and important to establish whether that might be related, for example, to lymphocytic PD-L2 staining and PD-1 expression, as both are significantly higher in these subtypes in the current study.

No link between OS or *GTF2I* mutations status and epithelial or lymphocytic H-scores was found (multivariate analysis). Examples of PD-L2 staining in a range of TETs can be found in Figure 7.13.

Table 7.6. Percentage of PD-L2 staining (≥5% of cells) in the epithelial and lymphocytic component of TETs

TET subtypes	PD-L2 E positive	PD-L2 L positive	PD-L2 E & L +ve	TOTAL
A	10 (59%)	11 (65%)	5 (29%)	17
MNT	3 (75%)	4 (100%)	3 (75%)	4
AB	12 (40%)	26 (87%)	8 (27%)	30
B1	13 (52%)	23 (92%)	12 (48%)	25
B2	21 (81%)	24 (92%)	20 (77%)	26
B3	4 (80%)	1 (20%)	0 (0%)	5
CA	10 (100%)	0 (0%)	0 (0%)	10
NETT	8 (100%)	0 (0%)	0 (0%)	8
TET TOTAL	***83 (66%)***	***89 (71%)***	***48 (38%)***	***125***
Hyperplasia	1 (6%)	18 (100%)	1 (6%)	18

E = Epithelial, L = Lymphocytic. Percentage of PD-L1 positive TETs was identical when using 1% or 5% as the positivity threshold as there were no samples that were positive in over 1% but fewer than 5% of cells.

Table 7.7. PD-L1 and PD-L2 co-expression in TET epithelia

TET subtypes	PD-L1 (E) +ve	PD-L2 (E) +ve	PD-1 +ve	PD-L1 and PD-L2 +ve	PD-1, PD-L1 and PD-L2 +ve	TOTAL
A	11 (65%)	10 (59%)	8 (47%)	8 (47%)	5 (29%)	17
MNT	2 (50%)	3 (75%)	3 (75%)	1 (25%)	1 (25%)	4
AB	24 (80%)	12 (40%)	30 (100%)	11 (37%)	11 (37%)	30
B1	19 (76%)	13 (52%)	21 (84%)	9 (36%)	8 (32%)	25
B2	26 (100%)	21 (81%)	21 (81%)	21 (81%)	15 (58%)	26
B3	5 (100%)	4 (80%)	5 (100%)	4 (80%)	4 (80%)	5
CA	9 (90%)	10 (100%)	4 (40%)	9 (90%)	3 (30%)	10
NETT	8 (100%)	8 (100%)	2 (25%)	8 (100%)	2 (25%)	8
TET TOTAL	***104* (83%)**	***83 (66%)***	**94 (%)**	**71 (57%)**	***49 (39%)***	***125***
Hyperplasia	5 (28%)	1 (6%)	17 (94%)	0 (0%)	0 (0%)	18

E = Epithelial

Figure 7.9. Box plot showing the distribution of epithelial PD-L2 H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive

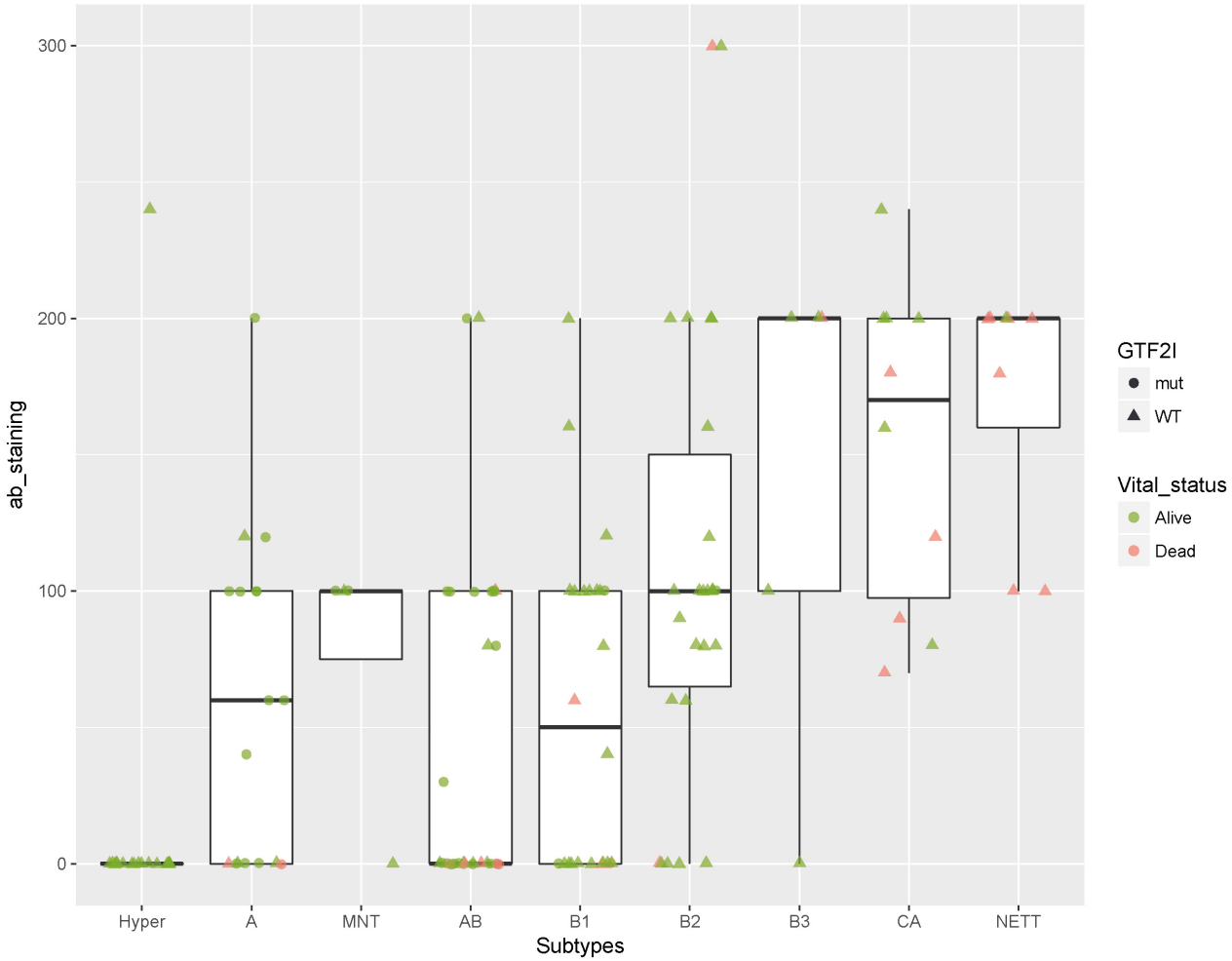


Figure 7.10. Box plot showing the distribution of lymphocytic PD-L2 H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive

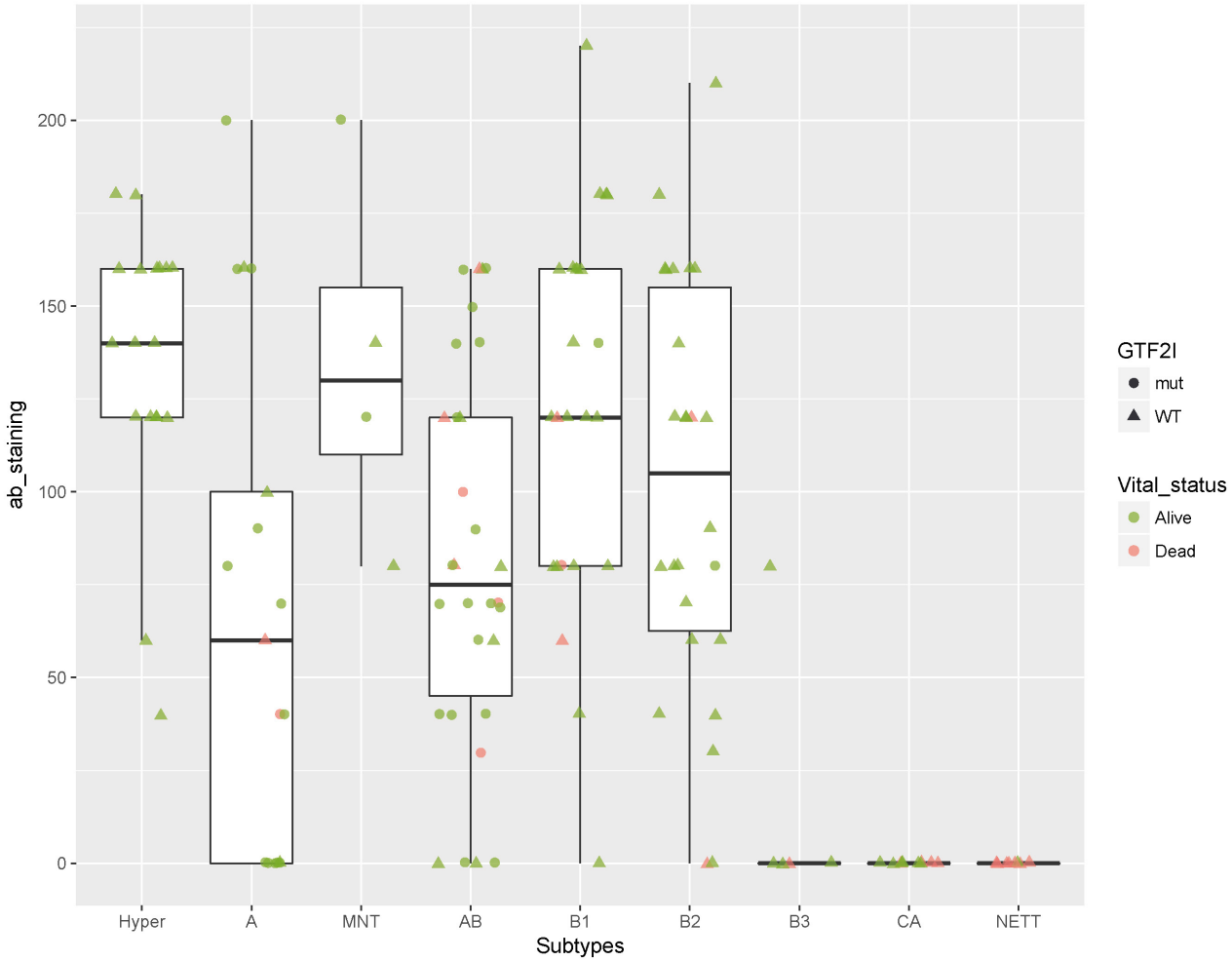


Table 7.8. TET subtype comparison matrix for epithelial PD-L2 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.33362	-	-	-	-	-	-	-
B1	0.86895	0.39922	-	-	-	-	-	-
B2	0.1310	0.00767	0.07486	-	-	-	-	-
B3	0.12786	0.04402	0.08167	0.41996	-	-	-	-
CA	0.00883	0.002	0.00651	0.1892	1	-	-	-
Hyper	0.00767	0.04387	0.00869	3.50E-04	0.00605	0.00035	-	-
MNT	0.71815	0.2717	0.5868	0.78004	0.33043	0.20114	0.0091	-
NETT	0.00605	0.00092	0.0022	0.04776	0.78004	0.51461	0.00036	0.04674

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.11. Heat map of TET subgroup epithelial PD-L2 H-score comparison

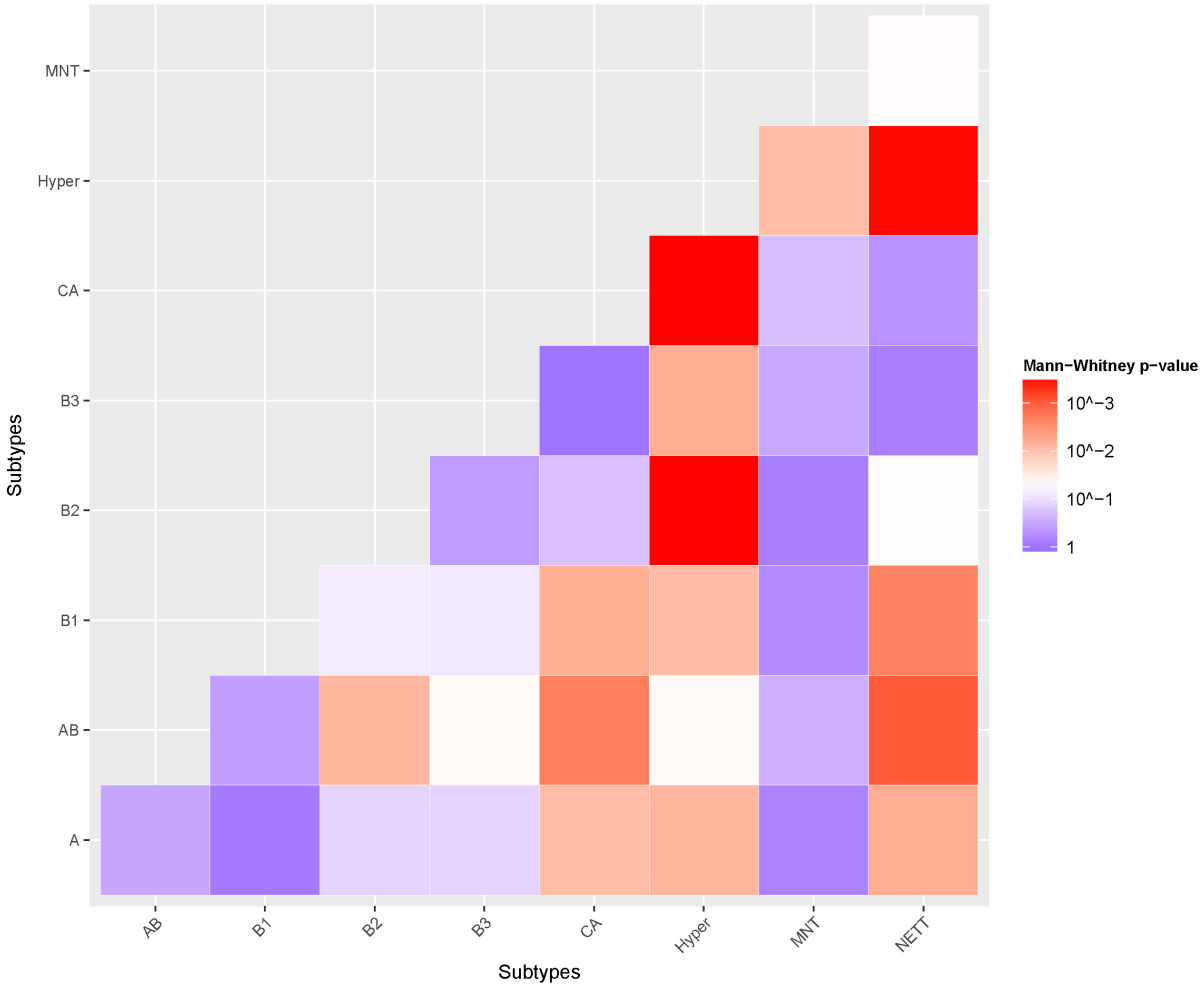


Table 7.9. TET subtype comparison matrix for lymphocytic PD-L2 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.41997	-	-	-	-	-	-	-
B1	0.02061	0.01284	-	-	-	-	-	-
B2	0.130	0.24511	0.21921	-	-	-	-	-
B3	0.14536	0.03048	0.00592	0.01268	-	-	-	-
CA	0.00563	0.0003	0.00017	0.00019	0.24511	-	-	-
Hyper	0.01044	0.0037	0.46385	0.0574	0.00402	0.00017	-	-
MNT	0.13015	0.13015	0.78706	0.34979	0.03453	0.00215	0.82754	-
NETT	0.01142	0.00071	0.0003	0.0004	0.31311	- *	0.0003	0.00506

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

*Note that in the NETT vs. CA comparison there is no p-value because both subtypes only had '0' values, which in effect means there is no difference at all (p = 1).

Figure 7.12. Heat map of TET subgroup lymphocytic PD-L2 H-score comparison

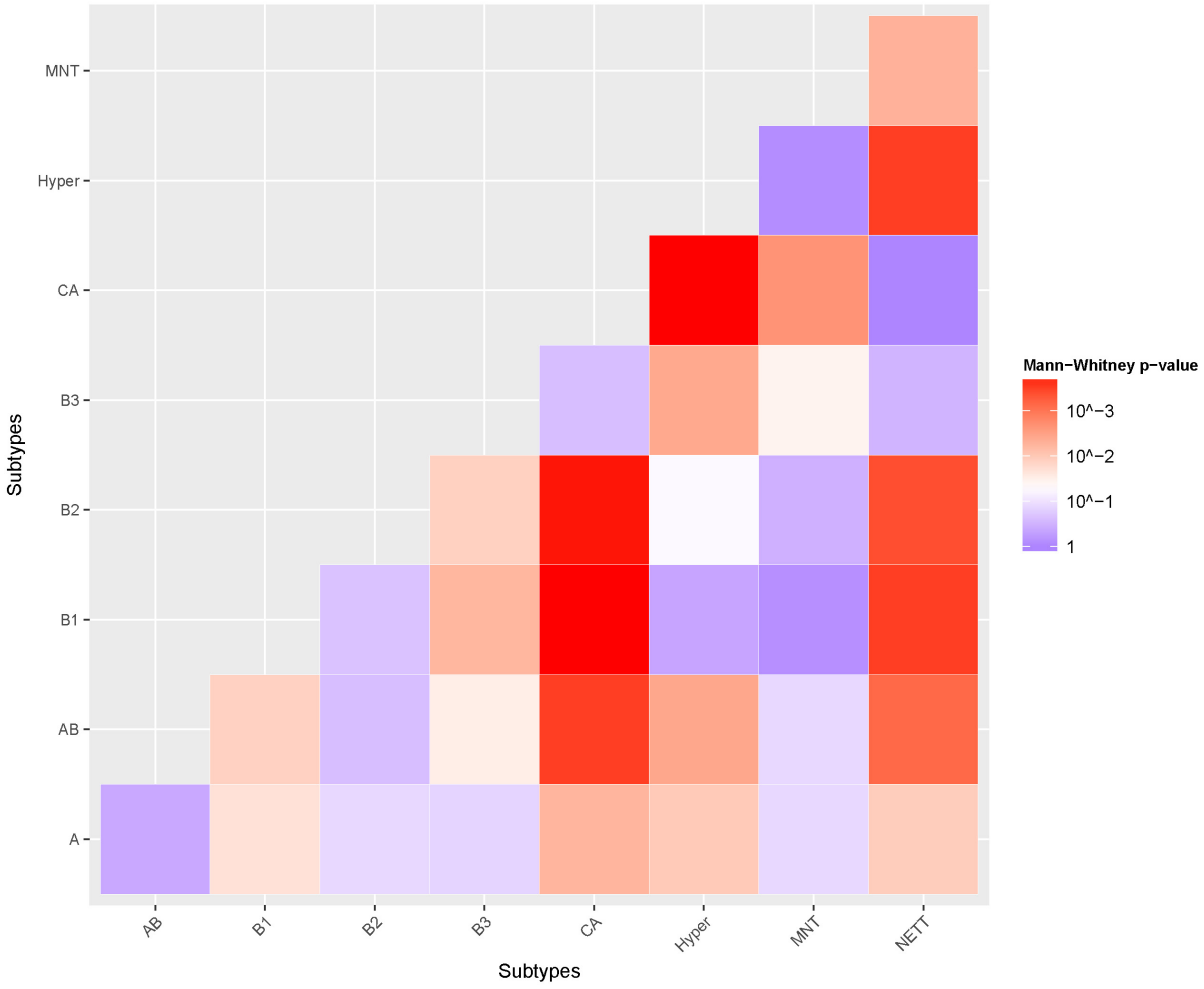
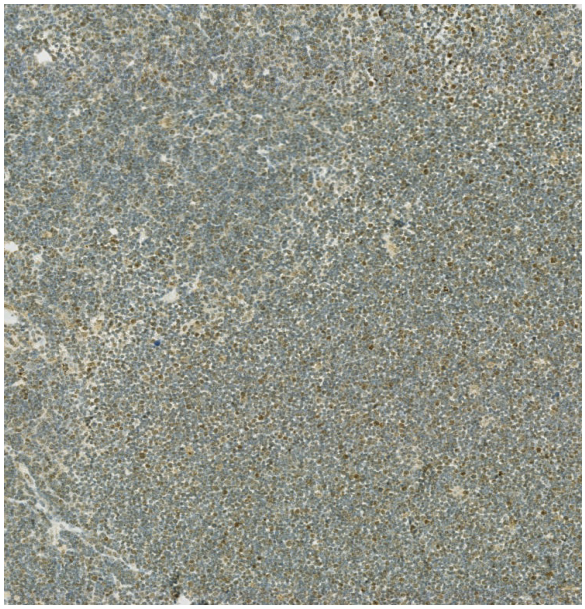


Figure 7.13. PD-L2 staining in TETs

a)



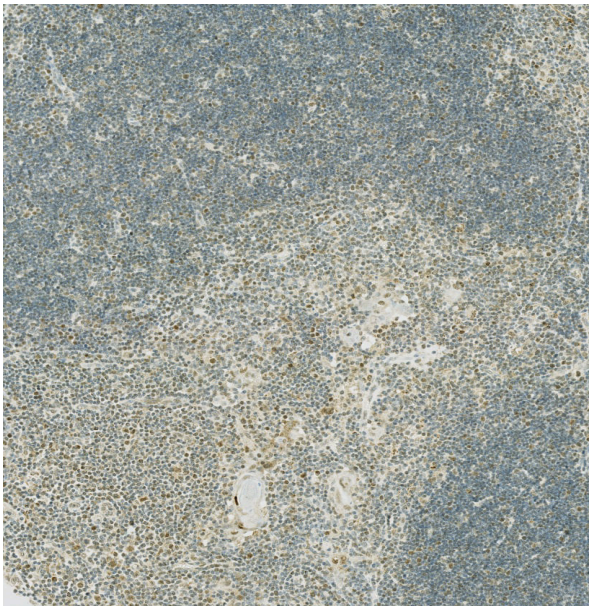
b)



c)



d)



a) NETT: E H-score: 100 x 2 = 200, L H-score: 0

b) B1: E H-score: 60 x 2 = 120, L H-score: 60 x 2 = 120

c) A: E H-score: 0, L H-score: 40 x 1 = 40

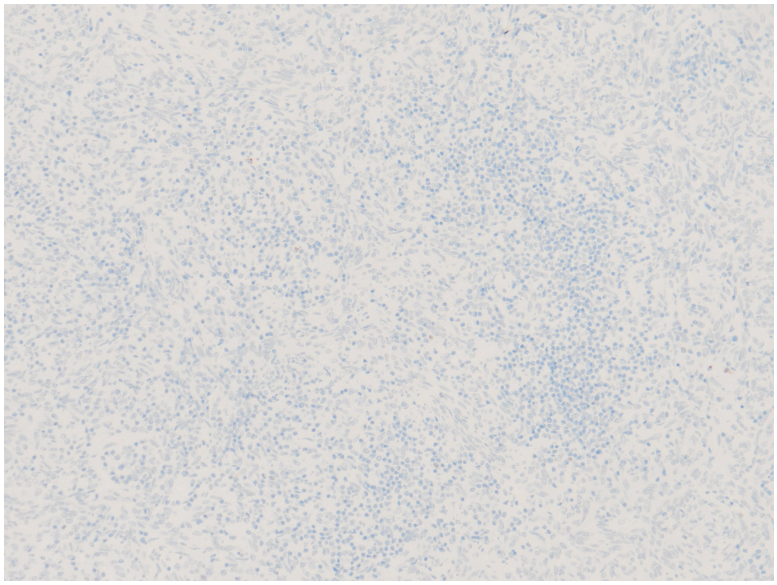
d) **Hyperplasia:** E H-score: 0, L H-score: 70 x 2 = 140

7.4.4 LAG-3

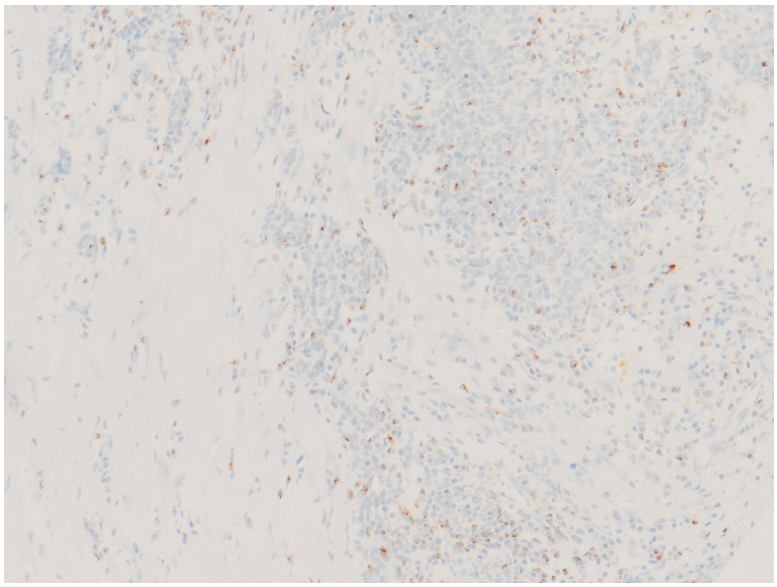
LAG-3 was negative across all TET subtypes (Figure 7.14). Epithelial staining as anticipated was zero in all cases and lymphocytic staining was minimal (H-score 20 x 1 = 20) in 1 of each of the following: A, MNT, B3 and CA. Further analyses were therefore not performed.

Figure 7.14. LAG-3 staining in TETs

a)



b)



a) Representative LAG-3 TET staining across all subtypes

b) A rare case of minimal lymphocytic LAG-3 staining (10 x 2 = 20) in a Type A thymoma

7.4.5 TIM-3

TIM-3 was detected in the epithelia of 18 of 125 TETs (14%) and in 0 of 18 hyperplasias (0%). These cases were found across the spectrum of all TET subtypes although the percentage was much higher in SCCs (4/10; 40%) and NETTs (6/8; 75%) vs. thymomas (9/107; 8%). The breakdown of TIM-3 staining in thymomas was the following: 1/17 A, 1/4 MNT, 2/30 AB, 2/25 B1, 2/26 B2 and 1/5 B3. All results are summarized in Figure 7.15. Of the nine thymomas that expressed epithelial TIM-3, six showed lymphocytic staining of TIM-3 (67%). Of the four SCCs with epithelial staining only two showed lymphocytic staining (50%) and none of the six NETTs did (0%). In a similar study in NSCLC, 18 of 132 (13.6%) of samples expressed TIM-3 (289). In that study TIM-3 was exclusively expressed by tumour infiltrating lymphocytes and not by the tumour cells and were significant differences in adenocarcinomas vs. non-adenocarcinomas despite the small numbers. In the same study there was a significant correlation between TIM-3 and PD-1 expression. In the current study all cases that were TIM-3 positive also expressed PD-1 apart from one of the four SCCs and four of the six NETTs. In another study TIM-3 expression in NSCLC was higher (17% in SCCs and 34% in adenocarcinomas) (290). In a mesothelioma study however TIM-3 was expressed by tumour cells as well as by lymphocytes and plasma cells (291). The only other study on TIM-3 expression in TETs has shown very different results where it was expressed in 18 of 24 cases (75%) of B2/3 thymomas (12) and SCCs (12) showed high TIM-3 expression (91). An immediate explanation for this is not obvious although the choice of TET subtypes in that study may have influenced the result. There could also have been confounders such as the type of antibody and the reagents/kit that were used.

The mean H-score value for each subtype was compared and the results are summarized in Table 7.10 and shown schematically in Figure 7.16. This score was significantly higher for NETTs compared to all thymomas, with the exception of MNT and B3, whose very small number of cases (four and five respectively) did not allow statistical significance to be reached. No link between OS or *GTF2I* mutations status and epithelial H-scores was found (multivariate analysis). Examples of TIM-3 staining in a range of TETs can be found in Figure 7.17.

Given these results and the observation that in preclinical as well as in mouse models, inhibition of TIM-3 with or without PD-1 inhibition appears to be effective, further research into TIM-3 in TETs is warranted. A phase I trial with an anti-TIM-3 monoclonal antibody with/without anti-PD-1 inhibition in patients with advanced solid tumours is already underway. Based on the results of the current study, evaluation of the TIM-3 status of patients with thymic SCCs and NETTs is warranted in case they can also benefit from trials with anti-TIM-3 therapies.

Figure 7.15. Box plot showing the distribution of TIM-3 epithelial H-scores per TET subtype, including GTF2I mutation status and whether dead or alive

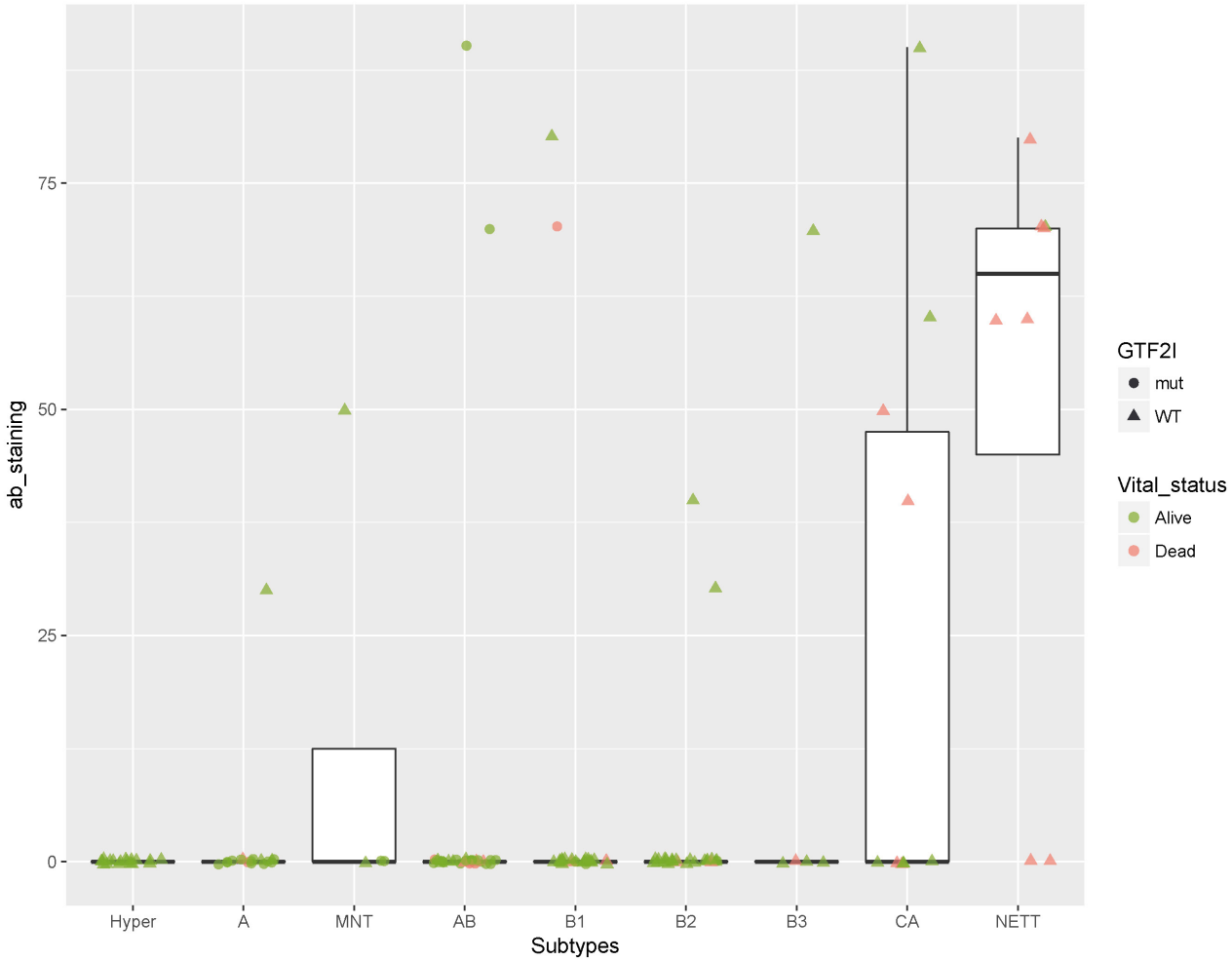


Table 7.10. TET subtype comparison matrix for epithelial TIM-3 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.9736	-	-	-	-	-	-	-
B1	0.929	0.9736	-	-	-	-	-	-
B2	0.955E	0.9833	0.9736	-	-	-	-	-
B3	0.5419	0.5504	0.6495	0.5504	-	-	-	-
CA	0.0954	0.0816	0.1335	0.0738	0.7702	-	-	-
Hyper	0.5419	0.5072	0.5035	0.504	0.2025	0.031	-	-
MNT	0.5035	0.5179	0.5572	0.5035	1	0.7702	0.1477	-
NETT	0.0024	0.0011	0.0027	0.001	0.2816	0.216	0.001	0.1858

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.16. Heat map of TET subgroup epithelial TIM-3 H-score comparison

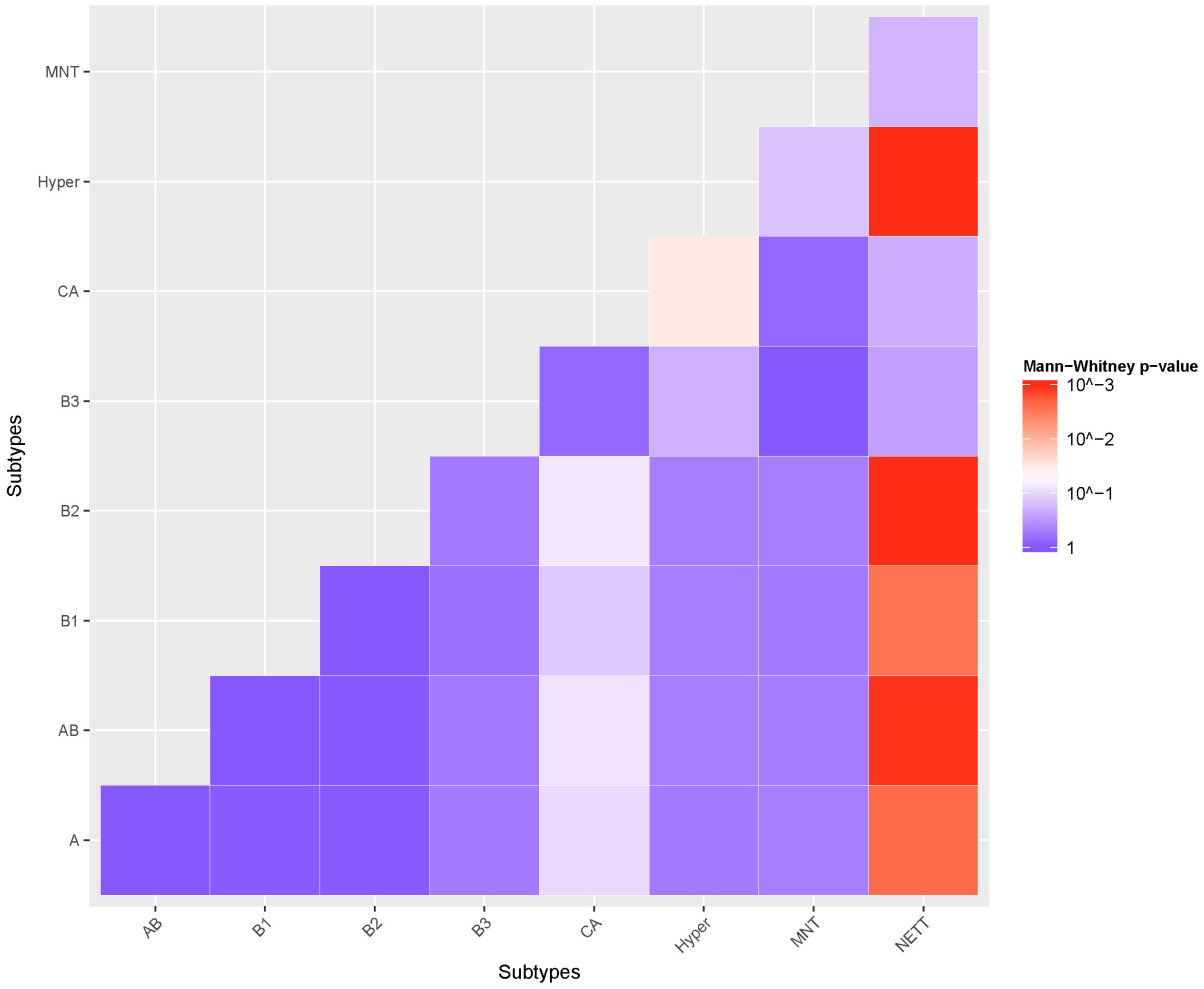
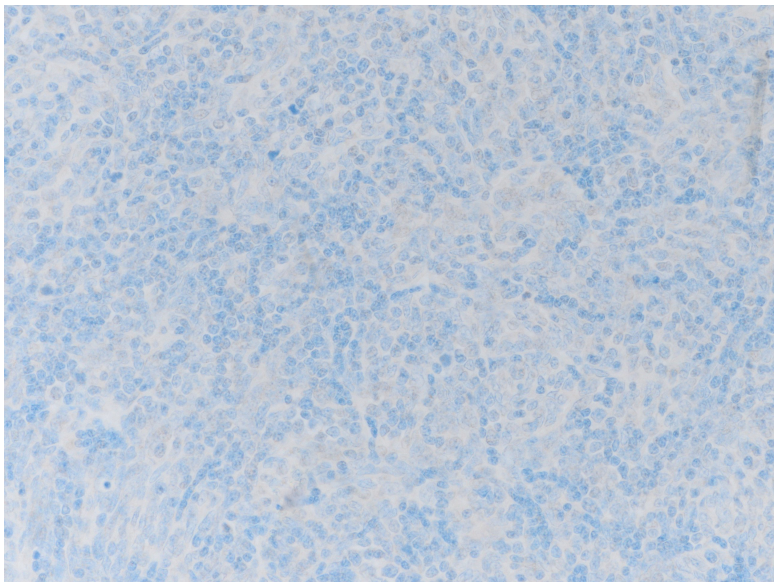
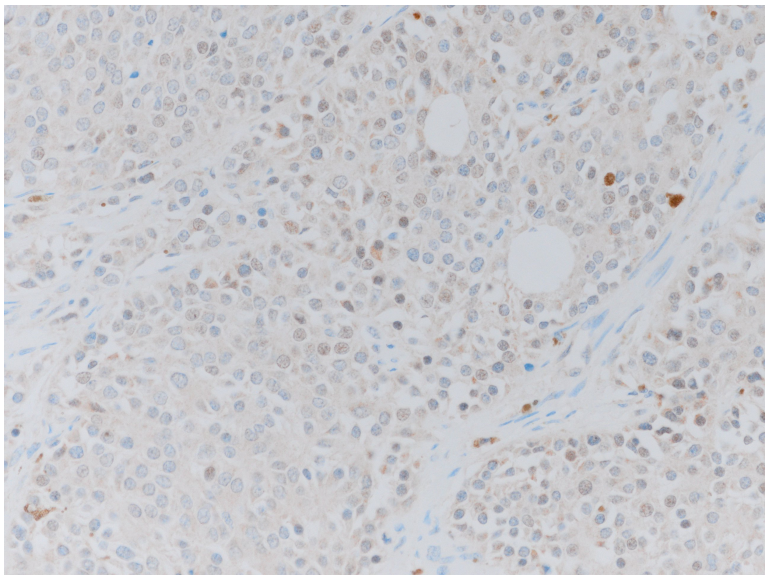


Figure 7.17. Epithelial TIM-3 staining in TETs

a)



b)



a) Representative epithelial TIM-3 staining across all TET subtypes

b) Epithelial TIM-3 staining of a representative thymic NETT (H-score = 70 x 1 = 70)

7.4.6 B7-H3

B7-H3 was detected in the epithelial component of 110 of 125 TETs (88%) and in 16 of 18 hyperplasias (89%). All type A (17/17) and MNT (4/4) thymomas expressed B7-H3 as did 27/30 AB, 18/25 B1, 22/26 B2 and 5/5 B3 thymomas and all 10 SCCs and 6/8 NETTs. Detailed scores for each sample and subgroup comparisons of the H-scores for each subtype as well as the *GTF2I* mutant/ wild-type cases and dead vs. alive status are summarized in Figure 7.18. The mean H-score value for each subtype was compared and the results are summarized in Table 7.11 and shown schematically in Figure 7.19. It is striking how the mean H-scores of Type A thymomas, that of AB thymomas and SCCs are significantly higher compared to B1 and B2 scores and that of hyperplasias. This is quite remarkable as such a pattern has not been observed with other proteins studied in TETs.

B7-H3 and PD-L1 were co-expressed in 90 of 125 TETs (72%). This is higher than what has been seen in NSCLC where co-expression was observed in only 17.6% of cases (306). Co-expression according to TET subtype can be found in Table 7.12. This was lowest in Type B1 (56%), A (65%) and AB (70%) thymomas and highest in B3 thymomas (100%), followed by SCCs (90%), B2 thymomas (85%) and NETTS (75%). Only two B1 TETs were negative for both. Fourteen of 107 (13%) thymomas, three of ten (30%) SCCs and one of eight (13%) NETTs also had lymphocytic B7-H3 expression. There were no cases with solely lymphocytic B7-H3 expression without epithelial expression. Despite the overall frequent and high expression of B7-H3 in TETs the range of scores is quite remarkable. In B1, B2 thymomas, NETTs and hyperplasias this was 0 – 200, for AB this was 0 – 300 and for A, B3, MNT and CA this was 50 – 300. Examples of B7-H3 staining in a range of TETs can be found in Figure 7.20.

No link between OS or *GTF2I* mutations status and H-scores was found (multivariate analysis). In other studies on B7-H3 in solid tumours, such as NSCLC, high B7-H3 expression has been linked with poor OS (306,307). In some studies this effect was only reported for heavy smokers (308). In the current study the high association between TET subtype and OS did not allow for such conclusions to be drawn.

B7-H3 has been studied in NSCLC to establish its relationship with PD-L1 expression, with B7-H4 and with clinical and pathological characteristics. 634 NSCLC cases in TMAs were assessed. Automated quantitative immunofluorescence (QIF) was used to assess the levels of B7-H3 (clone D9M2L, CST), PD-L1 (clone SP142, Spring), B7-H4 (Clone D1M8I, CST) CD3, CD8 and CD20. B7-H3 was expressed in 510 of 634 cases (80.4%) and levels were higher in tumour than in stroma. There was a significant association between high B7-H3 expression (top 10 percentile) and poor OS in two of the three patient cohorts examined ($p < 0.05$). High B7-H3 levels were

associated with a positive smoking history across all patients, but not with sex, age, clinical stage and histology. PD-L1 and B7-H3 were co-expressed in 112 of 634 cases (17.6%) (306).

