Driver Genes in CTCL
A Genomic Approach for Screening and a Detailed Analysis of a Candidate Gene

Woollard, Wesley Jackson

Awarding institution:
King's College London

Driver Genes in CTCL: A Genomic Approach for Screening and a Detailed Analysis of a Candidate Gene

Wesley J. Woollard

Genetics and Molecular Medicine PhD

Table of Contents

1 Introduction 12

1.1 Cutaneous T-cell lymphoma overview 12

1.2 Mycosis fungoides background 12

1.3 Sézary Syndrome background 16

1.4 Similarities and differences between MF and SS 18

1.5 Treatment of MF and SS 22

1.6 Early genomic studies of MF and SS 23

1.7 The driver mutation model of cancer 24

1.8 Molecular processes and signalling networks affected in MF and SS 26

1.8.1 NF-κB 26

1.8.2 JAK-STATs 29

1.8.3 PI3K-PKB/AKT 31

1.8.4 RAS/RAF 33

1.8.5 Phospholipase C gamma pathways 34

1.8.6 Programmed cell death 35

1.8.7 Global epigenetic regulation 41

1.8.8 Cell cycle control 44

1.8.9 DNA repair and genome maintenance processes 49

1.9 Case study of a common focal deletion: driver genes on chromosome 9p21.3 56

1.9.1 CDKN2A and CDKN2B 57

1.9.2 MTAP 57

1.10 The application of next generation sequencing to the study of cancer 61

1.10.1 Overview of NGS 61

1.10.2 NGS as a discovery tool 62

1.10.3 NGS as a molecular diagnostic tool 63

1.10.4 Challenges presented by NGS 64

1.11 Identifying cancer driver genes 65

1.11.1 Preparing libraries for next generation sequencing 65

1.11.2 Upstream bioinformatic analysis 66

1.11.3 Bioinformatic approaches and the interpretation bottleneck 68

1.11.4 Challenges for current methods of identifying cancer drivers 74

1.11.5 Functional approaches 75

1.12 Hypothesis 77

1.13 Aims of this thesis 77
1.13.1 Investigating the role of *MTAP* in CTCL: A candidate gene based approach 77
1.13.2 Identification of driver events in CTCL: A high throughput global approach 77
2 Materials and Methods 79
2.1 Collections and Processing of Samples 79
2.2 Isolation of PBMCs from Whole Blood 80
2.3 Enrichment of CD4+ T-Cells from PBMCs 80
2.4 Isolation of Fibroblasts 80
2.5 Mycoplasma testing 80
2.6 DNA Extraction 80
2.7 RNA Extraction 81
2.8 Synthesis of cDNA 81
2.9 PCR 81
2.10 QPCR 82
2.11 Whole Exome Capture and Sequencing 83
2.12 Targeted Capture and Sequencing 83
2.13 Preparation of Whole Cell Lysates 84
2.14 Immunoblotting 84
2.15 Stripping Membranes for Re-probing 85
2.16 Detection of Proteins 85
2.17 Thawing Frozen Cells 85
2.18 SeAx culture and passage 86
2.19 HEK293 culture and passage 86
2.20 Freezing Cells 86
2.21 Nucleofection 86
2.22 Transfection with Lipofectamine 2000 87
2.23 Micro-dissection 88
2.24 Variant calling 88
2.25 Selection of minimum allele frequency 89
2.26 Mutational pattern analysis. 89
2.27 Identification of SNV drivers 90
2.28 Gene copy number analysis 90
2.29 Pathway analysis 90
3 Independent loss of methylthioadenosine phosphorylase *(MTAP)* in Primary Cutaneous T-cell Lymphoma 92

3.1 Candidate contributions to chapter 3. 93
3.2 Additional statistical tests for chapter 3. 93
3.3 Additional figures and data for chapter 3 95
4 The effect of MTAP loss on type I PRMT activity 100
4.1 Backgound 101
4.2 Expression of PRMTs in SeAx and Jurkat cells 101
4.3 Transient knock-down of MTAP in SeAx cells 102
4.4 Proof of principle experiments in HEK293 cells 104
4.5 Expression of PRMTs in HEK293 cells 104
4.6 Transient knock-down of MTAP in HEK293 cells 104
4.6.1 Assessment of *MTAP* expression by RT-PCR 105
4.6.2 Assessment of MTAP expression by RT-QPCR 106
4.7 Assessment of MTAP protein expression 107
4.8 Assessment of pan asymmetric-dimethyle-arginine methylation in MTAP knock-down HEK293 cells 108
5 Candidate driver genes in Sézary syndrome: frequent perturbations of genes involved in genome maintenance and DNA repair 110
5.1 Candidate contributions to chapter 5 111
5.2 Introductory remarks for chapter 5 111
5.3 Additional figures for chapter 5 112
6 Discussion 114
6.1 Focal deletions of 9p21 in CTCL 114
6.2 The effects of MTAP loss on type I PRMT activity 115
6.3 High throughput screening for putative driver genes in CTCL 117
7 Conclusions 119
7.1 Future work 122
7.2 Concluding remarks 124
8 References 126

List of Figures

Figure 1.1 MF stages of progression 13
Figure 1.2 Pautrier micro-abscess. 14
Figure 1.3 A Sézary cell. 14
Figure 1.4 Erythroderma in a Sézary syndrome patient. 17
Figure 1.5 The lineage of mitotic cell division leading to neoplastic transformation. 24
Figure 1.6 Hallmarks of Cancer. 25
Figure 1.7 NF-κB pathway. 27
Figure 1.8 JAK-STAT pathways. 30
Figure 1.9 Signal transduction through RTKs and GPCRs. 32
Figure 1.10 PLCγ1 signal transduction. 35
Figure 1.11 Apoptosis. 37
Figure 1.12 Loss of MTAP may inhibit autophagy. 40
Figure 1.13 G1/S Regulation. 46
Figure 1.14 G2/M Regulation. 48
Figure 1.15 Summary of DNA damage and repair mechanisms. 53
Figure 1.16 The structure of human telomeres. 55
Figure 1.17 Schematic of genes in 9p21.3. 57
Figure 1.18 How MTAP may act as a tumour suppressor. 59
Figure 1.19 Summary of Library preparation for NGS. 66
Figure 1.20 Calling variants in NGS data. 67
Figure 1.21 Distribution of mutations in two oncogenes (PIK3CA and IDH1) and two tumor suppressor genes (RB1 and VHL). 71
Figure 1.22 Inter and intra-heterogeneity of mutational patterns across different cancer types. 73
Figure 3.1. Tukey's HSD for probe sets 1 and 2 for gene copy number analysis 95
Figure 3.2. Linear Regression analysis to determine *MTAP*/*CDKN2A* Copy Number 96
Figure 3.3. Micro-dissection of tumour rich tissue and epidermal regions from FFPE sections 97
Figure 3.4. Determination of *CDKN2A* and *MTAP* Copy Number in DNA from micro-dissected tissue 99
Figure 4.1 Expression of PRMTs in SeAx and Jurkat. 102
Figure 4.2 Nucleofection of SeAx cells. 103
Figure 4.3 Transfection of SeAx by Lipofectamine2000. 103
Figure 4.4 Expression of PRMTs in HEK293 cells. 104
Figure 4.5 HEK293 cells transfected with pmaxGFP. 105
Figure 4.6 Expression of MTAP and PPIA by RT-PCR in HEK293. 106
Figure 4.7 MTAP expression in HEK293 cells analysed by RT-QPCR. 107
Figure 4.8 Immunoblot analysis of MTAP, Asymetric dimethyl arginine methylation and β-actin. 108
Figure 5.1. Heatmap of tumours affected by perturbations in gene families. 112
Figure 5.2. Summary heatmap showing targeted capture gene alterations 113

List of Tables

Table 1. Summary of features of MF and SS 20
Table 2. Common chromosomal changes associated with MF and SS 21
Table 3. Genes reported as mutated in NF-KB in CTCL..... 27
Table 4. Genes with reported mutations in the JAK/STAT pathway in CTCL. 30
Table 5. Genes with mutations in RTK and GPCR pathways in CTCL. 32
Table 6. Genes involved in DNA methylation reported as mutated in CTCL. 43
Table 7. Genes with roles in chromatin regulation with mutations reported in CTCL. 44
Table 8. Genes reported as mutated or deleted in CTCL which affect the cell cycle and checkpoint activation. 48
Table 9 Taqman QPCR probes..... 82
Table 10 List of antibodies and optimized conditions for use. 85
Table 11 List of siRNAs. 87

Acknowledgements

I would like to express my greatest thanks to my supervisors Tracey Mitchell and Sean Whittaker for the mentoring and guidance you have both shown throughout my time at King's. I feel I have made much progress in my development as a scientist and particularly in my skills as a writer. I would also like to thank you both for the patience, flexibility, understanding and support you have shown as I have faced testing challenges in the adjustment from life as a first year PhD student living in halls to where I am now as a proud father, homeowner and (soon to be) post-doctoral researcher at UCL. I feel that just three years is a very short time for such a transition and I owe you my deepest gratitude for making sure my focus stayed fixed on the target.

I am very grateful to the British Skin Foundation for funding the project and in doing so, giving me the opportunity to pursue this PhD. I would also like to thank numerous other friends and colleagues at King's especially our group members including Silvia, Carl, Vidyha, Varsha and Rosie who have been invaluable sources of advice, information, research materials, social chit-chat and everything in between. I would also like to extend a great deal of gratitude to Bill Grey for giving me an essential leg up with western blotting. It is fair to say I would have been very lost without your clear guidance.

Analysing big genomic datasets is still far from being standardized. Developing creative ways to extract as much information as possible from an experiment requires a particular way of thinking. It is to that end that I would like to thank Michael Simpson as well as the BRC team for your input towards evolving an appropriate methodology for analysing our data. But more importantly, much of my insight into bioinformatics would probably never have developed without your inspiration. Before starting my PhD it hadn't occurred to me that I could one day program in python, make beautiful plots in R or analyse genomic data on a high performance computing cluster. These are life-long skills that have opened up numerous doors for me, thank you for giving me a new perspective.

Finally, I would like to thank my parents and siblings who have been there since the start and always embraced by choice of career. But most of all I would like to express my love and thanks to my more recent family members; my girlfriend Nicola and our daughter Aurelia for putting up with my constant absence and always reminding me that I have a beautiful life to come home to away from the lab and the computer. We have kept it together despite what has seemed at times to be an unrelenting and never ending pressure and now the extra time is ours to enjoy.

Abstract

The cutaneous T-cell lymphomas (CTCLs) mycosis fungoides (MF) and Sézary syndrome (SS) are T-cell malignancies affecting the skin. The heterogeneous genomic landscape of CTCL has hindered the identification of driver genes. However, maturing sequencing technologies and candidate gene studies of chromosomal hotspots such as 9p21 can be used to identify putative driver events. The aims of this thesis were to investigate: (i) if *MTAP* (found on 9p21) undergoes selective loss; (ii) if *MTAP* loss affects dimethyl-arginine status; and (iii) identify putative driver genes of CTCL. Tumour samples from 280 CTCL patients were analysed by QPCR for copy number changes across 9p21. There is a high propensity for *MTAP* loss (34%) occurring independently of *CDKN2A* loss (12%) across all CTCL stages. Expression of *MTAP* mRNA was measured by QPCR in 20 tumours, reduced *MTAP* mRNA coincides with *MTAP* loss. However, reduced mRNA also occurs in tumours without *MTAP* loss indicating selective pressure to silence *MTAP*. Genomic screening indicates that mutations are rare in *CDKN2A* and *MTAP* in CTCL suggesting that epigenetic mechanisms may be repressing *MTAP* expression. Promoter methylation studies in CTCL cell lines suggest aberrant methylation inhibits *MTAP* expression. Knock-down of *MTAP* by RNAi in HEK293 cells suggested protein-arginine methylation maybe reduced in MTAP absence; though further confirmation is required. A genomic screen of 101 tumours and 32 healthy controls, including 10 germline matched whole exomes and 91 tumours undergoing targeted sequencing was used to probe for driver events. Novel putative drivers identified include: genes which maintain genome integrity; *POT1* (14%) and *BRCA2* (14%), master epigenetic regulators; *ASXL3* (8%) and *KDM6A* (6%), and programmed cell death regulators; *PDCD11* (35%) and *TRPM3* (11%). In conclusion, putative driver events of CTCL have been identified using; (i) targeted analysis of a candidate region to identify selective pressure silencing *MTAP*; and (ii) a genomic sequencing approach to define a shortlist of genes.

Abbreviations

BCNU – Bis-chloroethylenitrosourea

BER – Base excision repair

CTCL – Cutaneous T-cell lymphoma

CNV – Copy number variations

DDR – DNA damage response

DISC – Death inducing signalling complex

DSB – Double stranded break

ECM – Extra-cellular matrix

ECP - Extracorporeal photophoresis

FP – False positive

FN – False negative

HC – Healthy control

HDAC – Histone deacetylase

HR – homologus recombination

INDEL – Insertion or deletion

LOH – Loss of heterozygousity

MCC – Mathews correlation coefficient

MF – Mycosis fungoides

MMR – Missmatch repair

MPS – Massively parallel sequencing

MSP – Methylation specific PCR

MTA – Methylthioadenosine

NER – Nucelotide excision repair

NHEJ – Non-homologus end joining

NGS – Next generation sequencing

PBMC – Peripheral blood mononucleocyte

PBS - Phosphate buffered saline

PCD – Programmed cell death

PCR – Polymerase chain reaction

PUVA – Psoralen and UVA

ROS – Reactive oxygen species

RPMI – Roswell Park Memorial Institute

SNV – Single nucleotide variation

SS – Sézary syndrome

TBST – Tris buffered saline and 1% tween

TC – Targeted capture

TCR – T-cell receptor

TN – True negative

TP – True positive

VAF – Variant allele frequency

WES – Whole exome sequencing

WT – Wild type

1 Introduction

1.1 Cutaneous T-cell lymphoma overview

Cutaneous T-cell lymphomas (CTCL) are a rare group of extra nodal, non-Hodgkin lymphomas (1). There are a variety of subgroups of CTCL which exhibit wide phenotypic heterogeneity varying by their histopathological, clinical, immunophenotypic and genetic features (1, 2). Though a heterogeneous disease group, all CTCLs are a malignancy of mature, skin homing T-lymphocytes which maintain a persistent effect on the skin (2). The overall incidence of CTCLs has been observed at 4.1/1,000,000 people per year (3), though discrepancies may appear in the literature which are likely due to differences in patient cohorts used to assess the incidence, as well as the demographic and genetic background of the population from which the cohort was described. The most common CTCL subgroups are mycosis fungoides (MF) and Sézary syndrome (SS) (4) which together will be the focus of this PhD thesis.

1.2 Mycosis fungoides background

The name 'mycosis fungoides' is related to the mushroom like appearance of the nodules in the tumour stage of the disease (**Figure 1.1**)(2) and was first recognized over 2 centuries ago by Alibert and Bazin (2, 4). The incidence of MF is estimated to be between 0.36-0.90 cases per 100,000 people per year (5, 6). The median age of patients at diagnosis is 57 years although MF can occur in children and adolescents (2, 4). There is a higher incidence in people of African descent (1) and a gender bias in MF with males at greater risk than females (2:1)(7). The clinical course of MF is generally indolent with a disease specific 5 year survival of 88% (4), although this is dependent on disease stage; patients with stages IA, IB and IIA typically survive over 15 years whereas IIB and more advanced stages have median survival less than 5 years (3). A small subset of MF patients progress to develop advanced stage disease which will be discussed in later sections. (See table 1 for summary of disease features)

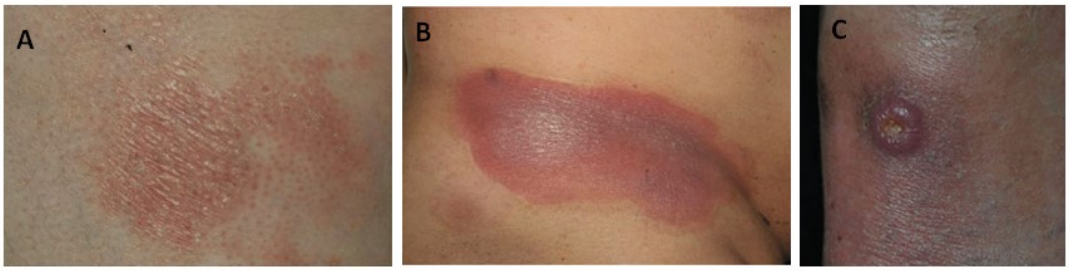


Figure 1.1 MF stages of progression
(A) Patch (B) Plaque (C) Tumour. Adapted from: An. Bras. Dermatol. vol.87 no.6 Rio de Janeiro Nov./Dec. 2012.

The early phenotype of MF presents scaly patches and plaques which occur in multiple sun protected regions of the skin such as the proximal extremities and the trunk (8)(**Figure 1.1**). Early MF bears a resemblance to benign skin disorders such as psoriasis, eczema or contact dermatitis and these features, in combination with its low incidence, means MF can remain undiagnosed for many years (2). Progression from stage I or II MF to stage III or IV disease occurs in 20% of patients (9). Those MF patients who show disease progression show a significantly reduced survival with median typically less than five years (3). The later stages of MF present with multiple patches and plaques as well as ulcerated tumours, erythroderma and leukemic involvement, with some patients showing involvement of the visceral organs and lymph nodes (4, 7).

Histological features of MF are dependent on disease stage, with the earlier patch stage showing band-like infiltrates of histiocytes and lymphocytes (4). The tumour cells are atypical, with a characteristic ceribreform nucleus and are restricted to the epidermis (epidermotrophism) colonizing the basal layer (10). In plaques, intra-epidermal aggregates of tumour cells form around Langerhans cells in a classic feature of MF known as Pautrier microabsesses (**Figure 1.2**)(2, 11). Though Pautrier microabsess formation does not occur in all cases, epidermotrophism is generally more pronounced (4).

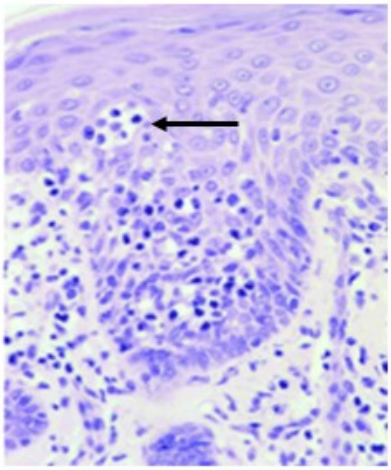


Figure 1.2 Pautrier micro-abscess.
Histological section showing MF T-cell infiltrates. The arrow indicates a Pautrier micro-abscess. Adapted from (Willemze et al 2005) WHO-EORTC classification for cutaneous lymphomas (**Willemze et al, 2005**).

A marked increase in the size and number of tumour cells occurs in tumour stage MF (12). In addition a range of cells with cerebriform nuclei may be present in the lesion along with blast cells with prominent nuclei (4). Cells with cerebriform nuclei are also known as Sézary cells and are a feature shared with Sézary syndrome (**Figure 1.3**). Epidermotrophism may be lost as aggregates of tumour cells become more diffuse (4).

Figure 1.3 A Sézary cell.
Electron micrograph showing the characteristic irregularly shaped 'cerebriform' nucleus. This is a hallmark feature of all stages of MF and SS, blast cells also become more numerous in advanced stages of mycosis fungoides.

The immunophenotype of MF tumour cells is somewhat variable but broadly consistent with a mature memory T-cell phenotype presenting CD2+, CD3+, CD5+ and either CD4+/CD8- or CD4-/CD8+ (4, 13). Neoplastic cells often lack the expression of T-cell markers such as CD7 and CD26 (14, 15). In MF plaques and tumours the normal CD4+/CD8+ T-cell ratio may become elevated from 2:1 to >4:1 (2). Large cell transformation with CD30+ or CD30- can occur and is associated with worse prognosis for life expectancy with median survival for those diagnosed with large cell transformation being 37 months verses 163 months for the untransformed group (12, 16).