Another study looked into B7-H3 expression in lung adenocarcinoma patients with a smoking history. TMAs of 270 cases were evaluated for B7-H3 with IHC. High expression of B7-H3 was associated with poorer OS in moderate and heavy smokers (hazard ratio (HR) = 3.07, 95% confidence interval (CI) =1.74-5.49, $p = 0.0001$) (log-rank test: $p < 0.0001$), but not in non- or light-smokers (HR = 2.24, 95% CI = 0.63-1.96, p = 0.64) (log-rank: p = 0.64). High expression of B7-H3 was also associated with wild-type EGFR (multivariable OR = 2.80, 95% CI = 1.38-5.84, $p = 0.0042$) (308).

In patients with clear cell renal cell carcinoma, 95% of cases were found to be positive for B7-H3-positive [56]. Its expression was linked significantly higher mortality risk compared to those who were B7-H3 negative (293).

A recent meta-analysis of 24 observational studies with a total of 4141 patients with a range of solid malignancies revealed that high B7-H3 levels are significantly correlated with poor OS (pooled HR = 2.09; 95% CI = 1.60–2.74; $p < 0.001$) but not PFS(307).

B7-H3 has also been linked to the metastatic potential of cancer cells, such as melanoma (293). shRNA silencing of B7-H3 was found to reduce the matrigel invasion capacity of the cells and mice and rats injected with these cells exhibited a prolonged PFS. Silencing of B7-H3 also led to a reduction in the expression of metastasis-associated proteins, including matrix metalloproteinase (MMP)-2, signal transducer and activator of transcription 3 (STAT3), and the level of secreted interleukin-8 (IL-8) (293).

There are no pre-clinical studies with anti-B7-H3 drugs.

A number of phase I studies are ongoing to investigate the action of enoblituzumab (also known as MGA271; anti-B7-H3 monoclonal antibody) in patients whose tumours and/or tumour-associated vasculature express B7-H3. It is being investigated in a range of refractory solid tumours as monotherapy (NCT01391143) and in combination with ipilimumab (anti-CTLA-4) (NCT02381314 – on NSCLC and melanoma) and in combination with pembrolizumab (NCT02475213). Preliminary results of the trial with enoblituzumab as monotherapy have shown good activity against heavily

pre-treated patients with a number of tumours, including prostate and bladder cancer and melanoma. Melanoma patients had all already received checkpoint inhibitors (anti-CTLA4 or anti-PD-1) (292). The findings reported here suggest that patients with TETs are likely to be good subjects for these trials. This will especially be the case if the side-effect profile is more acceptable compared to that of anti-PD-1 inhibitors. Side effects are a major limiting factor in administration of anti-PD-1 inhibitors to TET patients in particular (101,103,104,309) and there is a clinical need for other agents.

Also relevant to TETs, a study involving a drug screen on triple negative breast cancer has yielded very interesting results. Triple-negative cell lines, with different levels of B7-H3 expression were used to determine the effect of 22 anticancer compounds. Triciribidine and everolimus, which are both inhibitors of the PI3K/ AKT/mTOR pathway, were more active in B7-H3 knockdown tumour cells compared to cell lines expressing B7-H3 (100). In cells where B7-H3 was overexpressed, both drugs were less effective. This was replicated in vivo where xenografts over-expressing B7-H3 were less sensitive to everolimus than control tumors (310). As everolimus is already being used in some TET cases without a biomarker for response but with significant toxicity (100), replicating the findings of the breast cancer study in TETs would be very important.

Co-expression of immune checkpoints in solid tumours has only now started to be investigated with our increasing understanding of immunotherapy treatments and indicates the likelihood of needing combinations of these for improved outcomes. A recent trial examined the expression of multiple immune checkpoints in NSCLC. TMAs from 225 cases of stage I-III NSCLC were used (123 adenocarcinomas and 83 squamous cell carcinomas). Expression of the following checkpoint molecules was investigated: PD-L1, B7-H3, B7-H4, IDO1, ICOS, VISTA, TIM3, LAG3 and OX40, using quantitative image analysis system (Aperio). The median H-score of PD-L1, B7-H3, B7-H4 and IDO1, which are all expressed on tumours cells, was higher in squamous cell carcinomas vs. adenocarcinomas, with 3 out of 4 markers being significant ($p < 0.05$). On the contrary those expressed in inflammatory cells, ICOS, VISTA, OX40, LAG3 and TIM3 were higher in adenocarcinomas, with 3 of 5 being significant ($p < 0.05$). Only 11% of adenocarcinomas and 10% of squamous cell carcinomas co-expressed 8-9 markers. 45% of adenocarcinomas and 32% of squamous cell carcinomas co-expressed 6 to 7 markers and 28% of adenocarcinomas and 40% of squamous cell carcinomas co-expressed 4 to 5 markers(311).

Another crucial step in further understanding the importance and precise mechanism of function of B7-H3 will be to discover its receptor and other proteins with which it interacts.

Figure 7.18. Box plot showing the distribution of epithelial B7-H3 H-scores per TET subtype, including GTF2I mutation status and whether dead or alive

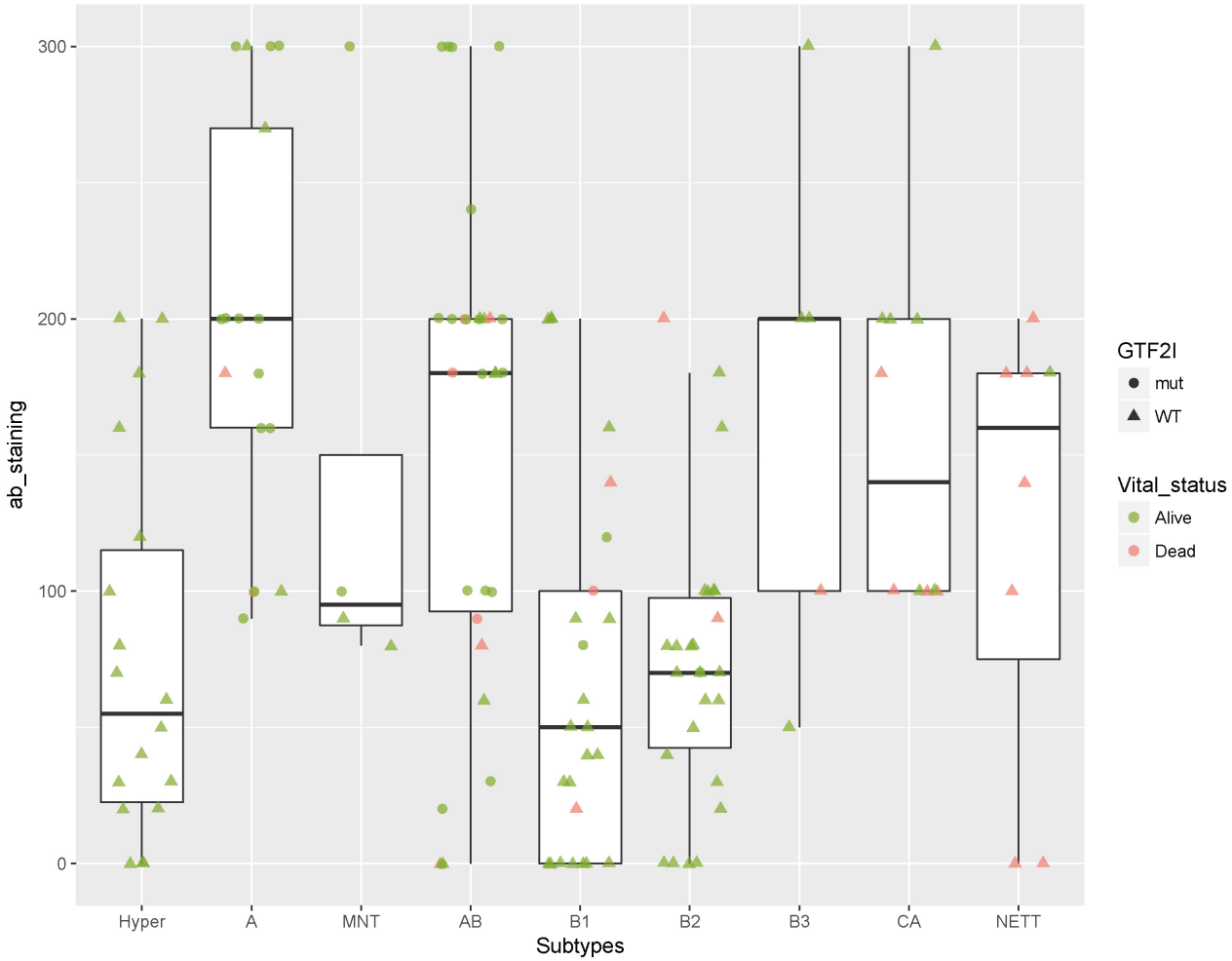


Table 7.11. TET subtype comparison matrix for epithelial B7-H3 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.41751	-	-	-	-	-	-	-
B1	0.00036	0.00375	-	-	-	-	-	-
B2	9.70E-05	0.00291	0.62792	-	-	-	-	-
B3	0.80919	0.80919	0.08243	0.10056	-	-	-	-
CA	0.43601	1	0.00949	0.00262	0.88839	-	-	-
Hyper	0.00188	0.01678	0.80919	0.81391	0.11851	0.02516	-	-
MNT	0.30151	0.80919	0.23358	0.1977	0.80919	0.41751	0.27361	-
NETT	0.1977	0.36511	0.27361	0.1977	0.43601	0.62792	0.41751	0.88839

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.19. Heat map of TET subgroup epithelial B7-H3 H-score comparison

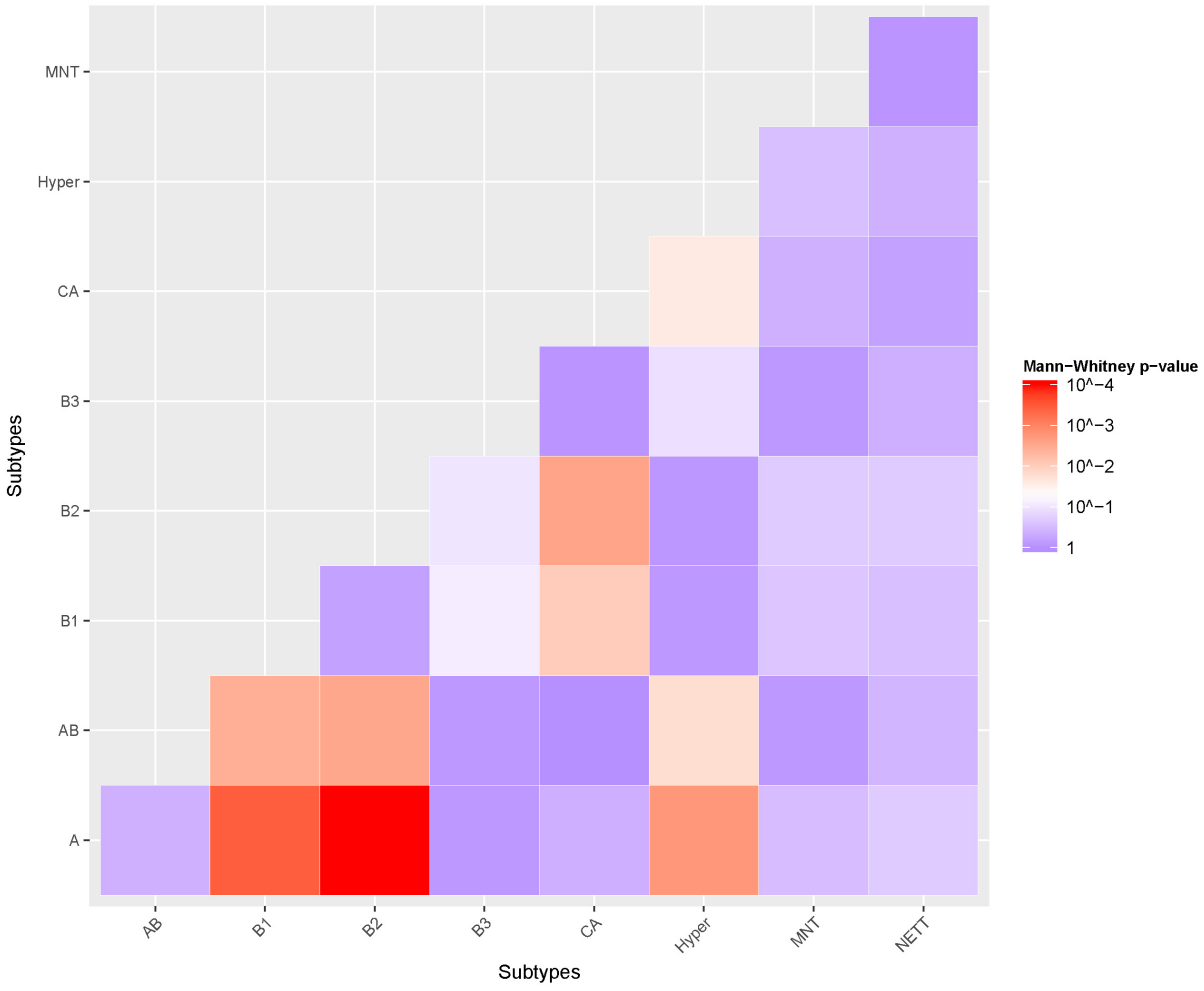


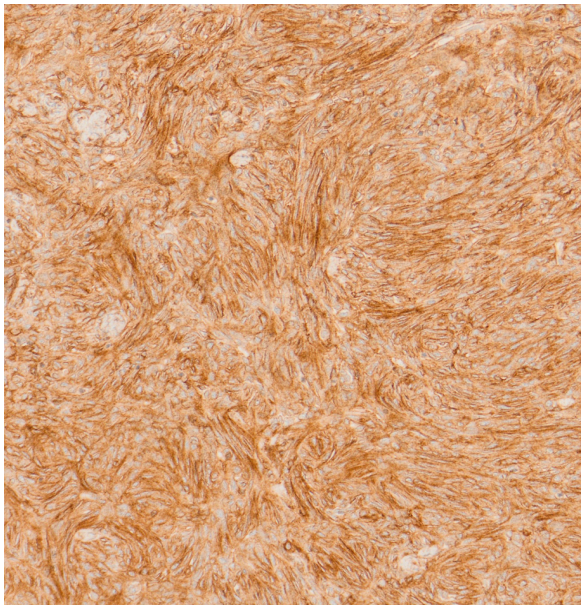
Table 7.12. Percentage of B7-H3 positive (≥5% of cells) TETs and co-expression with PD-L1 in TET epithelia.

TET subtypes	B7-H3 (E) positive	PD-L1 (E) positive	PD-L1 and B7-H3 +ve	TOTAL
A	17 (100%)	11 (65%)	11 (65%)	17
MNT	4 (100%)	2 (50%)	2 (50%)	4
AB	27 (90%)	24 (80%)	21 (70%)	30
B1	18 (72%)	19 (76%)	14 (56%)	25
B2	22 (85%)	26 (100%)	22 (85%)	26
B3	5 (100%)	5 (100%)	5 (100%)	5
CA	10 (100%)	9 (90%)	9 (90%)	10
NETT	6 (75%)	8 (100%)	6 (75%)	8
TET TOTAL	***110 (88%)***	***104* (83%)**	***90 (72%)***	***125***
Hyperplasia	16 (89%)	5 (28%)	4 (22%)	18

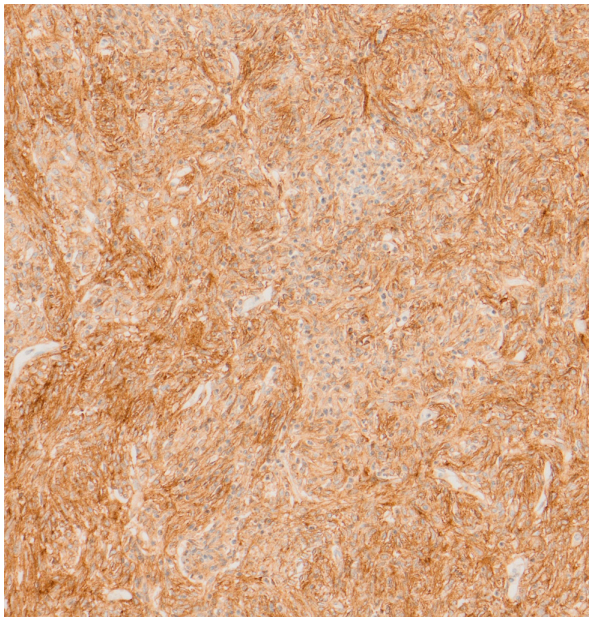
E = Epithelial. Percentage of B7-H3 and PD-L1 positive TETs was identical when using 1% or 5% as the positivity threshold as there were no samples that were positive in over 1% but fewer than 5% of cells.

Figure 7.20. Epithelial B7-H3 staining in TETs

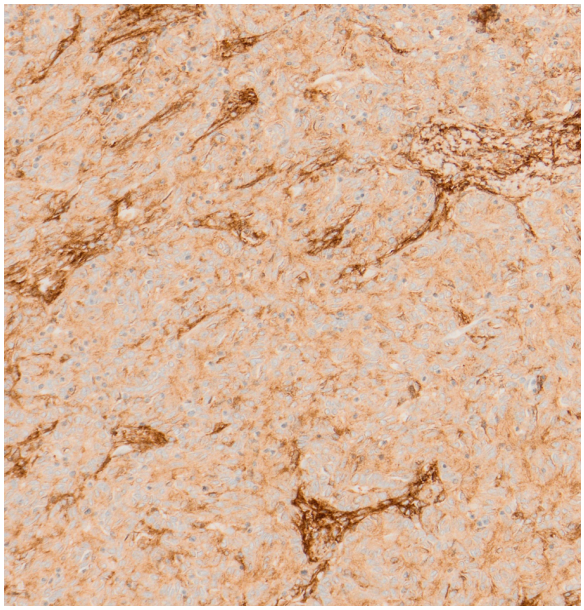
a)



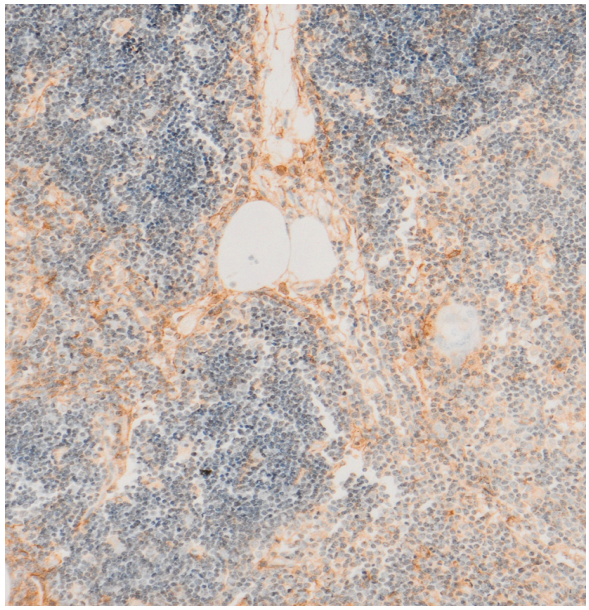
b)



c)



d)



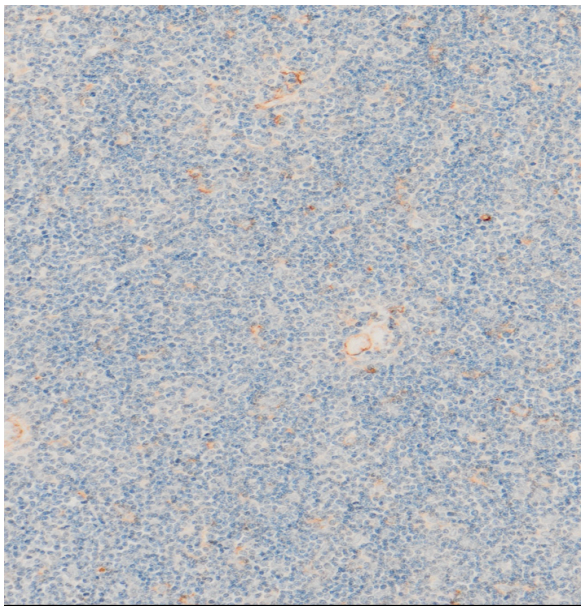
a) A: E H-score: 100 x 3 = 300, L H-score: 0

b) AB: E H-score: 100 x 2 = 200, L H-score: 0

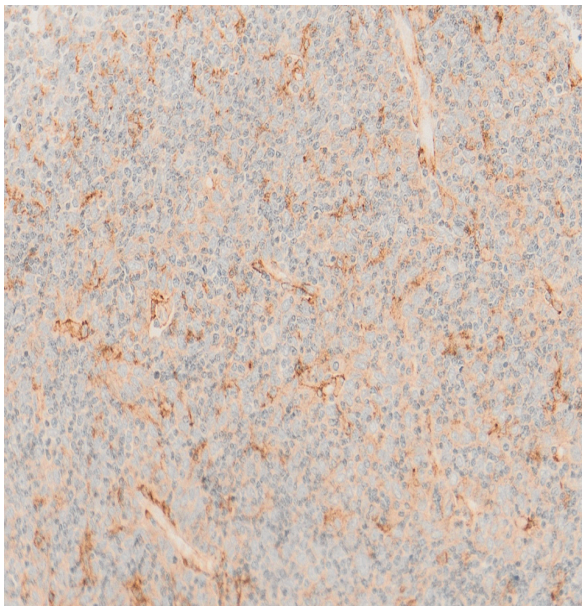
c): A: E H-score: 100 x 1 = 100, L H-score: 0

d) Hyperplasia: E H-score: 60 x 2 = 130, L H-score: 0

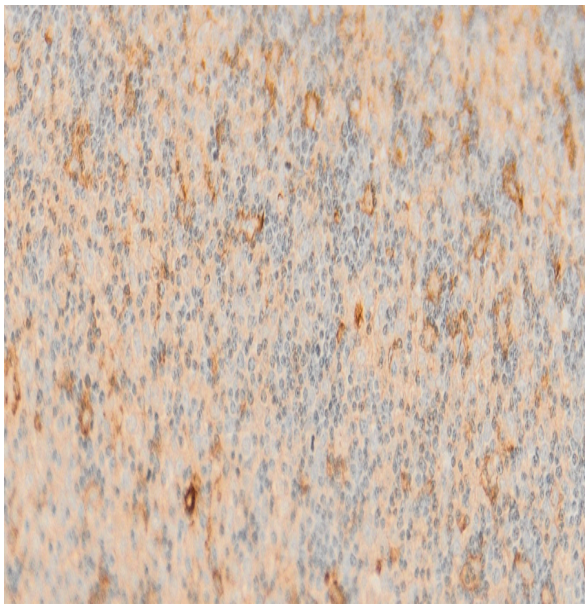
e)



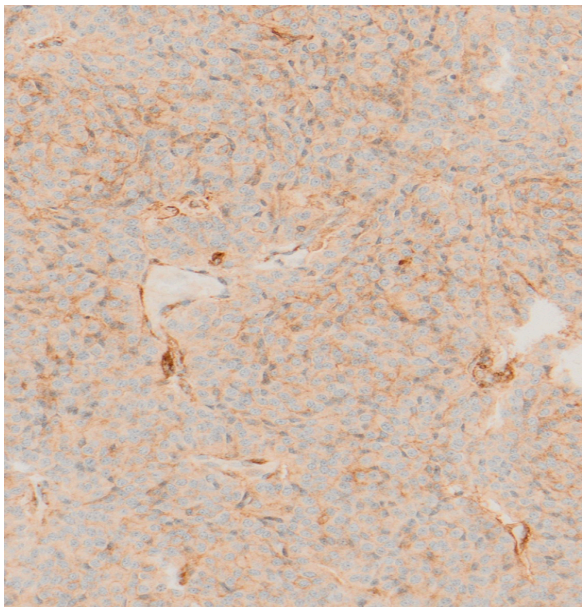
f)



g)



h)



e) B1: E H-score: 0, L H-score: 0

f) B2: E H-score: 100 x 1 = 100, L H-score: 0

g) CA: E H-score: 100 x 1 = 100, L H-score: 100 x 2 = 200

h) NETT: E H-score: 90 x 2 = 180, L H-score: 0

E = Epithelial

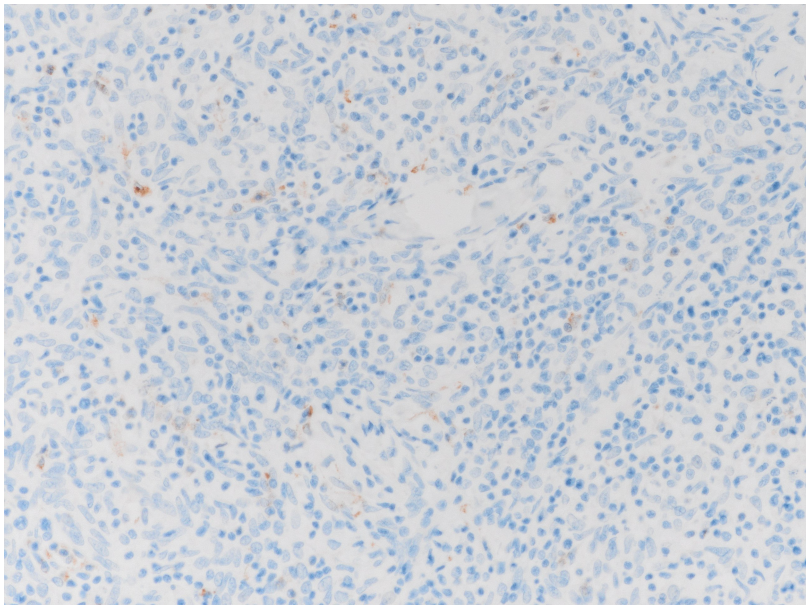
L = Lymphocytic

7.4.7 VISTA

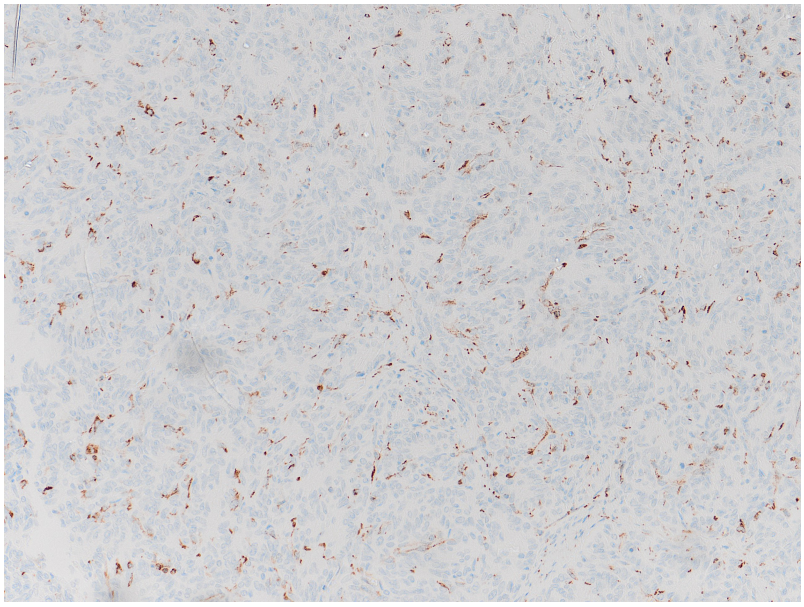
VISTA expression was largely absent from all TET subtype. There was no lymphocytic staining in any case and there was minimal expression by the epithelia in 5 of 17 Type A thymomas, 2 of 30 AB, 3 of 26 B2 and 2 of 5 B3 thymomas. Week staining was also seen in 3 of 10 SCCs whereas B1 thymomas and NETTs were completely negative. Overall expression in TETs was 17/125 (14%) with a mean H-score of 21 (range 5 – 60). Further analysis was therefore not performed. Typical VISTA staining and a weakly positive case can be found in Figure 7.21.

Figure 7.21. VISTA staining in TETs

a)



b)



a) Representative VISTA TET staining across all subtypes (H-score = 0)

b) A rare case of minimal epithelial VISTA staining in a Type A thymoma (H-score: 30 x 1 = 30)

7.4.8 4-1BB

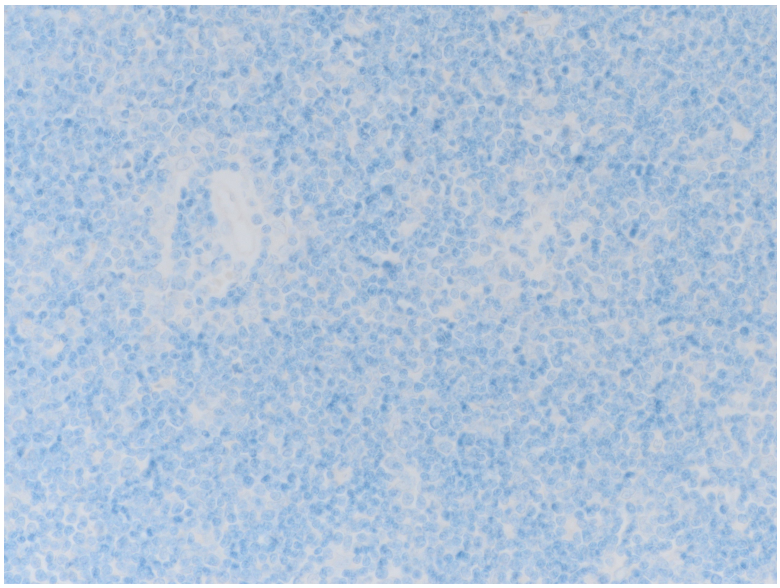
Epithelial staining with 4-1BB was negative across all thymomas and hyperplasias. It was also negative in all SCCs. However two (Pat 120 and 121) of the eight NETTs (25%) were positive with H-scores 80 x 2 = 160 (Pat120) and 90 x 1 = 90 (Pat121). Both were atypical carcinoids and lacked any clinical characteristics that made them stand out compared to the remaining NETTs. In terms of survival all NETT patients have passed away. In relation to the other antibody stains, what these two NETTs had in common was the absence of any TIM-3 staining whereas the remaining six NETTs all had varying degrees of positive staining for this antibody. The case with the highest score (Pat120) was also negative for B7-H3 and PDLIM1 staining and was an atypical carcinoid (Pat122). It was also negative for 4-1BB and was the only NETT that was negative for HGF (see chapter 7.4.10). Pat120 was one of the cases that was whole exome sequenced. Results have already been discussed but this case had a K-RAS hotspot mutation as well as somatic SLC2A3 and BRCA2 mutations amongst others.

Bearing in mind the difficulty in treating advanced stage disease in patients with thymic NETTs, even a minority expressing 4-1BB could be targeted with agonistic 4-1BB mAbs in clinical trials to establish the safety and efficacy in these rare and difficult to treat cases.

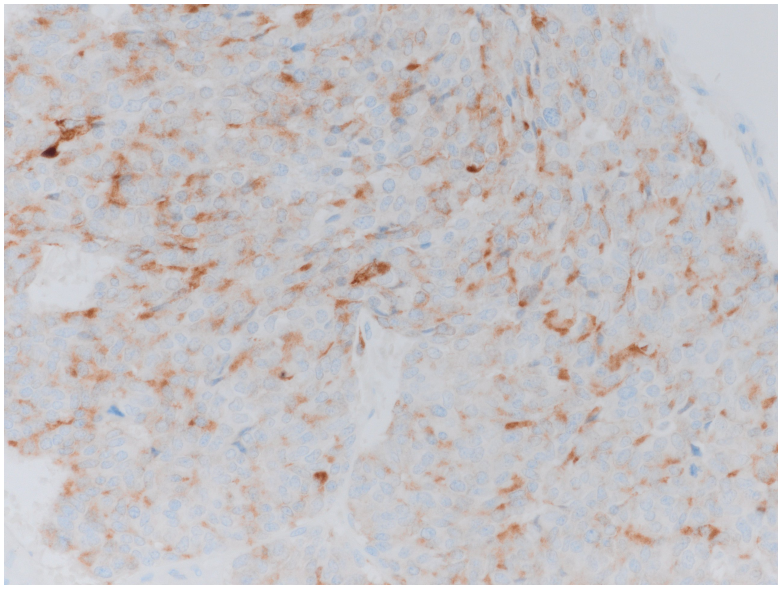
Figure 7.22 shows typical 4-1BB staining as seen in most TETs and an example of positive staining in one of 2 NETTs.

Figure 7.22. 4-1BB staining in TETs

a)



b)



a) Typical 4-1BB staining in all thymomas and thymic SCCs (H-score = 0)

b) Epithelial 4-1BB staining of 1 of 2 positive thymic NETTs (*Pat120*: H-score = 80 x 2 = 160)

7.4.9 PDLIM1

PDLIM1 was detected in the epithelia of 100 of 125 TETs (80%) and in 15 of 18 hyperplasias (83%). Expression was high throughout all TET subtypes being highest in the B subtypes and carcinomas, with 10/10 (100%) SCCs, 5/5 (100%) B3 thymomas, 24/26 (92%) B2 thymomas and 24/25 (96%) B1 thymomas staining positive. Expression dropped to 20/30 (67%) in ABs and 9/17 (53%) in A thymomas. MNT thymomas were all positive (4/4; 100%) and staining was seen in 4/8 (50%) NETTs. Detailed scores for each sample and subgroup comparisons of the H-scores for each subtype as well as the *GTF2I* mutant/ wild-type cases and dead vs. alive status are summarized in Figure 7.23.

The mean H-score value for each subtype was compared and the results are summarized in Table 7.13 and shown schematically in Figure 7.24. It is interesting how in the majority of TET subtypes the range of H-scores varies widely; from 0 to 300 in B1, B2, AB and NETTs and from 0 to 200 in Type A thymomas. Apart from being expressed in fewer A and AB thymomas, PDLIM1 also had a significantly lower mean H-score in these subtypes compared to all B subtypes and SCCs. Examples of PDLIM1 staining in a range of TETs can be found in Figure 7.25. Some TET subtypes demonstrated vast variability in terms of the intensity and pattern of staining, which is not reflected in the H-score. Figures 7.25c and 7.25d show two AB thymomas. One has an H-score of 300 and the other 0. They also have a strikingly different appearance. In Figure 8.25e,f,g there are three different NETTs that again apart from having different H-scores also demonstrate a completely different morphology with PDLIM1 staining. Apart from the likely tumour heterogeneity, explanation for this observation is not clear. It is possible that TETs that are normally characterized as being of a particular subtype in H&E might be sub-grouped further using other stains.

21 of 107 (19%) thymomas, 2 of 10 (20%) SCCs and 2 of 8 (25%) NETTs also had very low lymphocytic PDLIM1 expression (range 5-40). There were only one AB and one NETT with weakly positive lymphocytic PDLIM expression (5 x 2 = 10 for both) that were negative in their epithelium.

Co-expression of PDLIM1 and PD-L1 in each TET subtype can be found in Table 7.14. Co-expression was overall very high at 71% in TETs, as expected being highest in SCCs (90%) and B thymomas (76%, 92% and 100%, in B1, B2 and B3 respectively) and lowest in A (47%) and AB (60%). Overall there were only four type A, four AB and one B1 thymomas that did not express either (9/125 TETs; 7%). No link between OS or *GTF2I* mutations status and H-scores was found (multivariate analysis).

As the role of PDLIM1 in cancer is still under investigation the implications of these findings will need to be validated in further studies. Studies of PDLIM1 expression in colorectal and breast

cancer have been conflicting with some studies associating down-regulation of PDLIM1 with increased metastatic potential (297) and others showing the opposite (298).

Figure 7.23. Box plot showing the distribution of epithelial PDLIM1 H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive

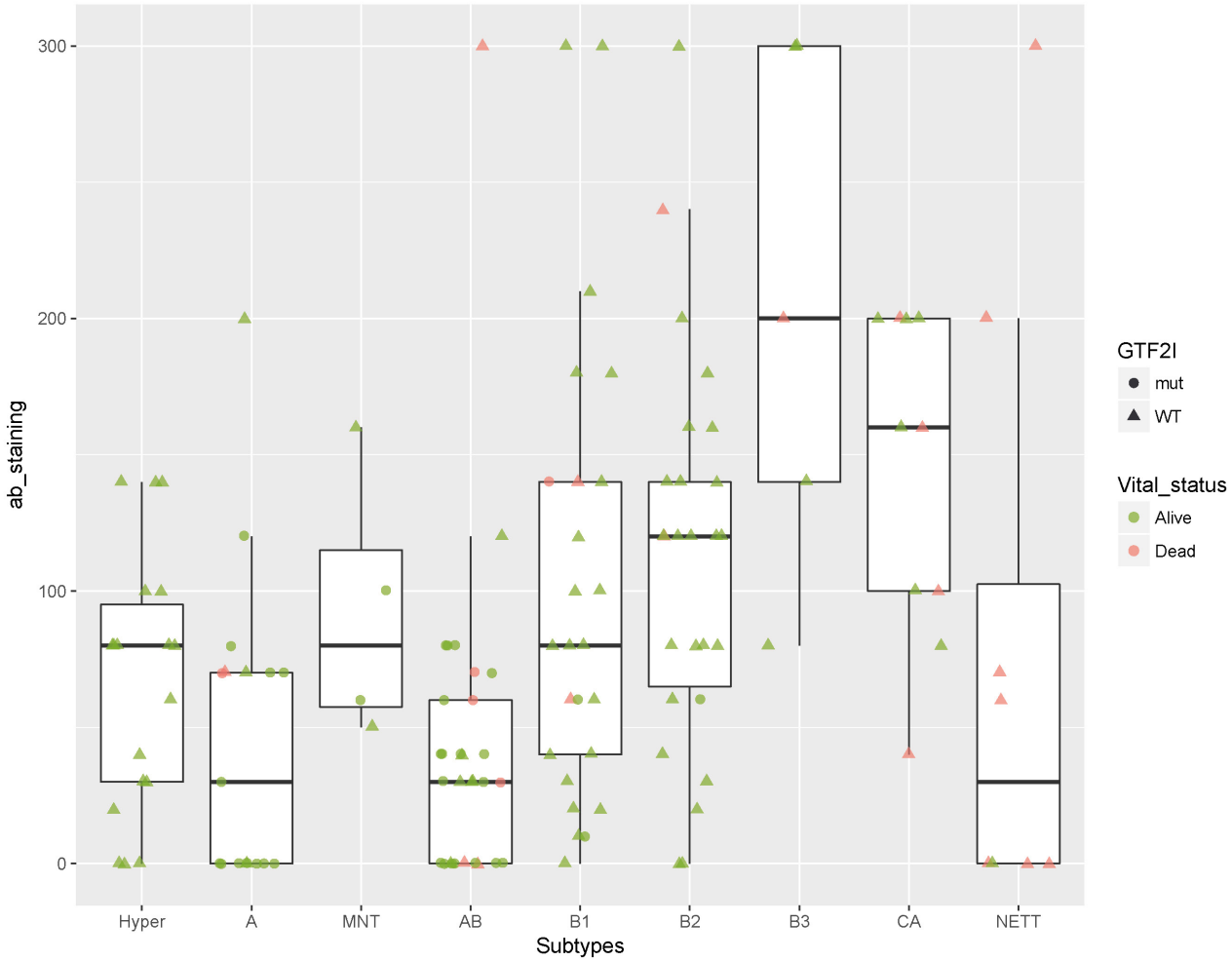


Table 7.13. TET subtype comparison matrix for epithelial PDLIM1 H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.9726	-	-	-	-	-	-	-
B1	0.046	0.011	-	-	-	-	-	-
B2	9.90E-03	0.0015	0.4842	-	-	-	-	-
B3	0.011	0.0099	0.0946	0.1159	-	-	-	-
CA	0.0087	0.0015	0.1202	0.2376	0.4221	-	-	-
Hyper	0.1513	0.0946	0.3894	0.0946	0.0294	0.0112	-	-
MNT	0.268	0.0946	0.9519	0.6769	0.1977	0.2667	0.4842	-
NETT	0.9521	0.9519	0.2751	0.2147	0.1169	0.1513	0.5161	0.5672

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.24. Heat map of TET subgroup epithelial PDLIM1 H-score comparison

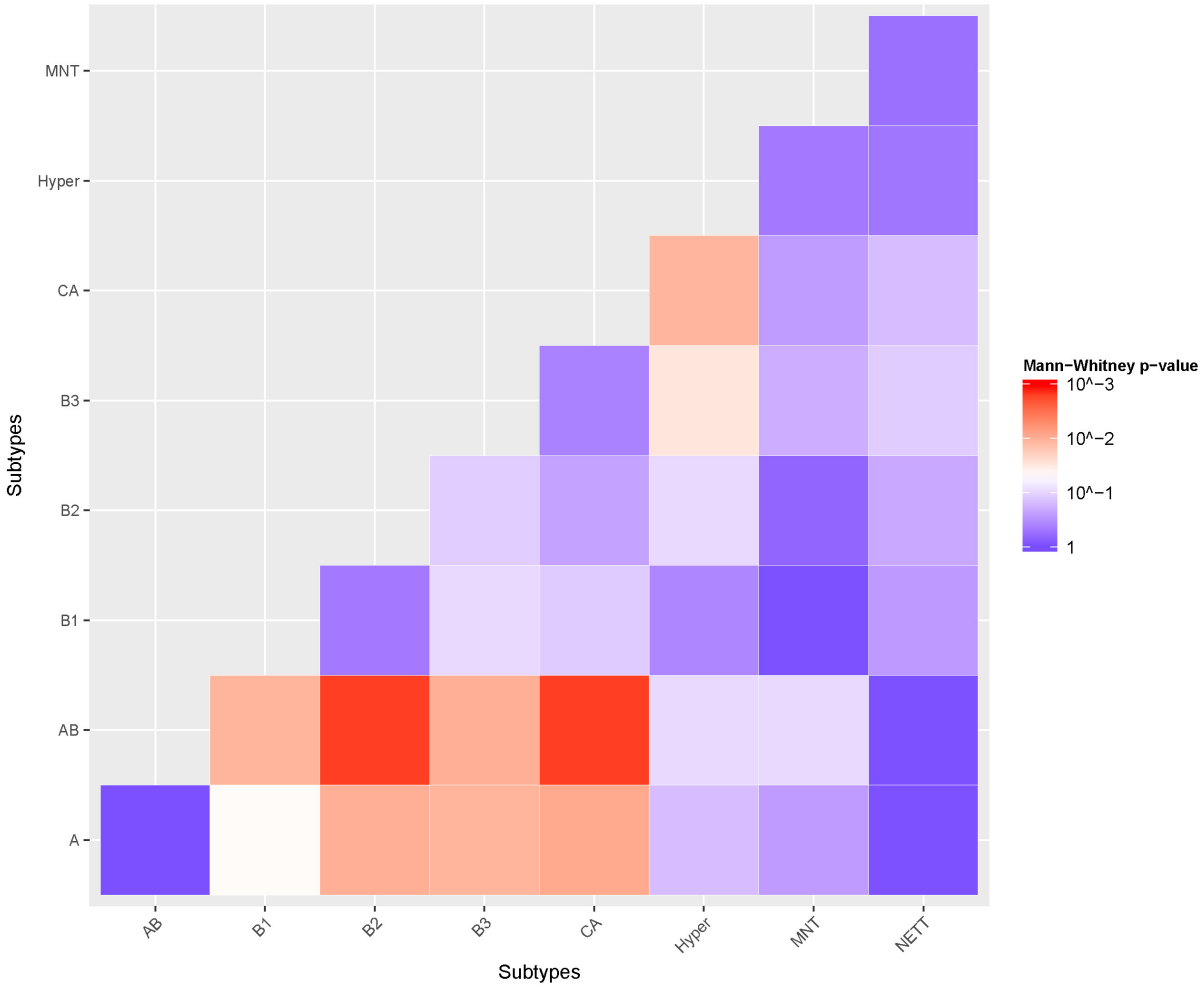


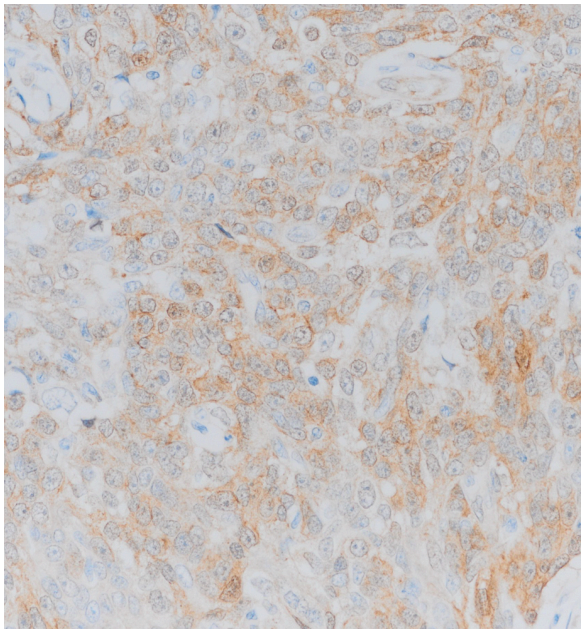
Table 7.14. PD-L1 and PDLIM co-expression in TET epithelia.

TET subtypes	PD-L1 (E) positive	PDLIM1 (E) positive	PD-L1 and PDLIM +ve	TOTAL
A	11 (65%)	9 (53%)	8 (47%)	17
MNT	2 (50%)	4 (100%)	2 (50%)	4
AB	24 (80%)	20 (67%)	18 (60%)	30
B1	19 (76%)	24 (96%)	19 (76%)	25
B2	26 (100%)	24 (92%)	24 (92%)	26
B3	5 (100%)	5 (100%)	5 (100%)	5
CA	9 (90%)	10 (100%)	9 (90%)	10
NETT	8 (100%)	4 (50%)	4 (50%)	8
TET TOTAL	***104* (83%)**	***100 (85%)***	***89 (71%)***	***125***
Hyperplasia	5 (28%)	15 (83%)	5 (28%)	18

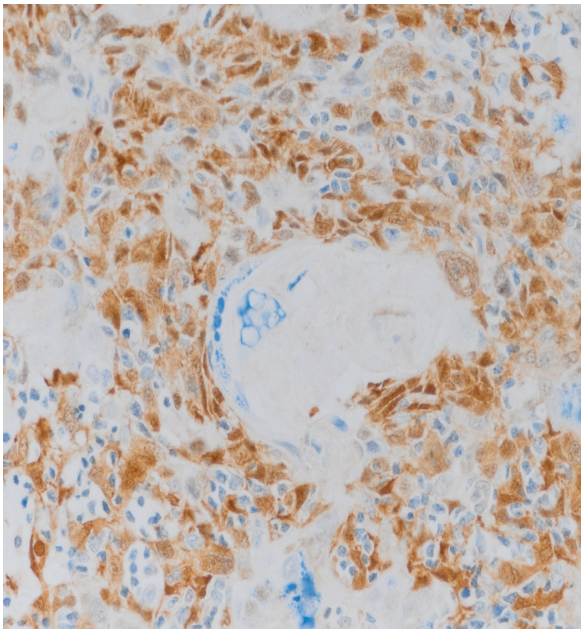
E = Epithelial

Figure 7.25. Epithelial PDLIM1 staining in TETs

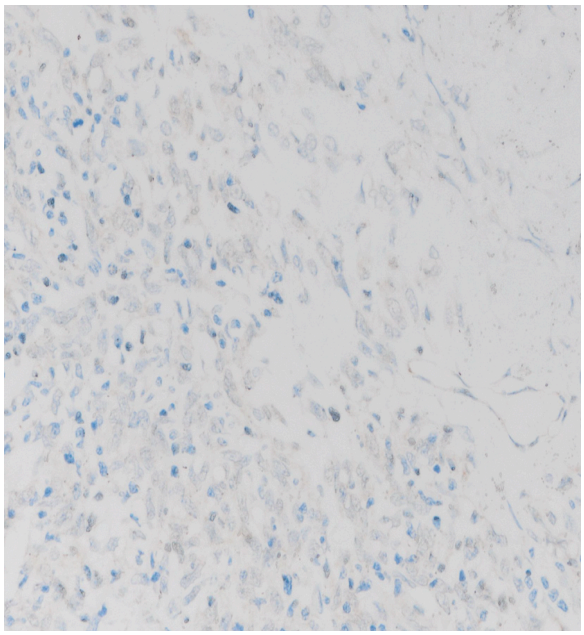
a)



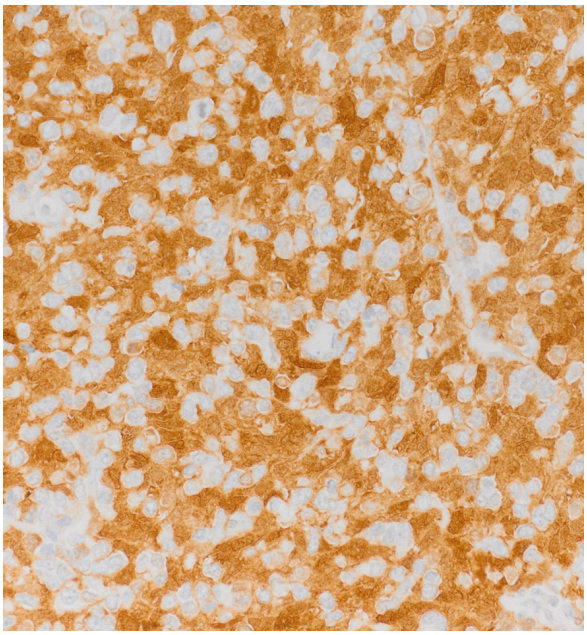
b)



c)



d)



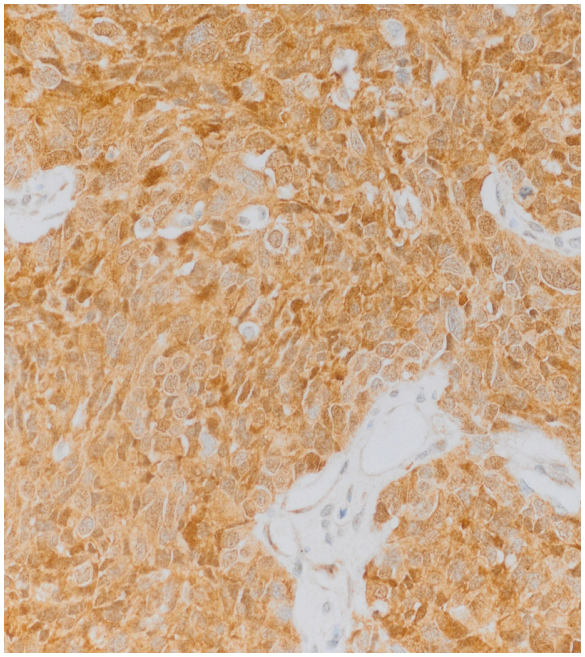
a) CA: E H-score: 100 x 2 = 200, L H-score: 20 x 1 = 20

b) B3: E H-score: 100 x 3 = 300, L H-score: 0

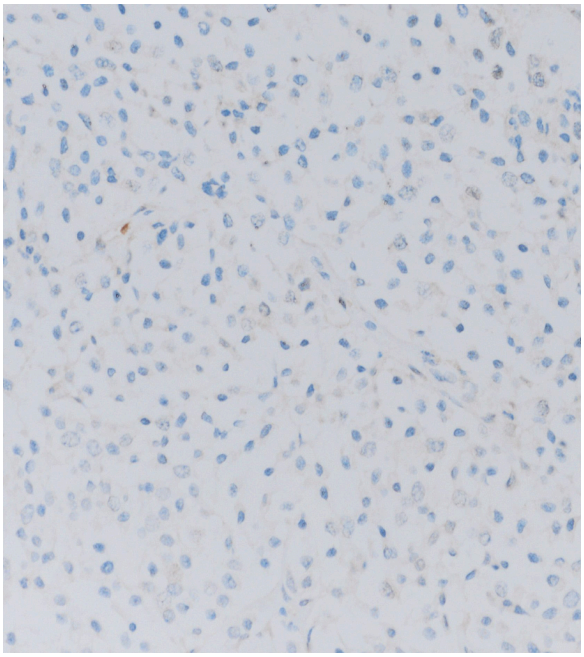
c) AB: E H-score: 0, L H-score: 0

d) AB: E H-score: 100 x 3 = 300, L H-score: 0

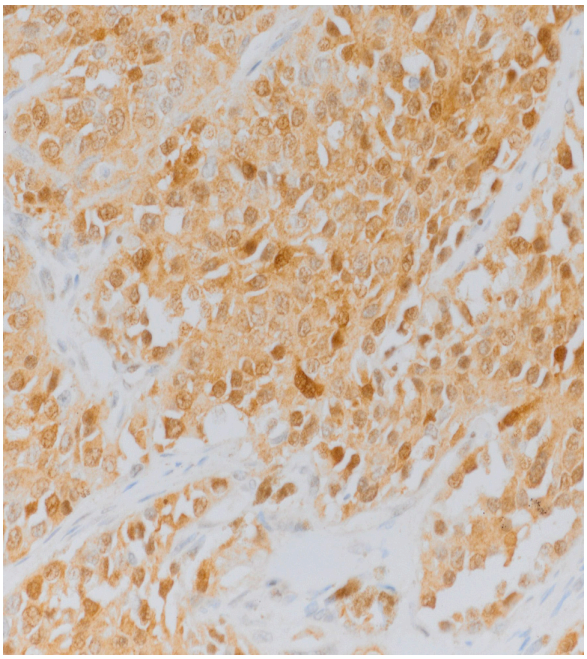
e)



f)



g)



e) NETT: E H-score: 100 x 3 = 300, L H-score: 0

f) NETT: E H-score: 0, L H-score: 10 x 1 = 10

g) NETT: E H-score: 100 x 2 = 200, L H-score: 0

E = Epithelial

L = Lymphocytic

7.4.10 HGF

HGF was detected in the epithelia of only 37 of 125 TETs (30%) and in six of 18 hyperplasias (33%). In terms of the TET subtypes, three of 17 (18%) Type A, five of 30 (17%) AB, three of 25 (12%) B1, ten of 26 (38%) B2 and two of 5 (40%) B3 thymomas expressed HGF. Surprisingly all four cases of MNT expressed HGF. In terms of carcinomas four of ten (40%) SCCs and seven of eight (88%) NETTs were positive. There was no lymphocytic staining in any TET. Detailed scores for each sample and subgroup comparisons of the H-scores for each subtype as well as the *GTF2I* mutant/ wild-type cases and dead vs. alive status are summarized in Figure 7.26. The mean H-score value for each subtype was compared and the results are summarized in Table 7.15 and shown schematically in Figure 7.27. NETTs and MNTs had significantly higher H-scores compared to all other subtypes with the exception of B3 that did not reach significance for either comparison and SCCs that did not reach significance in comparison to MNTs, all likely due to the small number of samples in these cohorts. Figure 7.28 illustrates a number of cases of HGF staining in TETs.

The case of MNT thymomas is very intriguing. All 4 cases were positive for HGF with a median H-score of 100 (range 40-100), which was significantly higher than all other thymomas. MNTs are very rare and have not been studied before in the context of expression of such markers. A recent review of the literature identified a total of only 64 cases of micronodular thymomas with lymphoid B-cell hyperplasia and 9 of micronodular thymic carcinoma with lymphoid hyperplasia. It is thought that there is a morphologic spectrum between these, with completely benign disease at the one end and a much more aggressive, malignant carcinoma at the other(30). In some cases there is diagnostic difficulty in establishing the exact nature of a TET. This is particularly true in the case of AB thymomas that sometimes have features very similar to MNT(10). If this finding can be replicated, HGF staining could be used as one of the stains to differentiate the two tumour types. Figure 7.28c demonstrates the most unexpected finding, which is an MNT TET with an H-score of 100.

In figure 7.28a appearances are typical of the majority of the TET subtypes that were completely negative for HGF. 7.28b and 7.28d are both cases of NETTs with an H-score of 200. What is striking is how different they look despite being categorized as the same TET subtype with routine IHC and scoring exactly the same. Although the clinical significance of these findings is still unknown such appearances could explain why certain cases behave differently even when their staining is described simply as ‘positive’ and even when it is scored as the same.

Many on-going clinical trials are using agents to disrupt the HGF/ MET pathway alone or in combination with other targeted agents or chemotherapy with promising results (300,302,303). In light of the findings of the current study, which is the first to demonstrate HGF expression in some

TET subtypes, particularly in MNT thymomas and NETTs, patients with advance disease that express HGF could benefit from entry into such trials following progression on standard treatment options.

Figure 7.26. Box plot showing the distribution of epithelial HGF H-scores per TET subtype, including GTF2I mutation status and whether dead or alive

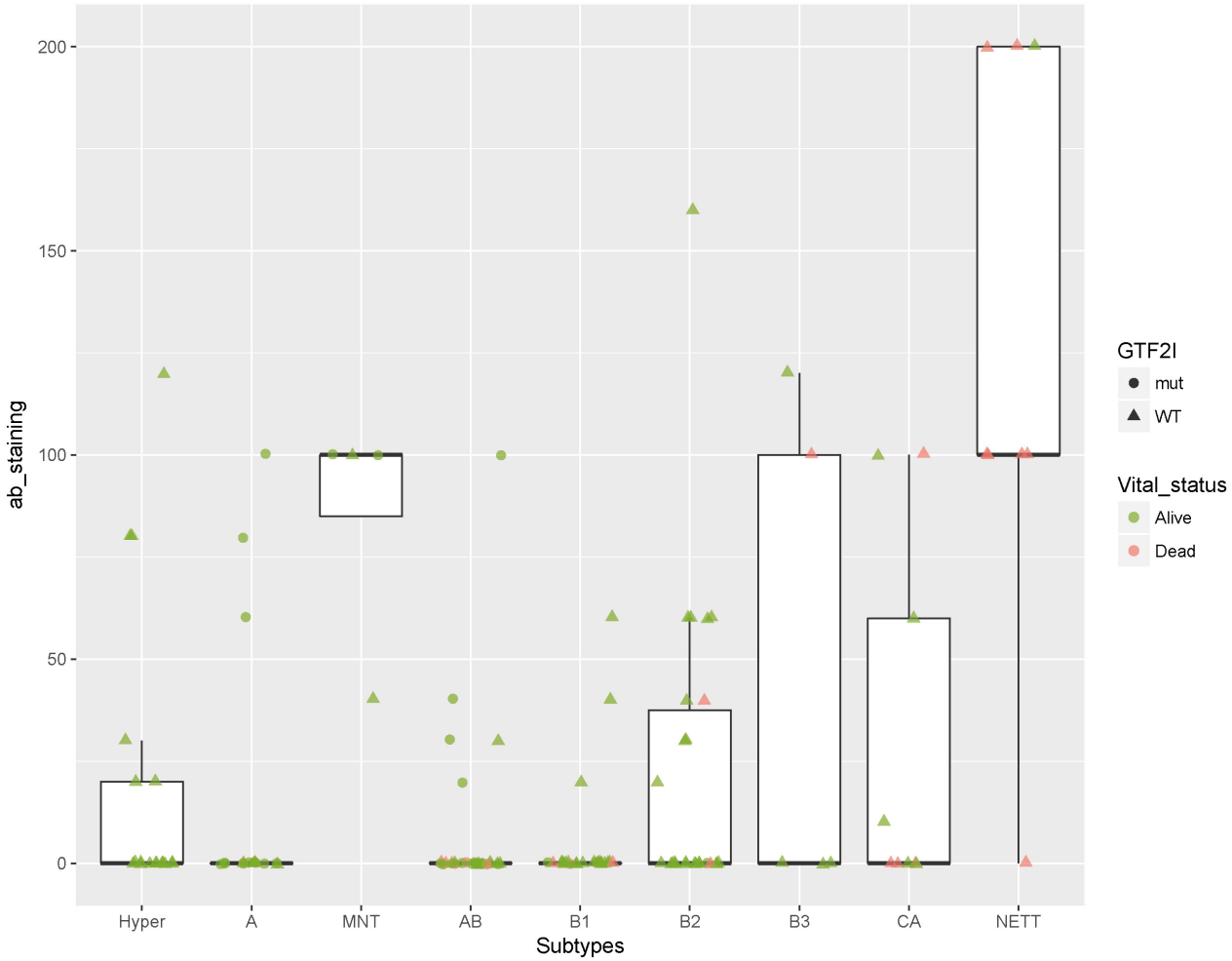


Table 7.15. TET subtype comparison matrix for epithelial HGF H-scores

	A	AB	B1	B2	B3	CA	Hyper	MNT
AB	0.82378	-	-	-	-	-	-	-
B1	0.67874	0.75631	-	-	-	-	-	-
B2	0.416	0.13459	0.10278	-	-	-	-	-
B3	0.34373	0.26425	0.19897	0.69951	-	-	-	-
CA	0.30497	0.15926	0.10798	0.70127	0.82378	-	-	-
Hyper	0.52373	0.30759	0.19897	0.794	0.67874	0.68713	-	-
MNT	0.00846	0.00088	0.00084	0.01912	0.57905	0.14399	0.02159	-
NETT	0.00173	0.00024	0.00024	0.00236	0.23531	0.03811	0.00488	0.4224

Subtype H-scores were compared with the Mann-Whitney test and p-values adjusted for multiple testing using the Benjamini-Hochberg method. Significant p – values (≤ 0.05) are in red colour.

Figure 7.27. Heat map of TET subgroup epithelial HGF H-score comparison

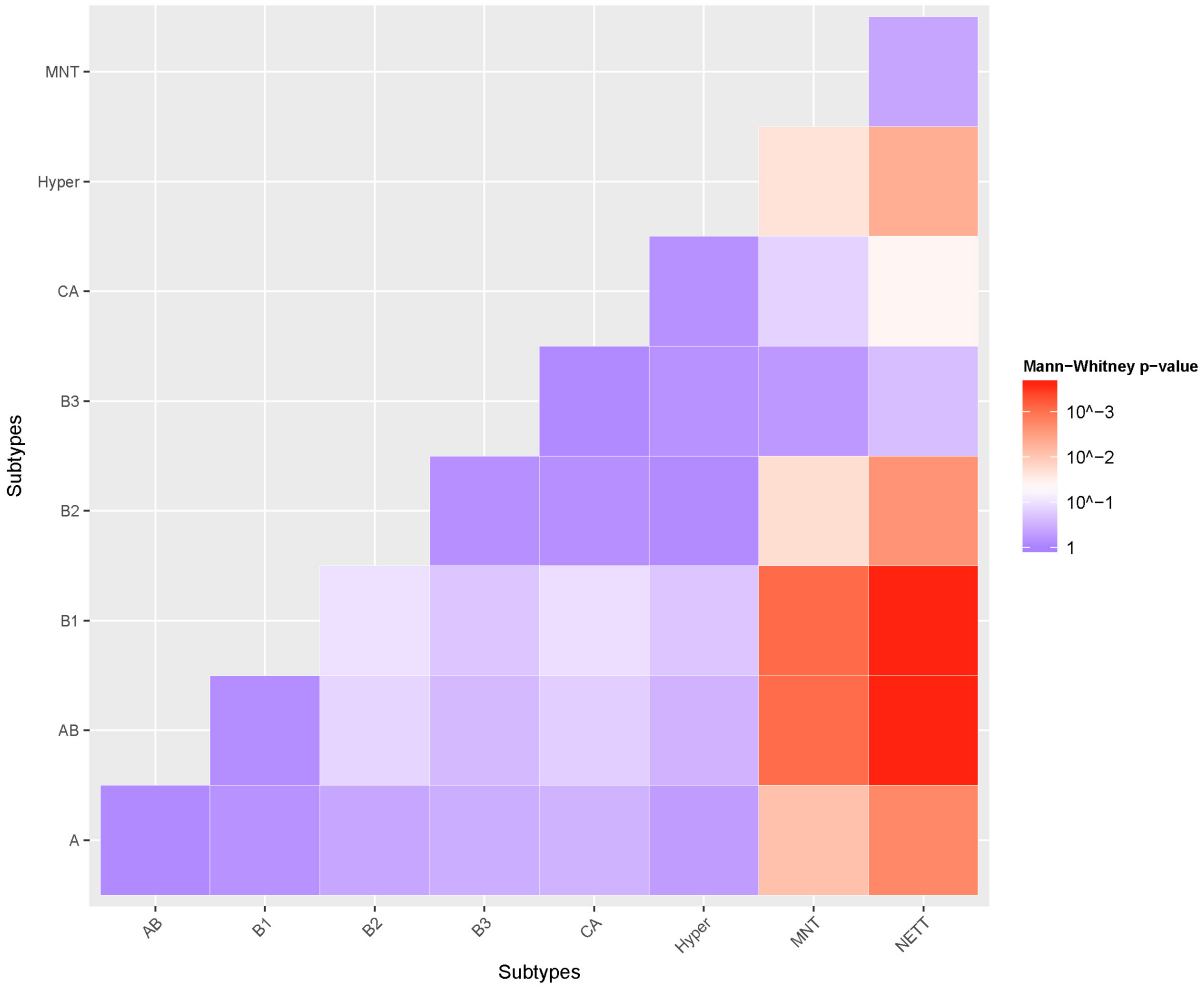
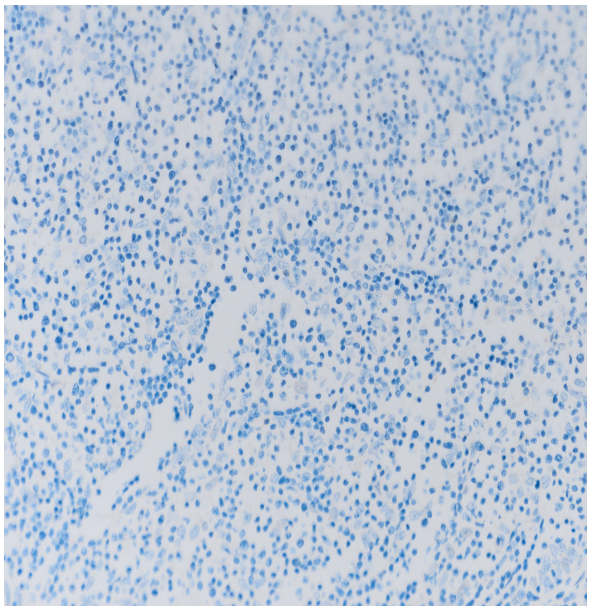
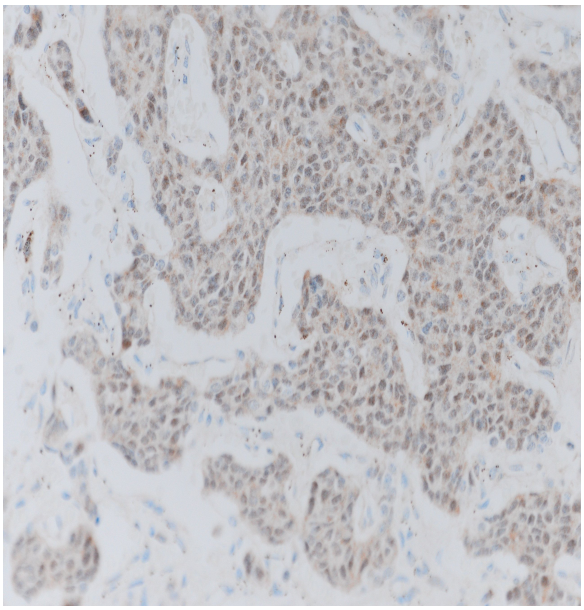


Figure 7.28. Epithelial HGF staining in TETs

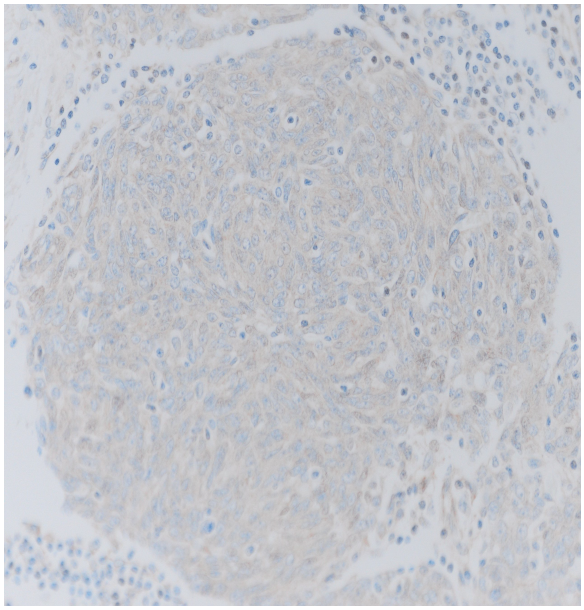
a)



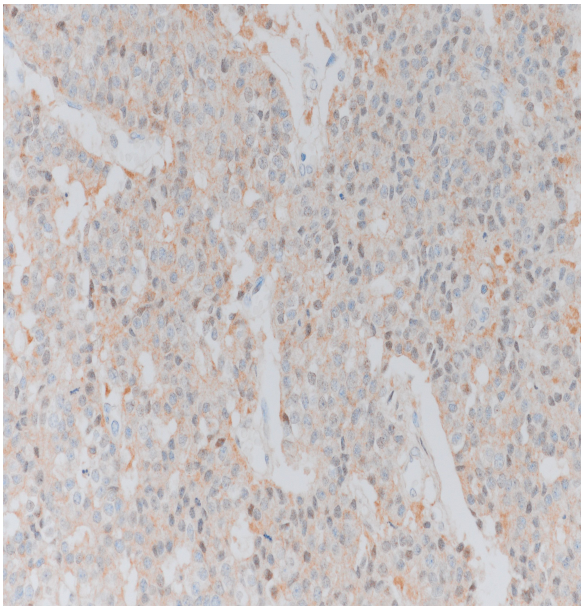
b)



c)



d)



a) Typical HGF staining across most TET subtypes: H-score = 0

b) NETT: H-score: 100 x 2 = 200

c) MNT: H-score: 100 x 1 = 100

d) NETT: H-score: 100 x 2 = 200

7.4.11 Staining of normal thymic tissue

Given the intriguing results from PD-L1, PD-L2 and B7-H3 staining of TETs, twenty cases of normal thymic tissue were stained with antibodies to these peptides. The mean percentage of cells staining positive is summarized in Table 7.16 For all three antibodies the epithelia of all twenty normal samples (100%) were positive (>5% of cells staining). This was very different in lymphocytes, which were all negative for PD-L1 and had a low percentage of lymphocytes staining positively for PD-L2 and B7-H3 (Table 7.16). Figure 7.29 shows three examples of normal thymic tissue staining with PD-L1, PD-L2 and B7-H3. The examples where chosen to exhibit the range of positivity of staining.

Table 7.16. Mean percentage of cells from normal thymic tissue staining for PD-L1, PD-L2 and B7-H3.

Antibodies to	**Epithelial cells**		**Lymphocytes**
PD-L1	Mean 60%	Range 50-75%	<1%
PD-L2	Mean 40%	Range 20-60%	1-10%
B7-H3	Mean 60%	Range 50-75%	<1%

What is striking from these findings is the absence of or low lymphocytic staining for PD-L1 and PD-L2. In TETs the lymphocytic component stained positively for PD-L1 and PD-L2 in 58% and 71% of samples respectively. What is even more intriguing is that 89% and 100% of thymic hyperplasia samples had lymphocytes positive for PD-L1 and PD-L2 respectively. This suggests a role for lymphocytic PD-L1 and PD-L2 expression not only in malignancy but also in the development of thymic hyperplasia. This requires validation in more samples. Also as lymphocytic staining is not uniform in normal thymus, in hyperplasia and TETs it suggests there is more than one lymphocyte subpopulations, which could be potentially clarified with the use of more specific lymphocytic stains. Single cell sequencing of TETs, normal thymus and thymic hyperplasias might be one of the best ways of pinpointing what cell types are expressing these molecules (e.g. CD8+ or CD4+ T cells, or double positive or double negative cells).

It is also interesting that the epithelia of all twenty (100%) normal thymic samples were positive for PD-L1, PD-L2 and B7-H3 although the percentage of cells staining varied considerably. Surprisingly only 28% and 6% of thymic hyperplasias were positive for staining with PD-L1 and PD-L2 antibodies respectively, which is the opposite trend from lymphocytic staining. For B7-H3, epithelial staining of thymic hyperplasias was seen in 89% of cases. These findings also require validation in a larger cohort of samples and may allow us to gain a better understanding of thymic hyperplasia, which so far has been significantly under-researched.

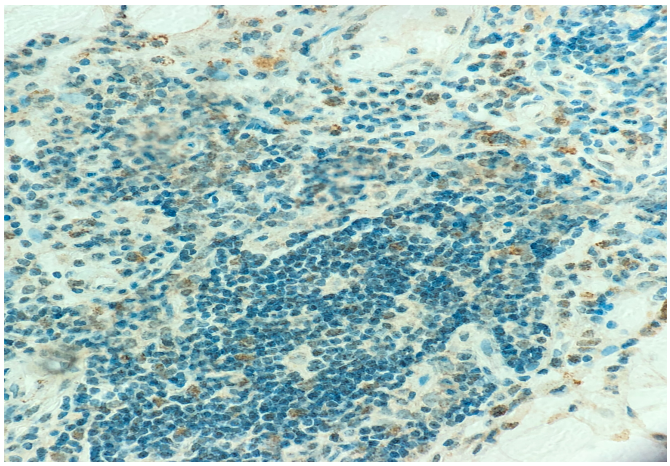
PD-L1 and PD-L2 transcripts are known to be expressed in many normal tissues with levels being higher in placenta, heart, lung and liver, with lower expression seen in spleen, lymph nodes and

thymus and no expression in the brain (312). The amount of transcripts does not always correspond to the amount of protein that is expressed (312).

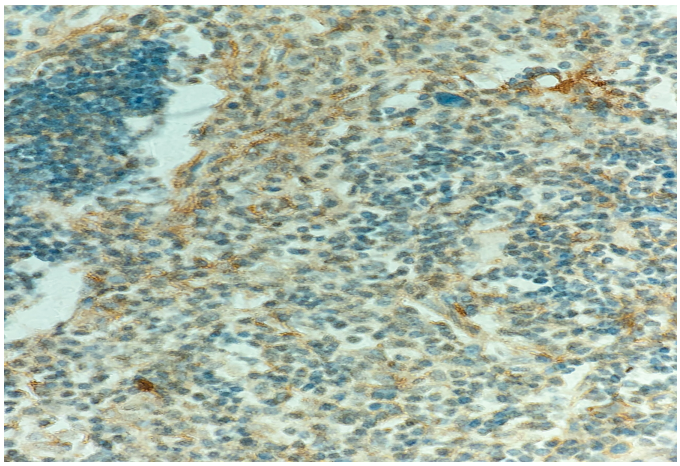
According to the Human Protein Atlas (https://www.proteinatlas.org/ENSG00000103855-CD276/tissue) there is low to medium B7-H3 expression in all solid organs, with no expression seen only in heart, smooth and skeletal muscle. Normal thymic tissue or TETs have not been included. The transcript of the B7-H3 gene has been found to be expressed in all normal tissues and solid tumours. As with PD-L1 and PD-L2, the protein however is preferentially expressed only in tumours (292).

Figure 7.29. PD-L1, PD-L2 and B7-H3 staining in normal thymic tissue

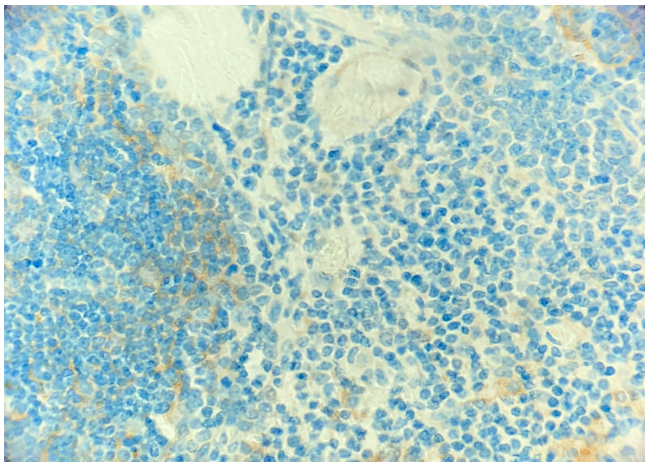
a)



b)



c)



a) PD-L1 scoring: 40% of epithelial cells positive

b) B7-H3 scoring: 75% of epithelial cells positive

c) PD-L2 scoring: 20% of epithelial cells positive

There was no lymphocytic staining in these cases.

7.4.12 Inter-rater agreement

With regards to the number of positive cases, whether a 1% or 5% threshold was used to define positivity, the concordance between the two observers was 100%. Therefore inter-observer weighted Cohen's kappa-scores were calculated based on the individual H-scores given for each TET for each stain. In the case of B7-H3 this was done for the epithelial staining only whereas for PD-L1 and PD-L2 it was done separately for the epithelial and the lymphocytic staining. The inter-observer correlation was confirmed using Spearman's rank correlation coefficient (Spearman's rho) on the same scores. Both methods reveal that the inter-observer variability was minimal.

Table 7.17. Inter-rater agreement scores for B7-H3, PD-L1 and PD-L2

	Weighted Cohen's Kappa	Spearman's
B7-H3 E	0.992153831	0.987536583
PD L1 E	0.993086532	0.979051337
PD L1 L	0.993825722	0.996261463
PD L2 E	0.973343787	0.979236514
PD L2 L	0.988694933	0.99244395

E = Epithelial, L = Lymphocytic

Using dedicated software that can be taught to differentiate between positive and negative cells one can optimize further scoring of these and other antibodies in the future. This can improve the speed of scoring and simplify the statistical analysis. It may also serve as a more reproducible means of obtaining such results. As the majority of the antibodies used have not yet been established in clinical practice extra caution needs to be given in the interpretation of the results and any such software will at least to start with need to be used alongside human scoring with weighted kappa-scores and Spearman's used between these too.

7.4.13 Co-expression of markers and clustering of TET subtypes according to these

Understanding patterns of expression and co-expression of immune checkpoint inhibitors and other important proteins in all solid tumours is a huge challenge but one that can potentially lead to important breakthroughs in the treatments of these malignancies either through the use of new

agents or with a combination of existing and/ or new therapies. This study reported findings from the investigation of expression of eight immune checkpoint markers (PD-1, PD-L1, PD-L2, LAG-3, TIM-3, B7-H3, 4-1BB, VISTA) as well as another two emerging targets (PDLIM1 and HGF). Findings were correlated with available clinicopathological data, including TET subtype, survival and *GTF2I* status. As already mentioned the histological breakdown of analyzed samples was 17 A, 4 MNT LS, 30 AB, 25 B1, 26 B2, 5 B3, 10 CA, 8 NETT and 18 hyperplasias. As well as performing an analysis of individual stains, a cluster analysis for all antibodies together was performed. The aim was to ascertain whether the current histological subtypes would predict the behaviour of these tumours in terms of their staining. It also provided another means of grouping these tumours and attempting to correlate these findings with clinical characteristics and, where available, WES data.

As LAG-3, VISTA and 4-1BB were largely negative in this cohort they were not included in the analysis although certain salient findings are noted. Also as in the cases of PD-L1 and PD-L2 there was a large proportion of lymphocytic as well as epithelial staining, these have been analysed separately and in terms of clustering have been used as independent markers, reaching a total of nine. Figure 8.29 is a heat map and cluster analysis of all 125 TETs and 18 hyperplasias that was created with these nine markers (The raw data for this figure can be found in Appendix 3). None of the 143 cases expressed all nine markers. Co-expression of eight markers was seen in twelve (8%), of seven markers in 23 (16%), of six in 46 (32%), of five in 35 (24%) and two - four in 27 cases (19%). Of the twelve with co-expression of eight, four were B2s, three ABs, and two MNT thymomas (50% of all). There was also one B1 thymoma, one SCC and one NETT. One other MNT expressed seven markers, and one expressed six markers. At the other end of the co-expression spectrum, of the 27 cases with co-expression of two - four markers, 8 were Type A thymomas (47% of all Type A thymomas in the study), eight were ABs, three were B1s, one was B2, two were hyperplasias, two were SCCs (20%) and two were NETTs (25%).

Looking at individual TET subtypes, NETTs are very interesting as they demonstrate enormous heterogeneity. Although the total number of available samples was only eight, the number of markers co-expressed varied considerably. Three NETTs co-expressed five markers (Pat 118, 121 and 123), two co-expressed only three (Pat 120 and 122) and the remainder expressed six (Pat 119), seven (Pat 124) and eight (Pat 125) markers. Looking at how these samples clustered and taking H-scores into account, Pat 118, 119, 123, 124 and 125 clustered very closely with each other. Not surprisingly very closely related to those were four of the ten SCCs, as were two ABs, two B1s, one B3 and one MNT (Pat 18). Pat 121 clustered closely with another group of TETs that was quite mixed, which apart from another two SCCs consisted of TETs of all subtypes. Pat 120 and 122 were completely different to the rest of the NETTs, clustering closely together with only an AB and B2 thymoma. It was however Pat 120 and 121, who were the only two NETTs (and only

two TETs) to express 4-1BB. Of these cases, only three were whole exome sequenced. Pat 120 had a KRAS hotspot mutation as well as further somatic mutations in SLC2A3 and BRCA2 amongst others, Pat 124 also had mutations in SLC2A3 and PARP4 and Pat 125 had a P53 hotspot mutation, as well as mutations in PARP4 and MSH2 amongst others. Thymic NETTs have not been included in many IHC studies therefore comparing these results with others was not possible.

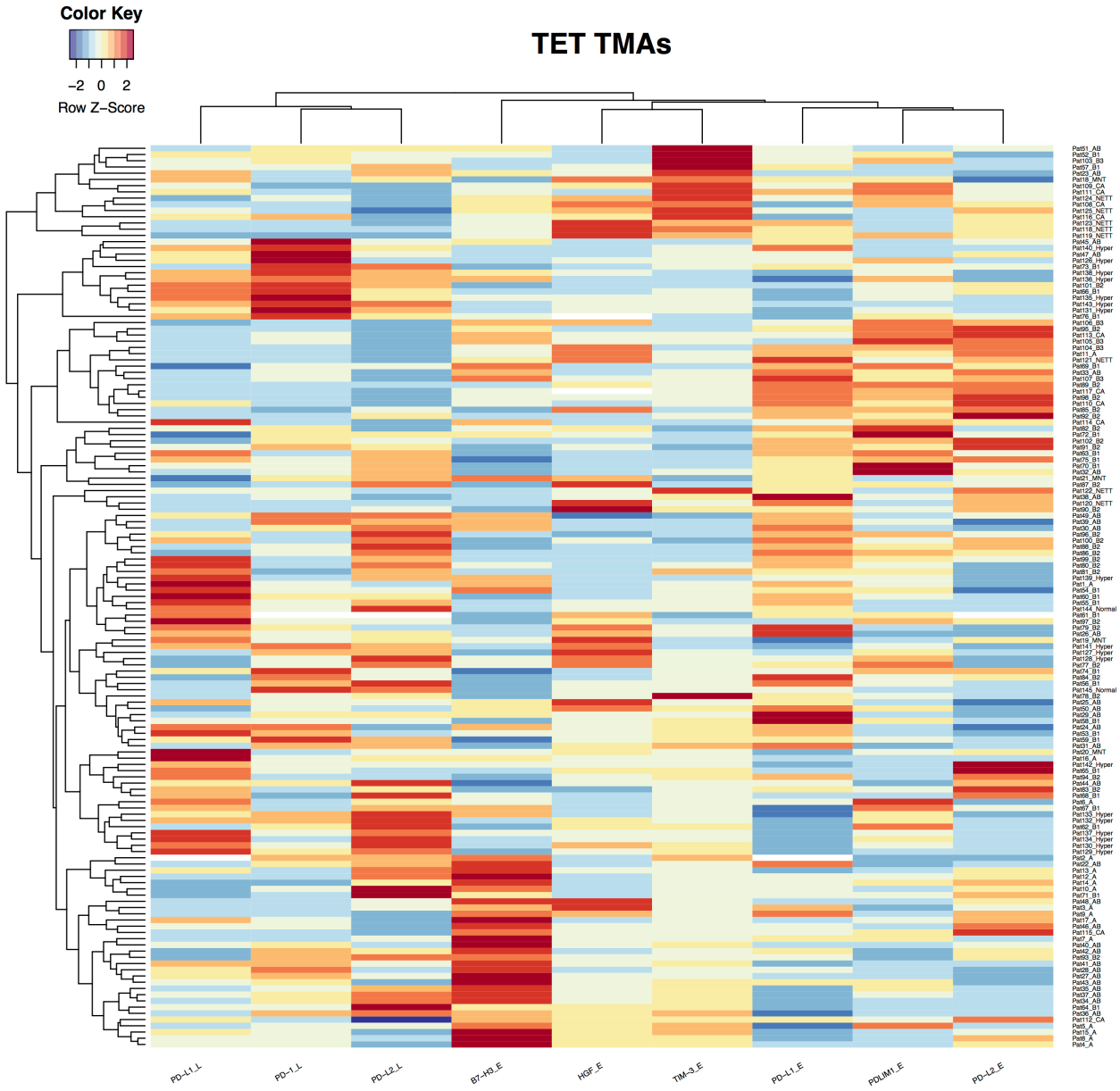
An unexpectedly interesting group was that of the 4 MNTs. As described overall these exhibited co-expression of most markers, with scores that were at the higher end of the spectrum for each (especially PD-L2 L, B7-H3, PDLIM1 and HGF). Their clustering with the rest of TETs was surprising: One (Pat 18) was as discussed above, closely related to a group of NETTs and SCCs. This was a *GTF2I* wild type MNT. Pat 21 (*GTF2I* mutant) clustered with a group of B1 and B2 thymomas but was also very closely related to Pat 120 and 122 (NETTs). Pat 19 (*GTF2I* mutant) clustered closely with a group of hyperplasias and Pat 20 (*GTF2I* wild type) clustered with another group of A, AB, B1 and B2 thymomas. Overall the four MNTs can be seen in Figure 8.29 to cluster very far away from one another. As it is known that components of other TET subtypes can co-exist with MNTs (30), this could explain why they behave so differently. Also as they are thought to be at one end of a spectrum of which the other end are micronodular thymic carcinomas with lymphoid hyperplasia (30) it is possible that one or two of these MNTs fall into that group.