A diagnosis of MF is dependent on a range of criteria and varies according to disease stage. Clinically, the presence of asymmetrical, persistent, irregular erythematous lesions on sun-protected areas must be observed along with failure to respond to conventional treatments (13). Histopathologically, superficial lymphoid infiltrates with epidermotrophism, without spongiosis and/or hyperchromatic, cerebriform nuclei and the accompanied detection of a T-cell clonal rearrangement (10, 17).

CTCLs express CD4 in the majority of cases (18).Other immunopathological diagnostic criteria include, CD2, 3, 5 expression in <50% T-cells, or CD7 expression in <10% T-cells, or epidermal discordance from expression of CD2, 3, 5 or 7 on dermal T-cells (3, 17). Reduction or loss of CD7 is often the first CD marker to lose expression in CTCL .(19). Diagnosis can be guided histologically by the presence of pautrier microabsesses which are strong criteria for diagnosis, however these occur infrequently (4). For cases of erythrodermic MF, multiple skin biopsies are usually required and additionally where progression has advanced to a leukemic state, biopsies of enlarged lymph nodes are often used as support (20, 21) with the detection of an identical T-cell clone in the circulating blood (10).

The genotype of MF is heterogeneous with complex karyotypes. The composite of gains and losses from studies investigating structural variation in MF indicate that large regions of the genome have been affected (22-26). (See table 2 for summary of common chromosomal

alterations in MF) These alterations overlap between studies and chromosomal 'hotspots' appear between cohorts but the foci and specific sites of structural alteration are somewhat variable. This indicates that there is no single underlying structural variation driving the disease and that genomic instability is a general characteristic. Chromosomal regions affected by CNVs often contain genes known to play a prominent role in cancer such as *TP53* and members of the JAK/STAT family on chromosome 17 (22, 24, 26), *CDKN2A* on chromosome 9 (22, 23, 25-27) and MYC on chromosome 8 (22, 23, 26) to name some of the better known candidates. However, numerous other genes, including many of unknown function can be found within MF structural genomic alterations.

Recent next generation sequencing (NGS) studies have confirmed and extended the earlier work and identified key mutations as well as CNVs (28-30) involved in MF. Heterogeneity is consistent and the most plausible explanation for how MF manifests is at the pathway level where alterations in key gene regulatory networks and signalling cascades drive the T-cell towards an MF phenotype. Considerable discussion will be devoted to signalling cascades and pathways later in the thesis.

1.3 Sézary Syndrome background

The French dermatologist Albert Sézary was the first to recognize SS in 1938 (31). This variant of CTCL has a significantly lower incidence than MF, occurring at a rate of just 0.3 cases per 100,000 (32). However, the age of onset is similar at 55-60 years and the higher risk of development in males is skewed similarly to MF at 2:1 (32). The prognosis of SS patients is considerably worse than MF with a median survival of just 3.13 years after diagnosis (3) and only 24% surviving more than 5 years (4). (See table 1 for summary)

A widespread, red, inflammation affects the entire skin in SS patients and is known as erythroderma (**Figure 1.4**) (4). This often co-occurs with exfoliation and lichenification and is intensely pruritic (4). Other common phenotypic features include; alopecia, onychodystrophy, palmoplantar hyperkeratosis, lymphadenopathy and edema (33).

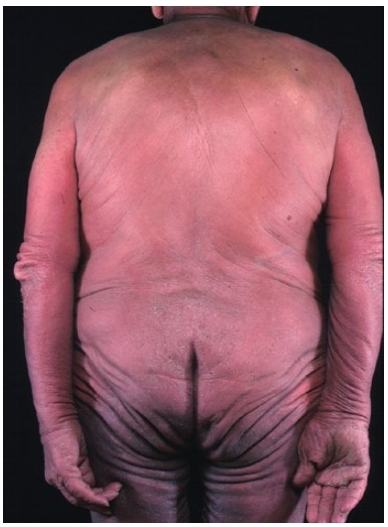


Figure 1.4 Erythroderma in a Sézary syndrome patient.
A hallmark feature of SS is that the entire skin is involved.

A key feature of SS is that, as a leukaemic tumour type, the neoplastic cells are present in the circulating blood (3, 4). However, the skin histopathology of SS shares considerable overlap with MF with both presenting atypical, hyperchromatic, ceribriform nuclei of T-cell origin (4). However, SS can show more monotonous cellular infiltrates and absent epidermotrophism (4). Biopsies from SS patients can be non-specific in up to 1/3 of cases (34, 35). Lymph node biopsies from affected nodes show effaced architecture with high numbers of infiltrated Sézary cells (36). Additionally, some patients present with sparse, interstitial infiltrates in the bone marrow (37).

Much like MF, the immunophenotype of SS tumour cells can be inconsistent but often show a CD3+ and CD4+ antigens but are CD8- (4). Furthermore the neoplastic cells usually do not express CD7 and CD26 (38). Expression of many common markers can vary although subsets of SS cells that express PLS3 are consistently associated with loss of CD26 expression (39).

For a diagnosis of SS, patients must present with erythroderma, generalized lymphadenopathy and an identical neoplastic T-cell clone in the peripheral blood, skin and lymph nodes (33). More recently additional criteria have been added to guide diagnosis including; a Sézary cell

count in the peripheral blood of >1000 cells per mm^3; a T-cell population with a CD4/CD8 of >10; loss of normal T-cell antigen expression including CD2, CD3 and CD5 (38).

Much like MF, complex and heterogeneous chromosome alterations are a hallmark genetic feature of SS (26, 40). (See table 2 for summary of common chromosomal alterations in SS) Some karyotypes are more common than others such as 8q gain (25, 41-43), 9p21 losses (27, 44), 10q losses (25, 26, 40-42, 45-47), 17p losses (25, 26, 41, 42, 47, 48) and 17q gains (25, 26, 40-42, 48) . Studies have shown alterations are common in these key chromosomal regions but a predictive pattern that is clearly driving disease is lacking. Recent NGS studies have shown that both CNVs and mutations affect genes that are likely drivers of the disease (16, 29, 49-51). However, these recent high resolution studies confirm that, much like MF, considerable heterogeneity occurs between different SS tumours, implying that there are multiple possible genetic routes that converge on an SS phenotype. It seems plausible that at the level of gene regulatory networks and signalling pathways, the diverse genotype seen in the disease group consolidates and manifests itself as perturbations in key pathways controlling the phenotype of the T-cell. Considerable discussion will be devoted to gene regulatory networks and signalling cascades later in the thesis.

1.4 Similarities and differences between MF and SS

Both MF and SS are both subgroups of CTCL (4) and between them represent 65% of CTCL cases (4). Both tumour types are of mature T-cell origin and express CD4 and CD5RO but have lost expression of other markers associated with mature T-cells (4, 52). Currently SS is regarded as a distinct erythrodermic disease entity to MF (34) and is leukaemic by definition (38). Later stages of MF can become leukaemic as well and meet clinical criteria to meet a diagnosis of SS but are classified separately (4, 52). A prior diagnosis of MF leading to leukaemic involvement is known as ‘secondary SS’ or ‘SS preceded by MF’ (53). There is some evidence that different cell surface markers can be expressed between MF and SS which appear to be consistent with different tissues of origin. One study demonstrated MF clones

lacked expression of CCR7, CD62L and the differentiation marker CD27 but strongly express CCR4 and CLA which is suggestive of cells derived from skin resident effector memory T-cells. In contrast clonal cells from SS patients universally expresses CCR7, CD62L and CD27 suggesting they are more likely to be derived from central memory T-cells (54). (Summarised in table 1)

It has been established that at the genomic level both MF and SS tumour cells feature complex chromosomal rearrangements (25, 28, 31, 50). Similar rearrangements are shared by both disease groups (26) although some rearrangements appear to be more common in MF than SS and vice versa. The heterogeneity at the genetic level of both diseases is ironically the most consistent feature, however genes rarely function as single entities and gene regulatory networks are commonly perturbed in both disease groups.

Table 1. Summary of features of MF and SS

Type of feature	Disease feature	Disease	Co-Occurences	Refs
Incidence	0.36-0.9 / 100,000	MF	NA	Chuang et al 1990, Weinstock et al 1999
	0.3 / 100,000	SS	NA	Cuneo et al 2005
Physical	Patches/plaques/tumours on skin	MF	NA	Kazakov et al 2004
	Erythroderma	Leukemic MF, SS	NA	Willemze et al 2005
Histological	Cerebreform nuclei	MF, SS	NA	Smoller et al 2003
	Pautrier microabses	MF	NA	Nickoloff et al 1988
	Epidermotrophism	MF	NA	Smoller et al 2003
	Lymphadenopathy	Leukemic MF, SS	NA	Russell-Jones et al 2005, Vonderheid et al 2006
Molecular	CD2+, CD3+, CD4+, CD5+, CD5RO+	MF, SS	NA	Willemze et al 2005, Burg et al 2005
	CD7-, CD26-	MF, SS	CD26-/PLS3+	Wood et al 1990, Jones et al 2012
	CCR7-, CD62L-, CD27- , CCR4+,CLA+	MF	Skin resident phenotype	Campbell et al 2010
	CCR7+, CD62L+, CD27+	SS	Central memory phenotype	Campbell et al 2010
	Identical T-Cell clone	MF, SS	NA	Olsen et al 2007
	Identical T-Cell clone in circulating blood	Leukemic MF, SS	NA	Vonderheid et al 2002

Table 2. Common chromosomal changes associated with MF and SS

Chromosome	Implicated genes	Maximum incidence	Disease	Co-Occurences	Refs
1q gain	MCL1, CLK2	41%	MF	Shorter survival	van Doorn et al 2009, Laharanne et al 2009, Salgado et al 2009
7q gain	AP1S1	59%	MF	NA	van Doorn et al 2009, Laharanne et al 2009, Salgado et al 2009
8q gain	MYC	75%	MF, SS	Shorter survival	Salgado et al 2010, van Doorn et al 2009, Mao et al 2002, Laharanne et al 2010, Vermeer et al 2008, Fischer et al 2004, Karenko et al 2007
9p21 loss	CDKN2A, CDKN2B, MTAP	41%	MF, SS	Shorter survival	Salgado et al 2010, Van Doorn et al 2009, Laharanne et al 2010, Scarrisbrick et al 2002
10p	ZEB1	40%	SS	NA	Vermeer et al 2008, Laharanne et al 2009, Caprini et al 2009
10q loss	PTEN, FAS, MXI1	68%	MF, SS,	NA	Wain et al 2005, Laharanne et al 2010, Mao et al 2002, Karenko et al 1999, Vermeer et al 2008, Fischer et al 2004, Barba et al 2008, Scarrisbrick et al 2000, Limon et al 1995
13q loss	RB1	25%	SS, MF	Shorter survival	Fischer et al 2004, Salgado et al 2009
17p loss	TP53, MNT	45%	MF, SS	NA	Salgado et al 2010, Carbonne et al 2008, Laharanne et al 2010, Mao et al 2002, Vermeer et al 2008, Fischer et al 2004, Limon et al 1995, Caprini et al 2009
17q gain	STAT genes	75%	MF, SS	NA	Prochazkova et al 2007, Laharanne et al 2010, Mao et al 2002, Karenko et al 1999, Vermeer et al 2008, Caprini et al 2009

1.5 Treatment of MF and SS

In MF, if disease is confined only to the skin then the most commonly used therapies are photochemo-therapy such as PUVA (psoralen and UVA), topical application of carmustine (BCNU) or nitrogen mustard, or radiotherapies such as total skin electron bean irradiation (55, 56). For patch stage MF treatment can include bexarotene gel or topical steroids, biological treatment including interferon alpha or IL12 or receptor targeted cytotoxic fusion proteins such as DAB_{389} IL2 (55, 57, 58). Multi-agent chemotherapy is generally used for later stage disease but not in early stages (59).

In SS treatment is of limited use but some patients respond to treatment such as extracorporeal photophoresis (ECP), often in combination with interferon alpha (60, 61). Other treatments used include PUVA, low dose chlorambucil, prednisone, methotrexate, bexarotene, alemtuzumab (55, 58, 62) but none appears to be of great benefit.

Other treatments for CTCL include HDAC inhibitors such as Romidepsin (63) and Vorinostat (64) which attempt to alter the epigenetic landscape and hence gene expression patterns. The most effective treatment option for late stage CTCLs to date remains to be allogenic haematopoietic stem cell transplantation (65). Whilst often effective, this approach tends to be reserved for younger patients due to risk factors involved.

As our understanding of the biological mechanisms underlying CTCL grows, there has been an increasing shift toward the use of targeted biological therapies; the development of HDAC inhibitors is an example of this. This approach is currently in its infancy in all malignancies but holds much promise. However, in order to maximise success, an increased understanding of malignancies at the molecular level must be sought including an understanding of genes, their normal functions and expression patterns and how CNVs and mutations can drive cancer.

1.6 Early genomic studies of MF and SS

Much of the constellation of genomic structural alterations in CTCL was observed prior to the more recent NGS studies. Lower resolution array based studies, including studies from our research group began to establish regions of gain and loss in a number of key regions such as chromosomes 1p, 4, 6q, 7, 8q, 9p21, 10, 13q, 17 and 19 (24, 26, 41, 42, 44, 66). Further studies using array based methods confirmed existing data (25, 48) demonstrating that losses and gains in these regions were reproducible between datasets and therefore fairly common. Some of the early work was able to correlate specific chromosomal alterations with prognosis such as 8q gain, 6q loss and 13q loss with shorter survival (42) and 9p21 loss and 1q21-1q22 gain with worse prognosis in MF (23). A more recent study was able to further refine the genomic regions that correlate with poor survival to 8q24.21, 9p21.3 and 10q26qter (22).

Though array based studies highlighted many regions of structural alterations and CNVs, what was predominantly lacking was resolution to identify individual genes as large regions contain far too many to all be considered a definitive list driving tumourigenesis, the large scale changes often observed on chromosome 17 are a good example of where it can be challenging to identify specific genes. Some areas are an exception to this and minimal affected regions were identified such as *cMYC* on chromosome 8q24, loss of MNT on 17p13 and loss of *MXI1* on 10q25 (25, 41). Another important area that seems to be regularly compromised in CTCL by a minimally affected region is 9p21 which was first reported in CTCL in 2002 (44). This region has been reduced to a minimal region of 9p21.3 and includes the genes *MTAP*, *CDKN2A* and *CDKN2B* and has come up in several major CTCL genomic studies since it was first connected to the disease (22, 23). The small number of genes in these regions, which are also linked to prognosis, is intriguing and highlights the importance of further investigation. The 9p21 region will be a key focus of investigation and discussion in this thesis.

1.7 The driver mutation model of cancer

The development of CTCLs, like all cancers seem to involve the accumulation of multiple rate limiting steps (67), progression through each step is similar in principle to Darwinian evolution. In this model surpassing each step is supported by the acquisition of genetic events which confer a selective growth advantage upon its host cell over neighbouring cells in an increasingly heterogeneous population. The genetic events range from simple point mutations to whole chromosome duplications (68). This process, like Darwinian evolution is more often passive or detrimental to the cell but occasionally an event will increase the viability of the cell and in this instance is known as a driver mutation (**Figure 1.5**)(68).

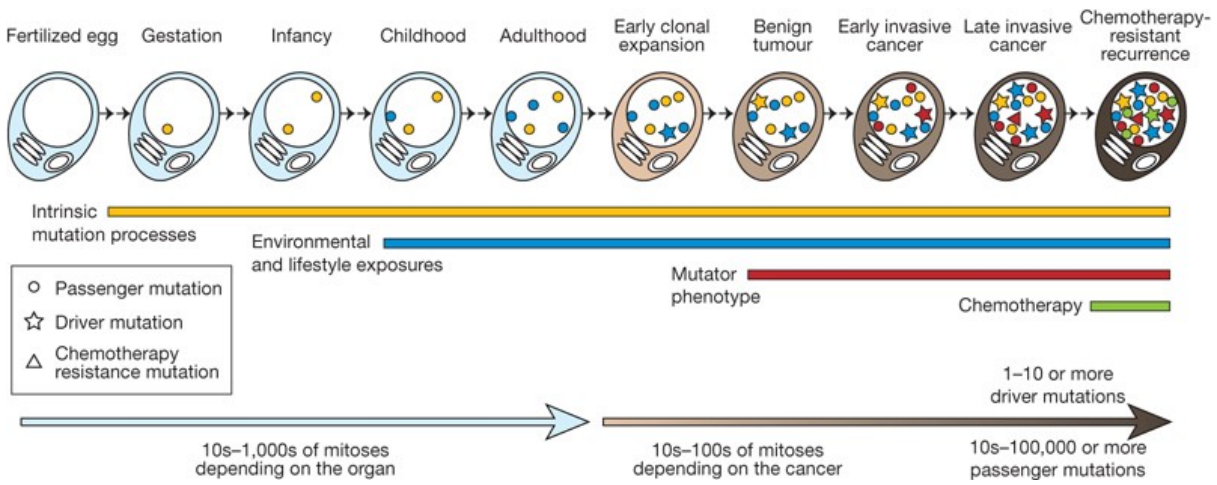


Figure 1.5 The lineage of mitotic cell division leading to neoplastic transformation.
Somatic mutations are acquired throughout all cell lineages. Some will confer a selective growth advantage and are known as driver mutations. Passenger mutations do not contribute to the process. Often during therapeutic intervention Darwinian processes shift the selection pressure towards mutations that protect the neoplasia from treatment effects. Illustration from **Stratton et al (2009)**.

Events that contribute to the enhancement of several cellular processes known as the hallmarks of cancer (**Figure 1.6A**)(68, 69) are termed driver mutations. These hallmark processes include; Sustaining proliferative signalling, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, resisting cell death and in the case of solid tumours inducing angiogenesis (68, 69). Several additional emerging features are also

likely contributors to tumourigenesis including; the ability to avoid immune-response mediated destruction, deregulated cellular energy metabolism, the presence of tumour promoting inflammation and general genomic instability (**Figure 1.6B**) (69).

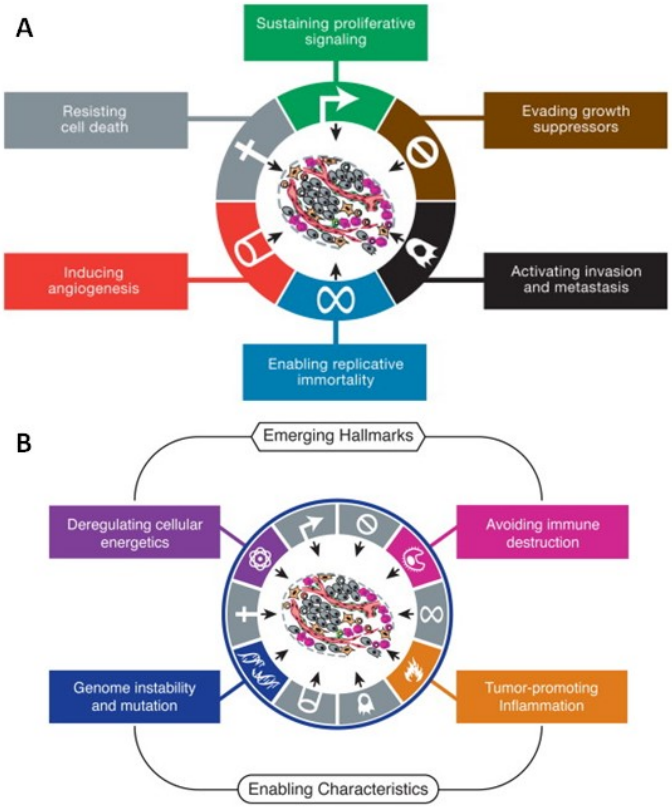


Figure 1.6 Hallmarks of Cancer.
(A) Original hallmarks of cancer **Hanahan and Weinberg 2000**. (B) Additional hallmarks of cancer highlighted in the revised hallmarks of cancer **Hanahan and Weinberg 2011**.

Identifying the genomic events that contribute to these hallmarks and features is a priority in all cancer related research, including CTCL. Identifying the genes will highlight targets for drug development and have important implications for the development of novel therapeutics and translational medicine. The challenge remains in separating driver events from the myriad of passenger events that co-occur in CTCL and other cancers.

1.8 Molecular processes and signalling networks affected in MF and SS

Driver events by definition, will contribute to the hallmarks of cancer, but different driver events can contribute to the same hallmark in ways which can be independent of each other as well as cumulative. Many genes contribute to each molecular process that takes place within a cell. Processes such as signalling cascades and gene regulatory networks are examples of multi-gene processes. As an example, driver events could occur within the same network or signalling cascade but in different genes and still produce very similar phenotypes. In this section, well known tumourigenic processes will be reviewed and examples of events contributing to CTCL in each molecular process will be highlighted. In many of these processes several genes have been discovered to be defective, highlighting the genetic heterogeneity of CTCL. Furthermore, it is important to consider that many of these processes exhibit considerable crosstalk and it remains an arduous challenge to fully elucidate their interactions.

Information from several large sequencing studies has been included in this section to show alterations in many of these pathways and it would be desirable to analyse co-occurrences of genes across CNV and SNV events. However, these studies do not report co-occurrences and this analysis would require a large meta-study which is beyond the scope of the work contributing to this thesis. For this reason, events have been presented as individual events though I would like the reader to consider that co-occurrences of both CNVs and SNVs are a likely feature of the pathways presented in this section.

1.8.1 NF-κB

The nuclear factor kappa-light-chain-enhancer of activated B-cells (NF-κB) pathway is a well-studied signalling network involved in inflammatory and innate immune responses (**Figure 1.7**) (70). Downstream targets of the NF-κB pathway include regulation of a number of cellular processes that function during tumourigenesis such as; cell survival, proliferation, differentiation, angiogenesis and metastasis (71). It therefore comes as no surprise that the

NF-κB pathway is dysregulated in numerous tumour types including leukaemic neoplasms (70) such as CTCL where the pathway is constitutively active (41, 71-73).

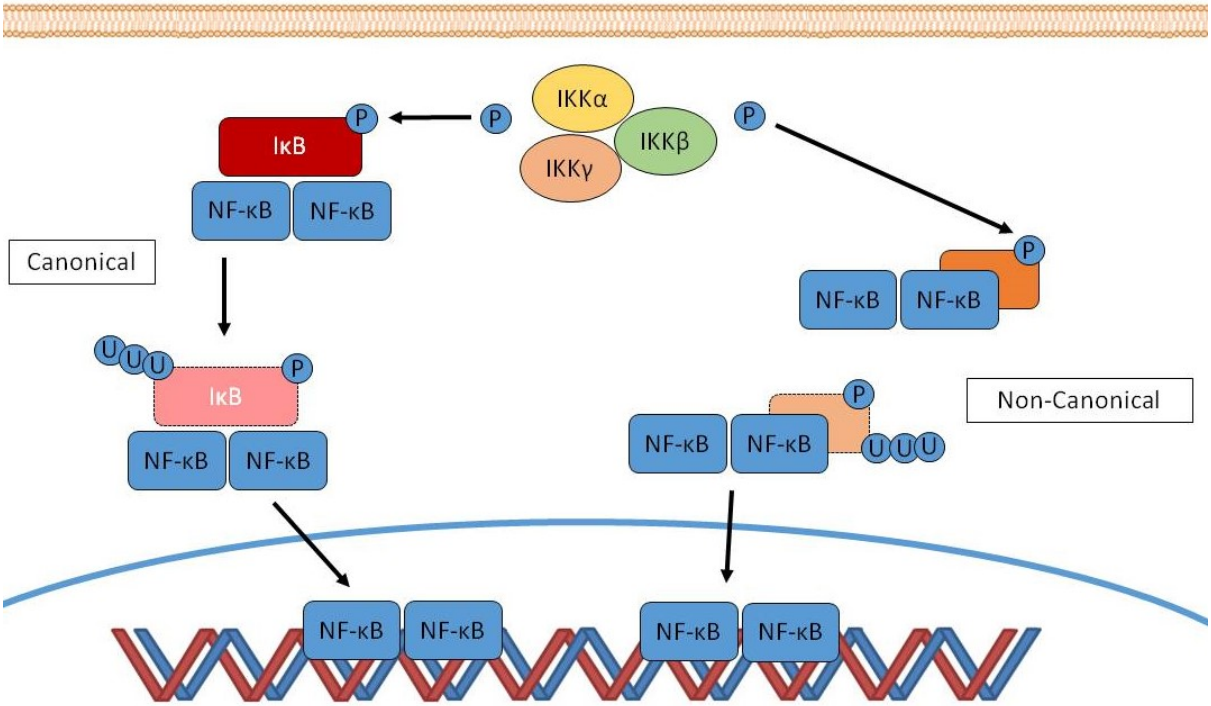


Figure 1.7 NF-κB pathway.
In the canonical NF-κB pathway IκB proteins bind NF-κB family members in the cytoplasm until IκB is phosphorylated by the IKK complex. IκB is then targeted for ubiquitin mediated degradation allowing NF-κB homo and heterodimers to enter the nucleus where they act as transcription factors. Some members of the NF-κB family such as p100 are not bound by IκB and require cleavage, after which the NF-κB dimmers can translocate to the nucleus and act as transcription factors. This is known as the non-canonical NF-κB pathway.

Genes with mutations MF and SS	Reference
CD28, CARD11, NF-κB2, RELB, MALT1, PRKCQ	Ungewickell et al 2015, da Silva Almeida et al 2015, Choi et al 2015, Wang et al 2015

Table 3. Genes reported as mutated in NF-KB in CTCL

The NF-κB pathway (**Figure 1.7**) is a complicated gene regulatory network, the core components are the NF-κB proteins which are bound by IκB proteins in the cytoplasm and sequestered preventing activation of target genes (70, 74). To activate the pathway, the IκB

component is phosphorylated by the IKK complex which consists of IKKα, IKKβ and IKKγ (75, 76). The NF-κB family members consist of 5 known members (RelA/p65, RelB, c-Rel, p50/p105/NF-κB1 and p52/p100/NF-κB2) (74). The IκB family consists of 7 known members (IκBα, IκBβ, IκBγ, IκBζ, Bcl3, p50/p105/NF-κB1 and p52/p100/NF-κB2)(74). Both p50 and p52 fit into both families. Upon activation, the NF-κB family members form up to 15 homo and hetero dimers in a cell type and context dependent manner (74). In addition, only the RelA, RelB and c-Rel have transcriptional activation potential via a domain called the TAD (carboxy-terminal transactivation domain) (74), RelB requires further addition of an amino-terminal leucine zipper to become fully active (77). Furthermore, the NF-κB family members are regulated by an array of post-translational modifications which allow the pathway to regulate different groups of target genes (78).

The NF-κB pathway has previously been shown in a study of 30 CTCL patient samples to be constitutively active in 100% of cases and cause resistance to apoptosis(79). In CTCL the NF-κB pathway directly regulates 3 sets of genes (70); survival genes (cIAP1, cIAP2, Bcl2), pro-inflammatory genes (IL-1β, IL-8, IL-17, TNFα, CCL2, CXCL5) and anti-inflamatory genes (IL-10, TGFβ) (70). The NF-κB mediated activation of inflammatory and survival genes contributes to inhibition of apoptosis and enhanced proliferation (70). This would lead to an increase in the T-cell pool. Activation of anti-inflamatory genes however, seems somewhat contradictory, especially as IL-10 and TGFβ are thought to repress the pro-inflamatory genes (70). There is speculation that this contributes to the immunosuppressive nature of CTCL. It is possible that different targets of the pathway are activated/repressed in different stages of the disease which may be at least partially facilitated by changes in the microenvironment, however it seems more plausible that specific subtleties of this complex network are yet to be fully established. Interestingly, mutations in *CARD11*, a positive mediator of NF-κB signalling have been identified in SS in 6% (50), 8% (51) and 15% (16) of cases. Focal copy number gains in the CARD11 activator PRKCQ have also been reported in 32% of SS tumours which would likely have a positive effect on the pathway if they are transcribed (49). Mutations in other

regulators of the pathway have also been detected such as CD28 in 10% of SS cases (49) and a CD28-CTLA4 fusion caused by chromosome alteration in 2 MF cases (29), the same study also reported alterations in 16 genes involved in NF-κB signalling, with 8 genes affecting just 1 tumour each across MF and SS cases which adds weight to the argument that alterations in pathways are more relevant to CTCL than alterations in individual genes. Together these findings highlight a selection pressure for mutagenesis of components of the NF-κB pathway.

1.8.2 JAK-STATs

Members of the JAK-STAT family play roles in a number of cellular processes and have well studied roles in inflammation and cancer (80, 81), including haematological malignancies (82). The Janus tyrosine kinase (JAK) family consists of 4 members (JAK1, JAK2, JAK3, Tyk2) (80) which form hetero and homodimers in the cell membrane (81). JAK family members transduce signals through several cytokine receptor types; type I receptors include the granulocyte colony stimulating factor (G-CSF) and the erythropoietin receptor (81), type II receptors are subdivided into type IIa and type IIb receptors such as granulocyte-macrophage colony stimulating factor (type IIa) and the IL2/6 and leukaemia-inhibitory factor (type IIb) (81). Upon activation of the receptor, JAK family members become active tyrosine kinases capable of phosphorylating members of the STAT family (83).

The Signal transducers and activators of transcription (STAT) family is made up of 7 members (*STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, STAT6*)(80). Upon phosphorylation, the STATS form dimers which translocate to the nucleus and act as transcription factors, this is known as the canonical JAK-STAT pathway (81). The JAKs and STATs can also act through a non-canonical method (84) where each has direct effects within the nucleus on the regulation of heterochromatin (85)(**Figure 1.8**).

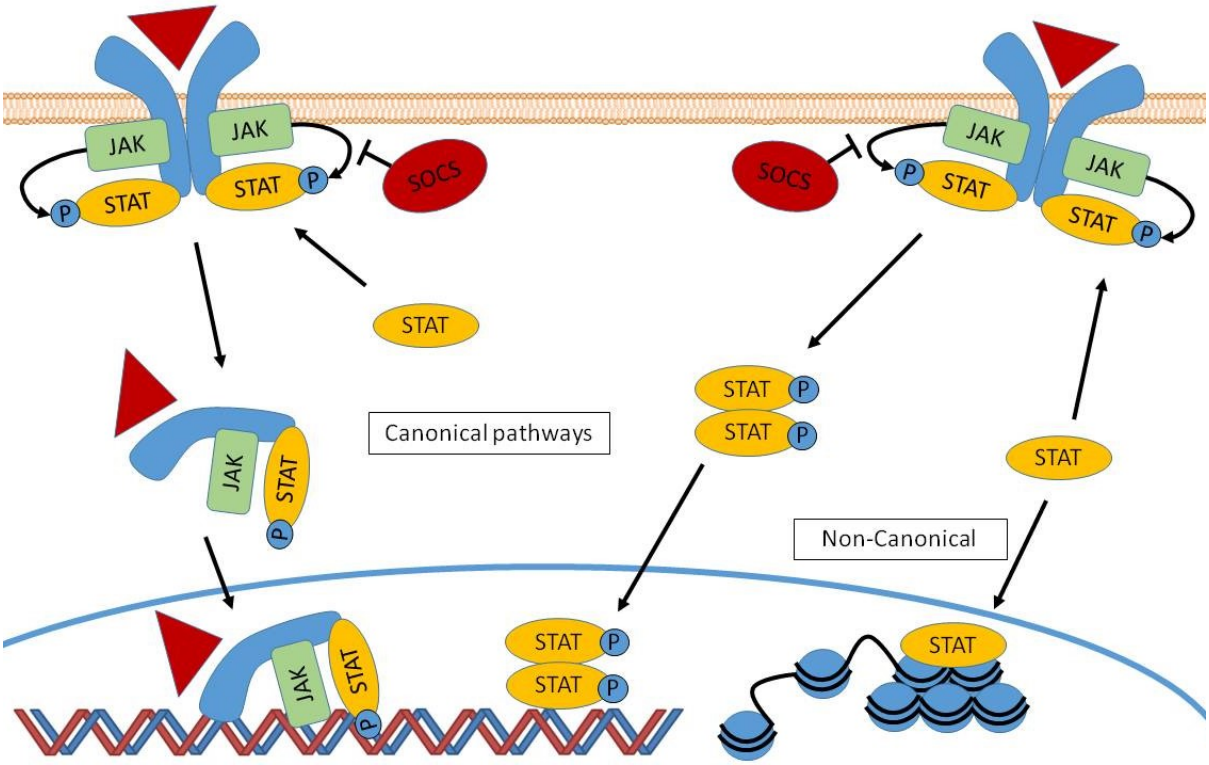


Figure 1.8 JAK-STAT pathways.
Cytokine receptors including type IIb, transduce signals though the JAK-STAT pathway. Upon ligand binding to the receptor, JAKs become activated and recruit and phosphorylate STATs. The STATs form homo and heterodimmers before transclocating to the nucleus to act as transcription factors or the receptors can themselves translocate. Signal transduction can be inhibited by the SOCS family. STATs can also act non-cannonically affecting or maintaining the state of hetero and euchromatin.

Genes with mutations MF and SS	Reference
JAK1, JAK3, STAT3, STAT5B, SOCS5	Kiel et al 2015, Choi et al 2015, Vaque et al 2015

Table 4. Genes with reported mutations in the JAK/STAT pathway in CTCL.

Perturbations in JAKs and STATs are well documented in numerous blood born neoplasms (82) as well as in MF and SS at the genetic level with recent studies identifying mutations in up to 11% of SS (50) and up to 20% of MF cases with low variant allele frequency (VAF) (28, 30, 50), single JAK1 and JAK3 variants have also been identified in MF and SS tumours in another study (30). Gains of STAT5B were also reported in 25/40 SS tumours and a mutation in 1 further patient in a recent NGS study (49). Some variants were found to make the pathways constitutively active (86-88). Differences in their activity state have also been reported in the absence of genetic variants (89). In SS, a study published by our lab has also indicated that

overactive STAT3 is driven by constitutive aberrant activation of JAK family members rather that by loss of inhibition (90). Earlier work from our group also suggests that alternate transcripts of *STAT5* are preferentially driving gene expression in SS compared with healthy controls (91). Overall perturbations in JAK-STAT signalling seem to affect a high proportion of CTCL tumour cells and likely constitute key driver events of the disease.

1.8.3 PI3K-PKB/AKT

A particularly well studied gene regulatory network is the PI3K-PKB/Akt pathway which is known to have a key role in regulating proliferation, survival and growth, reviewed in (92, 93). In this pathway growth and proliferation factors bind receptor tyrosine kinases (RTKs) such as; PDGFRs and FGFRs or G-protein coupled receptors (94). These receptors in turn activate class IA or class IB PI3Ks (respectively) in coordination with groups of co-factor molecules such insulin receptor substrates (IRS) and/or Ras (95). Active class I PI3Ks can then catalyse the conversion of phosphatidylinositol (3,4)-bisphosphate (PIP_2) lipids into phosphatidylinositol (3,4,5)-trisphosphate (PIP_3) (93) (**Figure 1.9**). There are additional class II and class III PI3Ks, but to date these have not been implicated in cancer (95).

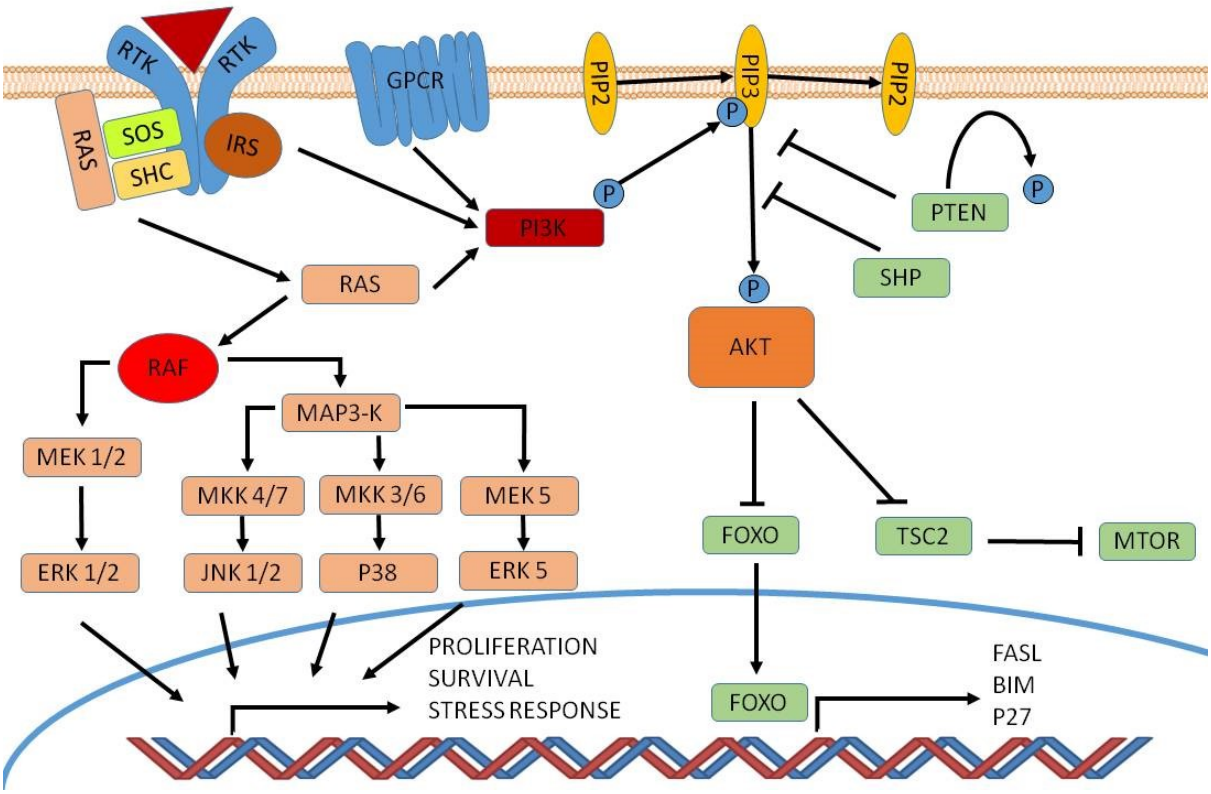


Figure 1.9 Signal transduction through RTKs and GPCRs.
Activated receptors can trigger transduction through the class I PI3K which promotes the conversion of PIP2 to PIP3 activating the AKT signalling hub. Signals can also be transduced through the RAS/RAF pathway, activating one or more of the MAP-kinase pathways. This network activates numerous oncogenic processes when active and is constitutively activated in many neoplasms.

Genes with mutations MF and SS	Reference
PTEN, RHOA, BRAF, KRAS, NRAS, PI3K, PI3KR	Scarrisbrick et al 2000, Caprini et al 2009, Choi et al 2015, Wang et al 2015, Kiessling et al 2011, Ungewickell et al 2015

Table 5. Genes with mutations in RTK and GPCR pathways in CTCL.