SCCs are another important and interesting group. Four (Pat 109, 111, 108 and 116) (40%) clustered very closely together with four NETTs (50%) as can be seen above. Another four (Pat 113, 117, 110 and 114) (40%) were closely linked to a group of TETs consisting of four of the five B3 thymomas (with the remaining one being linked with the first group of SCCs), Pat 121 (NETT), five B2 and one of each of A, AB, and B1 thymomas. Pat 95, which is one of the B2 thymomas, most closely linked to 2 B3 thymomas and one CA (Pat 113) was a rare case of anaplastic B2 thymoma that in fact had initially been characterized as a carcinoma and subsequently reclassified. The remaining two SCCs (pat 115 and 112) clustered very distantly from the remaining ones, in clusters almost exclusively of A and AB thymomas. In WES Pat 112 was found to have a germline BAP1 mutation with the patient having had an extensive history of previous malignancies consistent with BAP-1 predisposition syndrome (249). This case is discussed in detail in chapter 5.

Looking at the remaining TET subtypes and hyperplasias it is worth pointing out how although quite a few cluster closely together according to their subtype, there are cases in all subtypes that behave very differently in terms of their expression of these markers and have clustered more with other TET subtypes. Whilst it may reflect tumour heterogeneity to an extent, it does suggest that any thymoma subtype that behaves in a more aggressive fashion might therefore benefit from immnomodulatory therapy, independent of its histology. It can therefore be argued that any

subtype of thymoma should be considered for screening using appropriate antibody tests, if and when they become clinically available.

Figure 7.29 Clustering of TET subtypes based upon staining with all antibodies



Antibodies whose staining is included in the diagram: PD-1 (L), PD-L1 (E), PD-L1 (L), PD-L2 (E), PD-L2 (L), TIM-3 (E), B7-H3 (E), PDLIM1 (E) and HGF (E). LAG-3, VISTA and 4-1BB are not included, as they were largely negative. (E) = Epithelial component staining, (L) = Lymphocytic component staining.

7.5 Study limitations and future directions

Employing IHC by itself can show the presence or lack thereof of certain markers/ proteins in a tumour, and the therapeutic significance of staining patterns is largely unknown. Validation of these results and expansion to different cohorts of TETs is required. Regarding the clinical significance of these findings this becomes even harder and other studies of antibodies that block (or stimulate) these receptor in vitro or in vivo are required. As has already been the case with PD-1 inhibition in clinical trials, even treatments that work can generate unacceptable toxicities and therefore their clinical utility is minimized.

In terms of the TETs in this study, like all TETs, it is likely that there is some tumour heterogeneity within the tumours, although the majority of thymomas usually exhibit a single pattern. Variation is hopefully mitigated to an extent by reviewing 3 cores rather than a single one.

Operator/scorer/pathologist-dependent variability also needs to be stressed especially in cases of very weak staining. Assigning an H-score to TMA cores may also be problematic. For example in a sample with 100 x 3 epithelial staining and 20 x 2 = 40 lymphocytic staining, the area of epithelial cells can be only 20% of the whole core. Also even for more uniform staining it is not yet fully established, even for the more extensively studied markers, which characteristics matter the most. In the case of 100% of cells staining weakly (=1), giving an H-score of 100, and of 50% of cells staining intensely (=3), giving an H-score of 150, one would have to closely assess the effect of any treatment targeting these receptors. All these become even harder to assess as it is thought that expression of certain markers is dynamic and can change in time and almost definitely after treatments have been administered. This is another potential limitation to this study. Although there were hardly any patients that had received systemic treatment prior to resection of their tumour, samples were collected from 1996 – 2016, which makes preservation of proteins questionable, especially in the older samples. However, there was no evidence of weaker staining in the older specimens.

Despite these limitations IHC can be a very quick and efficient way of assessing expression of specific markers and can be a useful guide as to which tumours are likely to benefit from what therapies. Amongst others, the immune markers that will be stained for next include B7-H4, IDO1, ICOS and OX40 and also mesothelin, another emerging target.

Chapter 8
Overall conclusions and future work

TETs are rare and under-researched tumours. They are different from most other solid tumours in that the are very closely linked with a multitude of autoimmune conditions, MG being the commonest (48). The relationship with MG is complex in that 15-20% of MG patients will also develop a thymoma during their lifetime and patients with TETs have a 24-40% chance of developing MG(48). It is commoner in B2 and B3 thymomas compared to B1, AB and A (49). At the same time TETs have been shown, and the current study confirms that they exhibit a high expression of PD-L1 and PD-DL2 (84-87,89,90) amongst other immunotherapy targets. Clinical trials with anti-PD-1 agents in TETs are ongoing with encouraging results in terms of responses but so far have been hindered by significant immune-related toxicity, which has led to individuals with any history of autoimmune conditions to be excluded from current trials. (101,103,313). The observation that the same tumours that normally avoid immune surveillance and may respond to immunotherapies are also linked with autoimmunity is very interesting.

Despite increased interest in TETs over the last few years there is still a significant need to identify novel therapies for patients with advanced or recurrent disease. Currently there are no licensed targeted treatments for TETs, which is partly due to the lack of biomarkers of response to the few treatments that have shown modest results at best. Known drivers only account for a very small proportion of cases. This research project aimed to define TET subsets at the molecular level and to identify likely drivers that could be targeted with existing, emerging or new therapies. It also aimed to identify biomarkers that may differentiate different TET subtypes but also biomarkers of response to emerging immunotherapy and other treatments.

156 TETs that were collected over the last 30 years were used. A minority (27%) had been snap frozen with the remaining being FFPEs. Collecting FFPEs for this project increased the number of available samples and specifically of rarer subtypes that have surgery less frequently such as SCCs and NETTs. DNA extracted from FFPEs was of variable quality. However that was adequate to obtain good quality results from Sanger sequencing and satisfactory coverage in WES. Where the extracted DNA turned out to be suboptimal was SNP genotyping and subsequent analysis with ASCAT with good quality results available for 38% and 55% of samples with copy number gains and copy number losses respectively. Considerable effort was also put into obtaining matched normal tissue for as many cases as possible (76%) to be able to differentiate between somatic and germline changes in WES and CNA analysis.

8.1 The discovery and validation of the GTF2I mutation

The discovery of the recurrent missense mutation (T>A, chr7:74146970 (hg19)) in *GTF2I* was the first and so far biggest advance in the understanding of these tumours. We confirmed the presence of this mutation in our unique TET cohort (Table 8.1) at a high frequency in indolent thymomas (type A and AB). This was the first study to show it was not present in thymic hyperplasias or NETTs. This opens the road for the presence of this mutation to be used as a diagnostic test in difficult to diagnose cases, particularly when the differential diagnosis lies between type A or AB thymomas and benign lesions especially if tissue diagnosis is being sought using a biopsy.

Table 8.1. Prevalence of the *GTF2I* mutation in TETs according to subtype in our cohort Vs Petrini *et al.* (106), TCGA (107) and Feng *et al*. (212)

TET subtype	NHLI (n=141)	Petrini (n=274)	TCGA (n=117)	Feng (n=296)	ALL (n=828)
CA	0/12 (0%)	3/36 (8%)	0/9 (0%)	3/36 (8%)	6/93 (6%)
NETT	0/6 (0%)	0/4 (0%)	0/1 (0%)	N/A	0/11 (0%)
A	18/20 (90%)	46/56 (82%)	10/10 (100%)	20/23 (87%)	94/109 (86%)
MNT LS	2/4 (50%)	1/2 (50%)	1/2 (50%)	N/A	4/8 (50%)
AB	22/32 (69%)	40/54 (74%)	34/48 (71%)	70/89 (79%)	166/223 (74%)
B1	3/19 (16%)	9/28 (32%)	0/12 (0%)	17/58 (29%)	29/117 (25%)
B2	2/32 (6%)	7/32 (22%)	0/25 (0%)	8/40 (20%)	17/129 (13%)
B3	1/8 (13%)	13/62 (21%)	1/10 (10%)	3/30 (10%)	18/110 (16%)
Mixed B	N/A	N/A	N/A	3/17 (18%)	3/17 (18%)
Hyper	0/8 (0%)	N/A	N/A	N/A	0/8 (0%)
TOTAL	48/141 (34%)	119/274 (43%)	46/117 (39%)	124/296 (42%)	337/828 (41%)

NHLI: National Heart and Lung Institute; **TCGA:** The Cancer Genome Atlas; **CA:** Carcinoma; **NETT:** Neuroendocrine thymic tumour; **MNT LS:** Micronodular thymoma with lymphoid stroma; **Mixed B:** Cases of more than one B subtype; **Hyper:** Hyperplasia

8.2 WES and CNA of TETs

WES of seventeen TETs confirmed a very low mutation burden. Analysis of the unpublished TCGA WES data as well as those from our own cohort confirmed a very low recurrence of known cancer genes across all TET subtypes but particularly B thymomas. The presence of mutations in the MAP kinase pathway in approximately 25% of type and AB thymomas with the GTF2I mutation is very interesting especially as in our cohort in one of the cases the mutation was a germline GNAQ mutation raising the possibility that the co-mutation might be even higher if germline TCGA data are analysed. When combining all previous WES results and those from targeted capture by Petrini et al (106) and Zehir et al (105) the most commonly mutated gene in thymic SCCs is TP53, mutated in 14 of 46 cases (30%). It is interesting however how the frequency of mutations differs amongst groups. It is likely that the study by Zehir et al that included ten metastatic thymic SCCs had very aggressive cases overrepresented in their cohort (105). Many other know cancer genes were mutated in a small number of cases as can be seen in Tables 5.2, 5.3 and 5.4. A very similar picture was seen with germline mutations in our cohort. A noteworthy exception was a high impact insertion (chromosome 6 c29912395insGG) (p.Gly341fs) seen in *HLA-A* in four different TET subtypes (an SCC, a B3, a B2 and a B1 thymoma) making this the most frequently mutated gene in this cohort. Somatic loss-of-function mutations in HLA-A have been described in other tumour types (253). This finding needs to be validated with Sanger sequencing and expanded in the remaining TETs and cases of individuals without cancer to confirm its frequency and make sure it is not a commoner polymorphism than what is seen in ExAC.

There are a number of studies that have shown that orthogonal validation of NGS findings with Sanger sequencing is more likely to give false negative results that have been correctly identified as true positives from NGS rather than identify a false positive (186). However Sanger sequencing is still most often employed as the orthogonal technology for validation of findings from NGS. For testing involving low-percentage variants such as some somatic variants from tumour tissue, heteroplasmic mitochondrial variants, and germline mosaic variants, other approaches such as digital sequencing may be necessary (185).

CNA data was obtained through SNP genotyping of 143 TETs. Analysis of somatic copy numbers gains and losses, as expected from the WES data did not reveal many genes with recurrent CNAs. Chromosome 1q gains were seen in this study, which is in keeping with previous studies in TETs (106,122,165-170). As a likely oncogenic driver has not been identified as a cause for these gains, further analyses are required. In our cohort the known cancer gene with the most frequent copy number gains in 1q was *PDE4DIP*. Tools such as GISTIC can be used to identify 'peak' areas within chromosomes and although these have so far not given answers in this particular case may be more helpful if the entire cohort of TETs in the current study is included.

Using SNP arrays we determined that *MKI67* (chromosome 10q26.2) was the gene with the most recurrent copy number gains (30.91%) in TETs. Follow up studies would involve confirmation with a different experimental approach, such as Multiplex Ligation-dependent Probe Amplification (MLPA) (https://mlpa.com/WebForms/WebFormMain.aspx?Tag=_fNPBLedDVp38p-CxU2h0mQ..) and ideally would be performed in DNA from fresh frozen samples. Staining of TETs with antibodies to *MKI67* would also help to confirm these findings. If *MKI67* is specifically amplified in TETs it could also be a novel therapeutic target in a similar manner to HER2/neu in breast cancer (314).

Loss of *MUTYH* (chromosome 1p34.1) through homozygous deletion in TETs (18% of all TETs and in 25% of thymic SCCs) is also intriguing and this could also be confirmed with orthogonal methods to detect copy number changes such as MLPA. *MUTYH* is a human gene that encodes a DNA glycosylase, MUTYH glycosylase, and it is involved in base excision repair. Mutations in *MUTYH* in colorectal cancer leads to somatic G>T transversions in multiple genes, including APC and K-RAS (315-317). Hence, one could inspect TETs with homozygous deletions of this gene to see if these tumours had a similar increase in G>T transversions.

For validation of CNAs a number of methods have been used. These include fluorescence *in situ* hybridization (FISH), real-time quantitative PCR, CGH and genome-wide SNP arrays (318). Although they are still used they are inferior to SNP genotyping (318,319). MLPA is another very good alternative high-throughput method that allows the identification of CNAs in up to 50 genomic DNA sequences in one reaction with as little as 20ng of sample DNA {Os:2010ba}. Having gene expression data on these samples would also be really important in increasing our confidence that we have identified bona-fide driver genes {Camacho:2017fx}. This was unfortunately hindered by the fact that the majority of samples were FFPEs and hence not suitable for RNA sequencing.

8.3 Genes selected for targeted capture

In total there were 169 TETs that have been subject to WES when NHLI, Petrini et al and TCGA are considered (106,107). Mutations that were seen in targeted capture by Petrini et al were also taken into account. This included a further 26 samples taking the total to 195 samples.

Candidate genes were selected based on two main criteria: a) Known ‘cancer genes’ predicted to harbour damaging mutations in any of the three cohorts (TCGA, Petrini et al and NHLI) and b) Genes recurrently mutated across the three groups irrespective of whether they were known ‘cancer genes’ or not. Priority was given to genes that were recurrently mutated in carcinomas over thymomas. Eight of the 26 genes are known to have mutation hotspots that lead to

tumourigenesis, and for these only specific hotspot areas were targeted. A summary of the genes for targeted capture can be found in Tables 8.2 (whole genes) and 8.3 (genes with specific hotspots).

Targeted re-sequencing is being performed in additional 125 TETs of all subtypes, a list of which can be found in Appendix 2.

As this list was compiled in the absence of germline WES data and CNA data the gene panel will likely need to be expanded adding genes such as *HLA-A*, and *MUTYH*.

Table 8.2. Genes selected for Targeted Capture

#	Genes for TC	NHLI	Petrini et al	TCGA	Known cancer gene
1	TP53	x	x	x	x
2	CYLD	x	x	x	x
3	FAM230A	x		x	
4	AHNAK2	x	x	x	
5	PARP4	x		x	
6	KMT2C	x	x	x	x
7	PBRM1		x	x	
8	FOXD1			x	
9	CDKN2A		x	x	x
10	BAP1	x	x		x
11	BCOR			x	x
12	BRD4		x	x	x
13	NF1			x	x
14	BRCA2	x		x	x
15	ESRRA	x	x		
16	SLC2A3	x			x
17	GSE1	x	x		
18	EP300		x	x	x
19	MSH2	x		x	x
20	CSF1R	x	x		
21	ATRX	x	x		x
22	FANCM	x	x		
23	TRRAP		x	x	x
24	DHX37			x	
25	CHEK2	x		x	x
26	TSC1	x			x

X indicates that this group has seen mutations in that gene

Genes in red are known cancer genes

Table 8.3. Genes with targeted hotspot mutations

#	Gene	Genomic loc. of codon* affected by hotspot mut.	Hotspot_protein_loc
1	GTF2I	7:74146969-74146971	L424
2	HRAS	11:534287-534289	G12
		11:534284-534286	G13
		11:533876-533878	G60
		11:533873-533875	Q61
3	NRAS	1:115258746-115258748	G12
		1:115258743-115258745	G13
		1:115256528-115256530	Q61
4	KRAS	12:25398283-25398285	G12
		12:25398280-25398282	G13
		12:25380275-25380277	Q61
5	IDH1	2:209113111-209113113	R132
		2:209108315-209108317	V178
6	RET	10:43596062-43596064	R77
		10:43597792-43597794	R114
7	EGFR	7:55221821-55221823	A289
		7:55233042-55233044	G598
		7:55242415-55242513	Exon 19
		7:55249070-55249072	T790
		7:55259514-55259516	L858
		7:55259523-55259525	L861
8	SF3B1	2:198267482-198267484	R625
		2:198267359-198267361	K666
		2:198266832-198266834	K700

Genomic coordinates are according to GRCh37 genome version.

The coordinates indicate the location of the codon affected by the hotspot mutation. One *exception, indicated with an asterisk is the EGFR exon 19 which has multiple in-frame deletions and hence the whole exon's coordinate is noted* (target of recurrent InFrame deletions; G729 to D761). Genes in red are known cancer genes

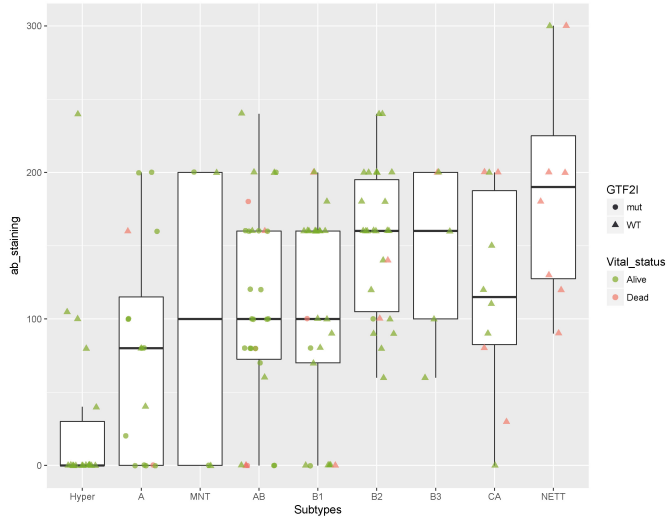
8.4 Immunohistochemistry

In the era of cancer immunotherapy multiple clinical trails are underway testing a variety of agents against established or emerging immunotherapy targets. For quite a few of those there have been no previous studies but the increasing understanding of the multiple and complex interactions between tumours and the immune system has led to IHC studies looking for the presence of specific molecules that can be targeted with potential antitumour effect. Anti-PD-1 inhibitors are the best example. In this study antibody staining on the emerging immunotherapy targets PD-1, PD-L1, PD-L2, LAG-3, TIM-3 B7-H3, VISTA and 4-BB as well another two emerging treatment targets, PDLIM1 and HGF was used on TMAs from 143 TET cases. Figure 8.1 and figure 8.2 show the epithelial and lymphocytic staining respectively for all antibodies that were positive across the TET subtypes. Epithelial PDL-1 and PD-L2 demonstrated a very similar pattern of staining with increasing scores from the more indolent to the more aggressive subtypes. The opposite trend was seen for the lymphocytic PD-1 PD-L1 and PD-L2 staining, which overall tended to be higher in the AB, B1 and B2 subtypes. These findings are likely to reflect that subtypes such as AB, B1 and B2 are made up of lymphocytes as part of the tumour whereas this in not the case for B3 thymomas and carcinomas. The significance of lymphocytic PD-L1 and PD-L2 staining vs. epithelial remains to be established. For the other antibodies trends are not as clear-cut although as discussed in detail in chapter 7, overall staining is highest in carcinomas for most antibodies even when it is completely negative in thymomas as is the case for TIM-3 and HGF. Salient findings such as the high HGF staining in MNT thymomas are worth following up in more samples.

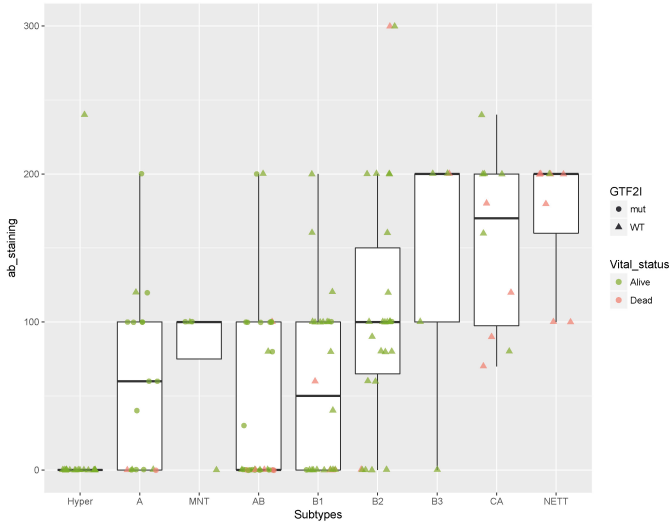
Clustering of all the samples together according to the co-expression pattern of all the positive markers gave very interesting results (Figure 7.29). Although quite a few members of the same subtype clustered together there are cases in all subtypes that behave very differently in terms of their expression of these markers and have clustered more with other TET subtypes. This is particularly important in those cases when a more indolent subtype behaves in a more aggressive manner. These findings mean that thymomas of any subtype should be screened with the appropriate antibodies if it behaves in a more aggressive fashion as it might benefit from Immunomodulatory therapy not otherwise indicated for that subtype.

Figure 8.1. Epithelial H-scores per TET subtype

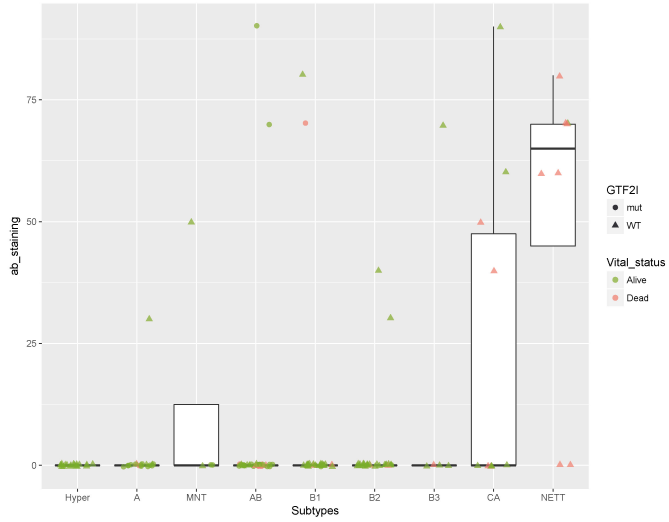
PD-L1



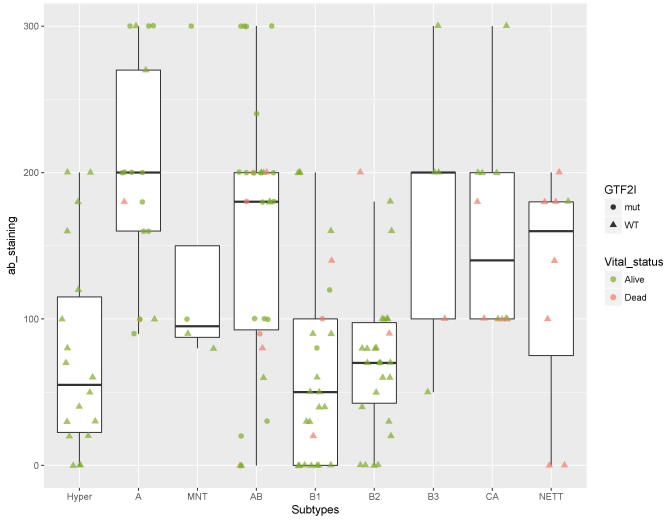
PD-L2



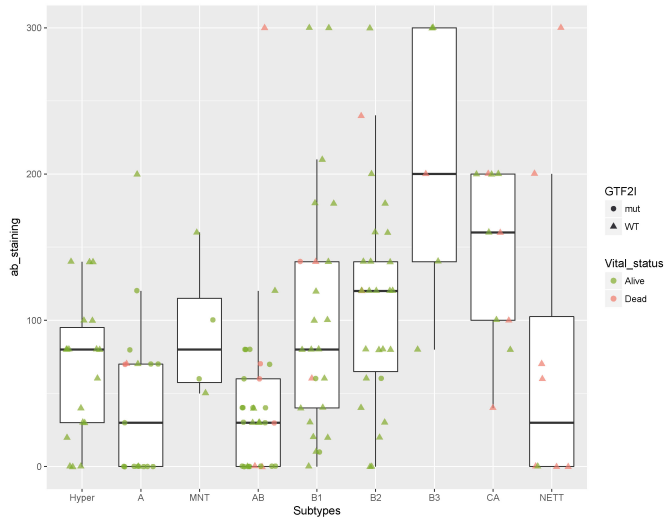
TIM-3



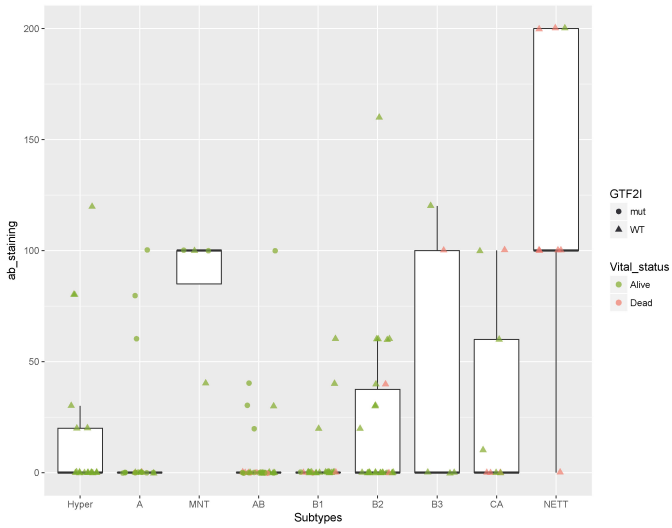
B7-H3



PDLIM1



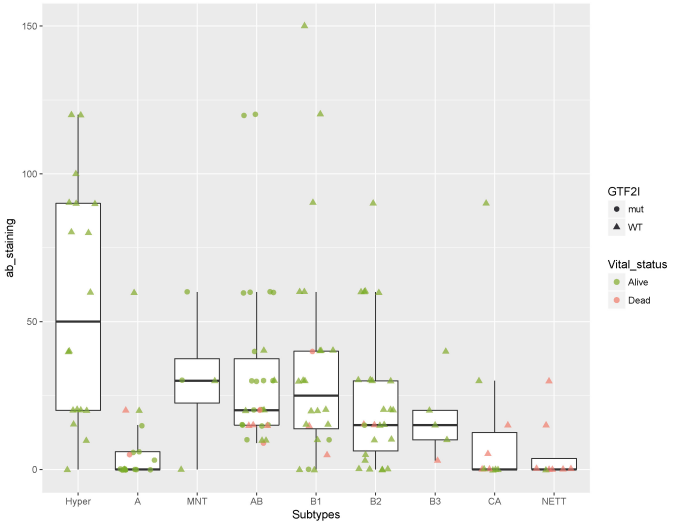
HGF



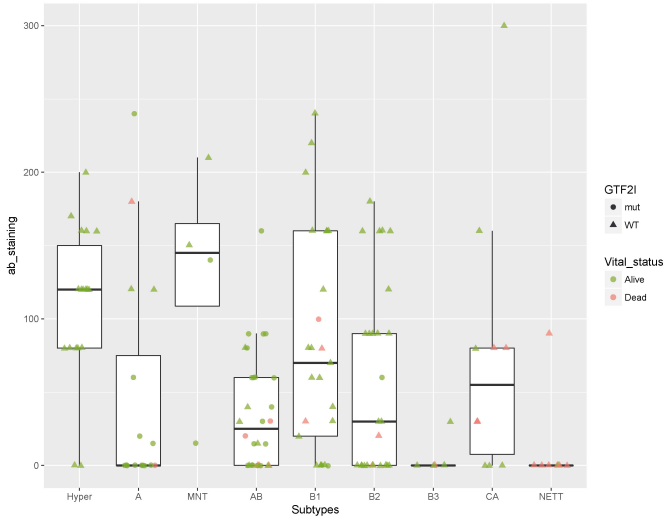
Box plots showing the distribution of epithelial H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive for PD-L1 and PD-L2, TIM-3, B7-H3, PDLIM1 and HGF

Figure 8.2. Lymphocytic H-scores per TET subtype

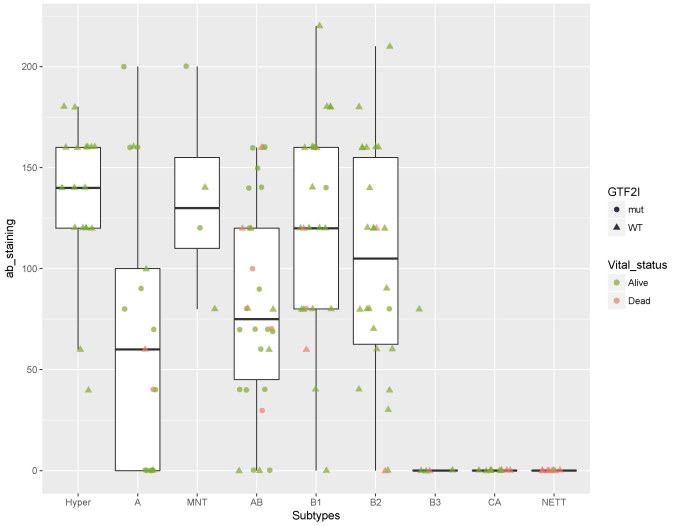
PD-1



PD-L1



PD-L2



Box plots showing the distribution of lymphocytic H-scores per TET subtype, including *GTF2I* mutation status and whether dead or alive for PD-1, PD-L1 and PD-L2

8.5 Future directions

In the future a number of functional studies will need to be performed to further validate and expand the current findings and to ensure their biological relevance. These can be done in a number of ways including the use of human-derived TET cell lines In vitro or animal models in vivo.

Cultured cancer cell lines are still widely used in vitro models and have significantly contributed to our understanding of human malignancy and cancer biology. Although cell-lines can have a significantly different mutations profile to their tissue of origin and they cannot capture the spectrum of tumour heterogeneity seen in most solid tumours, they can be used as an easier and faster means of investigating the function of candidate genes identified through molecular genetic approaches and provide a basis for evaluating targeted therapies (320). A small number of TET cell lines are already currently available commercially.

The effect of known gene mutations as well as the impact of inhibiting these effects (loss or gain of function of tumor suppressors or oncogenes respectively) with specific inhibitors can also be determined. Inducing mutations and blocking those can also give important information about the function of genes identified though WES, especially those not known to be driver mutations in other tumour types. Dominant negative mutations are particularly useful as in these cases a mutant gene abolishes that of the wild-type gene in the cell, which can be done using the antisense RNA strategy {Os:2010ba}.. This results in loss-of-function of that gene. Gain-of-function mutations can also be introduced to investigate their oncogenic potential {Os:2010ba}. Such mutations can also be introduced into the germline of many animals to allow analysis of the resultant phenotype, which are called transgenic organisms {Os:2010ba}.

Knocking down known potential tumour suppressor genes and looking for their effect on proliferation, invasion and metastasis is also key as it would be expected that these might be affected. This is of particular relevance in genes that are not fully established as tumour suppressor genes. A similar rationale can be used on potential oncogenes. Knock out mice which are missing specific genes are particularly important in establishing the impact of the loss of function of specific genes {Os:2010ba}.

The development of clustered regularly interspaced short palindromic repeats (CRISPR) and its CRISPR-associated protein (Cas9) has been a major breakthrough in regards to gene editing as it allows very precise genetic manipulations in an easy and cost-effective manner (321). These manipulations include correcting genetic mutations causing hereditary cancers. CRISPR-Cas9 is also allowing for large-scale drug screening and the generation of genetically engineered animal models that can then be used in drug studies (321). This technology is starting to be utilized in the world of immunotherapy and there is currently a clinical trial in which CRISPR/Cas9 is implemented to knock out the PD-1 gene to treat malignant melanoma (NIH Protocol #1604-1524,

“Title: Phase I Trial of Autologous T Cells Engineered to Express NYESO-1 TCR and Gene Edited to Eliminate Endogenous TCR and PD-1”; https://osp.od.nih.gov/mdocs-posts/protocol-1604-1524-title-phase-i-trial-of-autologous-t-cells-engineered-to-express-nyeso-1-tcr-and-gene-edited-to-eliminate-endogenous-tcr-and-pd-1/).

Chapter 9
Bibliography

1. Safieddine N, Keshavjee S. Anatomy of the thymus gland. Thorac Surg Clin [Internet]. 2011 May;21(2):191–5–viii. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=21477769&retmode=ref&cmd=prlinks

2. Takahama Y. Journey through the thymus: stromal guides for T-cell development and selection. Nat Rev Immunol. 2006;6(2):127–35.

3. Blackburn CC, Manley NR, Palmer DB, Boyd RL, Anderson G, Ritter MA. One for all and all for one: thymic epithelial stem cells and regeneration. Trends Immunol. 2002 Aug;23(8):391–5.

4. de Jong WK, Blaauwgeers J, Schaapveld M. Thymic epithelial tumours: a population-based study of the incidence, diagnostic procedures and therapy. European Journal of …. 2008;44(1):123–30.

5. Fornasiero A, Daniele O, Ghiotto C, Piazza M, Fiore-Donati L, Calabró F, et al. Chemotherapy for invasive thymoma. A 13-year experience. Cancer. 1991 Jul 1;68(1):30–3.

6. Ogawa K, Toita T, Uno T, Fuwa N, Kakinohana Y. Treatment and prognosis of thymic carcinoma. Cancer. 2002;94(12):3115–9.

7. Detterbeck FC, Asamura H, Crowley J. The IASLC/ITMIG thymic malignancies staging project: development of a stage classification for thymic malignancies. Journal of Thoracic …. 2013;8(12):1467–73.

8. Engels EA. Epidemiology of thymoma and associated malignancies. Journal of thoracic oncology: official publication of the …. 2010;5:S260–5.

9. Strollo DC, Rosado de Christenson ML, Jett JR. Primary mediastinal tumors. Part 1: tumors of the anterior mediastinum. Chest. American College of Chest Physicians; 1997 Aug;112(2):511–22.

10. Marx A, Ströbel P, Badve SS. ITMIG Consensus Statement on the Use of the WHO Histological Classification of Thymoma and Thymic Carcinoma: Refined Definitions, Histological Criteria, and …. Journal of Thoracic Oncology. 2014;9(5):596–611.

11. Travis WD, Brambilla E, Müller-Hermelink HK. Pathology and genetics of tumours of the lung, pleura, thymus and heart/edited by William D. Travis...[et al.]. 2004.

12. Kondo K, Monden Y. Therapy for thymic epithelial tumors: a clinical study of 1,320 patients from Japan. Ann Thorac Surg. 2003 Sep;76(3):878–84–discussion884–5.

13. Girard N, Ruffini E, Marx A, Faivre-Finn C, Peters S, ESMO Guidelines Committee. Thymic epithelial tumours: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Vol. 26 Suppl 5, Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2015. pp. v40–55.

14. Liu WT, Kao KP, Liu YC, Chang KS. Human foamy virus genome in the thymus of myasthenia gravis patients. Zhonghua Min Guo Wei Sheng Wu Ji Mian Yi Xue Za Zhi. 1996 Aug;29(3):162–5.

15. Saïb A, Canivet M, Giron ML, Bolgert F, Valla J, Lagaye S, et al. Human foamy virus infection in myasthenia gravis. Lancet. 1994 Mar 12;343(8898):666.

16. Weis C-A, Yao X, Deng Y, Detterbeck FC, Marino M, Nicholson AG, et al. The impact of thymoma histotype on prognosis in a worldwide database. J Thorac Oncol. 2015 Feb;10(2):367–72.

17. Moran CA, Weissferdt A, Kalhor N, Solis LM, Behrens C, Wistuba II, et al. Thymomas I: a clinicopathologic correlation of 250 cases with emphasis on the World Health Organization schema. Am J Clin Pathol [Internet]. 2012 Mar;137(3):444–50. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22338057&retmode=ref&cmd=prlinks

18. Shelly S, Agmon-Levin N, Altman A. Thymoma and autoimmunity. Cellular & molecular …. 2011;8(3):199–202.

19. Wang W, Zhang H, Lu R, Smith J, Conces L, Badve SS. Paraneoplastic syndrome and survival in thymic epithelial tumors (TET): The Indiana University experience. 2017.

20. Girard N. Thymic malignancies: Twisting between autoimmunity and immunotherapy. Lung Cancer. Elsevier; 2017 May 11;0(0):68–70.

21. Falkson CB, Bezjak A, Darling G, Gregg R, Malthaner R, Maziak DE, et al. The Management of Thymoma: A Systematic Review and Practice Guideline. Journal of Thoracic Oncology. Elsevier; 2009 Jul 1;4(7):911–9.

22. Marchevsky A, Marx A, Ströbel P, Suster S, Venuta F, Marino M, et al. Policies and reporting guidelines for small biopsy specimens of mediastinal masses. J Thorac Oncol. Elsevier; 2011 Jul;6(7 Suppl 3):S1724–9.

23. Rosai J, Sobin LH. Histological Typing of Tumours of the Thymus - Springer. 1999.

24. Okumura M, Ohta M, Tateyama H, Nakagawa K. The World Health Organization histologic classification system reflects the oncologic behavior of thymoma. Cancer. 2002;94(2):581–3.

25. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol. Elsevier; 2015 Sep;10(9):1243–60.

26. Marx A, Chan JKC, Coindre J-M, Detterbeck F, Girard N, Harris NL, et al. The 2015 World Health Organization Classification of Tumors of the Thymus: Continuity and Changes. J Thorac Oncol. 2015 Oct;10(10):1383–95.

27. International Agency for Research on Cancer. Who Classification of Tumours of the Lung, Pleura, Thymus and Heart. World Health Organization; 2015. 1 p.

28. Chen G, Marx A, Hu CW, Yong J, Puppe B. New WHO histologic classification predicts prognosis of thymic epithelial tumors. Cancer. 2002;95(2):420–9.

29. Bürger T, Schaefer I-M, Küffer S, Bohnenberger H, Reuter-Jessen K, Chan JK-C, et al. Metastatic type A thymoma: morphological and genetic correlation. Histopathology. 2016 Dec 7.

30. Mneimneh WS, Gökmen-Polar Y, Kesler KA, Loehrer PJ, Badve S. Micronodular thymic neoplasms: case series and literature review with emphasis on the spectrum of differentiation. Mod Pathol. 2015 Nov;28(11):1415–27.

31. Weissferdt A, Moran CA. Micronodular thymic carcinoma with lymphoid hyperplasia: a clinicopathological and immunohistochemical study of five cases. Mod Pathol. 2012 Jul;25(7):993–9.

32. Jain RK, Mehta RJ, Henley JD, Kesler KA, Loehrer PJ, Badve S. WHO types A and AB thymomas: not always benign. Mod Pathol. 2010 Dec;23(12):1641–9.

33. Lamarca A, Moreno V, Feliu J. Thymoma and thymic carcinoma in the target therapies era. Cancer Treat Rev. 2013;39(5):413–20.

34. Girard N. Thymic epithelial tumours: from basic principles to individualised treatment strategies. European Respiratory Review. 2013;22(127):75–87.

35. Kondo K, Yoshizawa K, Tsuyuguchi M, Kimura S, Sumitomo M, Morita J, et al. WHO histologic classification is a prognostic indicator in thymoma. Ann Thorac Surg. 2004 Apr;77(4):1183–8.

36. Masaoka A, Monden Y, Nakahara K, Tanioka T. Follow-up study of thymomas with special reference to their clinical stages. Cancer. 1981 Dec 1;48(11):2485–92.

37. Nakagawa K, Asamura H, Matsuno Y, Suzuki K, Kondo H, Maeshima A, et al. Thymoma: a clinicopathologic study based on the new World Health Organization classification. The Journal of Thoracic and Cardiovascular Surgery. 2003 Oct;126(4):1134–40.

38. Koga K, Matsuno Y, Noguchi M, Mukai K, Asamura H, Goya T, et al. A review of 79 thymomas: modification of staging system and reappraisal of conventional division into invasive and non-invasive thymoma. Pathol Int. 1994 May;44(5):359–67.

39. Detterbeck FC, Nicholson AG, Kondo K. The Masaoka-Koga stage classification for thymic malignancies: clarification and definition of terms. Journal of Thoracic Oncology. 2011;6:S1710–6.

40. Ritter JH, Wick MR. Primary carcinomas of the thymus gland. 1999 Feb;16(1):18–31. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=10355651&retmode=ref&cmd=prlinks

41. Blumberg D, Burt ME, Bains MS, Downey RJ. Thymic Carcinoma: Current Staging Does Not Predict Prognosis. The Journal of thoracic …. 1998;115(2):303–9.

42. Kondo K, Monden Y. Thymoma and myasthenia gravis: a clinical study of 1,089 patients from Japan. Ann Thorac Surg. 2005;79(1):219–24.

43. Amin MB, Edge S, Greene FL, Byrd DR, Brookland RK, Washington MK, et al. AJCC Cancer Staging Manual. Springer; 2016. 1 p.

44. Detterbeck FC, Stratton K, Giroux D, Asamura H, Crowley J, Falkson C, et al. The IASLC/ITMIG Thymic Epithelial Tumors Staging Project: proposal for an evidence-based stage classification system for the forthcoming (8th) edition of the TNM classification of malignant tumors. J Thorac Oncol. Elsevier; 2014 Sep;9(9 Suppl 2):S65–72.

45. Detterbeck F. International thymic malignancies interest group: a way forward. Journal of Thoracic Oncology. 2010;5(10):S365–70.

46. Detterbeck F, Youssef S, Ruffini E, Okumura M. A review of prognostic factors in thymic malignancies. J Thorac Oncol. Elsevier; 2011 Jul;6(7 Suppl 3):S1698–704.

47. Ahmad U, Yao X, Detterbeck F, Huang J, Antonicelli A, Filosso PL, et al. Thymic carcinoma outcomes and prognosis: results of an international analysis. The Journal of Thoracic and Cardiovascular Surgery. 2015 Jan;149(1):95–100–101.e1–2.

48. Evoli A, Lancaster E. Paraneoplastic disorders in thymoma patients. J Thorac Oncol. Elsevier; 2014 Sep;9(9 Suppl 2):S143–7.

49. Marx A, Hohenberger P, Hoffmann H, Pfannschmidt J, Schnabel P, Hofmann H-S, et al. The autoimmune regulator AIRE in thymoma biology: autoimmunity and beyond. J Thorac Oncol. Elsevier; 2010 Oct;5(10 Suppl 4):S266–72.

50. Evoli A, Punzi C, Marsili F, Di Schino C, Cesario A, Galetta D, et al. Extrathymic malignancies in patients with thymoma. Ann Oncol. 2004 Apr;15(4):692–3.

51. Bruserud Ø, Oftedal BE, Wolff AB, Husebye ES. AIRE-mutations and autoimmune disease. Current Opinion in Immunology. 2016 Dec;43:8–15.

52. Burbelo PD, Browne SK, Sampaio EP. Anti-cytokine autoantibodies are associated with opportunistic infection in patients with thymic neoplasia. …. 2010;116(23):4848–58.

53. Anderson MS, Su MA. Aire and T cell development. Current Opinion in Immunology. 2011;23(2):198–206.

54. Avidan N, Le Panse R, Berrih-Aknin S, Miller A. Genetic basis of myasthenia gravis - a comprehensive review. J Autoimmun. 2014 Aug;52:146–53.

55. Zagoriti Z, Kambouris ME, Patrinos GP, Tzartos SJ, Poulas K. Recent advances in genetic predisposition of myasthenia gravis. Biomed Res Int. 2013;2013(11):404053–12.

56. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010 Aug 19;363(8):711–23.

57. Peggs KS, Quezada SA. Ipilimumab: attenuation of an inhibitory immune checkpoint improves survival in metastatic melanoma. Expert Rev Anticancer Ther. 2010 Nov;10(11):1697–701.

58. Weber JS, Kähler KC, Hauschild A. Management of immune-related adverse events and kinetics of response with ipilimumab. Journal of Clinical Oncology. 2012 Jul 20;30(21):2691–7.

59. Wolchok JD, Hoos A, O'Day S, Weber JS, Hamid O, Lebbé C, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res. 2009 Dec 1;15(23):7412–20.

60. Iwai Y, Hamanishi J, Chamoto K, Honjo T. Cancer immunotherapies targeting the PD-1 signaling pathway. J Biomed Sci. 2017 Apr 4;24(1):26.

61. Herbst RS, Baas P, Kim D-W, Felip E, Pérez-Gracia JL, Han J-Y, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet. 2016 Apr 9;387(10027):1540–50.

62. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. N Engl J Med. 2016 Nov 10;375(19):1823–33.

63. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. N Engl J Med. 2015 Oct 22;373(17):1627–39.

64. Brahmer J, Reckamp KL, Baas P, Crino L, Eberhardt WEE, Poddubskaya E, et al. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. N Engl J Med. 2015 Jul 9;373(2):123–35.

65. Fehrenbacher L, Spira A, Ballinger M, Kowanetz M, Vansteenkiste J, Mazieres J, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. Lancet. 2016 Mar 9.

66. Robert C, Schachter J, Long GV, Arance A, Grob JJ, Mortier L, et al. Pembrolizumab versus Ipilimumab in Advanced Melanoma. N Engl J Med [Internet]. 2015 Jun 25;372(26):2521–32. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=25891173&retmode=ref&cmd=prlinks

67. Nanda R, Chow LQM, Dees EC, Berger R, Gupta S, Geva R, et al. Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. Journal of Clinical Oncology. 2016 Jul 20;34(21):2460–7.

68. Muro K, Chung HC, Shankaran V, Geva R, Catenacci D, Gupta S, et al. Pembrolizumab for patients with PD-L1-positive advanced gastric cancer (KEYNOTE-012): a multicentre, open-label, phase 1b trial. Lancet Oncol. 2016 May 3.

69. Ott PA, Bang Y-J, Berton-Rigaud D, Elez E, Pishvaian MJ, Rugo HS, et al. Safety and Antitumor Activity of Pembrolizumab in Advanced Programmed Death Ligand 1-Positive Endometrial Cancer: Results From the KEYNOTE-028 Study. Journal of Clinical Oncology. 2017 May 10;35(22):JCO2017725952–2541.

70. Ott PA, Piha-Paul SA, Munster P, Pishvaian MJ, van Brummelen EMJ, Cohen RB, et al. Safety and antitumor activity of the anti-PD-1 antibody pembrolizumab in patients with recurrent carcinoma of the anal canal. Ann Oncol. 2017 May 1;28(5):1036–41.

71. Alley EW, Lopez J, Santoro A, Morosky A, Saraf S, Piperdi B, et al. Clinical safety and activity of pembrolizumab in patients with malignant pleural mesothelioma (KEYNOTE-028): preliminary results from a non-randomised, open-label, phase 1b trial. Lancet Oncol. 2017 May;18(5):623–30.

72. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. Cancer Cell. Elsevier; 2015 Apr 13;27(4):450–61.

73. Taube JM, Klein A, Brahmer JR, Xu H, Pan X, Kim JH, et al. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. Clin Cancer Res. American Association for Cancer Research; 2014 Oct 1;20(19):5064–74.

74. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med. 2014 Dec 4;371(23):2189–99.

75. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. American Association for the Advancement of Science; 2015 Apr 3;348(6230):124–8.

76. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. N Engl J Med. 2015 Jun 25;372(26):2509–20.

77. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, et al. Mismatch-repair deficiency predicts response of solid tumors to PD-1 blockade. Science. 2017 Jun 8;357(6349):409–13.

78. Inoue Y, Yoshimura K, Mori K, Kurabe N, Kahyo T, Mori H, et al. Clinical significance of PD-L1 and PD-L2 copy number gains in non-small-cell lung cancer. Oncotarget. 2016 May 31;7(22):32113–28.

79. Heineman TE, Widman A, Kuan EC, St John M. The genetic landscape of programmed death ligand-1 (PD-L1) alterations in head and neck cancer. Laryngoscope Investigative Otolaryngology. 2017 May 17;2(3):99–103.

80. Nguyen LT, Ohashi PS. Clinical blockade of PD1 and LAG3--potential mechanisms of action. Nat Rev Immunol [Internet]. 2015 Jan;15(1):45–56. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=25534622&retmode=ref&cmd=prlinks

81. Yearley JH, Gibson C, Yu N, Moon C, Murphy E, Juco J, et al. PD-L2 Expression in Human Tumors: Relevance to Anti-PD-1 Therapy in Cancer. Clin Cancer Res. American Association for Cancer Research; 2017 Jun 14;23(12):3158–67.

82. Katsuya Y, Fujita Y, Horinouchi H, Ohe Y, Watanabe S-I, Tsuta K. Immunohistochemical status of PD-L1 in thymoma and thymic carcinoma. Lung Cancer. 2015 May;88(2):154–9.

83. Padda SK, Riess JW, Schwartz EJ, Tian L, Kohrt HE, Neal JW, et al. Diffuse high intensity PD-L1 staining in thymic epithelial tumors. J Thorac Oncol [Internet]. 2015 Mar;10(3):500–8. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=25402569&retmode=ref&cmd=prlinks

84. Katsuya Y, Horinouchi H, Asao T, Kitahara S, Goto Y, Kanda S, et al. Expression of programmed death 1 (PD-1) and its ligand (PD-L1) in thymic epithelial tumors: Impact on treatment efficacy and alteration in expression after chemotherapy. Lung Cancer. Elsevier; 2016 Sep;99:4–10.

85. Yokoyama S, Miyoshi H, Nakashima K, Shimono J, Hashiguchi T, Mitsuoka M, et al. Prognostic Value of Programmed Death Ligand 1 and Programmed Death 1 Expression in Thymic Carcinoma. Clin Cancer Res. 2016 Sep 15;22(18):4727–34.

86. Owen D, Chu B, Lehman A, Annamalai L, Yearley J, Shilo K, et al. P2.04-020 Expression Patterns and Prognostic Value of PD-L1 and PD-1 in Thymoma and Thymic Carcinoma. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S1008–9.

87. Weissferdt A, Fujimoto J, Kalhor N, Rodriguez J, Bassett R, Wistuba II, et al. Expression of PD-1 and PD-L1 in thymic epithelial neoplasms. Mod Pathol. 2017 Jun;30(6):826–33.

88. Berardi R, Goteri G, Brunelli A, Pagliaretta S, Paolucci V, Caramanti M, et al. P2.04-017 Prognostic Relevance of PD-1/PD-L1 Pathway in Thymic Malignancies with Combined Immunohystochemical and Biomolecular Approach. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S1006–7.

89. Yokoyama S, Miyoshi H, Nishi T, Hashiguchi T, Mitsuoka M, Takamori S, et al. Clinicopathologic and Prognostic Implications of Programmed Death Ligand 1 Expression in Thymoma. Ann Thorac Surg. Elsevier; 2016 Apr;101(4):1361–9.

90. Marchevsky AM, Walts AE. PD-L1, PD-1, CD4, and CD8 expression in neoplastic and nonneoplastic thymus. Human Pathology. 2017 Feb;60:16–23.

91. Naidoo J, Steele K, Sima CS, Yuan J, Robbins PB. Defining the immunologic phenotype of thymic epithelial tumors. 2015.

92. Hirsch FR, McElhinny A, Stanforth D, Ranger-Moore J, Jansson M, Kulangara K, et al. PD-L1 Immunohistochemistry Assays for Lung Cancer: Results from Phase 1 of the Blueprint PD-L1 IHC Assay Comparison Project. J Thorac Oncol. 2017 Feb;12(2):208–22.

93. Scheel AH, Dietel M, Heukamp LC, Jöhrens K, Kirchner T, Reu S, et al. Harmonized PD-L1 immunohistochemistry for pulmonary squamous-cell and adenocarcinomas. Mod Pathol. 2016 Oct;29(10):1165–72.

94. Shimanovsky A. **Immunohistochemical staining for programmed cell-death ligand 1 (PD-L1) in malignant thymoma and thymic carcinoma.** 2017.

95. Ströbel P, Bauer A, Puppe B, Kraushaar T. Tumor recurrence and survival in patients treated for thymomas and thymic squamous cell carcinomas: a retrospective analysis. Journal of Clinical …. 2004;22(8):1501–9.

96. Detterbeck FC. Evaluation and treatment of stage I and II thymoma. J Thorac Oncol. 2010 Oct;5(10 Suppl 4):S318–22.

97. Girard N, Merveilleux Du Vignaux C. ED15.02 Chemotherapy and Targeted Therapies of Thymic Malignancies. Journal of Thoracic Oncology. 2017 Jan 1;12(1):S67–8.

98. Riely GJ, Huang J. Induction therapy for locally advanced thymoma. J Thorac Oncol. 2010 Oct;5(10 Suppl 4):S323–6.

99. Thomas A, Rajan A, Berman A, Tomita Y, Brzezniak C, Lee M-J, et al. Sunitinib in patients with chemotherapy-refractory thymoma and thymic carcinoma: an open-label phase 2 trial. Lancet Oncol. 2015 Feb;16(2):177–86.

100. Zucali PA, De Pas TM, Palmieri G, Favaretto AG. Phase II study of everolimus in patients with thymoma and thymic carcinoma previously treated with cisplatin-based chemotherapy. 2014.

101. Cho J, Ahn MJ, Yoo KH, Lee H, Kim HK, Heo MH. A phase II study of pembrolizumab for patients with previously treated advanced thymic epithelial tumor. 2017.

102. Giaccone G, Thompson J, McGuire C, Manning M. Pembrolizumab in patients with recurrent thymic carcinoma: Results of a phase II study. 2017.

103. Giaccone G, Thompson J, Crawford J, McGuire C. A phase II study of pembrolizumab in patients with recurrent thymic carcinoma. 2016.

104. Rajan A. Safety and clinical activity of anti-programmed death-ligand 1 (PD-L1) antibody (ab) avelumab (MSB0010718C) in advanced thymic epithelial tumors (TETs). | 2016 ASCO Annual Meeting | Abstracts | Meeting Library [Internet]. meetinglibrary.asco.org. 2016 [cited 2017 Apr 27]. Available from: http://meetinglibrary.asco.org/content/169986-176

105. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med. 2017 Jun;23(6):703–13.

106. Petrini I, Meltzer PS, Kim I-K, Lucchi M, Park K-S, Fontanini G, et al. A specific missense mutation in GTF2I occurs at high frequency in thymic epithelial tumors. Nat Genet. 2014 Aug;46(8):844–9.

107. Loehrer PJ, Tsao AS, Radovich M, Felau I. Molecular analysis of thymic epithelial tumors (TETs): A report by The Cancer Genome Atlas (TCGA) research network. 2017.

108. Enkner F, Pichlhöfer B, Zaharie AT, Krunic M, Holper TM, Janik S, et al. Molecular Profiling of Thymoma and Thymic Carcinoma: Genetic Differences and Potential Novel Therapeutic Targets. Pathol Oncol Res [Internet]. 2017 Jul;23(3):551–64. Available from: http://link.springer.com/10.1007/s12253-016-0144-8

109. Girard N. Chemotherapy and targeted agents for thymic malignancies. Expert Rev Anticancer Ther. 2012 May;12(5):685–95.

110. Henley JD, Cummings OW, Loehrer PJ Sr, Loehrer PJ. Tyrosine kinase receptor expression in thymomas. 2004 Apr 1;130(4):222–4.

111. Pan C-C, Chen PC-H, Chiang H. KIT (CD117) is frequently overexpressed in thymic carcinomas but is absent in thymomas. J Pathol. 2004 Mar;202(3):375–81.

112. Schirosi L, Nannini N, Nicoli D, Cavazza A. Activating c-KIT mutations in a subset of thymic carcinoma and response to different c-KIT inhibitors. Annals of …. 2012;23(9):2409–14.

113. Palmieri G, Marino M, Buonerba C, Federico P. Imatinib mesylate in thymic epithelial malignancies. Cancer chemotherapy …. 2012;69(2):309–15.

114. J T Salter DLCYPJLSLR, Chiorean EG. Imatinib for the treatment of thymic carcinoma -- Salter et al. 26 (15 Supplement): 8116 -- ASCO Meeting Abstracts [Internet]. [cited 2014 Jun 19]. Available from: http://meeting.ascopubs.org/cgi/content/short/26/15_suppl/8116

115. Giaccone G, Rajan A, Ruijter R, Smit E. Imatinib mesylate in patients with WHO B3 thymomas and thymic carcinomas. Journal of Thoracic Oncology. 2009;4(10):1270–3.

116. Heinrich MC, Maki RG, Corless CL. Primary and secondary kinase genotypes correlate with the biological and clinical activity of sunitinib in imatinib-resistant gastrointestinal stromal tumor. Journal of clinical …. 2008;26(33):5352–9.

117. Bisagni G, Rossi G, Cavazza A, Sartori G, Gardini G, Boni C. Long lasting response to the multikinase inhibitor bay 43-9006 (Sorafenib) in a heavily pretreated metastatic thymic carcinoma. J Thorac Oncol. 2009 Jun;4(6):773–5.

118. Lynch TJ, Bell DW, Sordella R, Gurubhagavatula S, Okimoto RA, Brannigan BW, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med. 2004 May 20;350(21):2129–39.

119. Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 2004 Jun 4;304(5676):1497–500.

120. Shepherd FA, Rodrigues Pereira J, Ciuleanu T, Tan EH, Hirsh V, Thongprasert S, et al. Erlotinib in previously treated non-small-cell lung cancer. N Engl J Med. 2005 Jul 14;353(2):123–32.

121. Yoh K, Nishiwaki Y, Ishii G, Goto K, Kubota K. Mutational status of< i> EGFR</i> and< i> KIT</i> in thymoma and thymic carcinoma. Lung Cancer. 2008;62(3):316–20.

122. Girard N, Shen R, Guo T, Zakowski MF, Heguy A, Riely GJ, et al. Comprehensive Genomic Analysis Reveals Clinically Relevant Molecular Distinctions between Thymic

Carcinomas and Thymomas. Clin Cancer Res. American Association for Cancer Research; 2009 Nov 15;15(22):6790–9.

123. Aisner SC, Dahlberg S, Hameed MR, Ettinger DS, Schiller JH, Johnson DH, et al. Epidermal growth factor receptor, C-kit, and Her2/neu immunostaining in advanced or recurrent thymic epithelial neoplasms staged according to the 2004 World Health Organization in patients treated with octreotide and prednisone: an Eastern Cooperative Oncology Group study. J Thorac Oncol. 2010 Jun;5(6):885–92.

124. Girard N. Thymic tumors: relevant molecular data in the clinic. Journal of Thoracic Oncology. 2010;5:S291–5.

125. Weissferdt A, Lin H, Woods D, Tang X, Fujimoto J. HER family receptor and ligand status in thymic carcinoma. Lung Cancer. 2012;77(3):515–21.

126. Meister M, Schirmacher P, Dienemann H. Mutational status of the epidermal growth factor receptor (EGFR) gene in thymomas and thymic carcinomas. Cancer Letters. 2007;248(2):186–91.

127. Suzuki E, Sasaki H, Kawano O, Endo K. Expression and mutation statuses of epidermal growth factor receptor in thymic epithelial tumors. Japanese journal of …. 2006;36(6):351–6.

128. Ionescu DN, Sasatomi E, Cieply K, Nola M, Dacic S. Protein expression and gene amplification of epidermal growth factor receptor in thymomas. Cancer. 2005;103(3):630–6.

129. Hirabayashi K, Zamboni G, Nishi T, Tanaka A, Kajiwara H, Nakamura N. Histopathology of gastrointestinal neuroendocrine neoplasms. Front Oncol. 2013;3:2.

130. Yamaguchi H, Soda H, Kitazaki T, Tsukamoto K. Thymic carcinoma with epidermal growth factor receptor gene mutations. Vol. 52, Lung cancer (Amsterdam, Netherlands). Lung Cancer; 2006. 2 p.

131. Rajan A, Giaccone G. Targeted therapy for advanced thymic tumors. J Thorac Oncol. 2010 Oct;5(10 Suppl 4):S361–4.

132. Kelly RJ, Petrini I, Rajan A, Wang Y. Thymic malignancies: from clinical management to targeted therapies. Journal of clinical …. 2011;29(36):4820–7.

133. Chen Y, Gharwan H, Thomas A. Novel biologic therapies for thymic epithelial tumors. Front Oncol. 2014;4.

134. Nakagiri T, Funaki S, Kadota Y, Takeuchi Y. Does Gefitinib Have Effects on EGFR Mutation-Positive Thymoma?—Case Report of Thymoma Recurrence. Annals of Thoracic and …. 2014.

135. Christodoulou C, Murray S, Dahabreh J. Response of malignant thymoma to erlotinib. Vol. 19, Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. Annals of …; 2008. 2 p.

136. Palmieri G, Marino M, Salvatore M, Budillon A, Meo G, Caraglia M, et al. Cetuximab is an active treatment of metastatic and chemorefractory thymoma. Front Biosci. 2007;12(1):757–61.

137. Farina G, Garassino MC, Gambacorta M, La Verde N, Gherardi G, Scanni A. Response of thymoma to cetuximab. Lancet Oncology. 2007 May;8(5):449–50.

138. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005 Oct 20;353(16):1659–72.

139. Ciardiello F, Tortora G. EGFR antagonists in cancer treatment. N Engl J Med. 2008 Mar 13;358(11):1160–74.

140. Pao W, Wang TY, Riely GJ, Miller VA, Pan Q. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. PLoS medicine. 2005;2(1):e17.

141. Tomita M, Matsuzaki Y, Edagawa M, Maeda M. Correlation between tumor angiogenesis and invasiveness in thymic epithelial tumors. The Journal of thoracic …. 2002;124(3):493–8.

142. Kaira K, Oriuchi N, Imai H, Shimizu K. L-type amino acid transporter 1 (LAT1) is frequently expressed in thymic carcinomas but is absent in thymomas. Journal of surgical …. 2009;99(7):433–8.

143. Li X, Chen Q, Huang W, Ye Y. Response to sorafenib in cisplatin-resistant thymic carcinoma: a case report. Med Oncol. 2009;26(2):157–60.

144. Ströbel P, Bargou R, Wolff A, Spitzer D, Manegold C, Dimitrakopoulou-Strauss A, et al. Sunitinib in metastatic thymic carcinomas: laboratory findings and initial clinical experience. Br J Cancer. 2010 Jul 13;103(2):196–200.

145. Fiedler W, Giaccone G, Lasch P. Phase I trial of SU14813 in patients with advanced solid malignancies. 2011;22(1):195–201.

146. Karube Y, Kobayashi S, Maeda S, Sado T, Ishihama H, Chida M. Tumor-related gene expression levels in thymic carcinoma and Type B3 thymoma. J Cardiothorac Surg. 7 ed. BioMed Central; 2016 May 26;11(1):85.

147. Wheler J, Hong D, Swisher SG, Falchook G, Tsimberidou AM, Helgason T, et al. Thymoma patients treated in a phase I clinic at MD Anderson Cancer Center: responses to mTOR inhibitors and molecular analyses. Oncotarget. 2013 Jun;4(6):890–8.

148. Steele NL, Plumb JA, Vidal L, Tjornelund J. A phase 1 pharmacokinetic and pharmacodynamic study of the histone deacetylase inhibitor belinostat in patients with advanced solid tumors. Clinical Cancer Research. 2008;14(3):804–10.

149. Giaccone G, Rajan A, Berman A, Kelly RJ. Phase II study of belinostat in patients with recurrent or refractory advanced thymic epithelial tumors. Journal of Clinical Oncology. 2011;29(15):2052–9.

150. Asamura H, Matsuno Y, SUZUKI K. Thymoma: a clinicopathologic study based on the new World Health Organization classification. The Journal of thoracic …. 2003;126(4):1134–40.

151. Gong Y, Yao E, Shen R, Goel A, Arcila M. High expression levels of total IGF-1R and sensitivity of NSCLC cells in vitro to an anti-IGF-1R antibody (R1507). Blagosklonny MV, editor. PLoS ONE. 2009;4(10):e7273.

152. Breinig M, Mayer P, Harjung A, Goeppert B, Malz M. Heat shock protein 90-sheltered overexpression of insulin-like growth factor 1 receptor contributes to malignancy of thymic epithelial tumors. Clinical Cancer …. 2011;17(8):2237–49.

153. Zucali PA, Petrini I, Lorenzi E, Merino M, Cao L. Insulin-like growth factor-1 receptor and phosphorylated AKT-serine 473 expression in 132 resected thymomas and thymic carcinomas. Cancer. 2010;116(20):4686–95.

154. Haluska P, Shaw HM, Batzel GN, Yin D, Molina JR. Phase I Dose Escalation Study of the Anti–Insulin-Like Growth Factor-I Receptor Monoclonal Antibody CP-751,871 in Patients with Refractory Solid Tumors. Clinical Cancer Research. 2007;13(19):5834–40.

155. Rajan A, Carter CA, Berman A, Cao L, Kelly RJ. Cixutumumab for patients with recurrent or refractory advanced thymic epithelial tumours: a multicentre, open-label, phase 2 trial. Lancet Oncol. 2014;15(2):191–200.

156. Petrini I, Meltzer PS, Zucali PA, Luo J, Lee C. Copy number aberrations of BCL2 and CDKN2A/B identified by array-CGH in thymic epithelial tumors. Cell death & …. 2012;3(7):e351.

157. Hirabayashi H, Fujii Y, Sakaguchi M, Tanaka H, Yoon HE, Komoto Y, et al. p16INK4, pRB, p53 and cyclin D1 expression and hypermethylation of CDKN2 gene in thymoma and thymic carcinoma. Int J Cancer. 1997 Nov 27;73(5):639–44.

158. Wakelee HA, Burns M, Gubens M, Barbeau S. A phase II study of saracatinib (AZD0530), a SRC inhibitor, administered orally daily to patients (pts) with advanced thymic malignancies (ATM). … Thymic Malignancy Interest …; 2010.

159. Hayashi Y, Ishii N, Obayashi C, Jinnai K, Hanioka K, Imai Y, et al. Thymoma: tumour type related to expression of epidermal growth factor (EGF), EGF-receptor, p53, v-erb B and ras p21. Virchows Arch. 1995;426(1):43–50.

160. Pan C-C, Chen PC-H, Wang L-S, Lee J-Y, Chiang H. Expression of apoptosis-related markers and HER-2/neu in thymic epithelial tumours. Histopathology. 2003 Aug;43(2):165–72.

161. Tateyama H, Eimoto T, Tada T, Mizuno T, Inagaki H, Hata A, et al. p53 protein expression and p53 gene mutation in thymic epithelial tumors. An immunohistochemical and DNA sequencing study. Am J Clin Pathol. 1995 Oct;104(4):375–81.

162. Kaira K, Serizawa M, Koh Y, Miura S, Kaira R, Abe M, et al. Expression of excision repair cross-complementation group 1, breast cancer susceptibility 1, and β III-tubulin in thymic epithelial tumors. J Thorac Oncol. 2011 Mar;6(3):606–13.

163. Kaira K, Serizawa M, Koh Y, Miura S, Kaira R, Abe M. Expression of thymidylate synthase, orotate phosphoribosyltransferase and dihydropyrimidine dehydrogenase in thymic epithelial tumors. Lung Cancer. 2011;74(3):419–25.

164. Weirich G, Schneider P, Fellbaum C, Brauch H, Nathrath W, Scholz M, et al. p53 alterations in thymic epithelial tumours. Virchows Arch. 1997 Jul;431(1):17–23.

165. Petrini I, Wang Y, Zucali PA, Lee HS, Pham T, Voeller D, et al. Copy number aberrations of genes regulating normal thymus development in thymic epithelial tumors. Clin Cancer Res. 2013 Apr 15;19(8):1960–71.

166. Lee GY, Yang WI, Jeung HC, Kim SC, Seo MY. Genome-wide genetic aberrations of thymoma using cDNA microarray based comparative genomic hybridization. BMC …. 2007;8(1):305.

167. Inoue M, Starostik P, Zettl A, Ströbel P, Schwarz S, Scaravilli F, et al. Correlating genetic aberrations with World Health Organization-defined histology and stage across the spectrum of thymomas. Cancer Res. 2003 Jul 1;63(13):3708–15.

168. Penzel R, Hoegel J, Schmitz W, Blaeker H, Morresi-Hauf A, Aulmann S, et al. Clusters of chromosomal imbalances in thymic epithelial tumours are associated with the WHO classification and the staging system according to Masaoka. International Journal of Cancer. Wiley Subscription Services, Inc., A Wiley Company; 2003 Jul 1;105(4):494–8.

169. Zettl A, Ströbel P, Wagner K, Katzenberger T. Recurrent genetic aberrations in thymoma and thymic carcinoma. The American journal of …. 2000;157(1):257–66.

170. Nakanishi T, Menju T, Miyata R, Nishikawa S, Takahashi K, Cho H, et al. P2.04-018 Comprehensive Copy Number Alteration and Gene Expression Analysis of Surgically Resected Thymic Carcinoma. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S1007.

171. Moreira AL, Won HH, McMillan R, Huang J, Riely GJ, Ladanyi M, et al. Massively parallel sequencing identifies recurrent mutations in TP53 in thymic carcinoma associated with poor prognosis. J Thorac Oncol. 2015 Feb;10(2):373–80.

172. Chen C, Yin N, Yin B, Lu Q. DNA methylation in thoracic neoplasms. Cancer letters. 2011;301(1):7–16.

173. Suzuki M, Chen H, Shigematsu H, Ando S, Iida T, Nakajima T, et al. Aberrant methylation: common in thymic carcinomas, rare in thymomas. Oncol Rep. 2005 Dec;14(6):1621–4.

174. Mokhtar M, Kondo K, Namura T, Ali A, Fujita Y. Methylation and expression profiles of MGMT gene in thymic epithelial tumors. Lung Cancer. 2014;83(2):279–87.

175. Badve S, Goswami C, Gökmen-Polar Y, Nelson RP, Henley J, Miller N, et al. Molecular analysis of thymoma. van Diest P, editor. PLoS ONE. 2012;7(8):e42669.

176. Gökmen-Polar Y, Cook RW, Goswami CP, Wilkinson J. A gene signature to determine metastatic behavior in thymomas. Guan X-Y, editor. PLoS ONE. 2013;8(7):e66047.

177. Lee H-S, Jang H-J, Shah R, Yoon D, Hamaji M, Wald O, et al. Genomic Analysis of Thymic Epithelial Tumors Identifies Novel Subtypes Associated with Distinct Clinical Features. Clin Cancer Res. 2017 Apr 11;:clincanres.0066.2017.

178. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer. 2006;6(11):857–66.

179. Ganci F, Vico C, Korita E, Sacconi A, Gallo E, Mori F. MicroRNA expression profiling of thymic epithelial tumors. Lung Cancer. 2014.

180. Radovich M, Solzak JP, Hancock BA, Conces ML, Atale R, Porter RF, et al. A large microRNA cluster on chromosome 19 is a transcriptional hallmark of WHO type A and AB thymomas. Br J Cancer. 2016 Feb 16;114(4):477–84.

181. Freidin MB, Bhudia N, Lim E, Nicholson AG, Cookson WO, Moffatt MF. Impact of collection and storage of lung tumor tissue on whole genome expression profiling. J Mol Diagn. 2012 Mar;14(2):140–8.

182. Zhang Y, Sun Y, Pan Y, Li C, Shen L, Li Y, et al. Frequency of Driver Mutations in Lung Adenocarcinoma from Female Never-Smokers Varies with Histologic Subtypes and Age at Diagnosis. Clin Cancer Res. American Association for Cancer Research; 2012 Apr 1;18(7):1947–53.

183. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA. National Academy of Sciences; 1977 Dec;74(12):5463–7.

184. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, et al. ACMG clinical laboratory standards for next-generation sequencing. Genet Med. 2013 Sep;15(9):733–47.

185. Mu W, Lu H-M, Chen J, Li S, Elliott AM. Sanger Confirmation Is Required to Achieve Optimal Sensitivity and Specificity in Next-Generation Sequencing Panel Testing. J Mol Diagn. 2016 Nov;18(6):923–32.

186. Beck TF, Mullikin JC, NISC Comparative Sequencing Program, Biesecker LG. Systematic Evaluation of Sanger Validation of Next-Generation Sequencing Variants. Clin Chem. 2016 Apr;62(4):647–54.

187. Performance comparison of exome DNA sequencing technologies. 2011 Sep 25;29(10):908–14.

188. Marco-Sola S, Sammeth M, Guigó R, Ribeca P. The GEM mapper: fast, accurate and versatile alignment by filtration. Nat Methods. 2012 Dec;9(12):1185–8.

189. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010 Sep;20(9):1297–303.

190. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011 May;43(5):491–8.

191. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug 15;25(16):2078–9.

192. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012 Apr;6(2):80–92.

193. Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Hum Mutat. 2011 Aug;32(8):894–9.

194. Liu X, Jian X, Boerwinkle E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum Mutat. 2013 Sep;34(9):E2393–402.

195. Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, et al. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet. Frontiers; 2012;3:35.

196. Van Loo P, Nordgard SH, Lingjærde OC, Russnes HG, Rye IH, Sun W, et al. Allele-specific copy number analysis of tumors. Proc Natl Acad Sci USA. National Acad Sciences; 2010 Sep 28;107(39):16910–5.

197. Diskin SJ, Li M, Hou C, Yang S, Glessner J, Hakonarson H, et al. Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. Nucleic Acids Res. 2008 Nov;36(19):e126.

198. Hirsch FR, Varella-Garcia M, Bunn PA, Di Maria MV, Veve R, Bremmes RM, et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol. 2003 Oct 15;21(20):3798–807.

199. John T, Liu G, Tsao MS. Overview of molecular testing in non-small-cell lung cancer: mutational analysis, gene copy number, protein expression and other biomarkers of EGFR for the prediction of response to tyrosine kinase inhibitors. Oncogene. 2009 Aug;28 Suppl 1:S14–23.

200. Hedegaard J, Thorsen K, Lund MK, Hein A-MK, Hamilton-Dutoit SJ, Vang S, et al. Next-generation sequencing of RNA and DNA isolated from paired fresh-frozen and formalin-fixed paraffin-embedded samples of human cancer and normal tissue. PLoS ONE. 2014;9(5):e98187.

201. Srinivasan M, Sedmak D, Jewell S. Effect of Fixatives and Tissue Processing on the Content and Integrity of Nucleic Acids. The American Journal of Pathology. American Society for Investigative Pathology; 2010 Dec 13;161(6):1961–71.

202. Frankel A. Formalin fixation in the “-omics” era: a primer for the surgeon-scientist. ANZ J Surg. Blackwell Publishing Asia; 2012 Jun;82(6):395–402.

203. Wood HM, Belvedere O, Conway C, Daly C, Chalkley R, Bickerdike M, et al. Using next-generation sequencing for high resolution multiplex analysis of copy number variation from nanogram quantities of DNA from formalin-fixed paraffin-embedded specimens. Nucleic Acids Res. 2010 Aug;38(14):e151.

204. Chai H, Brown RE. Field effect in cancer-an update. Ann Clin Lab Sci. 2009;39(4):331–7.

205. Huang X, Stern DF, Zhao H. Transcriptional Profiles from Paired Normal Samples Offer Complementary Information on Cancer Patient Survival--Evidence from TCGA Pan-Cancer Data. Sci Rep. 2016 Feb 3;6(1):20567.

206. Román-Pérez E, Casbas-Hernández P, Pirone JR, Rein J, Carey LA, Lubet RA, et al. Gene expression in extratumoral microenvironment predicts clinical outcome in breast cancer patients. Breast Cancer Res. BioMed Central; 2012 Mar 19;14(2):R51.

207. Lee H-Y, Ahn JB, Rha SY, Chung HC, Park KH, Kim TS, et al. High KLF4 level in normal tissue predicts poor survival in colorectal cancer patients. World J Surg Oncol. 2014 Jul 24;12:232.

208. Sadanandam A, Lal A, Benz SC, Eppenberger-Castori S, Scott G, Gray JW, et al. Genomic aberrations in normal tissue adjacent to HER2-amplified breast cancers: field cancerization or contaminating tumor cells? Breast Cancer Res Treat. 2012 Dec;136(3):693–703.

209. Heaphy CM, Griffith JK, Bisoffi M. Mammary field cancerization: molecular evidence and clinical importance. Breast Cancer Res Treat. 2009 Nov;118(2):229–39.

210. Troester MA, Lee MH, Carter M, Fan C, Cowan DW, Perez ER, et al. Activation of host wound responses in breast cancer microenvironment. Clin Cancer Res. 2009 Nov 15;15(22):7020–8.

211. Mao R, Liu J, Liu G, Jin S, Xue Q, Ma L, et al. Whole genome sequencing of matched tumor, adjacent non-tumor tissues and corresponding normal blood samples of hepatocellular carcinoma patients revealed dynamic changes of the mutations profiles during hepatocarcinogenesis. Oncotarget. Impact Journals; 2017 Apr 18;8(16):26185–99.

212. Feng Y, Lei Y, Wu X, Huang Y, Rao H, Zhang Y, et al. GTF2I mutation frequently occurs in more indolent thymic epithelial tumors and predicts better prognosis. Lung Cancer. Elsevier; 2017 Aug;110:48–52.

213. Bakker den MA, Roden AC, Marx A, Marino M. Histologic classification of thymoma: a practical guide for routine cases. J Thorac Oncol. 2014 Sep;9(9 Suppl 2):S125–30.

214. Bakker den MA, Oosterhuis JW. Tumours and tumour-like conditions of the thymus other than thymoma; a practical approach. Histopathology. 2009 Jan;54(1):69–89.

215. Zheng J, Huang R, Huang Q, Deng F, Chen Y. The GTF2I rs117026326 polymorphism is associated with anti-SSA-positive primary Sjögren's syndrome. …. 2014.

216. Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet [Internet]. 2010 Oct;11(10):685–96. Available from: http://dx.doi.org/10.1038/nrg2841

217. Behjati S, Tarpey PS. What is next generation sequencing? Archives of Disease in Childhood-Education and …. 2013;98(6):236–8.

218. Garraway LA, Lander ES. Lessons from the cancer genome. Cell [Internet]. 2013 Mar 28;153(1):17–37. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0092867413002882

219. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012 Mar 8;366(10):883–92.

220. Petrini I, Meltzer PS, Kim I-K, Lucchi M, Park K-S, Fontanini G, et al. A specific missense mutation in GTF2I occurs at high frequency in thymic epithelial tumors. Nat Genet. 2014 Aug;46(8):844–9.

221. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, et al. A census of human cancer genes. Nat Rev Cancer. 2004 Mar;4(3):177–83.

222. Kufe DW. Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer. 2009 Dec;9(12):874–85.

223. Valque H, Gouyer V, Gottrand F, Desseyn J-L. MUC5B leads to aggressive behavior of breast cancer MCF7 cells. PLoS ONE. 2012;7(10):e46699.

224. Network TCGAR. Comprehensive molecular profiling of lung adenocarcinoma. Nature. Nature Publishing Group; 2014 Jul 9;:1–8.

225. Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012 Sep 27;489(7417):519–25.

226. Qiu L, Wu J, Pan C, Tan X, Lin J, Liu R, et al. Downregulation of CDC27 inhibits the proliferation of colorectal cancer cells via the accumulation of p21Cip1/Waf1. Cell Death Dis. 2016 Jan 28;7(1):e2074.

227. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013 Jul 11;499(7457):214–8.

228. Leiserson MDM, Vandin F, Wu H-T, Dobson JR, Eldridge JV, Thomas JL, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. Nat Genet. 2015 Feb;47(2):106–14.

229. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science. 2013 Mar 29;339(6127):1546–58.

230. Tamborero D, Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Kandoth C, Reimand J, et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. Sci Rep. 2013 Oct 2;3:2650.

231. Wang Y, Thomas A, Lau C, Rajan A, Zhu Y, Killian JK, et al. Mutations of epigenetic regulatory genes are common in thymic carcinomas. Sci Rep. 2014;4:7336.

232. Yoon S, Seger R. The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. - PubMed - NCBI. Growth Factors. 2009 Jul 11;24(1):21–44.

233. Diaz-Flores E, Shannon K. Targeting oncogenic Ras. Genes Dev. 2007 Aug 15;21(16):1989–92.

234. Kelleher FC, McArthur GA. Targeting NRAS in Melanoma. The Cancer Journal. 2012;18(2):132–6.

235. Yoon Y-K, Kim H-P, Han S-W, Oh D-Y, Im S-A, Bang Y-J, et al. KRAS mutant lung cancer cells are differentially responsive to MEK inhibitor due to AKT or STAT3 activation: implication for combinatorial approach. - PubMed - NCBI. Mol Carcinog. 2010 Apr;49(4):353–62.

236. Grimaldi AM, Simeone E, Ascierto PA. The role of MEK inhibitors in the treatment of metastatic melanoma. Curr Opin Oncol. 2014 Mar;26(2):196–203.

237. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P, et al. Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. Hum Mutat. 2007 Jun;28(6):622–9.

238. Muller PAJ, Vousden KH. p53 mutations in cancer. Nat Cell Biol. 2013 Jan;15(1):2–8.

239. Parrales A, Iwakuma T. Targeting Oncogenic Mutant p53 for Cancer Therapy. Front Oncol. 2015;5:288.

240. Guo C, Chen LH, Huang Y, Chang C-C, Wang P, Pirozzi CJ, et al. KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. Oncotarget. Impact Journals; 2013 Nov;4(11):2144–53.

241. Scarpa A, Chang DK, Nones K, Corbo V, Patch A-M, Bailey P, et al. Whole-genome landscape of pancreatic neuroendocrine tumours. Nature. 2017 Mar 2;543(7643):65–71.

242. Yoshida K, Miki Y. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. Cancer Sci. 2004 Nov;95(11):866–71.

243. Antoniou A, Pharoah PDP, Narod S, Risch HA, Eyfjord JE, Hopper JL, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. 2003 May;72(5):1117–30.

244. Friedenson B. BRCA1 and BRCA2 pathways and the risk of cancers other than breast or ovarian. MedGenMed. WebMD/Medscape Health Network; 2005 Jun 29;7(2):60.

245. Brown JS, Kaye SB, Yap TA. PARP inhibitors: the race is on. British Journal of Cancer. Nature Publishing Group; 2016 Mar 18;114(7):713–5.

246. Livraghi L, Garber JE. PARP inhibitors in the management of breast cancer: current data and future prospects. BMC Med. BioMed Central; 2015 Aug 13;13(1):188.

247. Wiesner T, Obenauf AC, Murali R, Fried I, Griewank KG, Ulz P, et al. Germline mutations in BAP1 predispose to melanocytic tumors. Nat Genet. 2011 Oct;43(10):1018–21.

248. Farley MN, Schmidt LS, Mester JL, Pena-Llopis S, Pavia-Jimenez A, Christie A, et al. A Novel Germline Mutation in BAP1 Predisposes to Familial Clear-Cell Renal Cell Carcinoma. 2013 Sep 16;11(9):1061–71.

249. Rai K, Pilarski R, Cebulla CM, Abdel-Rahman MH. Comprehensive review of BAP1 tumor predisposition syndrome with report of two new cases. 2016 Mar;89(3):285–94.

250. Alakus H, Yost SE, Woo B, French R, Lin GY, Jepsen K, et al. BAP1 mutation is a frequent somatic event in peritoneal malignant mesothelioma. 2015 Apr 16;13(1):105–16.

251. Popova T, Hebert L, Jacquemin V, Gad S, Caux-Moncoutier V, Dubois-d'Enghien C, et al. REPOR T Germline BAP1 Mutations Predispose to Renal Cell Carcinomas. The American Society of Human Genetics; 2013 Jun 6;92(6):974–80.