The availability of PIP_3 is highly controlled by its reconversion to PIP_2 by the phosphatase and tensin homolog (PTEN), which acts as the primary negative regulator of PI3K (92). PIP_2 and PIP_3 are both bound to the cells outer membrane but PIP_3 is the active form which transduces signals by binding via a pleckstrin homology (PH) domain (92). A key downstream target of PIP_3 is PKB/AKT which is a central signalling node of numerous core cellular processes (95). Direct targets of PKB/AKT include FOXO and TSC2 (95). PKB/AKT retains FOXO in the cytoplasm which can increase survival and proliferation by preventing FOXO from regulating apoptotic machinery such as FASL and BIM and cell cycle regulators such as p27Kip1 (92). The effects of

PKB/AKT on TSC2 lead to activation of the mTOR complex which acts as a growth promoter by increasing protein synthesis via 4E-BP1 (96).

The *PTEN* gene is located on the long arm of chromosome 10, a region shown by our group and others to be deleted in up to 23% MF tumours (46) and up to 68% of SS tumours (48) (46, 97). PTEN is a known tumour suppressor gene and its loss leads to an increase in activity through the PI3K-PKB/AKT pathway (95). Furthermore, PTEN has been shown to be down-regulated in most SS tumours by a group of regulatory RNAs (miR-106b, miR-486, miR-21) which also leads to over-activation of AKT (97). The PI3K and PI3KR genes themselves have also been altered in 5-10% of MF and SS tumours (29). Miss-regulation of gene expression in the PI3K/AKT pathway has also been reported in both MF and SS in a transcriptome sequencing study (98).

1.8.4 RAS/RAF

Proliferative signals can be transduced by RTK receptors through a cascade which starts at the receptor by the recruitment of a multi-protein complex (99). The complex includes Src homology 2 domain containing protein (SHC) and the RAS-GTPase Son of sevenless (SOS), which together activate membrane bound RAS family members (100). The cascade progresses with the recruitment of RAF family members to the membrane, which themselves becomes activated and target one or more of the 4 major MAP-kinases (101) (**Figure 1.9**).

Several components of this signalling cascade are members of multi-gene families which show some degree of redundancy (102) as well as cell and context dependent functionality (99), i.e the downstream targets of members of the pathway may vary depending on the cell type or the 'state' it is currently in (stressed or dividing etc). The RTKS which have been discussed in the previous section are the first variable component, RAS-GTPases are another (94, 102). The RAS family has 3 members (K-RAS, N-RAS, H-RAS)(99), as does RAF (B-RAF, RAF-1/C-RAF, A-RAF) (103, 104). The 4 core MAP-kinase pathways include the MEK1/2 pathway which acts through ERK1/2, and MAP3K which acts through the remaining 3 pathways; JNK1/2 via

MKK4/7, p38 via MKK3/6 and ERK5 via MEK5 (101). The MAPK pathways have numerous upstream and downstream targets (over 60 are known for ERK1/2 alone) and their activity is under tight intrinsic and extrinsic control (99), i.e. internal feedback loops maintain fine tuning of the intracellular environment via the transcriptome and the proteome as well as other processes, these are further tuned and fed into by extracellular cues such as nutrient levels or signalling events. Broadly the pathway regulates (and is to some extent regulated by) processes involved in proliferation, cell transformation, stress responses and cell survival and apoptosis (105).

Mutations in *KRAS* and *NRAS* have been reported in CTCL, including SS and MF although at frequencies of less than 1% (106). However, targeting this pathway with antagonists in CTCL cell lines induced apoptosis which has implications for translational medicine (106). A recent study also reported activating mutations in a single patient in *BRAF* in SS (49) and another reported 2 mutations in RHOA in SS (16, 49). Further down the cascade, the p38 MAP kinase pathway has also shown promise as a therapeutic target in SS as its inhibition lead to cell death in CTCL cell lines and patient samples (107).

1.8.5 Phospholipase C gamma pathways

Phospholipase-C gamma 1 (PLCγ1) is involved in several signalling pathways activated by receptor tyrosine kinases including those that signal through PI3K (108) and several of the ras-activated MAP-kinase pathways (101). Upstream activators include EGFRs, FGFRs, VEGFRs, IGF-1R and PDGFRs (108). Activation of PLCγ1 activates second messengers DAG and IP3 which cooperate in the activation of protein kinase-C (PKC) and intracellular calcium mobilisation amongst other targets (108) (**Figure 1.10**). Mutations in *PLCγ1* have been identified in numerous cancers and are thought to primarily be activating mutations including several in CTCL in a host of recent publications (16) (30, 49, 50). The incidence of PLCγ1 mutations has been reported at 9% (50), 10% (49), 13.5% (16) and 19% in recent studies looking SS tumours (30) and at 27% in MF tumours (30). Mutations in *PLCγ1* are of particular interest because they

likely effect transduction through several known tumourigenic signalling networks previously discussed.

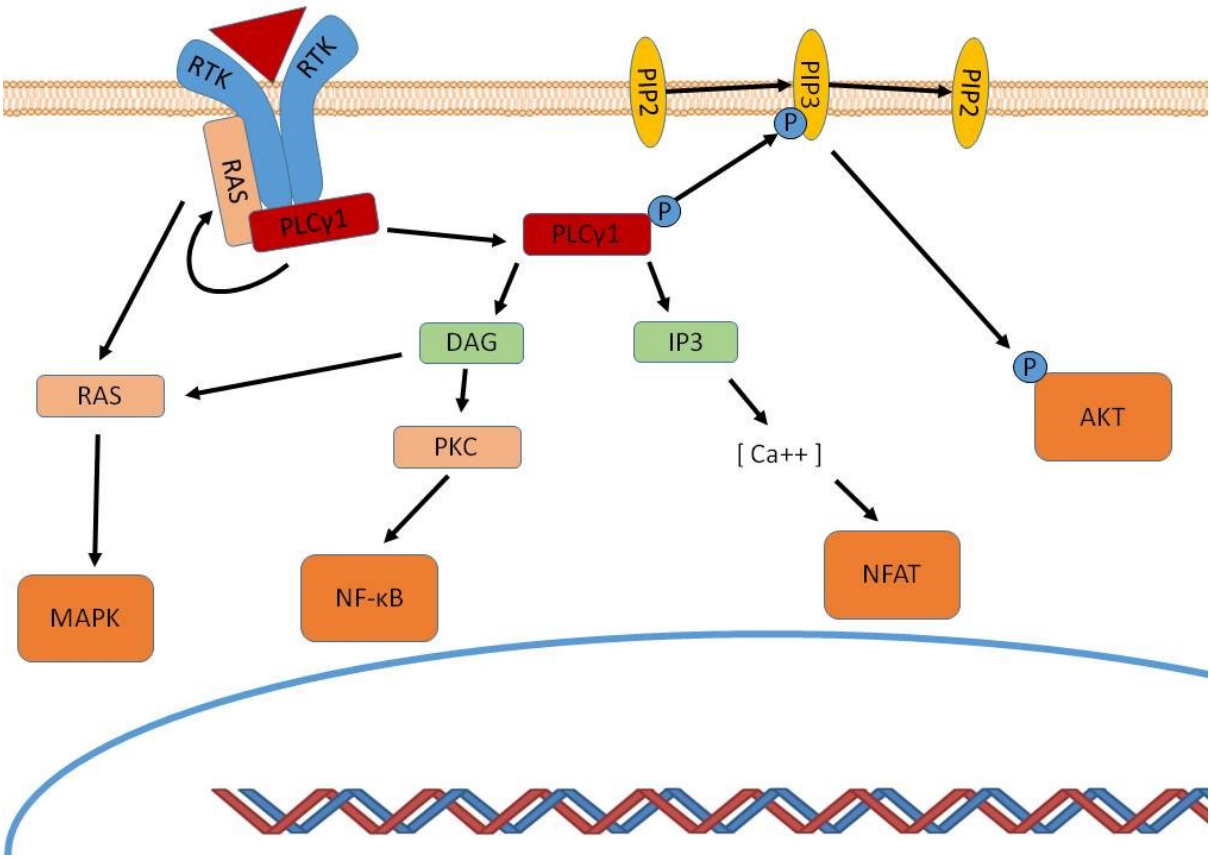


Figure 1.10 PLCγ1 signal transduction.
Signalling through this pathway is transduced by receptor tyrosine kinases and converges on several networks known to be perturbed in CTCL. These include MAP-kinase pathways, NF-κB, NFAT and AKT. Mutations have also been reported in PLCγ1 itself in CTCL and other lymphoid neoplasms.

1.8.6 Programmed cell death

Programmed cell death (PCD) is essential for normal developmental processes and homeostasis as it plays a key role in maintaining the equilibrium between cell survival and death (109). The breakdown of this process is a well-known contributing factor to the initiation and progression of all known cancers (69). The term PCD was originally used synonymously with apoptosis but has more recently been understood to encompass at least 3 varieties;

apoptosis, autophagy and programmed necrosis (110, 111). Each of these processes is morphologically distinct (112, 113).

1.8.6.1 Apoptosis

Apoptosis was first described microscopically in 1972 by Kerr et al (114), cells show shrinkage combined with nuclear compartmentalization and fragmentation and detachment from the surrounding cells and extracellular matrix (ECM). A process called blebbing orchestrates the controlled breakup of the cell into membrane bound compartments (115) and ultimately phagocytosis by surrounding cells (109).

Apoptosis can be induced by 2 main methods which both converge on activation of caspase 3, the penultimate step in the execution pathway (109). The extrinsic method requires membrane bound FAS receptors (including TNFR-α, FAS and TRAIL)(116) to bind an extracellular ligand (TNF-α, FAS-L), bound receptors recruit further members including FADD, TRADD and RIP (117, 118), then pro-caspase-8 to form the death inducing signalling complex (DISC). This complex then activates caspase 3 to trigger the execution pathway (114). Inducing apoptosis via the FAS pathway is the primary method of maintaining homeostasis in the immune system and is also used by cytotoxic T-cells to induce apoptosis in virally infected cells and other target cells (119) (**Figure 1.11**).

The other major apoptotic pathway is known as the intrinsic pathway (**Figure 1.11**) and is primarily induced by pathological levels of stress from a variety of sources including; hypoxia, toxins, high levels of reactive oxygen species (ROS) and irreparable DNA damage (109, 111). Under these conditions, pro-apoptotic members of the BCL2 gene family (such as BAX and BAK) become active (120, 121) and form pores in the mitochondrial outer membrane facilitating the release of cytochrome-c, amongst other apoptotic factors, into the cytosol (110). Cytosolic cytochrome-c combines with APAF1 and pro-caspase-9 to generate a large protein complex called the apoptosome which induces the execution pathway via caspase-3 (122).

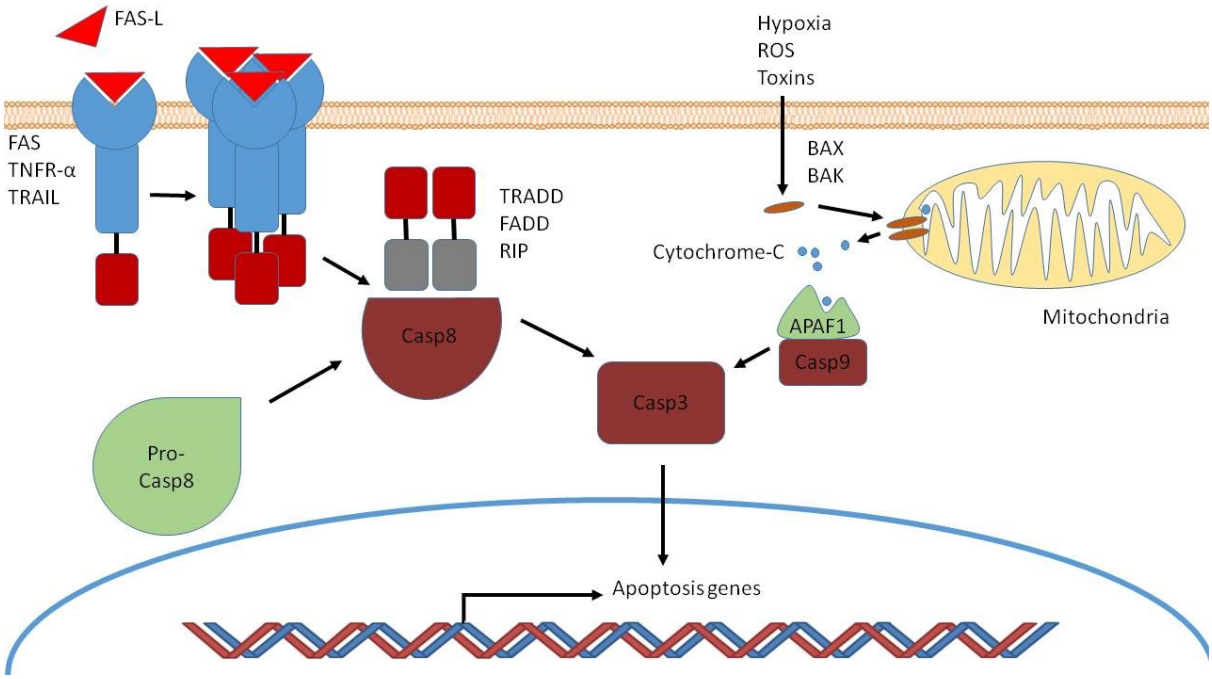


Figure 1.11 Apoptosis.
Extrinsic induction of apoptosis is mediated by signalling through FAS, TNFR-α and TRAIL. Further proteins TRADD, FADD and RIP are recruited through the intercellular domains along with pro-caspase-8 to produce the death inducing signalling complex (DISC). DISC signals through caspase-3 to switch on apoptotic genes. The intrinsic apoptosis pathway is activated by ROS, toxicity, irreversible DNA damage or other mechanisms which compromise the cell. BLC2 family members such as BAX and BAK cause the release of cytochrome-C from the inner mitochondrial membrane to the cytoplasm which are detected by APAF1. Pro-caspase-9 is then cleaved and binds APAF1 before activation of caspase-3 and induction of apoptosis.

T-cell mediated apoptosis can also be induced via the perforin/granzyme pathway. This is a special case pathway whereby perforin secreted by the T-cells forms pores in the target cell through which the T-cell releases cytoplasmic granules containing granzymes A and B which ultimately cause DNA cleavage, activation of caspase-10 and can even cleave BID to activate the intrinsic mitochondrial apoptosis pathway (109).

In CTCL there is evidence to suggest that a high proportion of tumours lack the ability to initiate apoptosis through the FAS pathway. Specifically, the presence of the FAS receptor has been shown to be down-regulated (123-126), in many cases as a result of deletion of all or part of chromosome 10q up to 68% in SS tumour cells (48) (26, 31, 40, 41, 127). Furthermore, point

mutations have been reported in a number of cases (128-130) alongside reduced sensitivity to FAS-L (128, 129). Downregulation of the FAS gene has also been shown by our group to be associated with promoter hypermethylation (131). These examples highlight the FAS pathway as a striking example of a pathway that is heavily compromised by selective pressure acting against several components of the pathway.

1.8.6.2 Autophagy

The process of autophagy is highly conserved by a core set of ~30 genes discovered in yeast termed the ATG genes (132). These genes function as part of several interrelated signalling cascades that regulate autophagy (133). In brief, autophagy is mediated by the formation of relatively large, double membraned organelles called autophagosomes which essentially assimilate other cell components that are damaged or otherwise need recycling (134, 135). The mature autophagosomes fuse with lysosomes and the components are reduced into building blocks for de-novo biosynthesis or used to generate ATP (134, 135).

Our understanding of the link between autophagy and tumourigenic processes has grown considerably in recent years. Autophagy was once considered to be primarily a survival function that becomes active in response to both extra and intracellular stress, it is now known to have complicated effects that vary depending on the circumstances (133, 136). Evidence is accumulating that over-activation of autophagy can have a tumour suppressive role by inducing a form of programmed cell death that is distinct from apoptosis and necrosis (113, 136, 137). Conversely, autophagy can also play an oncogenic role by regulating a number of other molecular signals involved in oncogenic processes such as BLC2, beclin-1, class I and III PI3K, mTORC1/C2 and P53 (138).

To date, little work has been carried out linking autophagy directly to MF or SS. However, several histone deacetylase (HDAC) inhibitors including Vorinostat and Romidepsin, which are used in the treatment of CTCL including MF and SS (139) are thought to act at least partly by

facilitating transcription of genes involved in caspase-independent, autophagy mediated cell death (139, 140).

Another possible link between autophagy and CTCL relates to the adenine biosynthesis pathway (141) Autophagy is known to be promoted by spermidine and increasing the bioavailability of spermidine promotes longevity and prevents necrosis (142). Interestingly, spermidine production has been shown to be inhibited in-vitro by Methyl-thioadenosine (MTA) (141) (**Figure 1.12**). Tumour cells of both SS and MF often have deletions of chromosomal region 9p21 (27, 44), a region where the Methyl-thioadenosine-phosphorylase (*MTAP*) gene is found. MTAP is the only known enzyme capable of degrading MTA (143, 144) and tumours cells lacking the *MTAP* gene show highly altered levels of spermidine and its downstream product spermine (141) suggesting a possible link between loss of MTAP and alterations in the regulation of autophagy. Loss of 9p21 will be discussed in more detail in later sections.

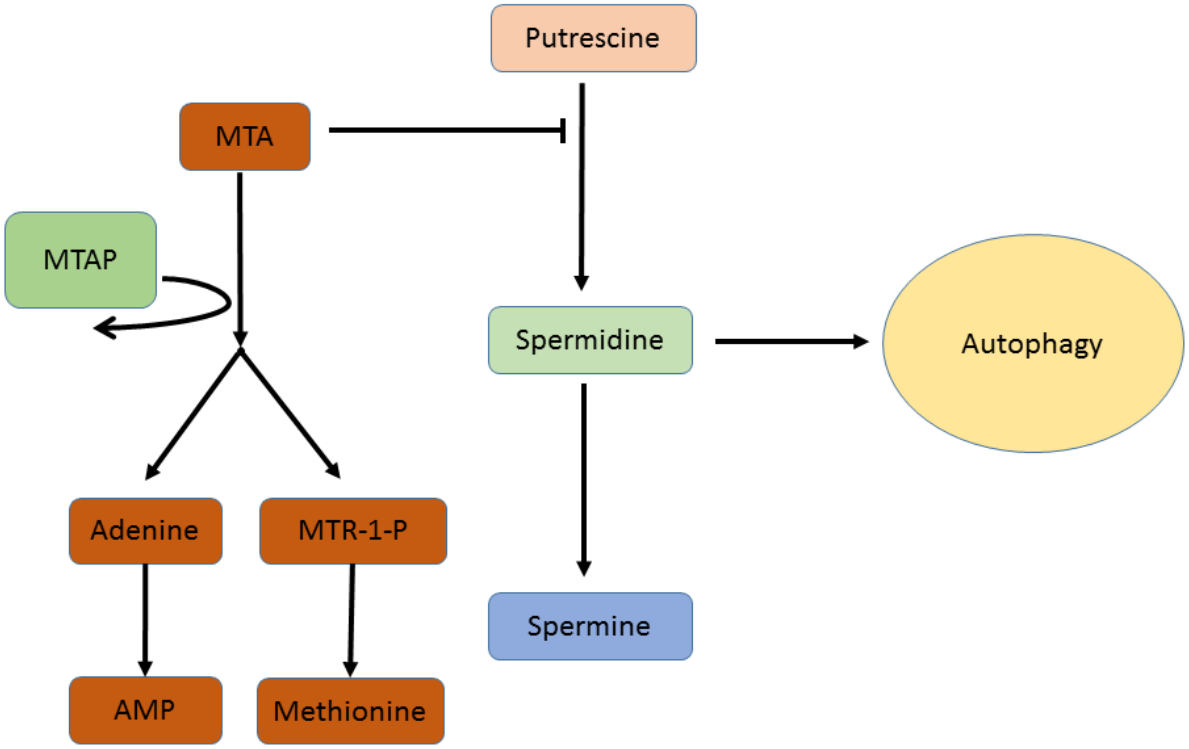


Figure 1.12 Loss of MTAP may inhibit autophagy.
MTAP is lost in a high number of CTCL tumours allowing MTA to accumulate. High concentrations of spermidine promote autophagy, however increased concentrations of MTA inhibit the production of spermidine suggesting that loss of *MTAP* may indirectly inhibit autophagy. The effect this may have on tumourigenic processes is not yet clear.