252. Testa JR, Cheung M, Pei J, Below JE, Tan Y, Sementino E, et al. Germline BAP1 mutations predispose to malignant mesothelioma. Nat Genet. 2011 Oct;43(10):1022–5.

253. Shukla SA, Rooney MS, Rajasagi M, Tiao G, Dixon PM, Lawrence MS, et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nat Biotechnol. 2015 Nov;33(11):1152–8.

254. Carbone M, Yang H, Pass HI, Krausz T, Testa JR, Gaudino G. BAP1 and cancer. Nat Rev Cancer [Internet]. 2013 Mar;13(3):153–9. Available from: http://link.aip.org/link/doi/10.1063/PT.4.2482

255. Pilarski R, Rai K, Cebulla C, Abdel-Rahman M. BAP1 Tumor Predisposition Syndrome. The Cancer Journal. Seattle (WA): University of Washington, Seattle; 2016 Oct 13;18(2):132–6.

256. Wu X, Bekker-Jensen IH, Christensen J, Rasmussen KD, Sidoli S, Qi Y, et al. Tumor suppressor ASXL1 is essential for the activation of INK4B expression in response to oncogene activity and anti-proliferative signals. Cell Res. 2015 Nov;25(11):1205–18.

257. Gelsi-Boyer V, Brecqueville M, Devillier R, Murati A, Mozziconacci M-J, Birnbaum D. Mutations in ASXL1 are associated with poor prognosis across the spectrum of malignant myeloid diseases. J Hematol Oncol. BioMed Central; 2012 Mar 21;5(1):12.

258. Abdel-Wahab O, Gao J, Adli M, Dey A, Trimarchi T, Chung YR, et al. Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. J Exp Med. 2013 Nov 18;210(12):2641–59.

259. Chan BA, Hughes BGM. Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. Transl Lung Cancer Res. 2015 Feb;4(1):36–54.

260. Zhao D, Chen X, Qin N, Su D, Zhou L, Zhang Q, et al. The prognostic role of EGFR-TKIs for patients with advanced non-small cell lung cancer. Sci Rep. Nature Publishing Group; 2017 Jan 12;7:40374.

261. Sullivan I, Planchard D. Next-Generation EGFR Tyrosine Kinase Inhibitors for Treating EGFR-Mutant Lung Cancer beyond First Line. Front Med (Lausanne). 2016;3:76.

262. Centeno I, Blay P, Santamaría I, Astudillo A, Pitiot AS, Osorio FG, et al. Germ-line mutations in epidermal growth factor receptor (EGFR) are rare but may contribute to oncogenesis: a novel germ-line mutation in EGFR detected in a patient with lung adenocarcinoma. BMC Cancer. BioMed Central; 2011 May 16;11(1):172.

263. Gazdar A, Robinson L, Oliver D, Xing C, Travis WD, Soh J, et al. Hereditary lung cancer syndrome targets never smokers with germline EGFR gene T790M mutations. J Thorac Oncol. 2014 Apr;9(4):456–63.

264. Yu HA, Arcila ME, Harlan Fleischut M, Stadler Z, Ladanyi M, Berger MF, et al. Germline EGFR T790M mutation found in multiple members of a familial cohort. J Thorac Oncol. 2014 Apr;9(4):554–8.

265. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature. 2016 Jun 2;534(7605):47–54.

266. LaFramboise T. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. Nucleic Acids Res. 2009 Jul;37(13):4181–93.

267. Cuylen S, Blaukopf C, Politi AZ, Müller-Reichert T, Neumann B, Poser I, et al. Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. Nature. 2016 Jul 14;535(7611):308–12.

268. Li L, Jiang G, Chen Q, Zheng J. Ki67 is a promising molecular target in the diagnosis of cancer (Review). Mol Med Report. Spandidos Publications; 2014 Nov 10;11(3):1–7.

269. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014 Jul 31;511(7511):543–50.

270. Adams JM, Cory S. The Bcl-2 apoptotic switch in cancer development and therapy. Oncogene. 2007 Feb 26;26(9):1324–37.

271. Yip KW, Reed JC. Bcl-2 family proteins and cancer. Oncogene. 2008 Oct 27;27(50):6398–406.

272. Williams MM, Cook RS. Bcl-2 family proteins in breast development and cancer: could Mcl-1 targeting overcome therapeutic resistance? Oncotarget. Impact Journals; 2015 Feb 28;6(6):3519–30.

273. Orsetti B, Nugoli M, Cervera N, Lasorsa L, Chuchana P, Rougé C, et al. Genetic profiling of chromosome 1 in breast cancer: mapping of regions of gains and losses and identification of candidate genes on 1q. British Journal of Cancer. 2006 Nov 20;95(10):1439–47.

274. Ströbel P, Zettl A, Shilo K, Chuang W-Y, Nicholson AG, Matsuno Y, et al. Tumor genetics and survival of thymic neuroendocrine neoplasms: a multi-institutional clinicopathologic study. 2014 Sep;53(9):738–49.

275. Zhao R, Choi BY, Lee M-H, Bode AM, Dong Z. Implications of Genetic and Epigenetic Alterations of CDKN2A (p16INK4a) in Cancer. EBIOM. The Authors; 2016 Jun 1;8(C):30–9.

276. Guo G, Chmielecki J, Goparaju C, Heguy A, Dolgalev I, Carbone M, et al. Whole-Exome Sequencing Reveals Frequent Genetic Alterations in BAP1, NF2, CDKN2A, and CUL1 in Malignant Pleural Mesothelioma. Cancer Res. 2014 Dec 8.

277. Xing X, Cai W, Shi H, Wang Y, Li M, Jiao J, et al. The prognostic value of CDKN2A hypermethylation in colorectal cancer: a meta-analysis. Br J Cancer. 2013 Jun 25;108(12):2542–8.

278. Fanourgakis G, Lesche M, Akpinar M, Dahl A, Jessberger R. Chromatoid Body Protein TDRD6 Supports Long 3' UTR Triggered Nonsense Mediated mRNA Decay. PLoS Genet. 2016 May;12(5):e1005857.

279. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Sander C, et al. The Cancer Genome Atlas Pan-Cancer analysis project. 2013 Oct;45(10):1113–20.

280. Sieber OM, Lipton L, Crabtree M, Heinimann K, Fidalgo P, Phillips RKS, et al. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. N Engl J Med. 2003 Feb 27;348(9):791–9.

281. Poulsen MLM, Bisgaard ML. MUTYH Associated Polyposis (MAP). Curr Genomics. 2008 Sep;9(6):420–35.

282. Bougatef K, Marrakchi R, Kourda N, Ben Lahely YB, Jileni SB, Khil El HK, et al. Somatic mutation of MutYH in Tunisian patients with sporadic colorectal cancer. J Clin Lab Anal. 2007;21(6):372–4.

283. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. 2013 Sep 26;45(10):1134–40.

284. Goldberg MV, Drake CG. LAG-3 in Cancer Immunotherapy. Curr Top Microbiol Immunol. 2011;344:269–78.

285. Khalil DN, Smith EL, Brentjens RJ, Wolchok JD. The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. Nat Rev Clin Oncol. 2016 May;13(5):273–90.

286. Marcq E, Pauwels P, van Meerbeeck JP, Smits ELJ. Targeting immune checkpoints: New opportunity for mesothelioma treatment? Cancer Treat Rev. 2015 Dec;41(10):914–24.

287. Anderson AC. Tim-3: an emerging target in the cancer immunotherapy landscape. Cancer Immunol Res. 2014 May;2(5):393–8.

288. Koyama S, Akbay EA, Li YY, Herter-Sprie GS, Buczkowski KA, Richards WG, et al. Adaptive resistance to therapeutic PD-1 blockade is associated with upregulation of alternative immune checkpoints. 2016;7:10501.

289. He Y, Rozeboom L, Rivard C, Dziadziuszko R, Ellison K, Yu H, et al. PUB044 TIM-3 Protein Expression in Non-Small Cell Lung Cancer and Its Relationship with PD-1/PD-L1 and Tumor-Infiltrating Lymphocytes. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S1473.

290. Lisberg A, Mckenna R, Dering J, Chen H-W, Hou D, Kamranpour N, et al. P2.01-078 Frequent High TIM-3 (HAVCR2) Expression in Resected NSCLC Specimens, Most Notably in Adenocarcinoma Histology. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S835.

291. Marcq E, Siozopoulou V, De Waele J, Van Audenaerde J, Zwaenepoel K, Santermans E, et al. OA02.07 Characterization of the Tumor Microenvironment and Investigation of Immune Checkpoint Expression in Malignant Pleural Mesothelioma. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S249–50.

292. Picarda E, Ohaegbulam KC, Zang X. Molecular Pathways: Targeting B7-H3 (CD276) for Human Cancer Immunotherapy. Clin Cancer Res. 2016 Jul 15;22(14):3425–31.

293. Castellanos JR, Purvis IJ, Labak CM, Guda MR, Tsung AJ, Velpula KK, et al. B7-H3 role in the immune landscape of cancer. Am J Clin Exp Immunol. 2017;6(4):66–75.

294. Lines JL, Sempere LF, Broughton T, Wang L, Noelle R. VISTA is a novel broad-spectrum negative checkpoint regulator for cancer immunotherapy. Cancer Immunol Res. 2014 Jun;2(6):510–7.

295. Liu J, Yuan Y, Chen W, Putra J, Suriawinata AA, Schenk AD, et al. Immune-checkpoint proteins VISTA and PD-1 nonredundantly regulate murine T-cell responses. Proc Natl Acad Sci USA. 2015 May 26;112(21):6682–7.

296. Chester C, Ambulkar S, Kohrt HE. 4-1BB agonism: adding the accelerator to cancer immunotherapy. Cancer Immunol Immunother. Springer Berlin Heidelberg; 2016 Oct;65(10):1243–8.

297. Chen H-N, Yuan K, Xie N, Wang K, Huang Z, Chen Y, et al. PDLIM1 Stabilizes the E-Cadherin/β-Catenin Complex to Prevent Epithelial-Mesenchymal Transition and Metastatic Potential of Colorectal Cancer Cells. Cancer Res. 2016 Mar 1;76(5):1122–34.

298. Liu Z, Zhan Y, Tu Y, Chen K, Wu C. PDZ and LIM domain protein 1(PDLIM1)/CLP36 promotes breast cancer cell migration, invasion and metastasis through interaction with α-actinin. Oncogene. 2015 Mar 5;34(10):1300–11.

299. Gherardi E, Birchmeier W, Birchmeier C, Vande Woude G. Targeting MET in cancer: rationale and progress. Nat Rev Cancer. 2012 Jan 24;12(2):89–103.

300. Ye S, Li J, Hao K, Yan J, Zhou H. The Efficacy and Risk Profile of c-Met inhibitors in Non-small Cell Lung Cancer: a Meta-analysis. Sci Rep. Nature Publishing Group; 2016 Oct 27;6(1):35770.

301. Cecchi F, Wright C, Bottaro D. Experimental Cancer Therapeutics Targeting the Hepatocyte Growth Factor/Met Signaling Pathway. Center for Cancer Research; 2010.

302. Bauman JE, Duvvuri U, Ferris RL, Ohr J, Gooding WE. Phase I study of the anti-HGF monoclonal antibody (mAb), ficlatuzumab, and cetuximab in cetuximab-resistant, recurrent/metastatic (R/M) head and neck squamous …. 2017.

303. Pothula SP, Xu Z, Goldstein D, Biankin AV, Pirola RC, Wilson JS, et al. Hepatocyte growth factor inhibition: a novel therapeutic approach in pancreatic cancer. Br J Cancer. 2016 Feb 2;114(3):269–80.

304. Mitsudomi T. SC04. 02 Management of Resistance to EGFR Tyrosine Kinase Inhibitors. Journal of Thoracic Oncology. 2017;12(1):S82–3.

305. Kim HR, Ha S-J, Hong MH, Heo SJ, Koh YW, Choi EC, et al. PD-L1 expression on immune cells, but not on tumor cells, is a favorable prognostic factor for head and neck cancer patients. Sci Rep. Nature Publishing Group; 2016 Nov 14;6(1):36956.

306. Altan M, Pelekanou V, Schalper K, Toki M, Herbst R, Rimm D. P2.01-046 Quantitative Measurement of B7-H3 Protein Expression and Its Association with B7-H4, PD-L1 and TILs in NSCLC. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S813–4.

307. Ye Z, Zheng Z, Li X, Zhu Y, Zhong Z, Peng L, et al. B7-H3 Overexpression Predicts Poor Survival of Cancer Patients: A Meta-Analysis. Cell Physiol Biochem. 2016;39(4):1568–80.

308. Inamura K, Yokouchi Y, Kobayashi M, Sakakibara R, Ninomiya H, Subat S, et al. P2.01-029 Tumor B7-H3 (CD276) Protein Expression, Smoking History, and Survival in Lung Adenocarcinoma Patients. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S802.

309. Giaccone G, Thompson J, Crawford J. A phase II study of pembrolizumab in patients with recurrent thymic carcinoma. J Clin Oncol; 2016.

310. Nunes-Xavier CE, Karlsen KF, Tekle C, Pedersen C, Øyjord T, Hongisto V, et al. Decreased expression of B7-H3 reduces the glycolytic capacity and sensitizes breast cancer cells to AKT/mTOR inhibitors. Oncotarget. 2016 Feb 9;7(6):6891–901.

311. Parra E, Villalobos P, Zhang J, Behrens C, Mino B, Moran C, et al. P2.01-061 Image Analysis-Based Expression of Nine Immune Checkpoints Identifies Distinct Immunoprofiling Patterns in Non-Small Cell Lung Carcinomas. Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S823–4.

312. Okazaki T, Honjo T. PD-1 and PD-1 ligands: from discovery to clinical application. Int Immunol. 2007 Jul;19(7):813–24.

313. Rajan A, Heery C, Mammen A, Pittaluga S, Lepone L, Donahue R, et al. OA18.03 Safety and Clinical Activity of Avelumab (MSB0010718C; Anti-PD-L1) in Patients with Advanced Thymic Epithelial Tumors (TETs). Journal of Thoracic Oncology. Elsevier; 2017 Jan 1;12(1):S314–5.

314. Wolff AC, Hammond MEH, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. Vol. 31, Journal of Clinical Oncology. 2013. pp. 3997–4013.

315. Syngal S, Brand RE, Church JM, Giardiello FM, Hampel HL, Burt RW. ACG Clinical Guideline: Genetic Testing and Management of Hereditary Gastrointestinal Cancer Syndromes. Am J Gastroenterol. 2015 Feb 3;110(2):223–62.

316. Boparai KS, Dekker E, Van Eeden S, Polak MM, Bartelsman JFWM, Mathus-Vliegen EMH, et al. Hyperplastic polyps and sessile serrated adenomas as a phenotypic expression of MYH-associated polyposis. Gastroenterology. 2008 Dec;135(6):2014–8.

317. Nielsen M, Morreau H, Vasen HFA, Hes FJ. MUTYH-associated polyposis (MAP). Crit Rev Oncol Hematol. 2011 Jul;79(1):1–16.

318. Lin C-F, Naj AC, Wang L-S. Analyzing copy number variation using SNP array data: protocols for calling CNV and association tests. Curr Protoc Hum Genet. 2013 Oct 18;79:Unit1.27.

319. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. J Mol Diagn. 2017 May;19(3):341–65.

320. Gillet JP, Varma S, Gottesman MM. The Clinical Relevance of Cancer Cell Lines. J Natl Cancer Inst. 2013 Apr 2;105(7):452–8.

321. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014 Nov 28;346(6213):1258096.

Chapter 10
Appendices

Appendix 1
Immunohistochemistry standard operating procedures

NIHR Respiratory Biomedical Research Unit
STANDARD OPERATING PROCEDURE

USER GROUPS	HIS
DOCUMENT TITLE	Construction of Tissue Microarray Blocks Using the Alphelys Minicore ®
INDEX CODE	rBRU-SOP-His-5
AREA OF APPLICATION	Histopathology Core Facility Royal Brompton and Harefield NHS Foundation Trust
AUTHOR	John Ronquillo
AUTHORISED BY	Alex Bowman – HCF Manager Clare Fulk – BRU Operations Manager Tracy Higgins – Respiratory BRU Manager
DATE OF ISSUE	
VERSION	4
REVIEW INTERVAL	One year

Contents **Page Number**

Summary and purpose	3
Introduction	3
Creating and Programming the TMA Template	3
TMA Construction	5
Safety and Precautions	10
Auditing	10
Duties	10
Review and revision arrangements	10
Glossary of key terms and acronyms	11
References	11

Summary and purpose

1. To construct TMA blocks so that analysis of multiple samples on a single slide is possible.

2. All staff members either employed to work within the HCF or making use of the HCF facility on an ad hoc basis, are required to comply fully with this SOP, where scientifically relevant.

Introduction

3. TMA allows the analysis of multiple specimens on a single slide and is a recent and upcoming technique in histopathology and molecular biology. The technique involves the collection of tissue cores from potentially hundreds of FFPE tissue blocks into recipient empty paraffin blocks using a machine called a Tissue Microarrayer. Due to having multiple specimens on one slide, there is a high throughput of results as well as standardised methodology as all specimens in that slide are treated under the same conditions when performing tests such as immunohistochemistry. It also has the advantage of being cost and time effective.

   Though the main disadvantage of TMAs is that the cores collected may not be representative of the tissue sample (e.g. tumour type) there are ways of overcoming this issue [(a)] but TMAs should only be used under suitable applications and certain considerations.

Creating and Programming the TMA Template

4. The Biobank team allocate the tissue blocks and slides as required by the project, based upon the specific tissue/pathological details required by the project.

5. Prior to the construction of a TMA block, appropriate FFPE tissue blocks and their corresponding H&E slides are sourced and collected by a member of the research team and the member of the Biobank department, respectively.

6. The H&E slides are analysed by a pathologist to mark the areas of the tissue to be cored that is most suitable for the research/study being undertaken. The number of cores to be taken is under the discretion of the pathologist.

7. Once all slides have been marked, the slides are then matched to the blocks to determine if they match and whether or not there is sufficient tissue and/or paraffin wax depth for a 4mm core to be taken out of it.

8. If there is insufficient thickness wax of the FFPE block, re-embed the tissue.

9. A TMA map/plan is created using Microsoft Excel and is saved under the project's allocated folder. Each cell will correspond to a core and its place on the blank recipient block; hence each cell must contain the block/slide identifier from which the core is to be collected from. Ensure that the map is asymmetric so the orientation of the TMA is clear. The map should also contain the number of cases that the TMA will contain. Print out the map once it has been completed.

10. The orientation and coordinates of the cores is under the discretion of the pathologist and/or researcher as well as the number of cores in the block.

11. Note: Apart from the TMA maps, all other TMA related software operations must be carried out on the laptop attached to the Microarrayer.

12. Using Microsoft Excel, create a block list by entering the block/slide IDs, working from top to bottom. The first cell, A1, is reserved for the block list title and is not counted when programming a new TMA project. First ID is entered in cell A2.

13. Save the block list as a '97-2003 Excel Workbook Project' in "TMA block lists" on the desktop

14. Note: The standard core diameter used is 1mm and the core depth is 4mm. The following programming steps are based on these measurements.

15. On the Desktop screen, open the TMADesigner® 2 software. The password is blank.

16. Click on ‘File’ → ‘New’ → ‘Recipient Block’ to create the size of the recipient paraffin block. The standard dimensions are (width x length) 20x20mm or 20x30mm with a margin of 2mm from the edge of the block and a depth of 5mm. Click save and return to the home screen by clicking on the purple door icon located in the bottom of the screen.

17. Click on ‘File’ → ‘New’ → ‘Tissue Array Project to set up the arrangement of the TMA. On here, the:

    a. Desired recipient block created is chosen
    b. Diameter of the cores is chosen (1000µm)
    c. Spacing between the cores is set (1700µm)
    d. Order/number in which the cores are to be positioned and placed. ‘Numerical’ is selected which positions the cores from left to right and top to bottom.

18. A screen will appear showing what the TMA block will look like with ‘spots’ representing the cores. On this screen, individual and/or multiple spots can be deleted and included to the final TMA design. It is recommended to delete spots to make the TMA asymmetrical which will help in the orientation of the TMA block. If not, another option is to add an orientation core that is placed outside the main TMA grid. Once the editing of the spots is complete, click on the pencil icon to draw around the cores to be included in the TMA. Save the tissue array and once again return to the home screen.

19. Click on ‘File’ → ‘New’ → ‘List of Tissue Types’. This is where the tissue types to be included in the TMA are added and are colour coded. Allocate a colour to each type at this stage if the TMA you are creating contains more than one type (i.e tumour and background lung). This will aid in the visualisation of the TMA when creating a new tissue array project. Save the tissue type list and once again return to the home screen.

20. A new tissue array project is now ready to be programmed. Click on ‘File’ →‘New’ → ‘Tissue Array Project’. This is where the block list is integrated into the template and is assigned to each core position. Complete each screen that appears, click next to proceed to the next screen.

21. Load ‘Donor Blocks’; list created in step (9)

22. Load ‘Tissue Types’; list created in step (18)

23. A screen will then appear that will allow each case to be matched with a tissue type and the number of replicates, ‘cores’, to be taken. Selection boxes will appear next to the cases numbers and the tissue types. Select the desired boxes and fill in the number of cores for the selected case. Click on the apply button.

24. Load Tissue Array Template, the filling mode selected should be ‘sorted’.

25. Click on the Gear Icon on the upper right corner of the screen. This will fill the tissue array. The case ID’s will then appear in the order the cores will be collected from them. If multiple cores are to be taken from one sample, then the case ID for that sample should appear multiple times in succession.

26. Save the project and close the program. The TMA is now ready to be constructed.

TMA Construction



Image 1: Alphelys Minicore®

27. ***Important - Before using the micro-arrayer, ensure that there will be no power and generator tests during the day. Check email as the Estates Department will notify the entire Trust via email, usually in the morning or the night prior to the tests being carried out. Avoid any TMA construction during power and generator tests as it will disrupt the procedure which may require restarting the entire process.***

28. ***Important - The recipient block is a standard sized cassette filled with wax in the same manner as embedding, but with no tissue. It is important to use the correct type of cassette, as certain cassette types may cause the tissue and wax to come away from the cassette during microtomy, which can cause major problems with TMAs as they cannot easily be re-embedded. It is best to avoid the mesh-style cassettes and also the ones with thin slits in the base. The best style to use are the ones with larger, circular holes***

29. The TMA blocks are constructed using the Alphelys Minicore® Microarrayer (Image 1).

30. The standard core needle diameter is 1mm with an adjusted depth of 4mm. No adjustment must be made unless it is necessary to do so and with approval from the laboratory manager.

31. Ensure that the USB cables are connected to the Minicore micro-arrayer. The USB port located on the right of the laptop should be connected to the actual micro-arrayer, while the port on the back should be connected to the camera that is built in into the micro-arrayer. Both USB ports (micro-arrayer and camera) on the Minicore are both located at the back of the machine and are appropriately labelled.

32. Upon booting up the laptop dedicated to the Minicore, follow these steps to reduce the chance of camera connection problems: Go to Control Panel>Device Manager>Imaging Devices>Integrated Webcam. Right click on this and disable it. This resets every time the laptop is turned on, so needs to be done every time.

    a. Troubleshooting: Sometimes the camera does not initialize and the message “Problem of connection with the camera” appears on the screen. If this occurs, save the project in Minicore Control Station (Project->Save), then press Quit, and turn off the TMA machine when the message appears stating it is safe to do so. Then turn it back on again, and reopen the project under “Uncompleted tissue array”. This may need repeating multiple times before the camera successfully connects.

b. Troubleshooting: If the camera is out of focus, cover it and take a photo, then uncover it and take another photo. This usually resolves focusing issues.

33. Clear out any obstruction that may interfere with the movement of the Minicore carousel (Image 2). Switch on the machine. The power switch is located on the back of the machine.

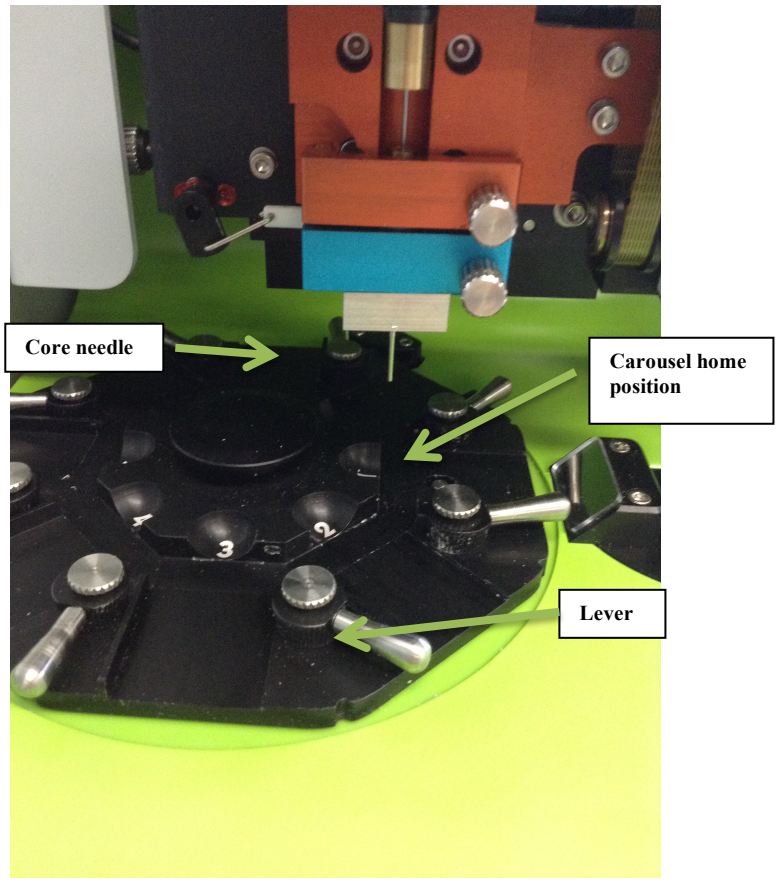


Image 2: Minicore carousel

34. On the desktop, open the MiniCore® Control Station software. The Minicore will then initialise, adjusting the carousel so that Position 1 is at the home position.

35. Click on 'New Tissue Array' and select the programmed Tissue Array Project. Click 'Valid' on the next two screens then a confirmation will ask if the core diameter to be is 1000µm. Click Yes.

36. The following screen is the allocation of the position of the blocks on the carousel. Click on Position 1 and allocate the recipient block to it. Click on Position 2 and allocate the donor blocks to it. Click next.

37. The carousel will then move so to allow the recipient block to be loaded into Position 1. Securely hold the block in place by turning the lever to the left. Click next. Position 1 will then move back to the home position. A picture of the recipient block will appear on the screen. Adjust the crosses as indicated to determine and focus on the position of the block. Click ok.

38. The screen will then confirm the adjustment of the depth of the needle. Click ok.

Collection of the recipient core

39. The needle will adjust to the first position of the array.

40. Turn the handle into the ‘Recipient’ position (Image 3).

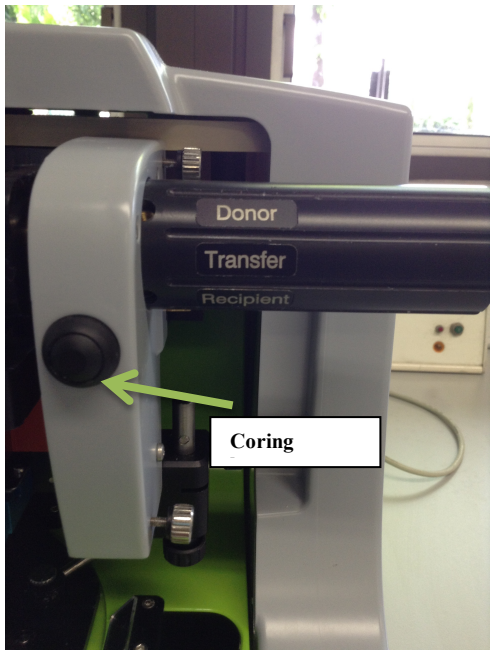


Image 3: Coring button and handle

41. Gently lower the needle until it penetrates the wax and cannot go any deeper. Do not force the needle any further as it may hit the plastic cassette and damage the needle.

42. Press the coring button (Image 3) while the needle is still in the wax. Allow the rotations of the needle to finish before lifting up the needle otherwise the coring process may not be successful.

43. Gently lift up the needle and turn the handle to the ‘Transfer’ position’ (Image 3).This will expose the wax core collected from the recipient block. Remove gently with a brush.

44. Click ‘Next’.

Collection of the donor core:

45. The carousel will then move so that the donor block can be placed into Position 2. Carefully match the case number shown on the screen to the printed TMA map making sure it is the right block in the right position in the array. Match the block and the slide ID’s then place the block into Position 2. Securely hold the block in place by turning the lever to the left. Click next.

46. The image of the donor block will then appear on screen which then requires placing markers on the tissue in which the cores will be collected. Precisely match the marks made on the slide to the marks being placed on the image of the donor tissue. Click next.

47. The needle will then move to the areas marked on the donor tissue ready for the donor core to be collected.

48. Turn the handle to the ‘Donor’ position (Image 3).Gently lower the needle until it penetrates the tissue. Do not force the needle if any resistance is met as it may cause damage to the needle.

49. Press the coring button while the needle is still in the tissue. Allow the rotations of the needle to finish before lifting up the needle as the coring process may not be successful.

50. Gently lift up the needle, leaving it in the donor position. Check the donor block if the coring was successful. If not, repeat. Click next.

Transferring the donor core:

51. The carousel will then return to the recipient block where the recipient core was taken.

52. Gently lower the needle until just the tip of needle touches the surface of the hole.

53. While holding down the needle, very slowly turn the handle back to the transfer position allowing the donor core to slowly fill the recipient hole.

54. Lift up the needle and click next.

55. A picture of the recipient block will then appear to confirm whether or not the transfer was successful. If not, click on the redo button which allows for another core to be taken from the donor block and the transfer to be repeated.

56. If successful, click next and the process will repeat until the final core is transferred.

57. Steps 39-56 involve the construction of the TMA and are repeated until the TMA is complete.

58. When the TMA construction is finished, the carousel will then move so that the TMA block can be removed. Click finish and then click quit on the main screen.

59. Wait until the carousel has moved back into the home position. Switch off the machine. Clean of the excess wax on the carousel and the needle gently with xylene then with IMS.

60. Place the completed TMA in the 37^{o}C oven for 10 minutes, cool then place it back in the oven for another 10 minutes to seal the wax around the tissue cores.

Cleaning

61. At the end of each day using the TMA machine, perform a routine clean of the punch. To do this, remove the punch from the holder by taking the front panel off the machine unlocking the gates and very carefully taking the punch out by holding two parts of it simultaneously so not to bend the thin metal. Then disassemble, clean each part down with a paper towel dabbed in xylene to remove paraffin wax deposits, then reassemble and carefully place back in the holder and use the top screw to screw in tightly.

Safety and Precautions

62. Laboratory coat must be worn though nitrile gloves are optional.

63. Xylene is an effective solvent in removing excess wax but it is quite ***harmful*** and should be used sparingly. Proper PPE must be worn when handling xylene (lab coat and nitrile gloves. Keep inhalation to a minimum).

Chemicals/ Solvents	Hazard/s	Safety Information
99% IMS	Highly flammable and harmful if ingested, inhaled and/or comes in contact with skin.	Refer to COSHH assessment sheet*
Xylene	Highly flammable and harmful if ingested, inhaled and/or comes in contact with skin.	Refer to COSHH assessment sheet*

*COSHH assessment sheets are located in Q-Pulse – Documents.

Auditing

64. Quarterly – the rBRU Clinical Governance Manager will carry out a full audit of the HCF to ensure compliance to all SOPs

    a. The rBRU Clinical Governance Manager will complete an audit report as per the BRU audit template

    b. The rBRU Clinical Governance Manager will report directly to the rBRU Operations Manager and the rBRU Manager

    c. The rBRU Operations Manager will be responsible for taking appropriate action following receipt of the audit

    d. RBHT and Imperial College London will audit the HCF as required for purposes including but not limited to, HTA compliance, H&S compliance, and Clinical Governance compliance

Duties

65. The HCF Manager is responsible for ensuring all HCF team members are trained on this SOP, completing necessary training logs and ensuring that any non-compliance with this SOP are reported to the rBRU Operations Manager.

Review and revision arrangements

66. This SOP will be reviewed annually or sooner if deemed necessary by the rBRU Management team.

Glossary of key terms and acronyms

NIHR	National Institute for Health Research
Respiratory Biomedical Research Unit (rBRU)	Research partnership between Royal Brompton Hospital and Imperial College awarded in 2008 and again in 2012
RBHT	Royal Brompton and Harefield NHS Foundation Trust
Histopathology Core Facility (HCF)	Laboratory space occupying room 2131 on the 2nd floor of Sydney Wing, Royal Brompton Hospital
FFPE	Formalin-fixed paraffin embedded

DOCUMENT TITLE Automated Immunohistochemistry Using The Ventana Ultra

H&E	Haematoxylin and eosin
H&S	Health and safety
HIS	Histopathology Q-Pulse user group
SOP	Standard operating procedure
HTA	Human Tissue Authority
IMS	Industrial methylated spirit
PPE	Personal protective equipment
TMA	Tissue microarray

References

a) Jawar, Nazar M.T. (2009) '*Tissue Microarray: A rapidly evolving diagnostic and research tool'*, Ann. Saudi Med. Mar-April pp. 123-12

INDEX CODE	rBRU-SOP-His-2
AREA OF APPLICATION	Histopathology Core Facility Royal Brompton and Harefield NHS Foundation Trust
AUTHOR	John Ronquillo
AUTHORISED BY	Alex Bowman – HCF Manager Clare Fulk – BRU Operations Manager Tracy Higgins – Respiratory BRU Manager
DATE OF ISSUE	
VERSION	3
REVIEW INTERVAL	One year
USER GROUPS	HIS

Contents	Page Number
Summary and purpose	3
Introduction	3
Antibody Handling	3
Antibody Optimisation	3
Logging in New Antibodies and Fillable Reagents	4
Setting up a Staining Protocol	6
Performing a Staining Run	8
Expected Results	12
Safety and Precautions	13
Auditing	13
Duties	13
Review and revision arrangements	13
Glossary of key terms and acronyms	14
References	15

Summary and purpose

67. To use the Ventana BenchMark Ultra for automated IHC staining of FFPE tissue sections.

68. To determine optimal staining protocols for newly added antibodies to the existing repertoire of the laboratory.

69. All staff members that are employed to work within the HCF are required to comply fully with this SOP.

Introduction

70. The use of automated IHC staining machines has become widespread and popular in both clinical and research laboratories. It has made a manually time consuming and delicate process more rapid and minimises the exposure to hazardous substances, most notably DAB chromogen which is a known carcinogen. The Ventana BenchMark Ultra has been designed to minimise as much of the manual staining process, even with an inbuilt de-waxing and antigen retrieval system, while still allowing the user great flexibility in regards with setting up staining protocols.

Antibody Handling

71. Antibodies are usually delivered in containers with some form of cooling implement as antibodies are prone to degradation, hence losing their functionality, if left at room temperature for elongated periods of time. Storage conditions will vary for different antibodies (usually indicated on the data sheet) so it is essential to deal with them as soon as they have been received.

72. Date and sign when received on the data sheet.

73. Check and place in appropriate storage. Most antibodies are stable in 4^{o}C and are placed in the 4^{o}C fridge located directly across the Ventana immunostainer. The other most common storage temperature is at -20^{o}C (freezer located adjacent to the microtome). It is then recommended to aliquot smaller volumes into smaller vials before initial freezing as repeated thawing and freezing of antibodies can result in degradation and loss of functionality.

Antibody Optimisation

74. The following optimisation procedure is the routine starting method to optimising an antibody, but can be altered if an actual protocol using the BenchMark Ultra has been provided by the antibody manufacturer.

75. Print out the 'Antibody Optimisation Sheet' which is located in Q-Pulse: rBRU-DOC-HIS-2. Fill in the necessary details at the top, and then each time an optimisation run is performed.

76. The manufacturer's recommended dilutions for their antibody can be found on the data sheet. Choice of dilutions and staining protocols at under the discretion of qualified staff.

77. Refer to '***Setting up a Staining Protocol'*** on how to set up a staining protocol and the recommended basic protocol using the BenchMark Ultra platform software as well as how to perform a staining run.

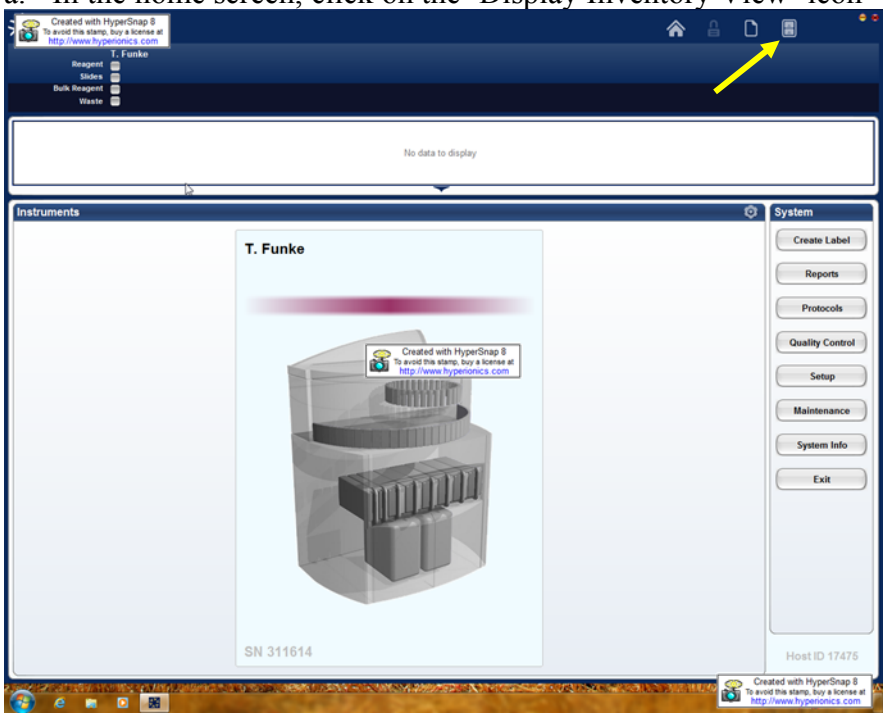
78. Make up at least 100µL diluted primary antibody per slide in Eppendorf tubes as the machine requires 100µL to be dispensed manually. Antibodies made up with diluents containing buffers, usually Phosphate Buffered Saline (PBS) and preservatives at neutral pH to stabilise the antibodies for prolonged periods of time.

79. Stained slides are checked by pathologists and/or researchers to determine if optimal staining has been achieved. Final protocol is chosen and signed off.