1.8.6.3 Programmed necrosis and inflammation

The hallmark feature that differentiates necrosis from other forms of cell death is the destruction of the plasma membrane (145). Necrosis has historically been considered as an unregulated, pathological process but more recently several genes have been shown to promote or inhibit necrosis suggesting that necrosis can be a programmed event (145). Several of these signals also promote apoptosis and autophagy but the intensity of the signal, as well as subtleties of the context can alter the balance and it has been speculated that programmed necrosis may be a last ditch attempt to induce cell death in cells that could otherwise cause catastrophic problems such as tumourigenesis (145).

Activation of several receptors including; TNF-R1, FAS, TOL-like receptors and TRAIL activated death receptors have been shown to converge through a necrosis signalling hub centred around the protein kinase RIP1 (146). RIP1 is an activator of NF-κB and can inhibit apoptosis, but under normal circumstances caspase-8 cleaves RIP1 allowing apoptosis if necessary (147). However, if apoptosis is suppressed or caspase-8 is disabled, RIP1 has been shown to induce a cascade of events that results in necrosis, it has been suggested that this may serve as a backup to more controlled forms of PCD (145, 148).

Necrosis is generally thought to elicit harmful effects through the tumour microenvirionment whereby the break-up of cells causes tumour promoting effects by inducing or enhancing general inflammatory processes (149). Indeed, numerous cancers are associated with chronic inflammation (150) and inflammatory processes, such as those mediated by IL-17 are thought to play a role in the development of CTCL (151, 152). Inflammation also contributes to the proliferation and survival of malignant cells and mediates genomic instability, creating an environment that can accelerate the evolution of tumour cells (153). There are several mechanisms of how this occurs; reactive oxygen species are increased in regions of inflammation and are known to be a factor contributing to DNA damage, the other key reason is that sites of inflammation require cells to undergo division to repair damaged tissue (154). This double edged sword underlies the association between inflammation and cancer. There is also increasing evidence that inflammation may promote tumourigenesis by altering the epigenetic landscape of cells which endure chronic inflammation as a feature of their microenvironment (155, 156).

1.8.7 Global epigenetic regulation

Epigenetics is an umbrella term for heritable changes to the regulatory machinery of DNA that do not alter the sequence of bases in the genome (157). Epigenetic changes can be a direct downstream consequence of driver events, or other processes such as inflammation (155) and can result in altered transcriptional patterns of many genes, including those involved in

tumourigenic processes (157). This section will summarize well studied global epigenetic regulatory mechanisms and discuss their relevance to CTCL.

1.8.7.1 DNA Modifications

There are 4 known epigenetic modifications that occur on mammalian DNA, of which methylation is by far the most widely studied and best understood (157). DNA Methylation occurs on cytosine bases when a methyl group is covalently added to the 5' carbon to produce 5-methylcytosine (5mC) (158). Methylated cytosine is found predominantly on cytosine-guanosine dinucleotides (CpGs) spanning ~4% of the genome (159). The CpG sites occur less than would be expected by chance but are generally clustered into small groups near gene promoters where they are known as CpG islands (160). The methyl groups generally have an inhibitory effect on transcription by projecting into the major groove of DNA and likely reducing the binding efficiency of the transcriptional machinery, however methylation is not enough to repress transcription by itself (159). Furthermore, methylation has been reported within the gene body of highly transcribed genes suggesting that the spatial distribution and context of methylation are important additional factors regulating transcription and it seems very plausible that these modifications are involved in the alternate splicing of genes (161).

Methylation is regulated in higher eukaryotes by the DNA-methyl-transferase family of genes. Making sure inheritance of DNA methylation occurs correctly during replication is the function of DNMT1 (162). DNMT3A and DNMT3B are mostly responsible for de novo DNA methylation although can reportedly maintain existing methylation as well (163), whereas DNMT2, despite the name, does not seem to be involved in DNA methylation (164).

Cytosine methylation can be further altered to produce 5-hydroxymethylcytosine (5hmC), a distinct epigenetic modification to 5mC (157). The TET family of DNA hydroxylases (TET1, TET2, TET3) are known to be responsible for this process as well as further oxidation of the residues (157). The purpose of these modifications appears to be complex with modifications in gene bodies and promoters of both transcribed and repressed genes mediated by TET genes (165).

In CTCL the DNA methyltransferase genes *DNMT3A* has been reported as both deleted in 14/66 and mutated in 5/66 (50) and *DNMT3B* which was deleted in 9/80nand mutated in 5/66 cases (50), another study reported 15 deletions and 4 mutations in DNMT3A (49, 50). Furthermore, recent NGS work has discovered focal deletions overlapping the *TET1* gene in 50% of SS tumours as well as nonsense mutations in SS tumours in 4/66 in *TET1* and 6/66 in *TET2* (50, 51). Perturbations in global DNA epigenome regulators would likely have significant effects on the epigenetic landscape of the disease and affect the regulation of many genes. Indeed, hypermethylation has been reported to affect several genes involved directly in key tumourigenic processes such as FAS (131) and T-plastin(39).

Genes with mutations MF and SS	Reference
DNMT3A, DNMT3B, TET1, TET2	Kiel et al 2015, Choi et al 2015,

Table 6. Genes involved in DNA methylation reported as mutated in CTCL.

1.8.7.2 Histone modifications

Post-translational modifications to histones are both diverse and dynamic mechanisms regulating the structure and function of chromatin. Different modifications show both repressive and activating effects on transcription, furthermore they are not necessarily mutually exclusive which can lead to so called 'bivalent domains' (157). Modifications across several histones exert a local combinatorial influence on transcription and other process, including signals which can lead to further modifications of other histones, this has been dubbed 'histone crosstalk' and is of great biological significance (166). The list of documented modifications to histones and their effects is extensive and detailed coverage is beyond the scope of this thesis; however, perturbations in several histone modifying genes have been identified in CTCL and are described below.

The chromatin modifiers *ARID1A* and *ARID5B* were found to be deleted or mutated in 25/62 and 23/80 SS tumours respectively in a recent study (50). Deletions and mutations in *ARID1A* have been observed in SS in 23/40 tumours by others as well (49). *SMARCC1*, an interacting

partner of the ARID family has also been reported as deleted 17/80 SS tumours (50). Members of several histone methyletransferase families including genes from SETD (*SETD1A, SETD1B, SETD6*) and *MLL2* have also been shown to be deleted or mutated in SS tumours (50), with MLL2 deleted in 11/80 and mutated in 7/66 tumours, MLL3 mutated in 39/66 tumours and MLL4 mutated in 12/66 tumours (50). The SETD1A and SETD1B genes were deleted in 14/80 and 12/80, and mutated in 4/66 and 11/66 respectively (50) Numerous other histone modifiers including the lysine-demethylase *KDM6B* and the histone deacetylase *NCOR1* were reported as deleted in the same study (50). Overall, a large variety of chromatin modifiers have been reported as mutated or deleted in CTCL, table 7 lists recurrent genes reported in recent studies.

Genes with mutations MF and SS	**Reference**
ARID1A, ARID1B, SMARCC1, SETD1A, SETD1B, SETD6, MLL2, MLL3, NCOR1, NCOR2, SMARCA4, CREBBP, CHD3, BRD9, ZEB1	Choi et al 2015, Kiel et al 2015, da Silva-Almeida et al 2015, Wang et al 2015

Table 7. Genes with roles in chromatin regulation with mutations reported in CTCL.

1.8.8 Cell cycle control

The cell cycle of all eukaryotes can be divided into 4 main sections; Growth 1/0 (G1/0), Synthesis (S), Growth 2 (G2) and Mitosis (M) (167). The most well studied mechanism which drives progression through the cell cycle is involves interactions between 2 key groups of genes; the cyclin-dependent kinases (CDKs) which are constitutively expressed, and the corresponding cyclins which are expressed in a specific temporal pattern (168). Without the presence of the corresponding cyclins, the CDKs remain inactive but when the cyclins are expressed, different cyclin-CDK complexes form and activate CDKs is a specific spactio-temporal manor driving progression through the cell cycle (168). A less well studied pathway

that can also push the cell through the cycle involves activation of the *C-MYC* gene by growth stimuli, C-MYC is thought to bypass the default cell cycle process and directly drive expression of cyclin-E and *CDC25* (168).

In G1, cyclin D's are the first cyclins to be expressed (168). The binding of cyclin-D family members (D1,D2,D3) to CDK4 or CDK6 activates the complex which can then phosphorylate retinoblastoma (RB), RB can then dissociate from a complex containing members of the E2F transcription factor family, as well as DP1 and RB (168). Once released, E2F members become active which results in the transcriptions of cyclin E, cyclin A, *CDC25A* and many other genes associated with the transition from G1 to S where DNA replication takes place (168). When the DNA has been completely replicated the cell is said to have completed S phase and be in G2 (168)(**Figure 1.13**).

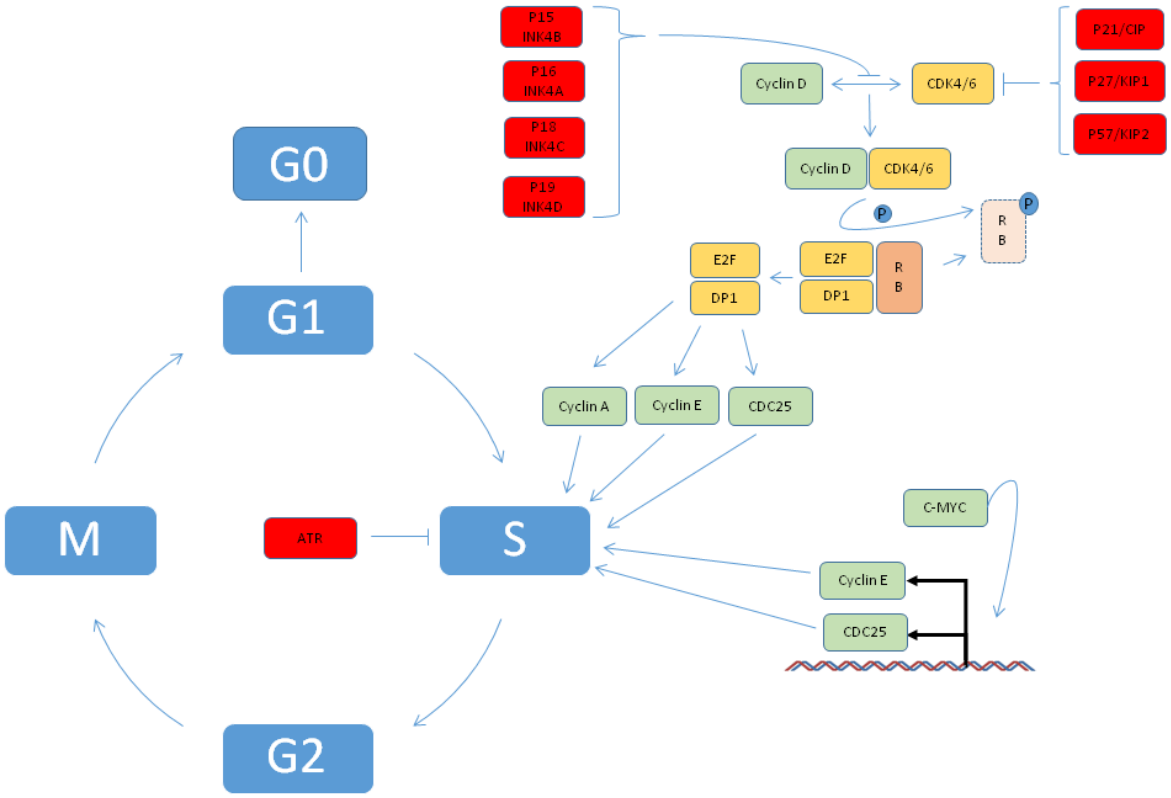


Figure 1.13 G1/S Regulation.
Progression from G1 to S phase is driven by increasing expression of CDC25 and Cyclin D/E combining with CDK4/6. The cylclin CDK complex formation can be inhibited by the INK4 genes and the CDK activity is negatively regulated by CIP/KIP genes. Sufficient cyclin/CDK activity phosphorylates RB, releasing E2F and DP1 to drive increased expression of Cyclin A/E and pushing the cell from G1 to S. This mechanism can be bypassed by C-MYC activity which can drive expression of *CDC25* and Cyclin E by itself. Further negative regulation is controlled by ATR signalling which can prevent progression through S-phase in response to replicative stress. Several layers of tight regulation control the balance of active cyclin/CDK complexes which promote G1/S transition. Green indicates cell cycle drivers, red genes indicate cell cycle repressors, genes in yellow are neutral.

The G1/S transition is negatively regulated by many factors; well-studied inhibitors include the INK4 family which interfere with the cyclin-D binding CDKs (CDK4 and CDK6). $P15^{INK4B}$ is transcribed by the *CDKN2B* gene (169). $P16^{INK4A}$ is also a product of the CDKN2A gene (170), the remaining members include $P18^{INK4C}$ and $P19^{INK4D}$ (171). The other CDK inhibitors are the CIP/KIP family; $P21^{CIP}$, $P27^{KIP1}$ and $P57^{KIP2}$, which inhibit all CDKs (171, 172). Many further factors can inhibit cell cycle progression through different mechanisms; the *CDKN2A* gene produces a functionally and structurally distinct transcript called $P14^{ARF}$ which mediates TP53

induced cell cycle arrest, reviewed in (173). Furthermore, the G1/S transition can be halted under replicated stress by ATR signalling (174).

Progress from G2 to M is dependent on the accumulation of active CDK1 (also called CDC2) bound by cyclin-B or cyclin-A (175, 176). CDK1 is initially rendered inactive by WEE1 and MYT1 phosphorylation, so to achieve full activation of the complex, further de-phosphorylation by CDC25 family members is necessary (177, 178). The B and A cyclins are degraded as the cell progresses from metaphase to anaphase during mitosis thus removing the activity of CDK1 (176). Further negative regulators of the G2/M transition include; GADD45, P21 and 14-3-3-sigma, all of which are targets of TP53 (175). In addition, the ATM/ATR genes can negatively regulate progression by directly activating CHK1 and CHK2 which themselves inactivate CDC25 (175) (**Figure 1.14**).

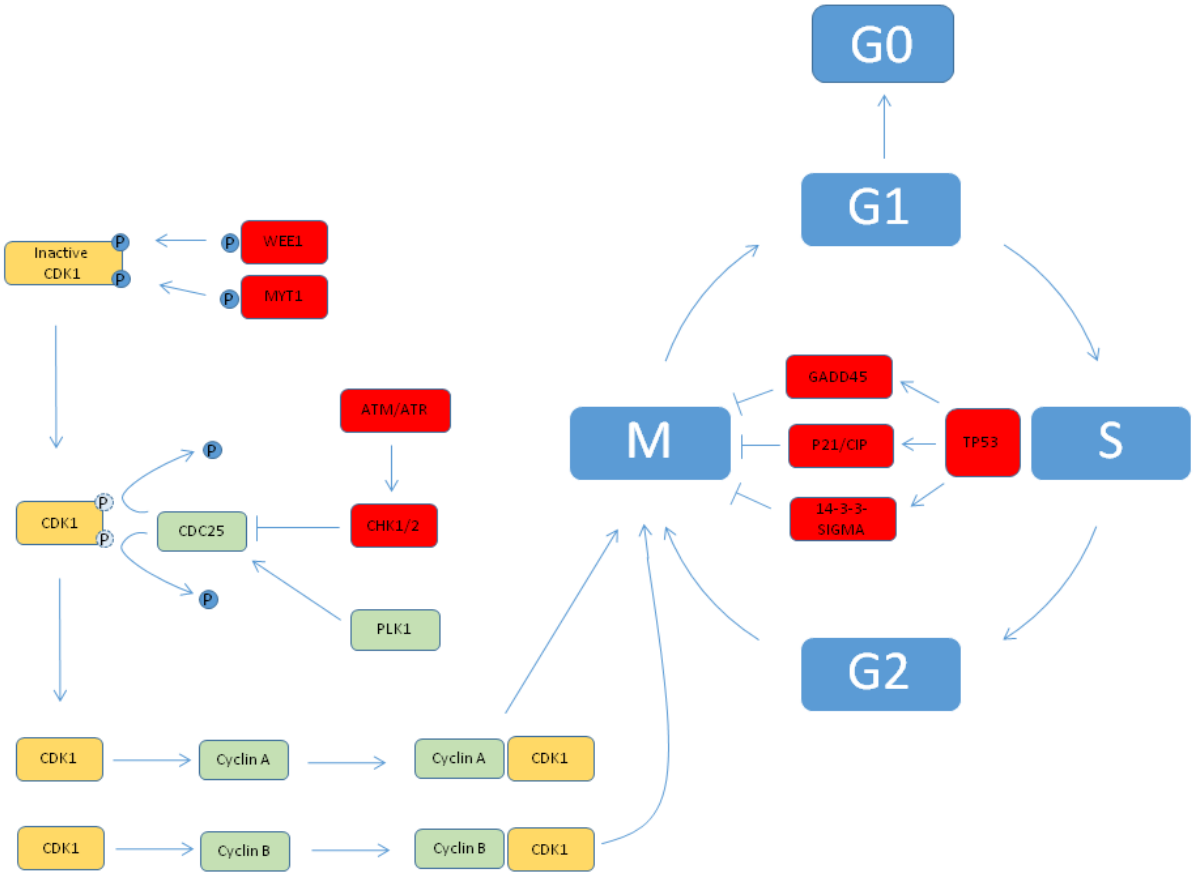


Figure 1.14 G2/M Regulation.
CDK1 activation by Cyclin A/B drives the G2/M transition. CDK1 is rendered inactive by phosphorylation via WEE1 and MYT1, however CDC25 activity reverses this phosphorylation allowing CDK1 to form active complexes with Cyclin A/B. CDC25 can itself be activated by PLK1 or inactivated by the ATM/ATR pathway via CHK1/2. TP53 can also prevent G2/M transition by activating GADD45, P21/CIP and/or 14-3-3-SIGMA. As with G1/S a series of tightly orchestrated events facilitate or inhibit the build-up of cyclin/CDK complexes which drive cell cycle progression. Green indicates cell cycle drivers, red genes indicate cell cycle repressors, genes in yellow are neutral.

Genes with mutations MF and SS	**Reference**
CDKN2A, CDKN2B, ATM, TP53	Salgado et al 2010, van Doorn et al 2009, Laharanne et al 2010, Laharanne et al 2010b, Scarrisbrick et al 2002, Caprini et al 2009, Wang et al 2015, da Silva Almeida et al 2015, Choi et al 2015.

Table 8. Genes reported as mutated or deleted in CTCL which affect the cell cycle and checkpoint activation.

In CTCL several core components of the cell cycle have been reported as disrupted, most prominent among these is loss of chromosomal region 9p21 in up to 41% of MF and SS tumours (16, 22, 23, 25, 27, 44, 48) which includes the genes *CDKN2A* and *CDKN2B* (Table8). Both of these genes express potent cell cycle inhibitors of the INK and ARF family (171). Members of the CIP/KIP family of cell cycle inhibitors also show loss of expression in CTCL (179). Cyclin-E has also been reported as constitutively expressed in CTCL cell lines (180). PLK1 is a serine/threonine kinase that is crucial in regulating the G2/M transition by initially activating CDC25 (181, 182). It has been reported as over-expressed in advanced CTCL and CTCL cell lines, RNAi knockdown resulted in increased cell cycle arrest and apoptosis suggesting that this gene may be a good target for translational investigation (181, 182).

1.8.9 DNA repair and genome maintenance processes

As previously discussed, complex chromosomal rearrangements and genomic instability are key features of MF and SS. Non-tumourigenic cells repair damaged DNA on an ongoing basis and remain healthy (183). Therefore, a sensible line of inquiry would be to examine how pathological DNA damage manifests and accumulates in CTCL cells. The DNA damage response (DDR) is the most obvious candidate network and includes ATM, TP53 and the checkpoint activation network (184). Problems with DNA repair processes themselves are also highly plausible candidates and problems with either or both would have catastrophic effects on genome maintenance.

1.8.9.1 Repair activation

Ataxia telangiectasia mutated (ATM) and the closely related ATR (ATM and RAD3-related) are protein kinases which become activated in response to DNA damage (185, 186). They have numerous targets and similar substrate specificity (187), with differing substrates depending on the phase of the cell cycle. ATR responds mainly to DNA damage caused by replication stress and DNA cross-linking whereas ATM is a principle activator of double stranded break (DSB) repair, cell cycle checkpoint induction and responses to radiation (188). ATM has been

shown to phosphorylate TP53, MDM3 and CHK2 and activate the G1 checkpoint (189-192). The S-phase checkpoint can be activated by ATM phosphorylation of NBS1, FANCD2, and BRCA1 (193-195). In G2, the checkpoint can be initiated by ATM phosphorylation of BRCA1 and RAD17 (195, 196). Numerous studies show that ATM can utilize cell cycle checkpoints in all phases to elicit the DDR. ATM has been reported as lost in 12/40 or mutated in 1/40 SS tumours (49) and this would likely result in problems initiating the DDR and contribute to problems relating to genome maintenance. See table 8 for a list of genes affected in CTCL involved in the cell cycle and checkpoint activation.

Tumour protein 53 (TP53) is one of the most widely studied tumour suppressors implicated in cancer. It operates as a master controller of many biological processes and activates genes involved in arrest of the cell cycle, activation of DNA repair, senescence and if necessary programmed cell death pathways (197). *TP53* is part of a family of related genes which includes *TP63* and *TP73*, which have separate roles but may also play a part in tumour suppression (188). Under normal conditions, TP53 is held in an inactive cytoplasmic complex with MDM2, MDM4, DAXX and HAUSP where it is targeted for ubiquitinylation and subsequent degradation by the proteasome complex (188). The complex is extremely sensitive to disruption by small molecules, stress stimuli and other activator pathways which release and activate TP53 (188). TP53 is directly activated and released from its cytosolic complex by ATM and ATR (188), the cell cycle regulator ARF also indirectly activates TP53 by sequestering its main binding partner MDM2 (198). Increasing activity in oncogenic signalling cascades, including MAPK pathways, can also raise the level of free TP53 by disrupting is sequestration by MDM2 (188). TP53 can then be activated by residual levels of ATR which is regarded as the classical route by which TP53 responds to hyper-proliferation (188, 199).

TP53 is located on the short arm of chromosome 17, a region frequently lost in CTCL in up to 45% of tumours (22, 24, 31, 46, 127). In addition, loss of function mutations have been recently reported in large sequencing studies in SS in 7/40, 7/37, and 4/25 tumours (16, 49, 51,

200), See table 8 for a list of genes affected in CTCL involved in the cell cycle and checkpoint activation. Loss of key tumour suppressor genes such as TP53 is likely to contribute significantly to the progression of CTCL as it causes reduced functionality of many tumour suppressing processes, key among them is the ability to initiate DNA repair or PCD. In addition, the 9p21 locus, which contains several tumour suppressor genes including *CDKN2A*, is often deleted in CTCL as has been previously discussed. *CDKN2A* produces several transcripts including $P14^{ARF}$, which normally sequesters MDM2 when cell cycle arrest is necessary. Loss of $P14^{ARF}$ indirectly leads to reduced active TP53 by increasing the availability of MDM2 to retain TP53 in its inactive state (188, 199).

1.8.9.2 DNA repair

There are several known mechanisms of DNA repair, each used to repair different types of damage (**Figure 1.15**). The main mechanisms include:

(1) Base excision repair (BER). This mechanism is used for repairing DNA lesions that do not alter the structure of the helix and involve the creation of scaffolding proteins and a single stranded DNA break as an intermediate step, reviewed in (201). Many of the same enzymes involved in the initial stages of the pathway are also involved in the repair of single stranded breaks that occur as a result of damage (202).

(2) Mismatch repair (MMR). The most common use of MMR is to replace mismatched bases after DNA replication that have not been corrected by the polymerase complex's initial proofreading, reviewed in (203). In addition, MMR also corrects INDEL loops that arise from polymerase slippage (203). The process is highly conserved between organisms as distant as humans and E.coli (204) and involves a recognition step, an excision step where a gap is generated, and finally a repair synthesis step to fill in the gap (205-207). Problems with the MMR process causes cells to by highly mutatable and display characteristic microsatellite instability (208), interestingly up to 27% of CTCL patients demonstrate microsatellite instability (209, 210). Promotor hyper-

methylation has been demonstrated in HMLH1 which is involved in MMR (210) suggesting a link between loss of MMR functionality and CTCL.

(3) Nucleotide excision repair (NER). DNA lesions that distort the helical structure are mainly repaired by NER. An example of such damage includes UV-induced cyclobutane pyrimidine dimers. Although NER is mechanistically similar to BER, it involves many steps and ~30 genes are known to play a role (211-213). The steps include; recognising the damaged DNA, unwinding ~30bp of the helix surrounding the lesion, excising a single-stranded DNA segment surrounding the lesion, repair synthesis and finally ligation of the repaired strand (183). A key gene involved in binding the unwound single-stranded helix prior to endonuclease cleavage is known as replication protein A (RPA), this gene is found on the short arm of chromosome 17 which is deleted in up to 45% of CTCL patients (25, 26, 41, 42, 47, 48). This suggests that loss of function of NER may contribute to genomic instability in CTCL.

(4) Homologous recombination (HR). Double stranded breaks (DSBs) are repaired in one of 2 ways, HR is the less error prone method because it uses the sister chromosome as a template (214). It occurs in several phases, the first of which is presynapsis which involves the conversion of both of the DSB ends into 3' overhangs. The synapsis phase involves aligning the 3' DSB ends to homologous regions on the sister chromosome and the formation of Holiday junctions. The postsynapsis phase involves elongation of the 3' end and subsequent re-joining of the DSB ends to form 2 separate chromosomes again (214). The HR method of DSB repair is mainly active during the late-S and G2 phases of the cell cycle (183, 215). The RPA gene, is involved in binding the 3' ends to stabilise the structure prior to binding by the RAD51 recombinase, BRCA2 and several other members of the HR repair process (183). Loss of the RPA gene from chromosome 17p is likely to affect HR in CTCL cells.

(5) Non-homologous end joining (NHEJ). NHEJ is the other method by which DSB repair occurs. It is considered an error prone process compared to other mechanisms of DNA

repair (216) but is able to act throughout all phases of the cell cycle(215). As the name suggests it involves the direct ligation of DSBs (217). A relatively small number of components facilitate the process by first identifying DSB termini, encapsulating it and recruiting additional components. These encapsulated ends are then able to form a synapse with other DSBs and allow the repair process to take place (218). The breaks may be of varying complexity and require clean-up steps to allow ligation to take place. Components involved in this process may vary according to the type of DSB but are thought to include components from other repair mechanisms such as lesion specific base excision repair (BER) enzymes including APEX1, TDP1 and PNKP (219). The final end processing is likely responsible for the error prone nature of NHEJ (220). A reduced level of Ku70/Ku80 has been reported in CD4+ T-cells in CTCL tumour sites (221). This heterodimer is an important component involved in the initial recognition and binding of DSBs in the NHEJ process which may confer a reduced ability to respond to DSBs in CTCL cells.

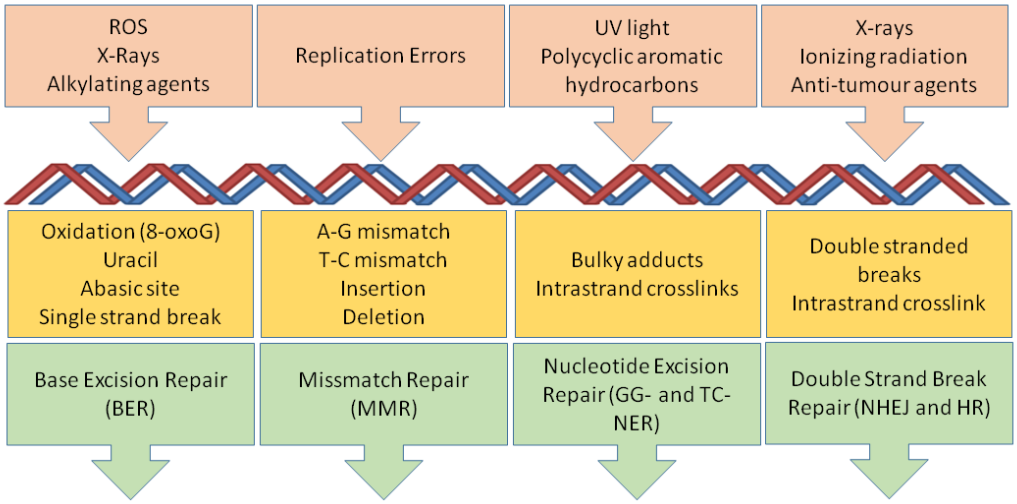


Figure 1.15 Summary of DNA damage and repair mechanisms.
Known sources of DNA damage are shown (top) with examples DNA lesions caused by each mutagen (middle). The relevant repair mechanism used for amending the damage is listed below (Adapted from **Mathews and Dexheimer 2013**).

1.8.9.3 Telomere maintenance

The structure of DNA and chromatin at the ends of the chromosomes is different from the rest of the chromosome arms and is known as a telomere (222, 223). Healthy telomeres are stable structures that differ to the standard DNA double-helix, predominantly being arranged in a formation known as a G-quadruplex (224). Telomeric sequences consist of repetitive '5 to 3' TTAGGG sequences ranging from a few Kb to >100Kb in length (222). The terminus of the telomere is separated into a complex single-stranded arrangement involving a large loop known as a T-loop and a smaller inner triple stranded loop known as a D-loop (222). The complex structure is coated by a 6-protein compound known as shelterin which consists of TRF1, TRF2, TIN2, RAP1, POT1 and TPP2 and serves to protect the telomere from being recognized by the DDR system and targeted for repair (225)(**Figure 1.16**).

Polymerase enzymes generally replicate DNA in a 3' to 5' direction and are initiated by an RNA primer (226). Replicating DNA in this way means that the 3' end cannot be fully replicated so the long repeated regions in telomeres are a solution to the problem, as small sections of the telomere are lost with each mitotic cycle (222). In terminally differentiated cells a division limit called the Hayflick limit (227) is reached when enough cell cycles have passed so that the telomere reaches a critically short size, upon which replicative senescence and exit from the cell cycle are triggered (228). The telomeres are maintained by a retrotranscriptase called telomerase which can elongate the repeat elements in the telomeres as it does in gametes, stem cells and many cancer cells (229), but is inactive in most cells that have undergone terminal differentiation (230, 231).

Paradoxically in CTCL, short telomeres have been observed in aggressive subgroups in parallel with active telomerase (232, 233), it has been suggested that telomerase may exert additional functions in CTCL besides telomere maintenance (232).

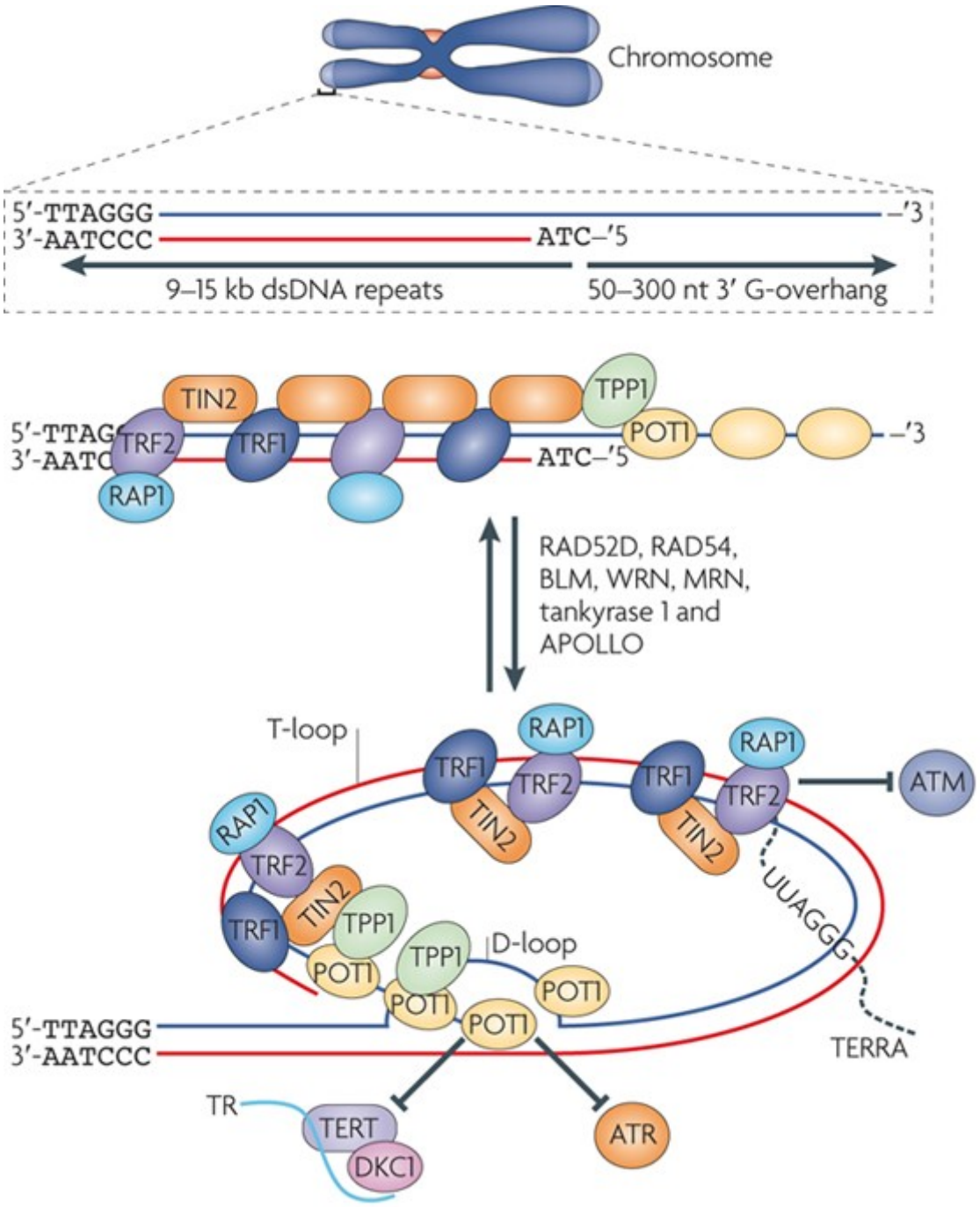


Figure 1.16 The structure of human telomeres.
Human telomers contain long repeats of TTAGGG with a 3' overhang. At the terminus, the telomere loops back on itself to form a protective structure coated in a protein complex called shelterin. This structure prevents the telomere from being recognised by ATM and components of the double-stranded break response. Shelterin is composed of the proteins TRF1, TRF2, TIN2, RAP1, POT1 and TPP2 and perturbations in these proteins are likely to compromise the structure of telomeres and hence the integrity of the chromosome ends. Ilustration from (**O'Sullivan and Karlseder 2010**).

1.9 Case study of a common focal deletion: driver genes on chromosome 9p21.3

CTCL tumours show a relatively high level of structural variation. This was detected in the earlier studies by us and others using low resolution genomic approaches (22-24, 26, 27, 31, 41, 48, 127, 234) and subsequently confirmed in recent NGS papers (16, 28, 29, 49-51). Many structural variations encompass large sections of chromosomes, most notably chromosome 17 which carries many known cancer driver genes such as *TP53* and likely even more passenger genes. However, some recurrently deleted chromosomal regions are much smaller and therefore one hypothesis is that these so called focal deletions are more likely to contain predominantly driver genes. One such region is chromosome 9p21.3 which sits ~500kb centromeric from the interferon-α cluster and contains the genes (from centromeric to telomeric) *CDKN2B-AS1*, *CDKN2B*, *CDKN2A* and *MTAP* (**Figure 1.21**). This region has been reported as deleted in many malignancies such as cutaneous B-cell lymphoma (235), non-melanoma skin cancer (236) and melanoma (237). Whilst CDKN2A and CDKN2B have well known roles in tumourigenesis, as has been previously discussed, the contribution of MTAP is less well known. In this section a recap of CDKN2A and CDKN2B and their established roles in cancer will be briefly reviewed, followed by a discussion of MTAP and evidence that MTAP may also act as a tumour suppressor.

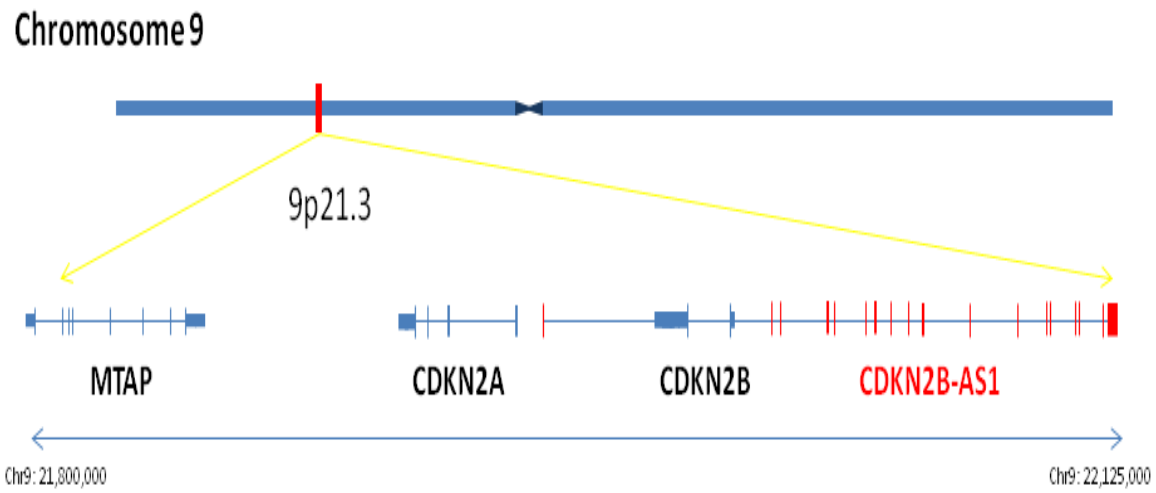


Figure 1.17 Schematic of genes in 9p21.3.
This small region on chromosome 9 is deleted or otherwise inactivated in many malignancies. The *CDKN2A* and *CDKN2B* genes expresses cell cycle regulators of the INK4 and ARF families (**Figures 1.13-1.14**) which have well known roles in tumourigenesis. The *CDKN2B-AS1* gene is a non-coding RNA involved in the regulation of *CDKN2B*. The *MTAP* gene is also selectively deleted in several malignancies but the mechanism by which it contributes to tumourigenesis is currently undefined.