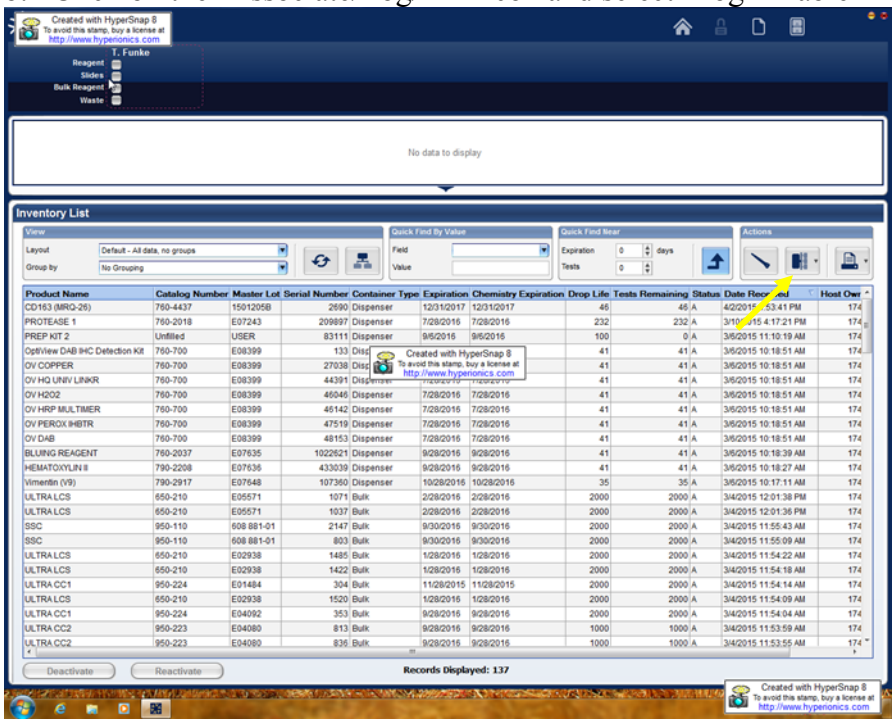
Logging in New Antibodies and Fillable Reagents

80. Once an antibody has been optimised, it can be diluted into a dispenser (Prep Kit) in greater volumes for regular use. The machine relies on a barcode system matching a barcode allocated to a dispenser and to the slide labels corresponding to the staining protocol set up by the user. The following method is applied, in the Ventana software:

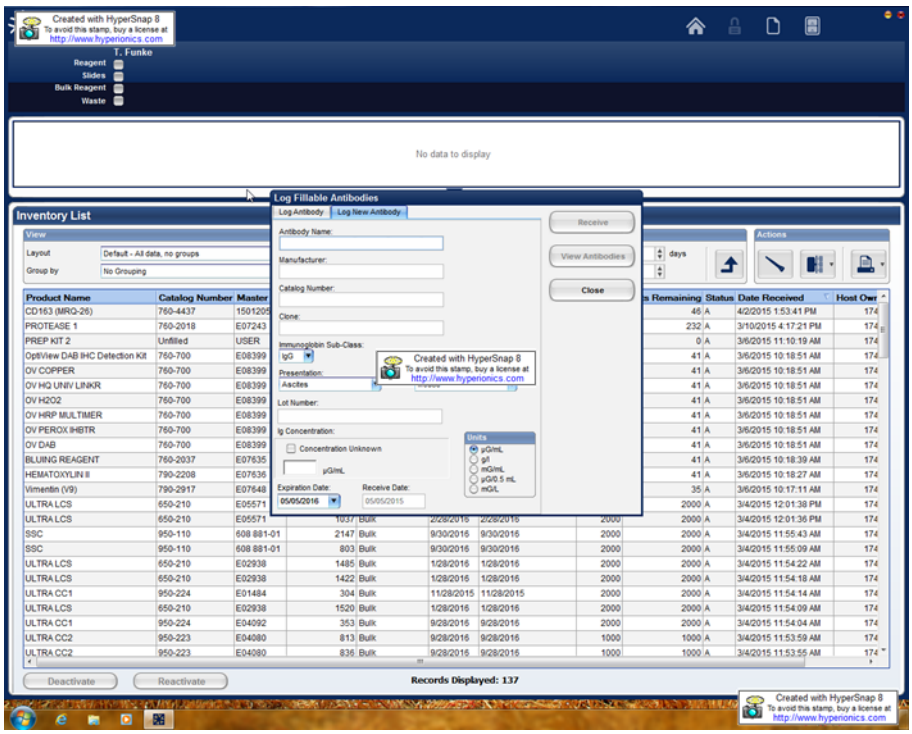
a. In the home screen, click on the 'Display Inventory View' icon



b. Click on the 'Associate/Log/Fill' icon and select 'Log Fillable Antibodies'

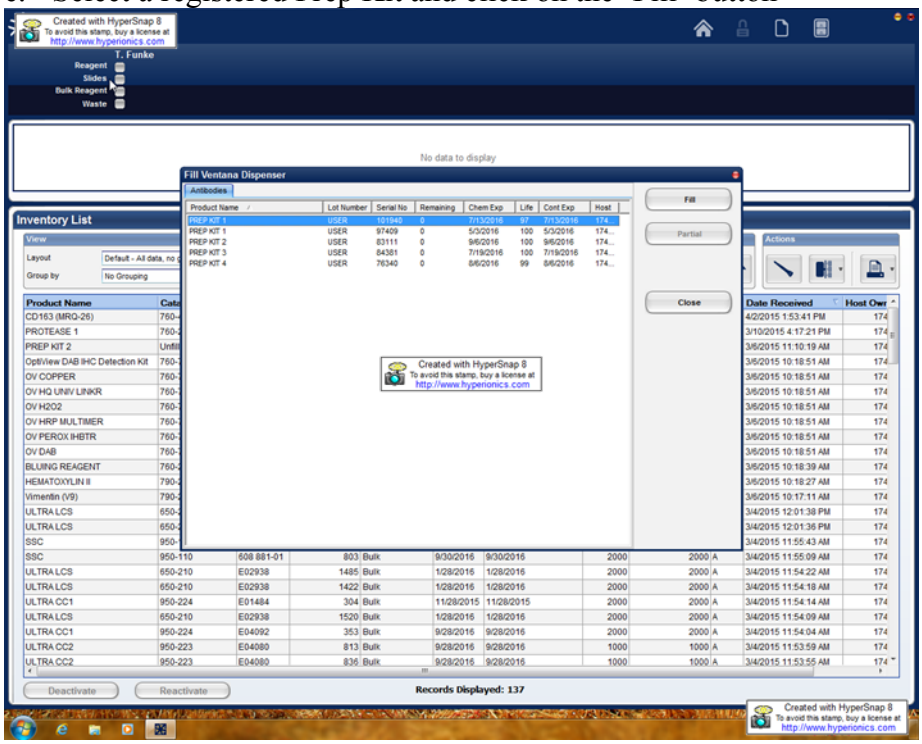


c. Click on the 'Log New Antibody' tab, fill in the details then click receive



d. Return to the Display Inventory View screen, select the Associate/Fill/Log icon and select 'Fill Ventana Dispenser'

e. Select a registered Prep Kit and click on the 'Fill' button



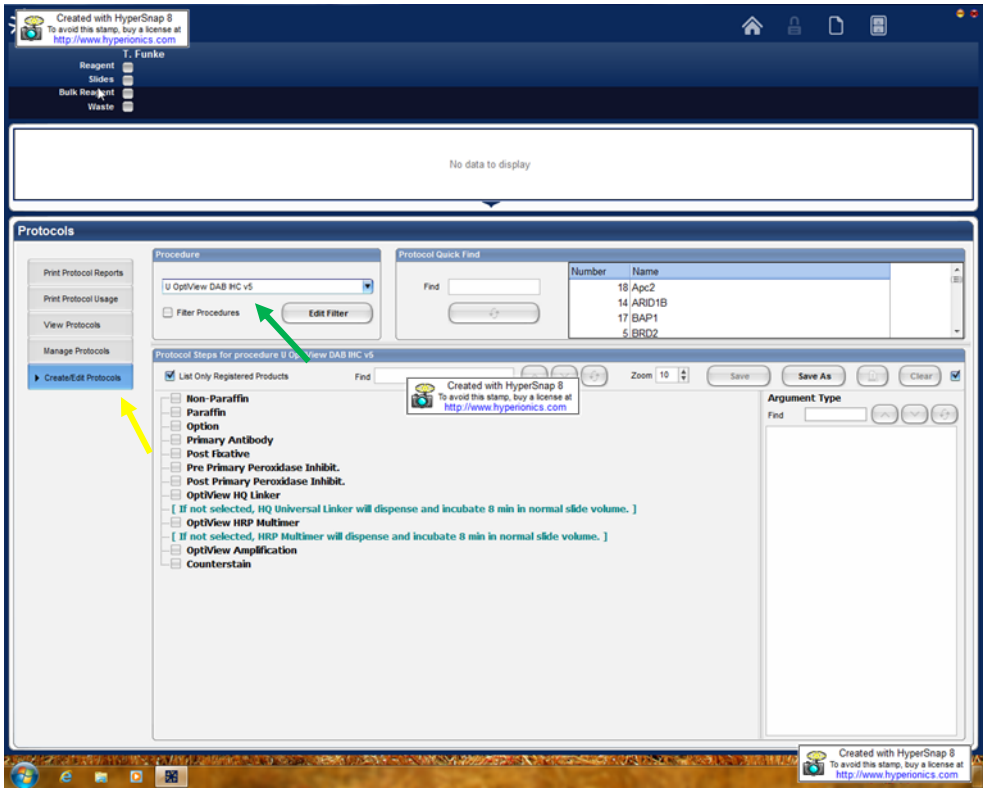
f. Select a logged antibody and enter 250 under 'Millilitre' located in the bottom of the screen. This ensures that the maximum dispenses of the Prep Kit will be used. Click 'Fill' once again.

81. The antibody chosen is now linked to the Prep Kit. Note down the Prep Kit number selected and stick the barcode on the corresponding dispenser properly labelled. ***Label must contain the antibody name, dilution factor and date of first use of dispenser***. It is recommended that antibody dilutions be done one at a time to avoid the possibility of diluting the wrong antibody in the wrong dispenser.

Setting Up a Staining Protocol

82. In the right hand panel in the home screen, click on the 'Protocol' button.

83. Click on the 'Create/Edit Protocols' button on the left side of the screen. Ensure that the procedure selected is 'U OptiView DAB IHC v5' (green arrow), or depending on the detection kit(s) being used. Select the options based on the optimised protocol.



84. Table 1 below shows the basic starting protocol recommended by Ventana Application Specialists, unless specified otherwise by the antibody manufacturer's specifications. Titration (manual application) of an antibody is usually chosen during optimisation of antibodies as it requires multiple dilutions to determine the final staining protocol.

Step	Time/Temperature
Paraffin	
Baking	8 min at 60°C
Deparaffinisation	Programmed time at 72°C
Cell Conditioning	CC1: 32 min at 100°C
Primary Antibody Incubation	36°C
Primary Antibody (Titration or Prep Kit)	32 min
Primary Peroxidase Inhibitor	Post Primary Peroxidase Inhibit.
OptiView HQ Linker	8 min
OptiView HRP Multimer	8 min
Haematoxylin II (Counterstain)	8 min
Blueing (Counterstain)	8 min

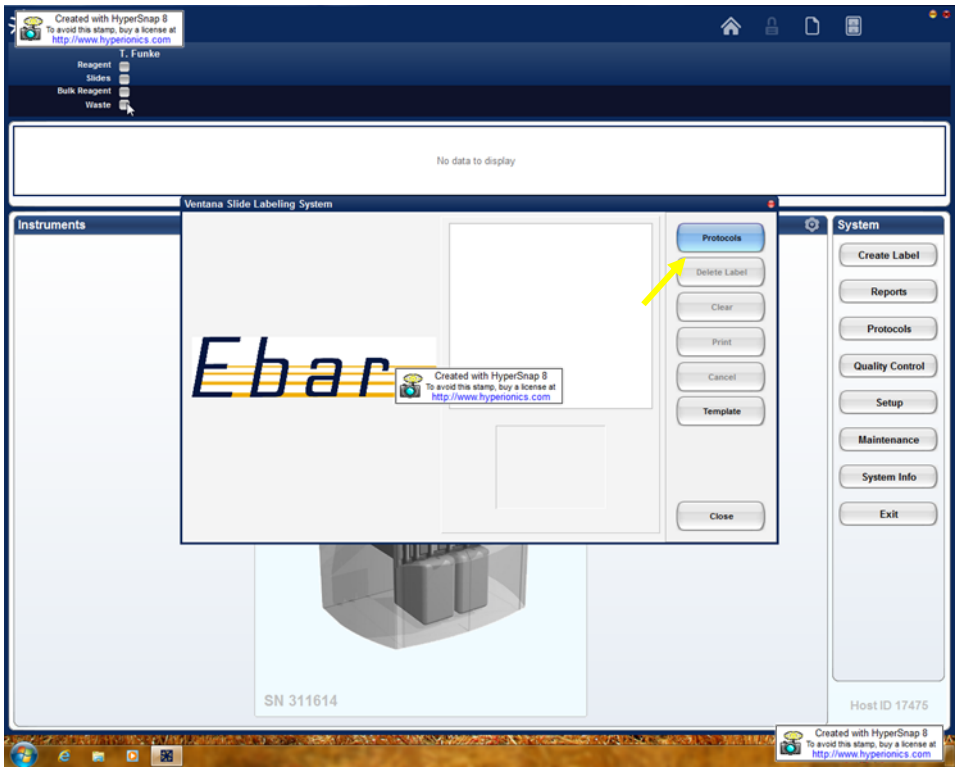
Table 1: Recommended starting protocol

85. Click ‘Save As’ and name the protocol. Assign the protocol with a number. Any number with an asterisk (*) next to it indicates that a protocol has been assigned with that number.

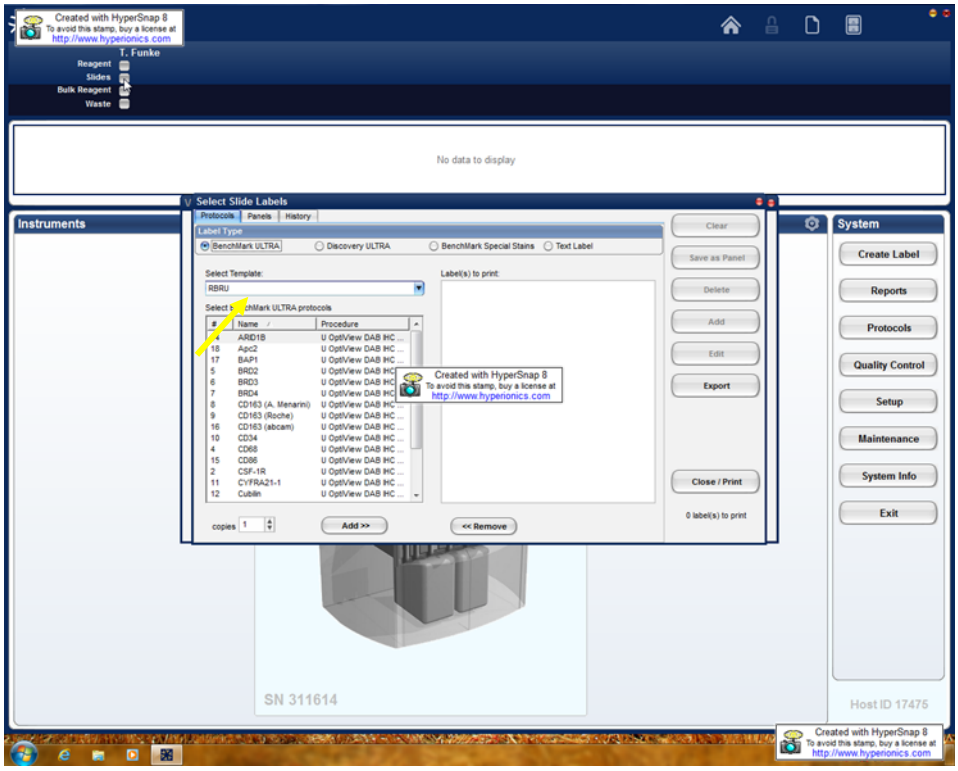
Performing a Staining Run

Ancillaries	Description
OptiView DAB Detection Kit	250 tests, brown staining
Haematoxylin II	250 tests, counterstain
Blueing Reagent	250 tests, counterstain
Protease 1	250 tests, enzymatic antigen retrieval depending on protocol
UltraView Universal Alkaline Phosphatase Detection Kit	250 tests, red staining. Usually used in conjunction with DAB kit for double staining protocols
Bulk Reagents	**Description**
EZ Prep	De-waxing solution (2L concentrate), diluted 1 in 10 with D.I water
Ultra LCS	2L predilute. Creates a stable environment for staining, prevents evaporation under high temperatures
SSC	2L concentrate. Used for stringency and rinsing washes during In-situ hybridisation (ISH) staining, diluted 1 in 5 with D.I water
Reaction Buffer	2L concentrate. A Tris based buffer at neutral pH used for rinsing between steps during staining, diluted 1 in 10 with D.I water
Ultra CC1	2L predilute. A Tris based buffer at pH 8.5 used for heat induced antigen retrieval
Ultra CC2	1L predilute. A citrate based buffer at pH 6.0 used for heat induced antigen retrieval

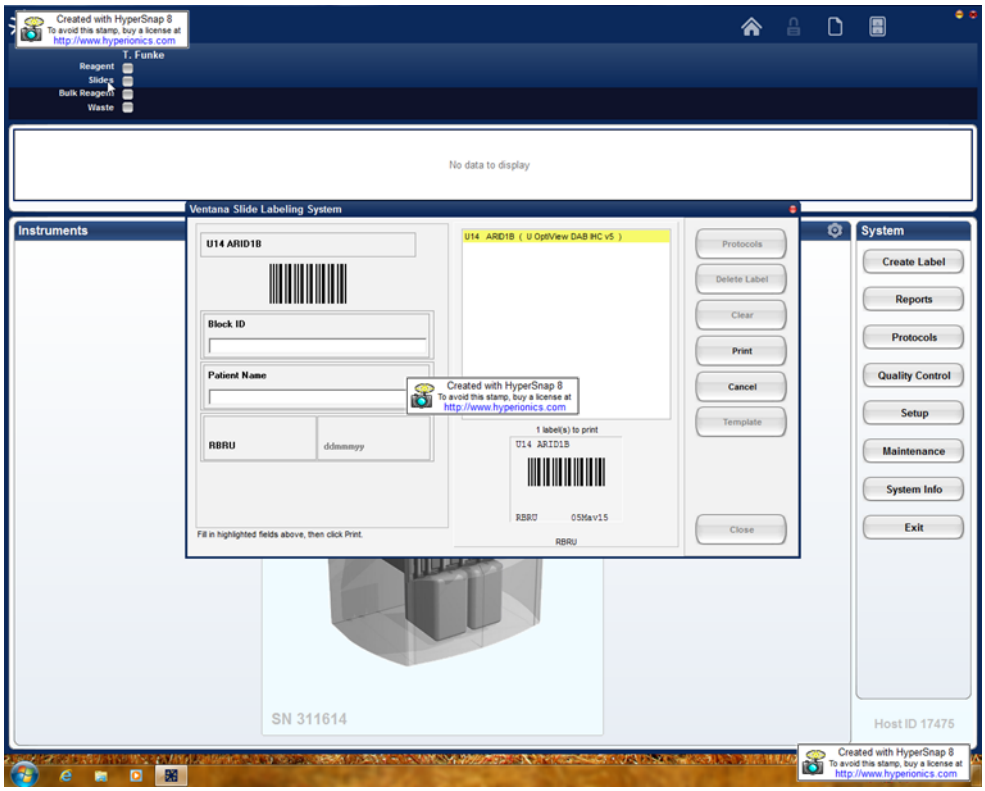
86. On the right hand panel in the home screen, click on ‘Create Label’ and then click on ‘Protocols’.



87. Select ‘RBRU’ as the template and select the protocol for which the slides are to be stained with:



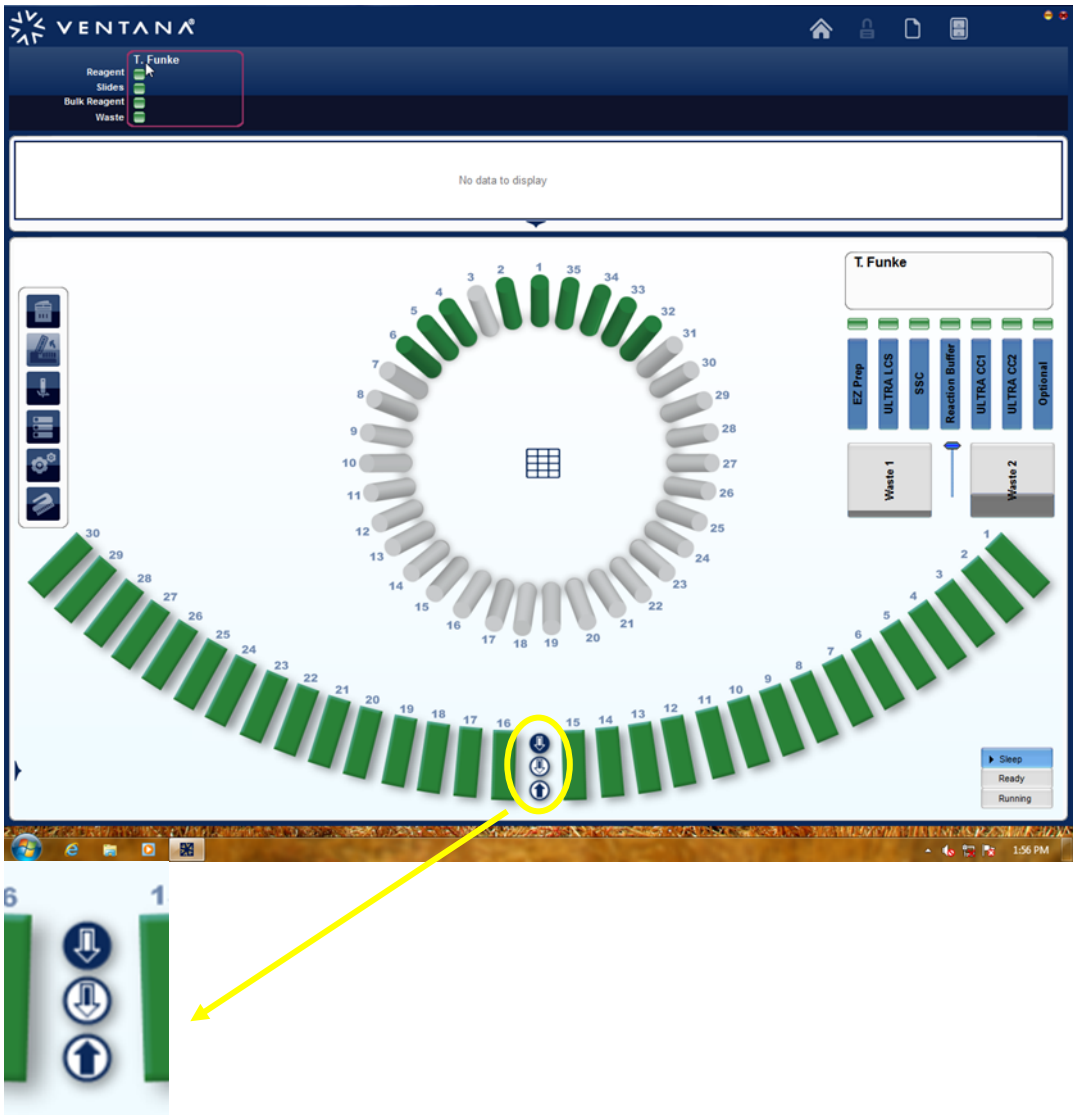
88. Enter the specimen/slide details and click on ‘Print’ to print out slide labels then carefully match the labels with the corresponding slides.



89. Prior to starting a run, it is good practice to check the volumes of the bulk reagents and the waste containers. The machine will not start a run if there is insufficient volume of at least one of the bulk reagents and if the waste containers are full.

90. Place the ancillary reagents and antibodies required for the staining run on reagent racks and onto the carousel, making sure that the dispenser caps are removed and that the reagent racks are securely fitted.

91. Switch the machine from ‘Sleep’ to ‘Ready’ mode on the main operation screen (yellow arrow).



92. Open the slide drawers by clicking on the top arrow. The middle arrow opens drawers of completed runs and the bottom arrow closes any open drawers. The same symbols can also be found on the machine and can be used alternatively

93. Load the slides with the slide label facing away from the user. Ensure that there is enough LCS on the heating platforms so that the slides are securely adhered to them. If not, place a drop of LCS before loading the slides.

94. Ensure that the positive control slide is placed last if running a batch of test slides

95. Close the slide drawers and switch the machine from ‘Ready’ to ‘Running’ to initiate the staining run. The lights on the slide drawers will indicate the status of each slide during the run:

    a. Green: run is finished or slide drawer is available.
    b. Amber: staining in process
    c. Red: an error has occurred, error message will appear on the operating screen.

96. The machine will alert the user prior and during the run if any problems occur in which error messages will also appear on the operating screen.

97. Once a run has finished the machine will alert the user and will switch from 'Running' to 'Ready'. Remove the slides and place them into a slide rack. Remove the ancillary reagents from the carousel. Close the drawers and switch the machine from 'Ready' to 'Sleep' then click on the brush icon located on the bottom of the left side panel to clean.



Wash the slides in warm water with dishwashing detergent to remove the excess oil. The slides can then either be air dried or dehydrated through increasing grades of IMS before clearing in xylene and coverslipping.

Expected results

98. Areas of expression of target protein: Brown.

99. Nuclei: Blue.

Safety and Precautions

100. Procedure involves dealing with ***toxic***, ***harmful*** and ***corrosive*** solvents/chemicals. PPE (lab coat and nitrile gloves) is to be worn at all times and procedure must be performed in a well-ventilated area.

101. Extreme care must be taken when handling DAB due to its known ***carcinogenic*** properties.

Chemicals/ Solvents	Hazard/s	Safety Information
DAB	Highly ***harmful*** and a known ***carcinogen.*** Has potential to cause cancer if ingested, inhaled and/or comes in contact with skin. Can also cause allergic reactions, irritations and tooth decay.	Refer to COSHH assessment sheet*
99% IMS	Highly ***flammable*** and ***harmful*** if ingested, inhaled and/or comes in contact with skin.	Refer to COSHH assessment sheet*
Xylene	Highly ***flammable*** and ***harmful*** if ingested, inhaled and/or comes in contact with skin.	Refer to COSHH assessment sheet*

*COSHH assessment sheets are located in Q-Pulse.

Auditing

102. Quarterly – the rBRU Clinical Governance Manager will carry out a full audit of the HCF to ensure compliance to all SOPs.

a. The rBRU Clinical Governance Manager will complete an audit report as per the rBRU audit template.

b. The rBRU Clinical Governance Manager will report directly to the rBRU Operations Manager and the rBRU Manager.

c. The rBRU Operations Manager will be responsible for taking appropriate action following receipt of the audit.

RBHT and Imperial College London will audit the HCF as required for purposes including but not limited to, HTA compliance, H&S compliance, and Clinical Governance compliance.

Duties

103. The HCF Manager is responsible for ensuring all HCF team members are trained on this SOP, completing necessary training logs and ensuring that any non-compliance with this SOP are reported to the rBRU Operations Manager.

Review and revision arrangements

104. This SOP will be reviewed annually or sooner if deemed necessary by the rBRU Management team.

Glossary of key terms and acronyms

NIHR	National Institute for Health Research
Respiratory Biomedical Research Unit (rBRU)	Research partnership between Royal Brompton Hospital and Imperial College awarded in 2008 and again in 2012
RBHT	Royal Brompton and Harefield NHS Foundation Trust
Histopathology Core Facility (HCF)	Laboratory space occupying room 2131 on the 2nd floor of Sydney Wing, Royal Brompton Hospital
Antibody	Immune proteins that are produced to target specific target proteins (antigens)
Antigen	In relation to IHC; expressed proteins within the cells and tissues in the body, including those expressed by host and foreign microbes
AP	Alkaline phosphatase
CC1	Cell conditioning 1; Tris-based buffer with a pH of 8.5 used in heat mediated antigen retrieval
CC2	Cell conditioning 2; Citrate-based buffer with a pH of 6 used in heat mediated antigen retrieval
DAB	Diaminobenzidine; chromogen used in IHC for visualisation. Produces a brown precipitate when reacts with hydrogen peroxide
FFPE	Formalin-fixed paraffin embedded
F/S	Frozen section
HTA	Human Tissue Authority
IHC	Immunohistochemistry
HIS	Histopathology user group
SOP	Standard operating procedure
IMS	Industrial methylated spirit
LCS	Liquid coverslip
SSC	Standard sodium citrate

References

a) Jacobsen, L; Nielsen, M; Månsson, S; Rudbeck, L (2015) *'Staining Protocol Optimisation'*, Dako UK Ltd.

b) Ramos-Vara, J.A; Miller, M.A. (2014) *'When Tissue Antigens and Antibodies Get Along, Revisiting the technical aspects of immunohistochemistry – the red, brown and blue technique'*

c) Zander, D.S; Farver, C.F (2008) *'Pulmonary Pathology, A volume in the series foundations in diagnostic pathology'* pp. 786-800

d) BenchMark Ultra Advanced Staining System Operating Guide; Issued September, 2012

DOCUMENT TITLE Sectioning of Formalin-Fixed Paraffin Embedded Tissues Using the Shandon® Finesse 325 Microtome

INDEX CODE	rBRU-SOP-His-3
AREA OF APPLICATION	Histopathology Core Facility Royal Brompton and Harefield NHS Foundation Trust
AUTHOR	John Ronquillo
AUTHORISED BY	Alex Bowman – HCF Manager Clare Fulk – BRU Operations Manager Tracy Higgins – Respiratory BRU Manager
DATE OF ISSUE	
VERSION	3
REVIEW INTERVAL	One year
USER GROUPS	HIS

Contents	Page Number
Summary and purpose	3
Introduction	3
Microtomy	3
Safety and Precautions	5
Auditing	5
Duties	5
Review and revision arrangements	6
Glossary of key terms and acronyms	7

Summary and purpose

105. To safely use a microtome.

106. To cut and prepare tissue sections from FFPE tissue blocks for subsequent staining and microscopic analysis of the tissue samples.

107. All staff members either employed to work within the HCF or making use of the HCF facility on an ad hoc basis, are required to comply fully with this SOP.

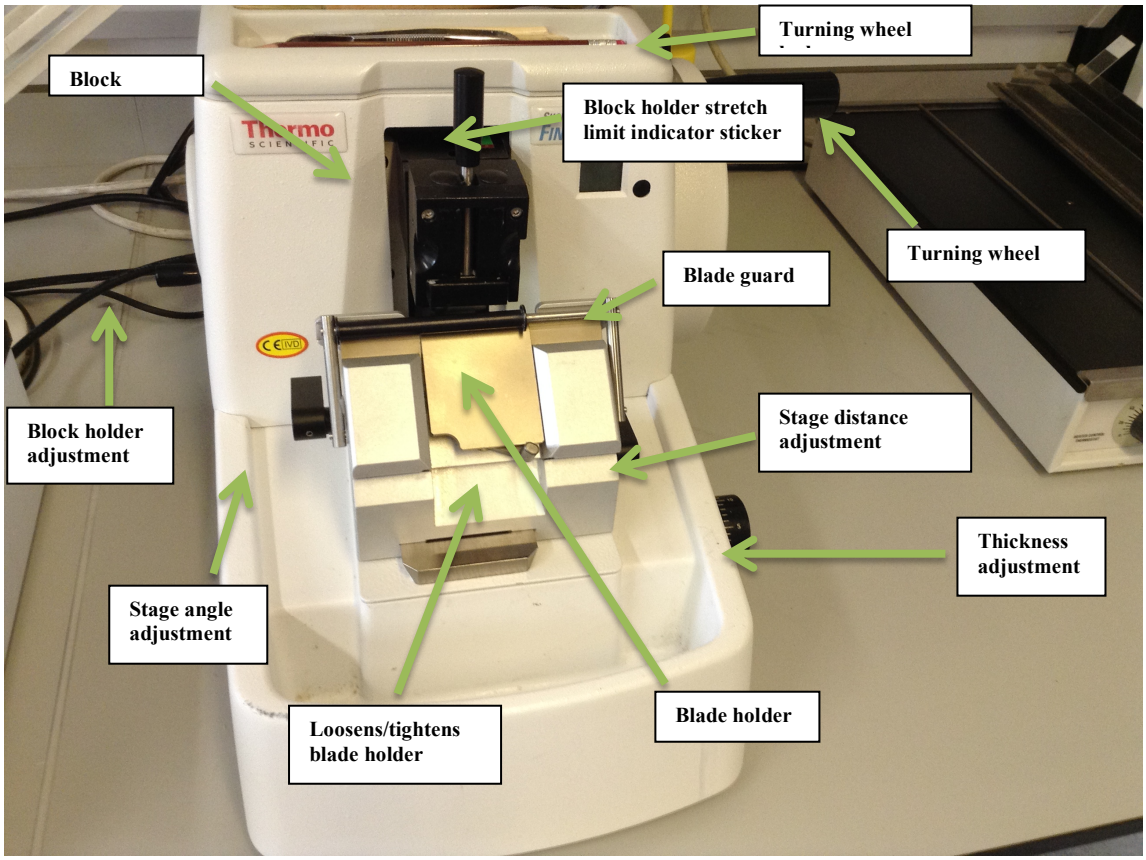
Introduction

108. FFPE sectioning (or microtomy) is a histological technique which allows embedded tissue to be sliced very thinly using a microtome, most commonly used are the rotary microtomes, which are then mounted onto glass slides. Extremely thin sections (usually 2-5µm) are needed so that light is able to pass through the tissue allowing visualisation under the light microscope.

Microtomy

109. Fill the water bath (Thermo Scientific, Section Flotation Bath) with tap water and switch it on, temperature should be kept at around 55°C. Switch on the cold plate.

110. Prior to cutting, several pre-checks must be done on the microtome (Thermo Scientific Shandon, Finesse 325, see picture on next page):



111. Ensure that there is no blade in the blade holder.

112. The turning wheel is locked.

113. The block holder is within its stretch limit, green area of sticker.

114. The angle of the stage is at 6.

115. The cutting thickness is at 2-5µm, depending on tissue type.

116. The alignment of the chuck and the angle of the stage must not be changed unless it is essential to do so with approval from senior staff.

117. Carefully load a blade into the blade holder then securely tighten it. Flip the blade guard back up to cover the blade.

118. Load the block into the block holder making sure that it is properly fitted. Adjust the stage and/or the block holder to give some space between the block and the blade.

119. Flip down the blade guard and unlock the turning wheel. Slowly turn the wheel on the first rotation to gauge the distance between the block and the blade, then gently start adjusting the chuck adjustment forward until the blade starts to cut the surface of the block.

120. Trim the block until full face of the tissue is revealed, or as close to as possible.

121. Once trimmed, brush the trimmed wax off the blade and flip up the guard.

122. Carefully remove the block from the block holder and place it on the cold plate to cool down.

123. Once cooled, replace the block back in the chuck and flip down the guard. Slightly reverse the block holder then carefully turn the wheel slowly to once again gauge the distance between the block and the blade. Turn the wheel until the blade reaches the surface of the block. Keep on turning the wheel slowly until a ribbon starts to form.

124. With forceps, grab the edge of the forming section gently pulling it while simultaneously slowly turning the wheel. Once a full section has been cut remove the other edge of the section from the blade with a brush and spread the section into the water bath.

125. ***Only sections from one block at time are to be spread onto the water bath to avoid mixing up specimens.***

126. Label slides with the corresponding block identification and slowly mount the sections onto the slide removing as much water as possible. Place the slides vertically on a slide rack to dry off and to drain any excess water. Sections for H&E and special stains are usually mounted on normal uncharged slides while charged/coated slides are used for sections for IHC staining.

127. Flip up guard and remove the block from the block holder.

128. Prior to staining, slides are placed into ovens to further dry out the slides ensuring that the sections stick to the slides. Slides can either be put into a 37^{o}C oven overnight or a 60^{o}C oven for 20 minutes (H&E/specials) or 60 minutes (IHC) minimum.

129. When finished, carefully remove blade and dispose of it carefully into the blade disposal hole in the bottom of the blade box. Once all the blades have been used and the blade box disposal part is full, put the box into a nearby yellow sharps bin. Brush of excess wax into a biohazard waste bin and wipe away excess wax with xylene. Turn off all other equipment and clean the working area.

Safety and Precautions

130. Microtomy is a technique which requires proper training and practice. Extreme care must be used when using the microtome as it utilises extremely sharp blades that have the potential to cause serious physical injuries. Laboratory coat must be worn though nitrile gloves are optional.

131. Xylene is an effective solvent in removing excess wax but it is quite ***harmful*** and should be used sparingly. Correct PPE must be worn when handling xylene. This includes a lab coat, nitrile gloves, and also eye protection if the user deems there to be a splash risk.

Chemicals/ Solvents	Hazard/s	Safety Information
Xylene	Highly ***flammable*** and ***harmful*** if ingested, inhaled and/or comes in contact with skin.	Refer to COSHH assessment sheet*

*COSHH assessment sheets are located in Q-Pulse - Documents

Auditing

132. Quarterly – the rBRU Clinical Governance Manager will carry out a full audit of the HCF to ensure compliance to all SOPs.

a. The rBRU Clinical Governance Manager will complete an audit report as per the rBRU audit template.

b. The rBRU Clinical Governance Manager will report directly to the rBRU Operations Manager and the rBRU Manager.

c. The rBRU Operations Manager will be responsible for taking appropriate action following receipt of the audit.

d. RBHT and Imperial College London will audit the HCF as required for purposes including but not limited to, HTA compliance, H&S compliance, and Clinical Governance compliance.

Duties

133. The HCF Manager is responsible for ensuring all HCF team members are trained on this SOP, completing necessary training logs and ensuring that any non-compliance with this SOP are reported to the rBRU Operations Manager.

Review and revision arrangements

134. This SOP will be reviewed annually or sooner if deemed necessary by the rBRU Management team.

Glossary of key terms and acronyms

NIHR	National Institute for Health Research
Respiratory Biomedical Research Unit (rBRU)	Research partnership between Royal Brompton Hospital and Imperial College awarded in 2008 and again in 2012
RBHT	Royal Brompton and Harefield NHS Foundation Trust
Histopathology Core Facility (HCF)	Laboratory space occupying room 2131 on the 2nd floor of Sydney Wing, Royal Brompton Hospital
FFPE	Formalin-fixed paraffin embedded
H&E	Haematoxylin and eosin
H&S	Health and safety
HTA	Human Tissue Authority

HIS	Histopathology Q-Pulse user group
SOP	Standard operating procedure
IHC	Immunohistochemistry
IMS	Industrial methylated spirit
PPE	Personal protective equipment

USER GROUPS	HIS

DOCUMENT TITLE	Sectioning of Formalin-Fixed Paraffin Embedded Tissue Rolls
INDEX CODE	rBRU-SOP-His-8
AREA OF APPLICATION	Histopathology Core Facility Royal Brompton and Harefield NHS Foundation Trust
AUTHOR	John Ronquillo
AUTHORISED BY	Alex Bowman – HCF Manager Clare Fulk – BRU Operations Manager Tracy Higgins – rBRU Manager
DATE OF ISSUE	
VERSION	3
REVIEW INTERVAL	One year

Contents	Page Number
Summary and purpose	3
Introduction	3
Microtomy	3
Safety and Precautions	5
Auditing	5
Duties	5
Review and revision arrangements	5
Glossary of key terms and acronyms	6

Summary and purpose

135. To safely use a microtome.

136. To cut and prepare tissue rolls suitable for subsequent molecular and/or genomic testing.

137. All staff members either employed to work within the HCF or making use of the HCF facility on an ad hoc basis, are required to comply fully with this SOP where scientifically relevant.

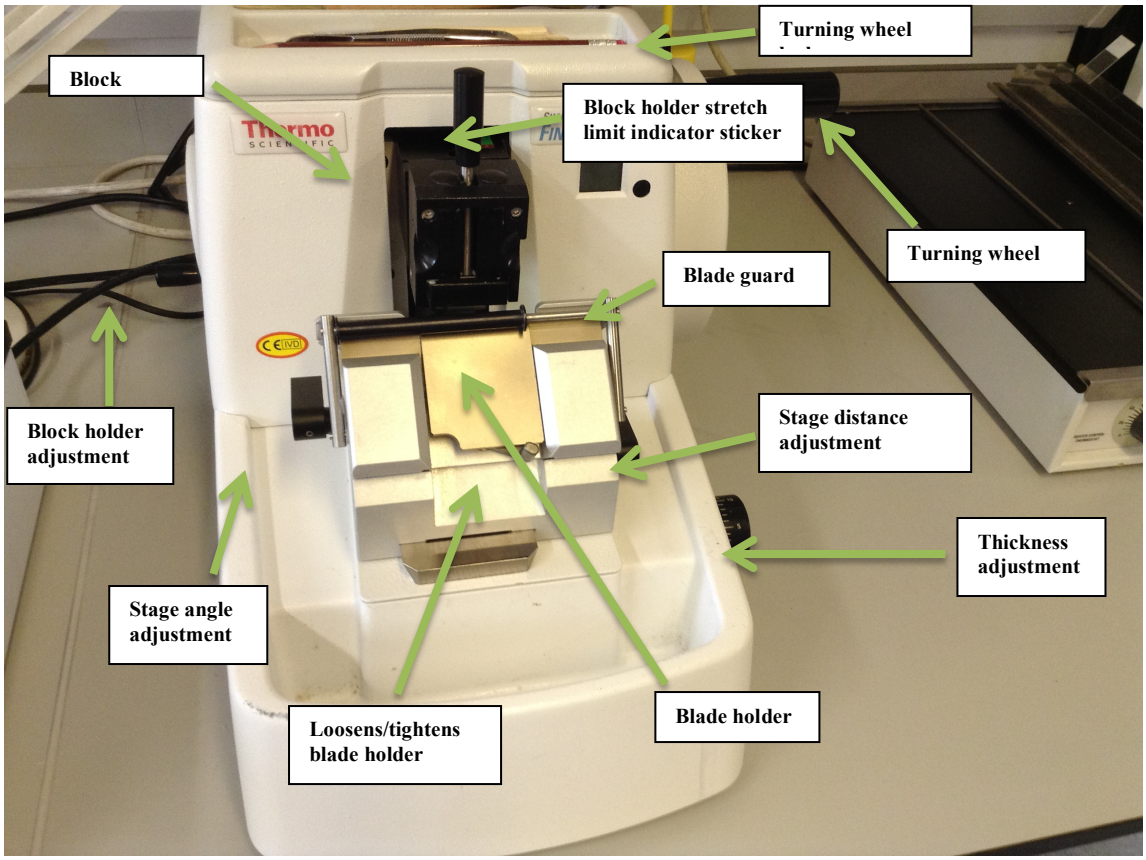
Introduction

138. FFPE sectioning (or microtomy) is a histological technique which allows embedded tissue to be sliced very thinly using a microtome, most commonly used are the rotary microtomes, which are then mounted onto glass slides. Extremely thin sections (usually 2-5µm) are needed so that light is able to pass through the tissue allowing visualisation under the light microscope.

139. Microtomy can also be used to obtain tissue with sufficient material (~10µm section rolls) from the FFPE block that can be used to extract protein, DNA and RNA for subsequent genomic testing.

Microtomy

140. Prior to cutting, several pre-checks must be done on the microtome (Thermo Scientific Shandon, Finesse 325, see image below):



a. Ensure that there is no blade in the blade holder.

b. The turning wheel is locked.

c. The block holder is within its stretch limit, green area of sticker.

d. The angle of the stage is at 6 .

e. The cutting thickness is at 2-5µm, depending on tissue type.

f. The alignment of the chuck and the angle of the stage must not be changed unless it is essential to do so and has approval from the laboratory manager.

141. Prior to cutting, the microtome and all other equipment involved must be cleaned with IMS or xylene. This is to minimise any contamination and cross-overs from previous blocks that may affect the results from the genomic testing. ***This must be done prior to cutting a new block***.

142. Carefully load a new blade into the blade holder then securely tighten it. Flip the blade guard back up to cover the blade. ***A new blade or a new section of the blade must be used for each block***.

143. Load the block into the block holder making sure that it is properly fitted. Adjust the stage and/or the block holder to give some space between the block and the blade.

144. Flip down the blade guard and unlock the turning wheel. Slowly turn the wheel on the first rotation to gauge the distance between the block and the blade, then gently start adjusting the chuck adjustment forward until the blade starts to cut the surface of the block.

145. Trim the block until full face of the tissue is revealed, or as close to as possible.

146. Once trimmed, brush the trimmed wax off the blade and flip up the guard. The block is best left at room temperature as it enable for rolls to easily form rather than producing sections.

147. Slightly reverse the chuck then carefully turn the wheel slowly to once again gauge the distance between the block and the blade. Turn the wheel until the blade reaches the surface of the block and a full section has been achieved.

148. Slowly the turn the wheel all the way down and allow sections to roll up.

149. With forceps, carefully grab the formed roll and place into a sterile container (i.e. Universal tubes) that has been labelled with the corresponding blocks' number and sub-number. Repeat until the desired number of rolls have been cut and collected.

150. Flip up guard and remove the block from the block holder.

151. When finished, carefully remove blade and dispose of it accordingly. Brush of excess wax into a biohazard waste bin and wipe away excess wax with xylene. Turn off all other equipment and clean the working area.

Safety and Precautions

152. Microtomy is a technique which requires full training and practice. Extreme care must be used when using the microtome as it utilises extremely sharp blades that have the potential to cause serious physical injuries. Laboratory coat must be worn though nitrile gloves are optional.

153. Xylene is an effective solvent in removing excess wax but it is quite ***harmful*** and should be used sparingly. Proper PPE must be worn when handling xylene (lab coat, nitrile gloves, and also eye protection if the user deems there to be a splash risk)..