1.9.1 CDKN2A and CDKN2B

There is a great deal of evidence to suggest that *CDKN2A* and *CDKN2A* have prominent roles in tumourigenesis. The genes encode several alternatively transcribed tumour suppressor proteins with functionally distinct roles named $p14^{ARF}$, $P16^{INK4A}$, $p15^{INK4B}$. The function of these genes has been previously discussed and they are known to have roles in cell cycle control and DNA repair activation (**Figures 1.13-1.14**). Further evidence for a tumour suppressing role includes; correlative deletion of the loci in tumour cells (235-237) and aberrant promoter hypermethylation leading to loss of expression (238-240), including in CTCL (44).

1.9.2 MTAP

MTAP, or 5'-methylthioadensine phosphorylase is the only known enzyme responsible for the breakdown of 5'-methylethioadenosine (MTA) (241). MTAP phosphorylates MTA which then

degrades into adenine and 5-methylthioribose-1-phosphate (MTR-1-P) (242). Adenine can then be further degraded into AMP whilst MTR-1-P is metabolized into methionine (242)(**Figure 1.12, 1.22**). Loss of *MTAP* expression is reported in a range of solid (243-245) and haematological malignancies (246-252) and it has been suggested that changes in MTA concentration and its downstream effects could cause a selective growth advantage which may explain the high frequency of MTAP deficiency in tumour cells (242).

There is increasing functional evidence that *MTAP* may act as a tumour suppressor. One of the strongest lines of evidence comes from an *MTAP* loss mouse model (253). Homozygous loss of *MTAP* is embryonic lethal in the mouse which highlights an essential developmental role. Perhaps more interestingly, mice heterozygous for *MTAP* loss are highly susceptible to developing mature T-cell lymphoma (253). Further evidence of MTAP acting as a tumour suppressor comes from the MCF-7 *MTAP*(-/-) cell line where reintroduction of MTAP inhibits tumour formation in xenographted in-vitro mouse models (141), the mechanism appears to be related to the enzymatic function of MTAP and its effect on intracellular polyamine pools as constructs engineered with defective catalytic regions show similar effects to *MTAP* loss (141).

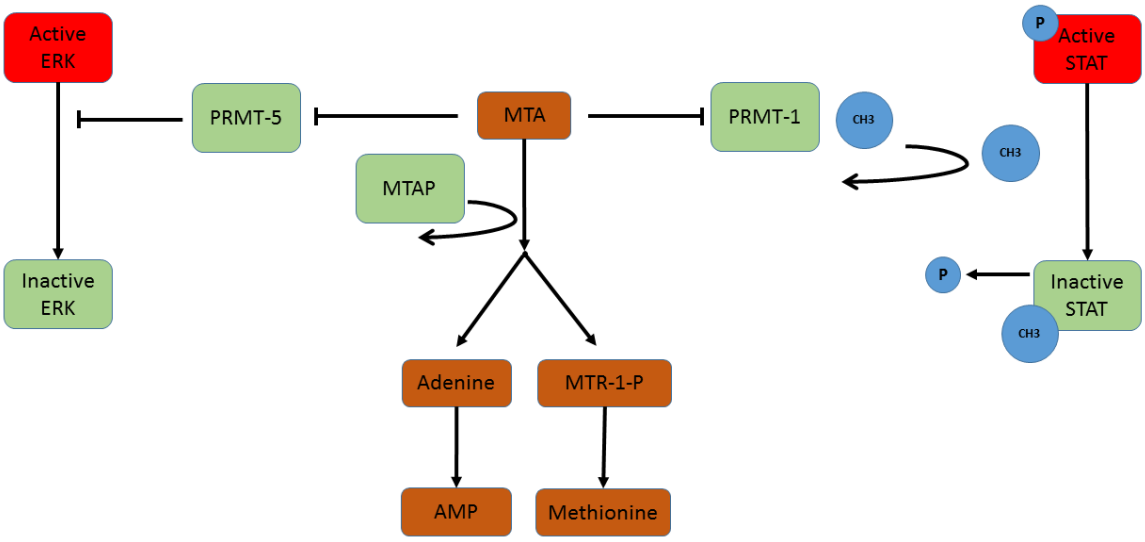


Figure 1.18 How MTAP may act as a tumour suppressor.
MTA is usually broken down by MTAP as the first step in the adenine and methionine salvage pathways. In the absence of functional MTAP, MTA is known to accumulate. High concentrations of MTA is thought to inhibit the PRMT family of methyltransferases. Several members of the PRMT family are known to reduce activity of the JAK/STAT and MAPK/ERK oncogenic pathways. It is possible that inhibition of PRMTs by increased MTA may lead to over-activity of these pathways and contribute to tumourigenesis.

Correlative genetic studies provide further support that MTAP is an active participant in tumourigenesis. In non-small cell lung cancer and astrocytomas *MTAP* loss has also been shown to occur without loss of CDKN2A indicating a possible selective pressure specifically for loss of MTAP in the development of these malignancies (254). Loss of *MTAP* is also recognized as a factor contributing to the difference between superficial spreading melanoma and nodular melanoma implying MTAP exerts a phenotypic effect on tumour subtype (255). In malignant melanoma, functional MTAP is established as a positive prognostic factor and its expression is predictive of response to adjuvant interferon therapy (238, 256-259). *MTAP* is also known to be silenced epigenetically via promoter methylation in gastric (260) and lung (261) cancers as well as hepatocellular carcinoma (262, 263) and melanoma (238).

The mechanism by which MTAP deficiency contributes to oncogenesis is likely to be via accumulation of MTA. Functional MTAP is essential for MTA breakdown and increased MTA is directly linked to increased tumourigenicity in several malignancies including hepatocellular

carcinoma (262, 263), melanoma (264) and head and neck cancers (265). In addition MTA has been shown to influence numerous critical cellular functions including regulation of gene expression, proliferation, differentiation and apoptosis (144).

How MTA affects these critical cellular functions is unclear but one possible mechanism is by inhibiting a key class of enzymes called protein-arginine methyltransferases (PRMTs). There are 9 of these enzymes and they are responsible for methylation of arginine and lysine residues on numerous target proteins which can have activating or repressive activity depending on the specific target (266). Importantly, STATs 1, 3 and 6 require methylation of the arginine-31 residue by PRMTs in order for dephosphorylation to occur, dephosphorylation is an essential step in the deactivation of activated STATs (267-269) suggesting that increased MTA as a result of *MTAP* loss may contribute to the constitutive activation of STAT family members (**Figure 1.22**). MTA has been shown to inhibit PRMT1 and thus prevent arginine methylation of STAT1 (267).Whilst activating mutations and copy number gains have been documented in JAKs and STATs in CTCL (16, 28, 30, 49, 50), JAK-STAT pathways are also known to be over-active in the absence of genetic anomalies, for example constitutive activation of STAT3 (90). High concentrations of MTA may contribute to inactivation of further PRMT family members and have downstream effects on STAT3 and/or others such as STAT5A/B.

A study in melanoma demonstrated loss of both symmetric and asymmetric arginine methylation correlated with loss of MTAP and increased MTA levels. In cell line models, restoring *MTAP* expression was able to restore protein methylation (270). Protein methylation by PRMT5 is also known to modulate ERK signalling (271). Interestingly, MTA levels also correlate with MAPK/ERK activity and restoration of *MTAP* expression was able to restore PRMT function and reduce ERK activity (270) suggesting another key cancer pathway may be indirectly attenuated by MTAP (**Figure 1.22**).

A recent study of MTAP in fibrosarcoma cells suggested the possibility that MTAP has tumour suppressing properties unrelated to its enzymatic function (272). Reintroducing *MTAP* into

MTAP negative cells resulting in loss of tumourigenic properties, however reintroducing a catalytically inactive version of MTAP also resulted in a similar loss of tumourigenic properties suggesting MTAP has other functions besides MTA metabolism (272), though it is not clear how this may work. Intriguingly, this is in contrast to earlier work which suggested the enzymatic function was responsible for the tumour suppressing properties of MTAP (141).

From a translational perspective, MTAP could be targeted therapeutically (242) as cells that lack MTAP cannot salvage adenine and methionine from MTA making them sensitive to de-novo purine biosynthesis inhibitors and methionine starvation (246, 252). Further work that supports a tumour suppressing role for MTAP will add weight to therapeutic strategies which target MTAP.

1.10 The application of next generation sequencing to the study of cancer

1.10.1 Overview of NGS

Next generation sequencing (NGS), also referred to as massively parallel sequencing, is a maturing technology that has kick-started a revolution in the way we can interrogate sequences of nucleic acids, primarily in terms of how many bases we can sequence per unit cost but also in how rapidly a genome can be sequenced. It expands on the earlier sequencing methodologies created in the late 1970s by Fred Sanger (273, 274). Whilst Sanger sequencing is still used as a gold standard, reliable method for detecting variants at known locations, its main limitation is the cost per base which currently stands at approximately $2400 per MB ($1 per 400-900bp read). The other key limitation is that it can only be used to interrogate specific target sites of the genome in any practical sense. The first NGS platform was the 454 pyrosequencing method which was developed commercially by Roche in 2005 and reduced costs and turnaround time considerably, allowing a full genome to be sequenced for between $10-25,000,000 (depending on size) in a matter of weeks (275). To give some perspective of the rate of improvement, the first complete publication of the human genome was published

in 2003 and took approximately 10 years to complete at an estimated cost of $2.7 billion (276). Since then the year on year reduction in price has drawn comparison to Moore's law (277), currently the Illumina HiSeq X10 is said to deliver sequencing at approximately $10 per Gb making it the first commercially available platform capable of sequencing a human genome for under $1000 (278). The emerging, so called third generation sequencing technologies such as Pacific Biosciences RS have several novel advantages over previous systems as they utilize a sequencing by synthesis approach that can read sequences in excess of 20,000bp, enabling the detection of complex structural variants and reducing error from duplicate PCR reads (278). The fall is price will undoubtedly mark the start of an era of personalized medicine from which society will greatly benefit.

1.10.2 NGS as a discovery tool

Over the past 10 years NGS has been used increasingly in genetic studies of malignancy. It has been used primarily in the discovery of variants and genes that are affected by mutations and copy number variations (CNVs). Recently a flurry of publications has discovered many genes perturbed in CTCL that were not previously known to be implicated. These studies used NGS as the main investigative tool and have galvanized some already known signalling and genomic perturbations as well as some which appear in other malignancies that were not previously implicated in CTCL. The Full list of genes is too numerous to list here but important groups include; the JAK-STAT pathway (28, 50), global epigenetic regulators (49, 50), PLCG1 and related pathways in SS and MF (29, 30, 49), the ATM DNA repair initiation pathway (49).

Rapid and cost effective discovery of numerous genes implicated in CTCL has confirmed the genetic heterogeneity of the tumours but also brought to light some commonalities. Gene families, signalling pathways, as well as genome maintenance and regulatory mechanisms are commonly perturbed and at this level some homology starts to emerge in terms of functional effects. A good example of this is that most tumours seem to show perturbations in global epigenetic regulators (50). In addition some signalling pathways appear to be shared with

other blood born malignancies including PLCG1 pathways which are dysfunctional in adult T-cell lymphoma (ATL) (279). This is particularly important as it raises the possibility of using treatment options that were originally developed for other malignancies. Whilst NGS has highlighted many genes as implicated in the disease, a massive amount of functional work is now needed in order to confirm these genes as driver genes, as well as characterisation of their roles in the initiation, maintenance and/or progression.

1.10.3 NGS as a molecular diagnostic tool

Whilst initial use of NGS has been predominantly restricted to research, it has begun to be used in a clinical setting as a molecular diagnostic tool and will undoubtedly feature more prominently in this area in the coming years. Haematological malignancies have historically been diagnosed using a combination of cytogenetics, immunology and PCR based methods (280). Histological criteria have also been used to guide diagnosis, including in CTCL (4). However, NGS will likely supersede many of these methods in the near future as it combines the ability to detect specific variants with single base resolution as well as structural variants for an entire tumour in a single experiment (281).

Characterizing the entire genome of a tumour has implications well beyond diagnosis, indeed it is likely to change the way that diagnosis and classification of tumours occurs in general. The reason for this is because accumulating whole genomic data of a tumour type will likely at some point allow that tumour type to be categorized, very accurately, into distinct tumour subtypes which can be treated differently if necessary and given different prognostic outcomes using multivariate models. This has already started to happen in other cancer types such as breast cancer (282) and glioblastoma where over 500 tumours have undergone NGS studies and been separated into proneural, neural, classical and mesenchymal subtypes (283-285).

In a further somewhat unexpected development it seems that underlying genetic similarities can even be drawn between different tumour types suggesting that the biology of some tumours may in some cases be more similar to completely different tumour types than to

superficially related sub-categories (278). An example of this is squamous cell lung cancer, which has a similar mutation and variant profile to non-HPV head and neck carcinomas (286). Overall this implies that NGS will be a necessary tool in the design of targeted therapies as tumours initially thought of as similar may require radically different treatment between different patients.

1.10.4 Challenges presented by NGS

NGS studies produce massive amounts of data, often on the order of gigabytes to terabytes per sample for deep whole genome sequencing (278). The vast majority of this data is not likely to be relevant to the disease and storage of large data sets remains one of the major challenges of NGS studies.

Another major drawback is the large start-up costs associated with NGS sequencing, ranging from the ION torrent PGM which is currently the cheapest at $80,000 plus reagents, up to the illumina HiSeqX10 currently priced at $10 million (278). However, the large start-up costs can be mostly offset by the cost per base of sequencing.

Resolving different subclones within the same tumour sample will remain a challenge for some time, although this is not strictly an NGS problem as it is equally related to tumour sample collection. The Pacific Biosciences platform is currently the only platform capable of long read sequencing but has a relatively high operational cost, improving the cost effectiveness of long read sequencing should contribute to improving the ability of NGS to resolve subclones in a single sample.

High performance computing infrastructure is currently required to analyse NGS datasets with a minimum of 8 quad core, 32 gig RAM and 10 terabytes of hard disk space (280). Additional costs of dedicated ICT and bioinformatics staff required to operate and maintain the computing infrastructure can be prohibitive for most research and diagnostic labs (287). However modest improvements to IT and bioinformatics training for bioscience graduates

should make this challenge easily surmountable in the near future. This coupled with cloud based analysis pipelines will likely generate considerable advancement in the field (288)

1.11 Identifying cancer driver genes

The first mutation identified as a bona-fide cancer driver was a single base G>T which converts a glycine to a valine in codon 12 of the HRAS gene (289, 290). The confirmation that genetic events cause tumourigenesis motivated the race to find more genes and events that contribute. This enthusiasm continues today and many more have been discovered with 518 driver genes confirmed through functional studies and a further 1053 strong candidate genes identified through bioinformatic approaches (291). As previously discussed, NGS has been widely applied to the study of cancer, particularly in the pursuit of identifying cancer driver genes. In this section an overview of the general workflow and initial analysis pipeline and will be presented. Then the different bioinformatic approaches used downstream to screen for driver variants will be discussed. Finally, a brief general discussion of functional methods that can be used to validate candidates will be mentioned.

1.11.1 Preparing libraries for next generation sequencing

Staring from genomic DNA, libraries are prepared for sequencing by a series of steps beginning with fragmentation by ultrasound to uniform lengths. Fragments are then processed to repair the ends and ligate universal adaptors to each fragment. At this stage samples can be sequenced if whole genome sequencing is required of further target enrichment can be undertaken such as whole exome capture. A popular method of target enrichment is to hybridize the libraries with streptavidin coated RNA probes which are complimentary to the target region. Various biotinylated surfaces can be used to capture and isolate the target regions from the non-target regions (292)(**Figure 1.17**).

The universal adaptors ligated previously are used to anneal the ends of the remaining fragments to a solid platform where the sequential addition of fluorescent dNTPs, followed by image capture is performed. This is the mechanism used in current Agilent and Illumina

technologies which make up the bulk of large scale sequencing studies and were used to generate sequence data for experiments contributing to this thesis.

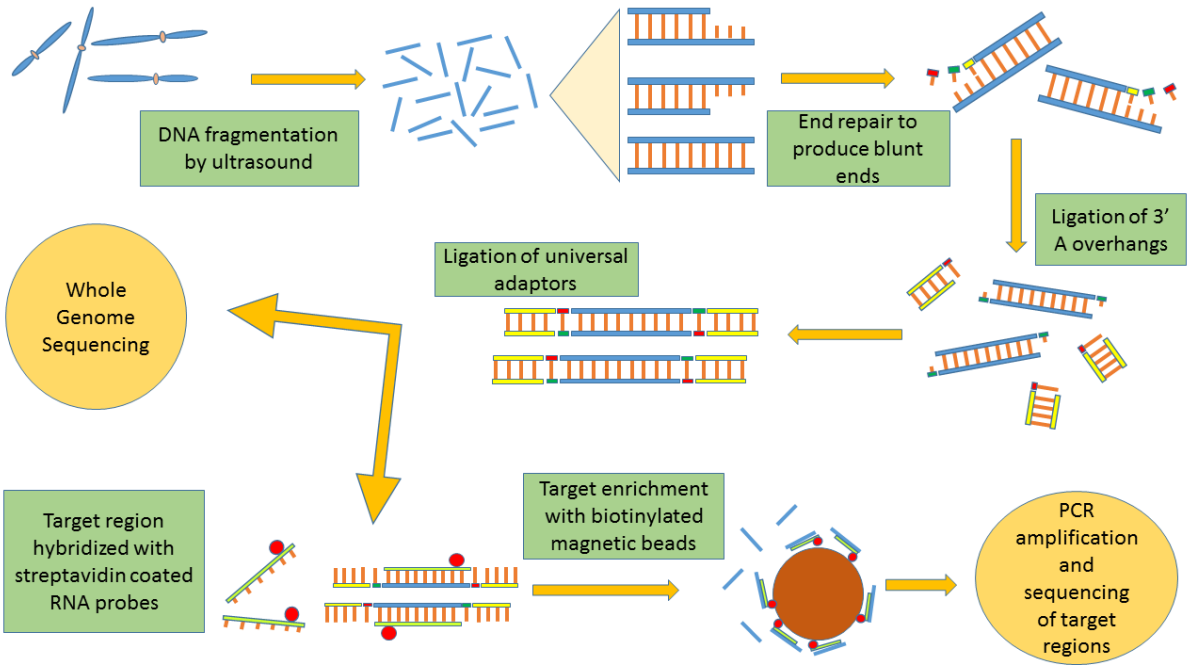


Figure 1.19 Summary of Library preparation for NGS.
Genomic DNA is sheared by ultrasound to 170-190bp. Fragments can have overhangs on the 5' and 3' ends or can be blunt ended, all fragments are repaired to be blunt ended. After repair, 3'-'A' overhangs are ligated followed by universal adaptors which can be used to amplify the successfully ligated fragments. Libraries are then ready for whole genome sequencing if required, or additional target enrichment can be carried out to select exomes or other parts of the genome that may be of interest. Target regions are isolated by hybridization with streptavidin coated RNA probes, followed by subsequent binding to biotinylated magnetic beads. Non-target regions can then be simply washed away and the bound targets amplified by low cycle PCR and finally sequenced.

1.11.2 Upstream bioinformatic analysis

After FASTQ reads have been retrieved from the sequencing platform, initial analysis steps are typically performed by bioinformatic pipelines. These steps generally include; quality checks, removal of duplicates and low quality reads, trimming of reads, alignment/mapping, SNP and small indel calling, the removal of common variants and the identification of structural variants. The choice of QC thresholds, alignment and variant calling software can strongly influence the final output. Common aligners used include BWA, Novoalign, Bowtie2, and SOAP (293). The more commonly used variant callers include SAMtools m-pileup, and GATK (294).