Chemicals/ Solvents	Hazard/s	Safety Information
99% IMS	Highly ***flammable*** and ***harmful*** if ingested, inhaled and/or comes in contact with skin.	Refer to COSHH assessment sheet*
Xylene	Highly ***flammable*** and ***harmful*** if ingested, inhaled and/or comes in contact with skin.	Refer to COSHH assessment sheet*

*COSHH assessment sheets are located in Q-Pulse – Documents.

Auditing

154. Quarterly – the rBRU Clinical Governance Manager will carry out a full audit of the HCF to ensure compliance to all SOPs.

a. The rBRU Clinical Governance Manager will complete an audit report as per the rBRU audit template.

b. The rBRU Clinical Governance Manager will report directly to the rBRU Operations Manager and the rBRU Manager.

c. The rBRU Operations Manager will be responsible for taking appropriate action following receipt of the audit.

d. RBHT and Imperial College London will audit the HCF as required for purposes including but not limited to, HTA compliance, H&S compliance, Clinical Governance compliance.

Duties

155. The HCF Manager is responsible for ensuring all HCF team members are trained on this SOP, completing necessary training logs and ensuring that any non-compliance with this SOP are reported to the rBRU Operations Manager.

Review and revision arrangements

156. This SOP will be reviewed annually or sooner if deemed necessary by the rBRU Management team.

Glossary of key terms and acronyms

NIHR	National Institute for Health Research
Respiratory Biomedical Research Unit (rBRU)	Research partnership between Royal Brompton Hospital and Imperial College awarded in 2008 and again in 2012
RBHT	Royal Brompton and Harefield NHS Foundation Trust
Histopathology Core Facility (HCF)	Laboratory space occupying room 2131 on the 2nd floor of Sydney Wing, Royal Brompton Hospital
FFPE	Formalin-fixed paraffin embedded
H&E	Haematoxylin and eosin
HIS	Histopathology Q-Pulse user group
SOP	Standard operating procedure
H&E	Health & Safety
HTA	Human Tissue Authority
IMS	Industrial methylated spirit
PPE	Personal protective equipment

DOCUMENT TITLE Manual Coverslipping

NIHR Respiratory Biomedical Research Unit
STANDARD OPERATING PROCEDURE

INDEX CODE	rBRU-SOP-His-17
AREA OF APPLICATION	Histopathology Core Facility Royal Brompton and Harefield NHS Foundation Trust
AUTHOR	John Ronquillo
AUTHORISED BY	Alex Bowman – HCF Manager Clare Fulk – BRU Operations Manager Tracy Higgins – Respiratory BRU Manager
DATE OF ISSUE	
VERSION	3
REVIEW INTERVAL	One year
USER GROUPS	HIS

Contents	Page Number
Summary and purpose	3
Introduction	3
Coverslipping procedure	3
Re-coverslipping	4
Safety and Precautions	4
Auditing	4
Duties	5
Review and revision arrangements	5
Glossary of key terms and acronyms	6
References	6

Summary and purpose

157. To coverslip tissue sections post-staining for examination under light-microscopy.

158. All staff members either employed to work within the HCF or making use of the HCF facility on an ad hoc basis, are required to comply fully with this SOP, where scientifically relevant.

Introduction

159. Coverslipping is a procedure in which a thin piece of glass is mounted onto a stained tissue section. The piece of glass is held onto the slide by an adhesive or a 'mountant' that has an appropriate refractive index that will allow light from the microscope to pass through the tissue sections for visualisation. Due to most stains being water based, resin based mountants are usually used to preserve the tissues and stains for a long period of time.

160. Other than preserving the stained tissue section, coverslipping also serves the purpose of protecting both the tissue sections from foreign substances and the microscope lenses by preventing direct contact with the tissue section which can potentially damage the lenses.

Coverslipping procedure

161. Once sections have gone through the dehydration steps, transport the slides whilst in the final container of xylene, to the **chemical fume hood** which steps b-j are performed in.

162. Turn on the light and exhaust fan of the fume hood prior to coverslipping.

163. Lay out the appropriate size coverslip (22x22mm square or 24x50mm rectangle), with regards to the size of the tissue section, onto a clean piece of filter paper so to absorb excess xylene.

164. Place a small amount DePeX mounting medium onto the coverslip.

165. With forceps, gently take a slide from the slide rack in xylene still making sure the slide is still wet with xylene. Xylene will help the mountant to spread along the coverslip.

166. Gently place the section face-down onto the coverslip.

167. Gently press the back of the slide to ensure that the mountant spreads along the coverslip.

168. Flip the slide over and remove any air bubbles present by gently pressing the surface of the coverslip.

169. Wipe away excess xylene on the back of the slide.

170. Place the slide onto a slide tray to air dry for several minutes before microscopy. Leave the slides to dry for a minimum of 12 hours before filing away.

Re-coverslipping

171. The quality of coverslipping is important as improper preparation for and the actual coverslipping may hinder or make it difficult to analyse the tissue sections during microscopy.

172. Some common issues seen with coverslipping are air bubbles due to insufficient mountant, water droplets due to insufficient dehydration, multiple coverslips on a section and a dried out section due to the section being allowed to dry out of xylene directly prior to coverslipping.

173. If there are issues discovered with the initial coverslipping, soak the slide in xylene until the mountant has been dissolved so that the coverslip can be removed easily.

174. Repeat the coverslipping procedure.

175. ***NOTE: Do NOT force the coverslip away if the mountant is still intact as doing so can both damage the tissue section and the coverslip.***

Safety and Precautions

176. Procedure involves dealing with **toxic**, **harmful** and **corrosive** solvents/chemicals. Appropriate PPE (lab coat and nitrile gloves) are to be worn at all times and the procedure must be performed in the chemical fume hood.

177. Glass coverslips are thin and are easily broken. Care must then be taken when dealing with coverslips as they are prone to cause physical harm and must be treated as a **sharps** hazard, even intact as the edges and corners of the coverslips are able to penetrate the skin.

Chemicals/ Solvents	Hazard(s)	Safety Information
DePeX Mounting Medium	Contains xylene. ***Harmful*** if ingested, inhaled and/or comes in contact with skin and eyes. ***Flammable***.	Refer to COSHH assessment sheet*
Xylene	Highly ***flammable*** and ***harmful*** if ingested, inhaled and/or comes in contact with skin and eyes.	Refer to COSHH assessment sheet*

*COSHH assessment sheets are located in Q-Pulse (rBRU-DOC-HIS-1).

Auditing

178. Quarterly – the rBRU Clinical Governance Manager will carry out a full audit of the HCF to ensure compliance to all SOPs.

a. The rBRU Clinical Governance Manager will complete an audit report as per the rBRU audit template.

b. The rBRU Clinical Governance Manager will report directly to the rBRU Operations Manager and the rBRU Manager.

c. The rBRU Operations Manager will be responsible for taking appropriate action following receipt of the audit.

d. RBHT and Imperial College London will audit the HCF as required for purposes including but not limited to, HTA compliance, H&S compliance, and Clinical Governance compliance.

Duties

179. The HCF Manager is responsible for ensuring all HCF team members are trained on this SOP, completing necessary training logs and ensuring that any non-compliance with this SOP are reported to the rBRU Operations Manager.

Review and revision arrangements

180. This SOP will be reviewed annually or sooner if deemed necessary by the rBRU Management team.

Glossary of key terms and acronyms

NIHR	National Institute for Health Research
Respiratory Biomedical Research Unit (rBRU)	Research partnership between Royal Brompton Hospital and Imperial College awarded in 2008 and again in 2012
RBHT	Royal Brompton and Harefield NHS Foundation Trust
Histopathology Core Facility (HCF)	Laboratory space occupying room 2131 on the 2nd floor of Sydney Wing, Royal Brompton Hospital
H&S	Health and safety
HIS	Histopathology Q-Pulse user group
SOP	Standard operating procedure
HTA	Human Tissue Authority
PPE	Personal protective equipment

References

a) National Diagnostics: Histology (2011); *Mounting Tissue Sections* https://www.nationaldiagnostics.com/histology/article/mounting-tissue-sections

Appendix 2

Baseline clinicopathological data for all TETs including the downstream applications of the extracted DNA

#	Sample collection	Tumour Samples (n=156)	Sample Type	Tumour content (%)	Matched Normal Tissue (n=119)	Sample Type	WHO histotype	Sex	Stage	Age at Diagnosis	Resection	Survival (months / or alive)	MG	Other autoimmune condition	Significant Hx of CA (self or 1st degree relative)	WES	GTF2I	SNP genotyping	Target re-sequencing
1	>2010	150098PT	FFPE				A	M	I	81	R0	A			Bowel CA - S & C		MUT*	✓	✓
2	>2010	150113PT	FFPE				A	F	IIB	43	R0	A					MUT	✓	✓
3	2000-9	THRBF04707T1	FFPE	100	✓	FFPE	A	M	I	77	R0	18					MUT	✓	✓
4	2000-9	THRBF05579T1	FFPE	100			A	M	I	63	R0	A					MUT	✓	✓
5	2000-9	THRBZ07321T2	Frozen	90			A	F	IIA	69	R0	A			Breast CA		MUT	✓	✓
6	>2010	THRBF11200T2	FFPE	80	✓	FFPE	A	F	I	55	R0	A					MUT	✓	✓
7	2000-9	THRBF13203T1	FFPE	80	✓	FFPE	A	F	I	69	R0	A	✓				MUT	✓	✓
8	2000-9	THRBF16927T1	FFPE	100	✓	FFPE	A	M	I	57	R0	A	✓				MUT	✓	✓
9	2000-9	THRBF17788T1	FFPE	75	✓	FFPE	A	F	II	74	R0	A					WT	✓	✓
10	>2010	THRBF20533T1	FFPE	100	✓	FFPE	A	M	II	65	R1	A					MUT	✓	✓
11	2000-9	THRBF20683T1	FFPE	100	✓	FFPE	A	M	IIA	62	R0				Thymoma Recurrence at 65		MUT	✓	✓
12	2000-9	THRBF21048T1	FFPE	100	✓	FFPE	A	M	IIA	47	R0	A					MUT*	✓	✓
13	<2000	THRBZ23369T2	Frozen	100	✓	FFPE	A	F	I	55	R0	A					MUT*	✓	✓
14	>2010	THRBZ23415T1	Frozen	100	✓	FFPE	A	M	IIA	75	R0	A					MUT	✓	✓
15	>2010	THRBF26625T1	FFPE	90	✓	FFPE	A	M	I	72	R0			Hypothyroidism			MUT	✓	✓
16	2000-9	THRBF2887T1	FFPE	90	✓	FFPE	A	M	I	62	R0	A					MUT	✓	✓
17	>2010	THRBF30253T1	FFPE	90	✓	FFPE	A	F	IIA	53	R0	A	✓	Hypothyroidism			MUT	✓	✓
18	2000-9	THRBZ57149T2	Frozen	100	✓	FFPE	A	F	I	85	R0	A					MUT	✓	✓
19	2000-9	THRBZ85801T2	Frozen	100	✓	Frozen	A	F	I	63	R0	A					MUT	✓	✓
20	<2000	THRBF94124T1	FFPE	100	✓	FFPE	A	M	I	57	R0	0					WT	✓	✓
21	>2010	THRBZ61001T1	Frozen	100	✓	FFPE	A	M	I	77	R0	A						✓	✓
22	>2010	THRBZ80427T1	Frozen	100	✓	Frozen	A	M	IIA	75		11			Metastatic Transitional RCC			✓	✓
23	2000-9	THRBZ03917T2	Frozen	100	✓	FFPE	A (MNT LS)	M	IIA	72	R0	A					MUT*	✓	✓
24	<2000	THRBF15012T1	FFPE	95	✓	FFPE	A (MNT LS)	M	IIA	66	R0	A					WT	✓	✓
25	>2010	THRBF20233T2	FFPE	100	✓	FFPE	A (MNT LS)	M	IIA	75	R1	A					MUT*	✓	✓
26	<2000	THRBF86319T1	FFPE	90	✓	FFPE	A (MNT LS)	M	IIB	60	R0	A					WT	✓	✓
27	>2010	150077PT	FFPE				AB	F	I	74	R0	A					MUT	✓	✓
28	>2010	150101PT	FFPE				AB	F	II	56	R0	A			Concurrent Breast CA		MUT*	✓	✓
29	<2000	THRBF00959T1	FFPE	100	✓	FFPE	AB	F	I	61	R0	A					WT	✓	✓
30	2000-9	THRBZ02130T2	Frozen	100	✓	Frozen	AB	M		60		A			Multiple recurrences in the 70s (ie in his 30s)		MUT*	✓	✓
31	2000-9	THRBF05377T1	FFPE	100	✓	FFPE	AB	F	I	62	R0	88			Breast CA w Recurrence		WT	✓	✓
32	>2010	THRBZ00852T2	Frozen	100	✓	FFPE	AB										MUT*	✓	✓
33	>2010	THRBF09409T1	FFPE	100			AB	M	I	67	R0	A			Myxofibrosarcoma		MUT*	✓	✓
34	>2010	THRBF13694T1	FFPE	80	✓	FFPE	AB	M	IIA	60		A					MUT*	✓	✓
35	>2010	THRBZ14059T1	Frozen	100	✓	Frozen	AB	F	I	53	R0	A					MUT*	✓	✓
36	>2010	THRBF14859T1	FFPE	80	✓	FFPE	AB	F	IIA	64	R0	A		1ry hyperparathyroidism			MUT*	✓	✓
37	>2010	THRBF15330T1	FFPE	100	✓	FFPE	AB	F	I	61	R0	A					MUT*	✓	✓
38	>2010	THRBZ16097T3	Frozen	100	✓	Frozen	AB	M	I	72	R0	A					WT	✓	✓
39	2000-9	THRBF19872T1	FFPE	100	✓	FFPE	AB	F	I	72	R0	44					WT	✓	✓
40	<2000	THRBZ21732T2	Frozen	100	✓	FFPE	AB	M	I	57	R0	A					MUT*	✓	✓
41	2000-9	THRBF24855T2	FFPE	80	✓	FFPE	AB	M	I	70	R0	A				✓	MUT*	✓	✓
42	2000-9	THRBF25392T1	FFPE	100	✓	FFPE	AB	F		31		A	✓				WT	✓	✓
43	>2010	THRBZ25604T1	Frozen	100	✓	FFPE	AB	F	I	66	R0	A					MUT	✓	✓
44	>2010	THRBF25774T1	FFPE	100	✓	FFPE	AB	F	II	67	R1	A		Hypothyroidism	Atrial myxoma		MUT*	✓	✓
45	2000-9	THRBF28521T1	FFPE	100	✓	FFPE	AB	F	I	44	R0	A		Aplastic Anaemia			WT	✓	✓
46	<2000	THRBZ32960T2	Frozen	100	✓	FFPE	AB	F	IIA	66	R0	A			Colon CA		MUT	✓	✓
47	<2000	THRBZ33408T2	Frozen	100			AB	M	I	74	R0	A					MUT*	✓	✓
48	<2000	THRBZ42901T2	Frozen	100			AB	M	I	68	R0	134					MUT*	✓	✓
49	2000-9	THRBF54962T1	FFPE	80	✓	FFPE	AB	F	I	50	R0	A	✓	Good's Syndrome		✓	MUT*	✓	✓
50	2000-9	THRBF55773T1	FFPE	100	✓	FFPE	AB	F	IIA	41	R0	A	✓				WT	✓	✓
51	<2000	THRBF73550T1	FFPE	100	✓	FFPE	AB	F	I	60	R0	A					WT	✓	✓
52	2000-9	THRBZ73761T2	Frozen	100	✓	FFPE	AB	F	IIA	45	R0	A	✓				WT	✓	✓

53	2000-9	THRBF79419T1	FFPE	80	✓	FFPE	AB	F	I	75	R0	42					**MUT***	✓	✓
54	<2000	THRBZ8209T2	Frozen	100	✓	FFPE	AB	M	IIA	72	R0	A					**MUT***	✓	✓
55	2000-9	THRBF84054T1	FFPE	80			AB	F	IVA	49	R2	98			Thymoma Recurrence at 54		WT	✓	✓
56	2000-9	THRBZ95591T2	Frozen	95	✓	Frozen	AB	M	I	53	R0	61			Lung CA		**MUT***	✓	✓
57	2000-9	THRBZ95929T2	Frozen	90	✓	Frozen	AB	F	2A	73	R0	A					**MUT***	✓	✓
58	2000-9	THRBZ96635T2	Frozen	90	✓	Frozen	AB	F	I	56	R0	A					**MUT***	✓	✓
59	>2010	THRBZ00546T2	Frozen	100	✓	Frozen	AB	M	I	70	R0	A						✓	✓
60	>2010	THRBZ08002T1	Frozen	100	✓	Frozen	AB	F	IIA	77	R0	A						✓	✓
61	2000-9	THRBZ67094T1	Frozen				AB	F	I	54	R0	A						✓	✓
62	>2010	THRBF28180T1	FFPE	60	✓	FFPE	B1	F	IV	65	R0	A			Thymoma at 55		WT	✓	✓
63	>2010	150092PT	FFPE		✓	FFPE	B1	F		60	R1	A	✓				WT	✓	✓
64	>2010	150099PT	FFPE				B1	F	II	60	R0	A					WT	✓	✓
65	<2000	THBRF20178T1	FFPE	90	✓	FFPE	B1	M	IIA	49	R0	A				✓	WT	✓	✓
66	>2010	THRBF00036T2	FFPE	90	✓	FFPE	B1	F	IVA	43							WT	✓	✓
67	2000-9	THRBF10445T1	FFPE	100	✓	FFPE	B1	F	I	46	R0	40	✓				**MUT**	✓	✓
68	>2010	THRBF12588T1	FFPE	100	✓	FFPE	B1	F	II	70	R1	A					**MUT***	✓	✓
69	>2010	THRBF12695T1	FFPE	95	✓	FFPE	B1	F	IIA	54	R0	A					**MUT***	✓	✓
70	2000-9	THRBF13228T1	FFPE	80			B1	F	III	58	R0	A					WT	✓	✓
71	2000-9	THRBF15530T1	FFPE	75	✓	FFPE	B1	F	I	53	R0	A					WT	✓	✓
72	2000-9	THRBF16249T1	FFPE	100	✓	FFPE	B1	F	IVA	48	R0	A					WT	✓	✓
73	>2010	THRBF19516T2	FFPE	100	✓	FFPE	B1	F	IIA	64	R0	A					WT	✓	✓
74	>2010	THRBF23148T1	FFPE	100	✓	FFPE	B1	F	IIA	57		A			Hypothyroidism		WT	✓	✓
75	2000-9	THRBF25622T1	FFPE	100			B1	M		26		A					WT	✓	✓
76	2000-9	THRBF25781T1	FFPE	100			B1	F	III	53	R0	A			Thymoma Recurrence at 64		WT	✓	✓
77	2000-9	THRBZ56585T2	Frozen	100	✓	FFPE	B1	F	I	41	R0	A					WT	✓	✓
78	<2000	THRBF87472T1	FFPE	100	✓	FFPE	B1	M	IVB	37		114					WT	✓	✓
79	>2010	THRBZ02570T1	Frozen	100	✓	Frozen	B1	M	IIA	73	R0	A	✓					✓	✓
80	>2010	THRBZ03029T1	Frozen	100	✓	Frozen	B1	M	II	46	R1							✓	✓
81	2000-9	THRBZ84946T1	Frozen	90	✓	Frozen	B1	F	III	69	R0				RA		WT	✓	✓
82	<2000	THRBF59802T1	FFPE	100	✓	FFPE	B1	M	IVA	44	R0	A	✓				WT		✓
83	>2010	THRBF17723T1	FFPE		✓	FFPE	B1	M	II	60	R0	A							✓
84	2000-9	THRBZ74105T1	Frozen	100	✓	FFPE	B2	M	IVA	54	R0	A					**MUT***	✓	✓
85	>2010	10LW00453TIS1	Frozen	95			B2	F	IVB	56		28	✓		Previous Thymoma at 50			✓	✓
86	>2010	150095PT	FFPE				B2	F		55	R0	A					WT	✓	✓
87	<2000	THRBF00151T1	FFPE	100	✓	FFPE	B2	F	III	48	R0	A					WT	✓	✓
88	>2010	THRBF02631T1	FFPE	100	✓	FFPE	B2	M	II	59	R0	A					WT	✓	✓
89	>2010	THRBF05018T1	FFPE	90	✓	FFPE	B2	M	III	53	R0	A			Ankylosing spondylitis		WT	✓	✓
90	2000-9	THRBF05109T2	FFPE	30	✓	FFPE	B2	M	IVA	68	R1						WT	✓	✓
91	2000-9	THRBF06854T1	FFPE	100			B2	M	IIB	64	R0	A				✓	WT	✓	✓
92	2000-9	THRBF08261T1	FFPE		✓	FFPE	B2	M	III	66	R0	A					WT	✓	✓
93	>2010	THRBF08297T1	FFPE	100	✓	FFPE	B2	M	III	76	R0	A	✓				WT	✓	✓
94	2000-9	THRBF08609T1	FFPE	80	✓	FFPE	B2	F	IVA	42	R0		✓		Thymoma Recurrence at 45		WT	✓	✓
95	2000-9	THRBF13057T1	FFPE		✓	FFPE	B2	M	IIB	58	R0	A	✓				WT	✓	✓
96	>2010	THRBF15321T1	FFPE	100	✓	FFPE	B2	F	II	67		A					WT	✓	✓
97	>2010	THRBF19801T2	FFPE	50	✓	FFPE	B2	M	IIB	67		A					WT	✓	✓
98	2000-9	THRBF20036T1	FFPE	95			B2	F	IVA	40	R1	65			Thymoma Recurrence at 44		WT	✓	✓
99	>2010	THRBZ21401T1	Frozen	100	✓	Frozen	B2	F	IIA	60	R1	A	✓				WT	✓	✓
100	<2000	THRBF2161T1	FFPE	80			B2	F	IVA	65	N/A						WT	✓	✓
101	>2010	THRBF21890T1	FFPE	80	✓	FFPE	B2	M	II	60	R0	A			Basal Cell CA		WT	✓	✓
102	>2010	THRBF23542T1	FFPE	100	✓	FFPE	B2	M	III	63	R2						WT	✓	✓
103	2000-9	THRBF25074T1B	FFPE	90	✓	FFPE	B2	M	IIB	40	R0				Colorectal CA	✓	WT	✓	✓
104	<2000	THRBF25324T1	FFPE	100	✓	FFPE	B2	M	IIA	61	R0	A	✓				WT	✓	✓

105	>2010	THRBF27768T2	FFPE	100	✓	FFPE	B2	F	IVB	41	R0						WT	✓	✓
106	<2000	THRBF34401T1	FFPE	100	✓	FFPE	B2	M	IVA	70	R2	69					WT	✓	✓
107	<2000	THRBZ36836T2	Frozen		✓	FFPE	B2	M	IVA	30							WT	✓	✓
108	2000-9	THRBF47489T2	FFPE	90	✓	FFPE	B2	F	IVA	21	R0						WT	✓	✓
109	2000-9	THRBZ48563	Frozen	100	✓	FFPE	B2	F	IVA	68	R1						WT	✓	✓
110	2000-9	THRBF49184T1	FFPE	95			B2	F		57		A	✓		Resection 2000 and recurrence 2001		WT	✓	✓
111	2000-9	THRBF60143T1	FFPE	90	✓	FFPE	B2	F	IVA	66	R0	A	✓				WT	✓	✓
112	2000-9	THRBF62382T1	FFPE	50	✓	FFPE	B2	F	IIA	38	R0	A	✓				**MUT**	✓	✓
113	2000-9	THRBF89748T2	FFPE	20	✓	FFPE	B2	F	I	18	R0	A	✓				WT	✓	✓
114	<2000	THRBF99775T1	FFPE	100	✓	FFPE	B2	F	III	56	R1	A	✓				WT	✓	✓
115	<2000	THRBF99782T1	FFPE	100	✓	FFPE	B2	M	IIB	78	R0	A					WT	✓	✓
116	>2010	THRBZ11462T2	Frozen	100	✓	FFPE	B2	F	IIA	71	R0	3	✓						✓
117	<2000	TCRBZ16327T2	Frozen	100	✓	FFPE	B2 (!)	F	IVA	62	R0	15	✓			✓	WT	✓	✓
118	>2010	150108PT	FFPE				B2&B3	M	IV (Rec)	47	R0	A	✓		Thymoma resected at 42		WT	✓	✓
119	>2010	150089PT	FFPE				B3	M	III	73		A					WT	✓	✓
120	>2010	150096PT	FFPE				B3	M		77		A					**MUT**	✓	✓
121	>2010	THRBF10175T1	FFPE	100	✓	FFPE	B3	F	II	48	R0	A	✓				WT	✓	✓
122	2000-9	THRBF13799T1	FFPE		✓	FFPE	B3	F	IIA	63	R0	A				✓	WT	✓	✓
123	>2010	THRBF23949T2	FFPE	80	✓	FFPE	B3	M	II	74	R0	A					WT	✓	✓
124	2000-9	THRBF70621T1	FFPE	95	✓	FFPE	B3	F	IVB	30	R0	25				✓	WT	✓	✓
125	>2010	THRBZ16503T2	Frozen	100			B3	M	IIA	58	R0	A			Lymphoblastic leukemia			✓	✓
126	>2010	TCRBF12844T1	FFPE	100	✓	FFPE	B3 (!)	M	IVB	59	R1	A					WT	✓	✓
127	<2000	TCRBZ03870T1	Frozen	50	✓	FFPE	CA	M	IVB	48	R0	160	✓		Renal Cell Carcinoma same year	✓	WT	✓	✓
128	>2010	TCRBF06753T1	FFPE	100			CA	M	III	64	R1			Coelial Disease	Gastric CA at 63		WT	✓	✓
129	>2010	TCRBF08757T1	FFPE	80	✓	FFPE	CA	F		63							WT	✓	✓
130	2000-9	TCRBF15505T3	FFPE	70	✓	FFPE	CA	F	IVB	58	R0	46		RA		✓	WT	✓	✓
131	2000-9	TCRBF15586T1B	FFPE	50	✓	FFPE	CA	F	IIB	50	R0	A			Uterine Dysgerminoma at 34	✓	WT	✓	✓
132	<2000	TCRBF16264T1	FFPE	60	✓	FFPE	CA	M	III	60	R0	75				✓	WT	✓	✓
133	>2010	TCRBF19400T1	FFPE	100	✓	FFPE	CA	M		73	R0	A			Thymic CA at 71		WT	✓	✓
134	2000-9	TCRBF21432T3	FFPE	80	✓	FFPE	CA	M	III	64	R2					✓	WT	✓	✓
135	>2010	TCRBF25163T3	FFPE	30	✓	FFPE	CA		III		R0						WT	✓	✓
136	>2010	TCRBF26901T1	FFPE	100			CA	M	III	69	R1	A					WT	✓	✓
137	2000-9	TCRBF31499T1	FFPE	100	✓	FFPE	CA										WT	✓	✓
138	<2000	TCRBF91244T4	FFPE	90			CA	M	IIB	51	R0	108			RCC, Mesothelioma, Melanoma*	✓	WT	✓	✓
139	>2010	150086PB	FFPE	N/A			Hyperplasia	F	N/A	29	R0	A			Cervical adenoCA		WT	✓	✓
140	2000-9	TPRBF97974P1	FFPE	N/A			Hyperplasia	F	N/A	23	R0	A					WT	✓	✓
141	2000-9	TPRBF26458	FFPE	N/A	✓	FFPE	Hyperplasia	F	N/A	40		A		Hyperthyroidism			WT		
142	2000-9	TPRBF018	FFPE	N/A	✓	FFPE	Hyperplasia		N/A			A							
143	2000-9	TPRBF1519	FFPE	N/A			Hyperplasia	F	N/A	28		A	✓						
144	2000-9	TPRBF04051	FFPE	N/A			Hyperplasia		N/A										
145	2000-9	TPRBF02155	FFPE	N/A			Hyperplasia		N/A			A							
146	2000-9	TPRBF26464	Frozen	N/A			Hyperplasia	F	N/A	23		A					WT		
147	>2010	150075FB	Frozen	N/A	✓	FFPE	Hyperplasia	F	N/A	39	R0	A			Ovarian CA		WT		
148	>2010	150087PB	FFPE	N/A			Hyperplasia	M	N/A	27	R0	A					WT		
149	>2010	150104PB	FFPE	N/A			Hyperplasia	F	N/A	44	R0	A					WT		
150	>2010	150114PB	FFPE	N/A			Hyperplasia	F	N/A	34	R0	A	✓				WT		
151	2000-9	TNRBZ63876T2	Frozen	100	✓	FFPE	NETT	M	IVB	52	R0						WT	✓	✓
152	2000-9	TCRBZ64925T1	Frozen	100	✓	FFPE	NETT	F				108			Thymoma recurrence 8 years later		WT	✓	✓
153	2000-9	TNRBF73531T1	FFPE	90	✓	FFPE	NETT	M	IIA	58	R0	81				✓	WT	✓	✓
154	<2000	TNRBF76261T1	FFPE		✓	FFPE	NETT	F	IVB	49	R0	46					WT	✓	✓
155	2000-9	TNRBF92784T1	FFPE	100	✓	FFPE	NETT	F	IVB	57	R0					✓	WT	✓	✓
156	2000-9	TNRBF95414T1	FFPE	70	✓	FFPE	NETT	F	IVB	67	R0	7				✓	WT	✓	✓

Appendix 3

Clustering of TET subtypes based upon staining with all antibodies – Raw data for figure 7.29

PD-L1 E	PD-L1 L	PD-1 L	PD-L2 E	PD-L2 L	B7-H3 E	PDLI M1 E	TIM-3 E	HGF E	Subtypes	Pat	Co-exp
200	150	30		140	90	160	50	100	MNT	18	8
200	15	60	100	200	300	100		100	MNT	21	8
80	160	15		160	200	30	70	20	AB	23	8
70	15	10	100	70	300	30		100	AB	48	8
160	30	60	100	160	200	60	90		AB	51	8
200	200	15	80	180	40	210		20	B1	63	8
240	90	60	80	60	180	300		60	B2	82	8
180	90	60	200	120	70	120		30	B2	91	8
90	160	30	100	80	70	120		40	B2	97	8
160	90	20	100	120	80	80		20	B2	100	8
30	30	15	120		180	160	40	100	CA	108	8
120	90	15	200		200	60	70	100	NETT	125	8
160	20	6	60	80	160			80	A	3	7
	140	30	100	120	100	60		100	MNT	19	7
100	60	20		70	100	40		40	AB	25	7
160	90	120	100	90	200	80			AB	45	7
80	90	120	100	70	30	40			AB	47	7
160	90	60	100	150	200	80			AB	49	7
100	100	40		80	100	140	70		B1	52	7
160	30	20		160	50	20	80		B1	57	7
160		60	100	160	200	300		40	B1	72	7
80	60	60	100	80	30	100			B1	74	7
200	120	30		60	80	30		60	B2	79	7
160	160	5		160	100	140	30		B2	81	7
100	60	15	100	80	70	60			B2	83	7
140	30		160	80	30	140		60	B2	85	7
160		15	100	210	40	160		160	B2	87	7
160	30	30	120	180	50	120			B2	88	7
160	90	20	80	160	100	120			B2	96	7
180	180	15	80	140	80	120			B2	99	7
100	30	40		80	50	140	70		B3	103	7
120	80		80		100	200	60	60	CA	109	7
200	80	5	90		100	200	50		CA	111	7
	160	90	200		200	80	90	100	CA	116	7
130		30	100		180	200	70	100	NETT	124	7
160	180	20		60	180	70			A	1	6
80		15	100	200	200	30			A	10	6
40	120	20	120		300	70			A	17	6
	210		100	80	80	50		40	MNT	20	6

100	60	30		40	100	40			AB	24	6
160	80	10		80	60			30	AB	26	6
60	40	30		60	200	30			AB	27	6
100	60	60		40	240	40			AB	28	6
180	20	20		70	90	30			AB	29	6
200	30	40		160	180	30			AB	30	6
160	40	60		120	180	80			AB	39	6
80	30	20		30	180	60			AB	43	6
120		15		40	100	70		30	AB	50	6
100	80	20		40	60	20			B1	53	6
100	120	20		80	160	80			B1	54	6
160	160	10		120	50	10			B1	55	6
160	20	40		180	30	40			B1	56	6
200	30	15		60	20	60			B1	58	6
	160	90	120	120	40	80			B1	66	6
	80	5	60	120	140	140			B1	67	6
70	70		100	120	90	40			B1	68	6
160		30	100	80	200	140			B1	69	6
160	60	30	100	180		300			B1	70	6
80		10	100	160	120	60			B1	71	6
160	160	30	200	160		180			B1	75	6
	240	150	100	160	200	180			B1	76	6
100		15		120	90	120		40	B2	77	6
160	160	3		160	100	80			B2	80	6
200		60	90	80	20	60			B2	84	6
200		20	100	160	60	140			B2	86	6
200			200	40	60	200		30	B2	89	6
140	20	15	300		200	240			B2	95	6
60	160	90	100	160		80			B2	101	6
240		60	300	120	100	180			B2	102	6
200		3	200		100	200		100	B3	104	6
160		10	200		300	300		120	B3	106	6
110		30	240		200	200		10	CA	113	6
200			200		140	300	60	200	NETT	119	6
100	80	120		40	20	140			Hyper	126	6
80		20		180	120	140		80	Hyper	128	6
	80	10		120	20	30		20	Hyper	130	6
	80	40		160	30	80		20	Hyper	132	6
	160	100		160	180	100		30	Hyper	138	6
105	200	20		160	200	100			Hyper	139	6
240	120	90		120	50	20			Hyper	140	6
	120	80		140	80	80		80	Hyper	141	6
80	120			100	100	200			A	6	5
200		3	120		200			60	A	9	5
200			200		90	120		100	A	11	5
100			100	160	300	70			A	12	5

80			100	90	200	80			A	14	5
160		15	100	160		300			AB	32	5
200		20	200		200	120			AB	33	5
	15	20		120	180	30			AB	37	5
120		40	80	80	200				AB	42	5
80	60	30	100	140					AB	44	5
80	15	15	200		300				AB	46	5
90	40	40		80		30			B1	59	5
180	220	40		120		60			B1	60	5
160	160				90	120		60	B1	61	5
160		120	40	220		80			B1	73	5
100		10		70		40	40		B2	78	5
90			60	40		20		60	B2	90	5
160			200	90	70	80			B2	92	5
80		30	60	120	160				B2	93	5
200		10	200		80	160			B2	98	5
60		20	200		200	300			B3	105	5
200		15	100		200	80			B3	107	5
200	80		180		100	100			CA	110	5
80	30		70		100	40			CA	112	5
150	300		160		300	200			CA	114	5
200			200		100		80	200	NETT	118	5
300			180		180	70		100	NETT	121	5
300			200		180		70	200	NETT	123	5
40		60		140		80		120	Hyper	127	5
	80	90		120	40	30			Hyper	131	5
	80	40		160	160	80			Hyper	133	5
	160	120		120	100	40			Hyper	135	5
	120	80		140	200	140			Hyper	136	5
	160	20		160	70	60			Hyper	137	5
	120	20	240	60	60				Hyper	142	5
-	-	60		160	270		30		A	2	4
20		6	40		180				A	4	4
		5		40	100	70			A	5	4
100				40	300	70			A	7	4
	60		100	160	300				A	13	4
200		30		140	300				AB	22	4
100		15		69	20				AB	31	4
		9		100	200	70			AB	35	4
240		10	80			40			AB	38	4
	80	30		60	180				AB	41	4
120	90		100	30					B2	94	4
90			200		200	100			CA	115	4
200			200		100	160			CA	117	4
	170			180	30	80			Hyper	134	4
	240			70	100				A	16	3

		15		120	200				AB	34	3
		15		80	80				AB	36	3
		20	30		300				AB	40	3
		15		140		100			B1	62	3
				140	80	10			B1	64	3
180			100					100	NETT	120	3
90			200				60		NETT	122	3
	120	15		120					Hyper	129	3
	120	90		160					Hyper	143	3
			60		200				A	8	2
	15				160				A	15	2
	80		160						B1	65	2

Appendix 4

Copy of licence to use two illustrations from Elsevier

ELSEVIER LICENSE
TERMS AND CONDITIONS
Jul 10, 2017

This Agreement between Spyridon Gennatas ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number 4145650684424
License date Jul 10, 2017
Licensed Content Publisher Elsevier
Licensed Content Publication Thoracic Surgery Clinics
Licensed Content Title Anatomy of the Thymus Gland
Licensed Content Author Najib Safieddine,Shaf Keshavjee
Licensed Content Date May 1, 2011
Licensed Content Volume 21
Licensed Content Issue 2
Licensed Content Pages 5
Start Page 191
End Page 195
Type of Use reuse in a thesis/dissertation
Portion figures/tables/illustrations
Number of figures/tables/illustrations
2
Format both print and electronic
Are you the author of this Elsevier article?
No
Will you be translating? No
Order reference number
Original figure numbers 1, 2
Title of your thesis/dissertation
Genetic alterations leading to thymic epithelial tumours
Expected completion date Aug 2017
Estimated size (number of pages)
250
Elsevier VAT number GB 494 6272 12
Requestor Location Spyridon Gennatas
Flat 5, 32 Nevern Square
London, SW5 9PE
United Kingdom

Attn: Spyridon Gennatas
Publisher Tax ID GB 494 6272 12
Total 0.00 USD
Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at

http://myaccount.copyright.com).
GENERAL TERMS
2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to
the terms and conditions indicated.
3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:
"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."
4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.
5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.
6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.
7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed
use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be
deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement

and publisher reserves the right to take any and all action to protect its copyright in the materials.
9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.
11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.
12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).
13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.
14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described
in this License at their sole discretion, for any reason or no reason, with a full refund payable
to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.
LIMITED LICENSE
The following terms and conditions apply only to specific license types:
15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.
16. Posting licensed content on any Website: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must
maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.
Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier
homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.
Posting licensed content on Electronic reserve: In addition to the above the following

clauses are applicable: The web site must be password-protected and made available only to

bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.
17. For journal authors: the following clauses are applicable in addition to the above:
Preprints:
A preprint is an author's own write-up of research results and analysis, it has not been peerreviewed,
nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).
Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).
If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.
Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes authorincorporated
changes suggested during submission, peer review and editor-author communications.
Authors can share their accepted author manuscript:
immediately
via their non-commercial person homepage or blog
by updating a preprint in arXiv or RePEc with the accepted manuscript
via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
directly by providing copies to their students or to research collaborators for their personal use
for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
After the embargo period
via non-commercial hosting platforms such as their institutional repository
via commercial sites with which Elsevier has an agreement
In all cases accepted manuscripts should:
link to the formal publication via its DOI
bear a CC-BY-NC-ND license - this is easy to do
if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.
Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:
Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect,

and so links will help your users to find, access, cite, and use the best available version.
Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.
If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.
Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.
Please refer to Elsevier's posting policy for further information.
18. For book authors the following clauses are applicable in addition to the above:
Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. Posting to a repository: Authors are permitted to post a summary of their chapter only in their institution's repository.
19. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.
Elsevier Open Access Terms and Conditions
You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.
Terms & Conditions applicable to all Open Access articles published with Elsevier:
Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.
The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.
If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.
Additional Terms & Conditions applicable to each Creative Commons user license:
CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new

works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.
CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts,

abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.
CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of
the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0.
Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.
Commercial reuse includes:
Associating advertising with the full text of the Article
Charging fees for document delivery or access
Article aggregation
Systematic distribution via e-mail lists or share buttons
Posting or linking by commercial companies for use by customers of those companies.
20. Other Conditions:
v1.9

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or
+1-978-646-2777.

Appendix 5
Presentations and publications

Oral presentations

'Genetic alterations leading to thymic epithelial malignancies'
'Hormones & Cancer' conference (Attendance awarded with 6 CPD points), Areteion Hospital, Athens, Greece. 14 June 2015.

'Thymic Epithelial Tumours'
Speaker and member of the Scientific Committee of the '3rd International Symposium on Thoracic and Upper Aerodigestive Malignancies'. Hilton Hotel, Athens, Greece, 5-7 February 2015.

'Validation of a specific missense GTF2I mutation in more indolent thymic epithelial tumours'. 16th World Conference on Lung Cancer. (International Association for the Study of Lung Cancer). Denver, Colorado, USA. September 6 - 9 2015.

'Comprehensive characterization of thymic epithelial tumour subtypes through an analysis of somatic mutations and copy number alterations'
Gennatas S, Mandal A, Nicholson A.G, Popat S, Bowcock AM
Accepted as an oral presentation at IASLC WCLC 2017 - IASLC 18th World Conference on Lung Cancer on 16/10/2017

Poster presentations

'Somatic and germline mutations in thymic epithelial tumours and their correlation with histological and clinical phenotypes'. '15th Annual British Thoracic Oncology Group Conference 2017'. Doubletree Dublin by Hilton - Burlington Road, Upper Leeson St, Dublin 4, Ireland. 25th – 27th January 2017.

'B7-H3 protein expression in Thymic Epithelial Tumour subtypes and its association with PD-L1 and clinical characteristics'
Gennatas S1, Mandal A1, Robertus JL2, Bowman A3, Nicholson A.G2, Popat S4, Bowcock AM1
Accepted for a poster presentation at IASLC WCLC 2017 - IASLC 18th World Conference on Lung Cancer on 16/10/2017

Published abstracts

'Validation of a Specific Missense GTF2I Mutation in More Indolent Thymic Epithelial Tumours'
Gennatas S, Anbunathan H, Montero FA, Nicholson AG, Popat S, Bowcock, AM
Journal of Thoracic Oncology, 2015;10 (MINI38.03). ISSN 1556-0864.

'Expression of PD-L1, PD-L2 and PD-1 in thymic epithelial tumours (TETs)'.
Gennatas S, Montero MA, Bowman A, Ronquillo J, Nicholson A, Popat S, Bowcock AM
ASCO 2016 Annual Meeting (Publication)

'Somatic and germline mutations in thymic epithelial tumours and their correlation with histological and clinical phenotypes'.
Gennatas S., Anbunathan H., Bowman, A., Montero A., Nicholson A.G., Popat S., Bowcock A.M
Poster Abstracts of the 15th Annual British Thoracic Oncology Group Conference 2017, S52(116) Volume 103, Supplement 1 (2017).
DOI: http://dx.doi.org/10.1016/S0169-5002(17)30166-6

Publications

Submitted to Annals of Oncology:
***GTF2I* mutations in indolent thymomas as a diagnostic aid in difficult to diagnose mediastinal masses**
S. Gennatas, A. Mandal, A. Bowman, E. Lim, A. G. Nicholson, S. Popat, A. M. Bowcock