Functional understanding about the algorithms behind aligners and variant callers encompass bioinformatic and mathematical disciplines beyond the scope of this thesis so discussion will be limited to some of the more superficial analytical choices that can be applied to variant calling. The most obvious filter that can be applied relates to the depth of the sequence, i.e. the number of reads covering a given base and the proportion of wild-type to mutant calls represented throughout the depth. For example, to call standard germline variants it is typical to require a minimum depth of 20 reads and ~50% of those reads should show the mutant allele to call a heterozygote. Analysis of impure tumour samples can be much more complicated due to the presence of tumour sub-clones as well as healthy host tissue contaminating the sample (**Figure 1.18**). As a result, variants are often called if only represented by a few reads and require validation by follow-up experiments. It stands to reason that the greater the depth is then the higher confidence one has in accurately calling potential variants.

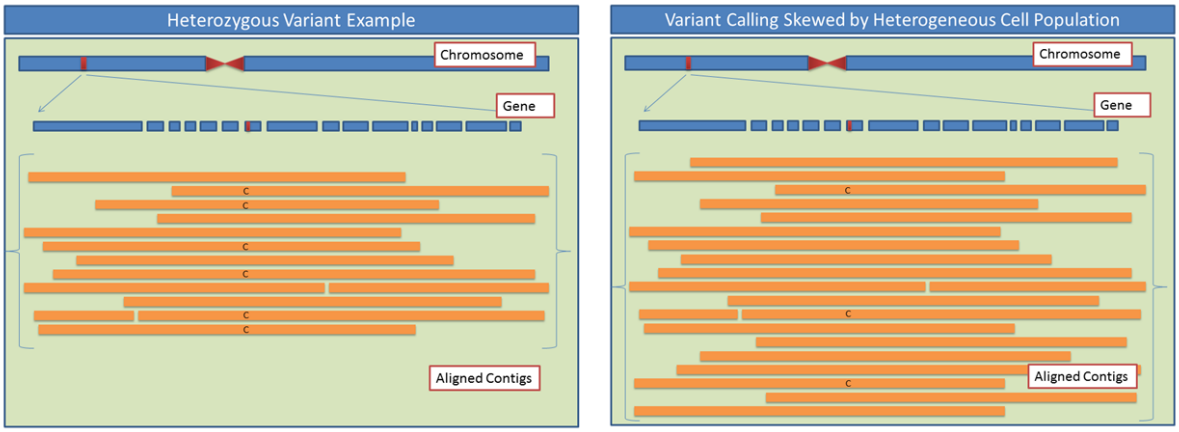


Figure 1.20 Calling variants in NGS data.
Standard heterozygous variants, such as many of those detected in healthy germline DNA, are relatively easy to call because contigs from each allele make up approximately 50% of reads. In contrast, variants from impure tumour DNA or tumour sub-clones can be much more dilute within a sample leading to lower confidence when calling variants. Sequencing target genes in more depth can partially address this, as can enriching tumour populations wherever possible.

Variants can also be filtered based on strand bias, during pair-end sequencing the reads are sequenced from both the 5' and 3' ends. If a variant is called and most of the non-wild type

(WT) reads appear to only occur on one strand (often referred to as the forward or reverse strands) then it is often reasonable to exclude this variant as the bias has likely arisen due to a sequencing artefact or a mistake arising from low complexity sequence.

Calling structural variants is considerably more developed and reliable for whole genome data as libraries are less biased by the target enrichment process and coverage tends to be more uniform. WES or targeted capture studies tend to be limited to calling losses or gains of genes or exons rather than mapping the exact location of structural variants. Usually additional software is required to call structural variants and these are technically more challenging and often ambiguous to detect but strategies fit into several main approaches. These include; read-count methods - whereby detection is based on the number of reads at a given location relative to control samples, read-pair based methods - where differences between reads and sections of the reference genome are visible during alignment, split-read approaches which tether part of a read to the reference genome and then allow the other part of the read to the align independently. De-novo assembly based methods can also in principle be used to detect structural variation independently of a reference genome (295).

1.11.3 Bioinformatic approaches and the interpretation bottleneck

Once a list of variants has been generated the challenge remains to identify true driver variants from the myriad of passenger events. A human exome contains ~12,500 non silent coding variants, excluding structural variation (296). It has been estimated that between 2 and 8 driver events are required to cause a malignancy (297). If we accept these estimates then mere superficial observation of most NGS cancer datasets would suggest that most of the variants in a dataset are passenger events. However, there are some exceptions with relatively low numbers of mutations such as cancers that occur in children (298). Several approaches have been developed in order to identify driver events and are discussed in this section.

1.11.3.1 Frequency based methods

As the name implies, this method simply looks for the genes that contain the highest number of novel mutations in a dataset. This is a relatively crude approach that can easily be confounded by several factors. The most obvious confounding factor is that the larger a gene is then just by probability the gene is more likely to harbour a mutation in any given sample. However, this method requires little bioinformatics expertise and can be useful for quickly identifying recurrent variants or candidate genes that had been previously hypothesized as being drivers.

1.11.3.2 Enrichment methods

Beyond basic frequency approaches, more sophisticated methods have been developed that model the expected mutation rate and look for enrichment. Genes that appear to be mutated above the background rate are more likely to be drivers because a selection pressure must be driving a rate of mutation above what would be expected. These methods rely on empirical Bayesian statistics and various methodologies have been developed which rank genes according to the number and nature of mutations (299).

Synonymous and non-synonymous mutations can be calculated to occur at a certain ratio at each given codon. This can be used as a type of enrichment method to calculate if a selection pressure is occurring by observing whether or not non-synonymous mutations are occurring at a rate higher than expected by chance. Similar approaches have been widely used to identify selection during studies of evolution (299-301). The selection pressure that drives mutagenesis in cancer can be thought of as a similar process to Darwinian selection as previously mentioned (68, 69).

The frequency of non-synonymous coding mutations has been shown to be approximately half of that found in non-coding regions (302, 303). This can be used as a guide to identify selection pressure and therefore aid in identifying driver genes (299). To make full use of this approach

it is important to correct for (or exclude) regions where loss of heterozygosity (LOH) has occurred (299).

1.11.3.3 Mutation clustering as a predictor

Cancer driver genes are often classified as oncogenes or tumour suppressor genes, the former can also be thought of as dominant and the latter recessive. Each are known to be positioned non-randomly throughout their respective genes with oncogene mutations often clustering to specific regions and tumour suppressor genes more likely to be truncating or spread throughout the gene (297) (**Figure 1.19**), although it is noteworthy that some cancer driver genes can be oncogenes or tumour suppressors depending on the tumour type (297), for example NOTCH1 can function as a tumour suppressor gene in solid tumours but in lymphomas tends to be an oncogene (304-306). Several bioinformatic software packages have been developed in order to predict cancer driver genes based on these mutation localization differences including SomInaClust (307) and OncodriveCLUST (308).

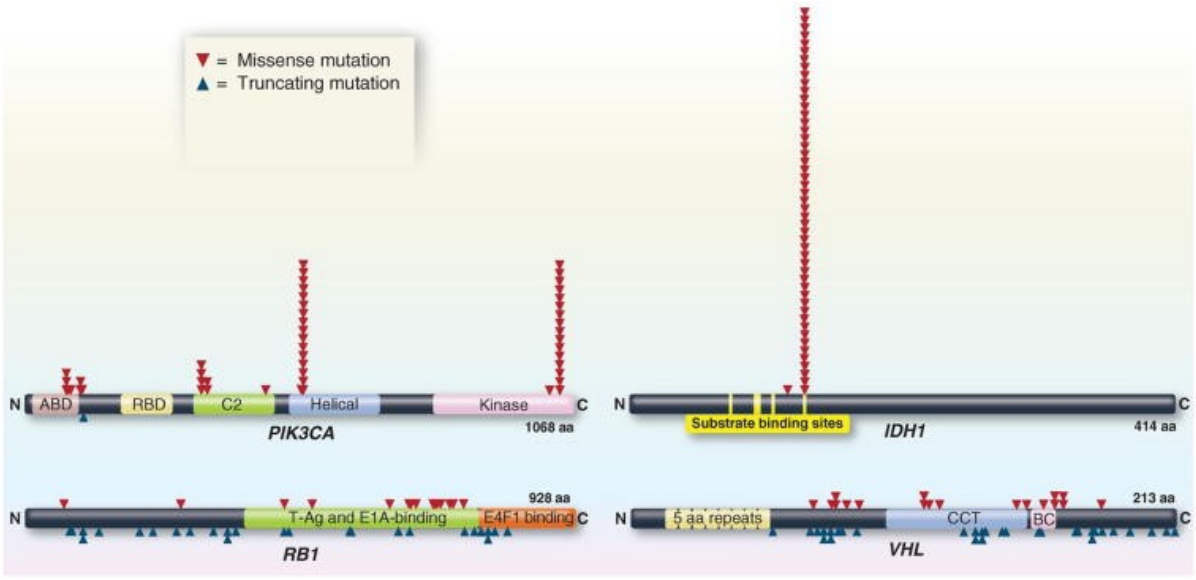


Figure 1.21 Distribution of mutations in two oncogenes (*PIK3CA* and *IDH1*) and two tumor suppressor genes (RB1 and VHL).
The distribution of missense mutations (red arrowheads) and truncating mutations (blue arrowheads) in representative oncogenes and tumor suppressor genes are shown. The data were collected from genome-wide studies annotated in the COSMIC database (release version 61). For *PIK3CA* and *IDH1*, mutations obtained from the COSMIC database were randomized by the Excel RAND function, and the first 50 are shown. For *RB1* and *VHL*, all mutations recorded in COSMIC are plotted. aa,= amino acids. (**Vogelstein et al 2013**).

Ratio and spatiometric methods have been applied to the Catalogue of Somatic Mutations in Cancer (COSMIC) database (309, 310) in a recent large study in order to identify and classify mutation driver genes (297). The COSMIC database is a large composite bank of mutations reported in numerous cancer types. At the time of the original analysis COSMIC contained 404,863 mutations across 18,306 different genes (297). In this study the so called 20/20 rule was proposed whereby genes are classified as oncogenes if over 20% of mutations occur at recurrent sites and are missense. Alternatively genes are classified as tumour suppressor genes if over 20% of mutations are inactivating (297). This study reported 125 mutation based driver genes of which 71 are tumour suppressors and 54 are oncogenes (297). Whilst it is clear that there are likely many more cancer driver genes mutated less frequently such as those reported in the network of cancer genes (291), this serves are a useful approach that can be improved upon with more recent builds of COSMIC and on individual datasets.

1.11.3.4 Composite methods

The enrichment methods of identifying driver genes discussed earlier in this section rely on the assumption that as sample sizes grow, so will the power to detect driver genes that occur less frequently. However, there have been growing signs that the picture is more complicated than this as many genes which are implausible as candidates, such as olfactory receptors, show enrichment.

False positives largely stem from mutational heterogeneity across different samples within a cohort as most enrichment methods use assumptions about the expected rates of different types of mutations (298). Individual samples within a cohort which carry more passenger mutations will contribute to the enrichment of a pool of false positives if the heterogeneity between the samples is not accounted for (311). This is particularly important for cancers which show a higher range of mutations across samples such as melanoma and lung cancers where the sample specific mutation rate can vary by several orders of magnitude (0.1-100/Mb)(298)(**Figure 1.20**). Recent NGS studies have reported mutation ranges approaching the higher end for CTCL (49-51), indicating that inter-sample heterogeneity may partially confound simple methods for calling drivers in this cancer type.

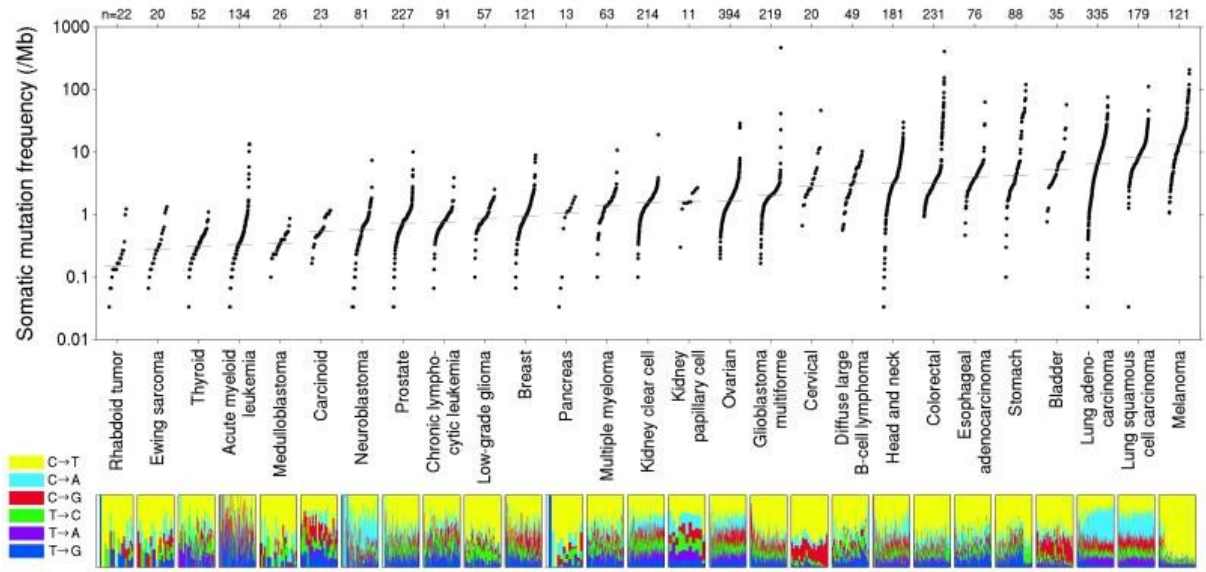


Figure 1.22 Inter and intra-heterogeneity of mutational patterns across different cancer types.
Somatic mutation frequencies observed in exomes from 3,083 tumour-normal pairs. Each dot corresponds to a tumour-normal pair, with vertical position indicating the total frequency of somatic mutations in the exome. Tumour types are ordered by their median somatic mutation frequency, with the lowest frequencies (left) found in haematological and paediatric tumours, and the highest (right) in tumours induced by carcinogens such as tobacco smoke and UV light. Mutation frequencies vary more than 1000-fold between lowest and highest mutation rates across cancer and also within several tumor types. The lower panel shows the relative proportions of the six different possible base-pair substitutions, as indicated in the legend on the left. (**Lawrence et al 2013**).

Another type of mutational heterogeneity contributing to false positives is the occurrence of mutational hotspots across the genome which can vary by up to five fold (298). Several factors seem to account for most of this heterogeneity the first being expression level. Highly expressed genes tend to be less prone to mutations most likely as a result of transcription coupled nucleotide excision repair (312). The other main contributor appears to be the time during S-phase that replication occurs for a particular gene. Late replicating genes are known to be more prone to germline mutations (313-315), possibly as a result of depletion of free dNTPs (313). The increased rate of mutation in late replicating genes has been shown to extend to cancer cells (298).

Any method of identifying driver genes that doesn't account for each of these areas of heterogeneity will inevitably be flawed at least to some extent and will call some false positives as well as overlook some genes that only affect a modest number of tumours (311), the so called hills (297). One method that attempts to take account of differing mutations rates within a cohort as well as expression and timing of replication is MutSigCV (298). Whilst it is not possible to fully implement without coupled transcriptomic data as well as mutational data MutSigCV accounts for many sources of variation that standard enrichment techniques do not and could therefore act as a very effective screening tool.

Another method that uses multiple factors to predict driver mutations in a dataset is MuSic (mutational significance in cancer) (316). This algorithm applies a number of statistical methods as well as making predictions based on the known pathway interactions that potential drivers are involved in, furthermore clinical data associated with the individual tumours can also be used by the algorithm (316).

1.11.4 Challenges for current methods of identifying cancer drivers

One key challenge for current methods are that they show little in terms of overlap with each other (317). Whilst it would be unreasonable to expect any algorithm to flawlessly predict all drivers in a dataset, it is of some concern that predictions vary by a significant margin with less than 25% overlap in some composite approaches applied to the same datasets (317, 318). Although selection of the features that are used to identify drivers will play a large role in how each algorithm will perform (317) the research community has yet to reach consensus about what these features should be, furthermore it is likely to vary between different malignancies.

Another significant challenge relates to the changing roles of driver genes throughout the development of a cancer. Some drivers may be passengers in the early development of a cancer but play a significant role in later stages of the disease (or vice-versa), or when the environment changes, for instance in response to treatment (317). No strategies currently exist to address this, however sequencing efforts tend to concentrate on late stage diseases. If

efforts were made to collect samples and sequence tumours from all disease stages in the same patients then identifying drivers that operate at different disease stages may be feasible. However, it's acknowledged that it may be challenging to isolate enriched tumour tissue from early stage tumours in many cancers.

The cumulative effect of multiple small-effect or rare drivers is a challenge that is very difficult to measure or investigate algorithmically with current datasets. These genes/variants are the so called hills that are infrequently affected in small proportions of tumours or have a very small effect on tumour progression (298). By far the most effective way to identify them is to increase the size of the datasets which will increase statistical power whilst also considering the caveats identified in the previous section. The next generation of machine learning algorithms, particularly deep learning neural networks, may also prove to be an excellent tool for identifying small effect or rare driver genes.

1.11.5 Functional approaches

Hypothesis driven, functional investigations are the traditional approach in experimental genetics. Whilst bioinformatic approaches are a powerful tool for short listing candidate genes for consideration as cancer driver genes, they are essentially screening tools and candidates need to be followed up by functional studies. This section will describe in general the main experimental approaches that can be taken to discern gene function.

1.11.5.1 Correlative studies

These are the first step towards functional characterisation; they are theoretically not dissimilar to bioinformatic screening. The main difference is that a range of methods can be used and analyzing single genes is often considerably more cost effective than analyzing large numbers of genes. The main disadvantage is that a good rational or hypothesis is required to investigate specific genes. Approaches such as QPCR can be used to analyze both gene copy number changes and changes in expression. Sanger sequencing can be used to investigate SNVs and/or small indels and methylation specific PCR (MSP) can be used to investigate

differentially methylated nucleotides in comparison with control samples. If significant differences are seen between cases and controls then further experiments can be designed around modelling the differences.

1.11.5.2 Loss and gain of function experiments

Often the most straightforward experiment that can be undertaken in order to investigate a gene of interest is a loss of function experiment, i.e. knock out the gene entirely or aim to reduce its expression. The opposite of this is a gain of function experiment, whereby the aim is to add extra copies of a gene or otherwise increase its expression. After the desired effect is conferred on the gene of interest then observations can be made about the effects (of either losing or gaining the gene) on the model system being used. In cancer studies loss of function is likely to be more appropriate for studies of tumour suppressor genes whereas gain of function is likely to be more appropriate for studying oncogenes. In the case of some SNVs the change may well be ambiguous prior to investigation but emerging tools such as the CRISPR-Cas9 system (319) are well placed to design studies investigating such variants. Both gain and loss of function experiments can be conducted in stable cell lines, model organisms such as mice or in wild type or tumour derived cells, although the latter two models are often technically challenging.

1.12 Hypothesis

The highly heterogeneous genomic landscape of CTCL contains a mix of driver and passenger events, recurrent genomic events are more likely to contain driver genes. In addition, driver events may occur in different genes in different tumours but affect related gene regulatory networks.

1.13 Aims of this thesis

1.13.1 Investigating the role of *MTAP* in CTCL: A candidate gene based approach

An initial aim of this thesis is to analyse the chromosome 9p21.3 focal deletion in CTCL patients. The *CDKN2A* and *CDKN2B* loci are well established cancer driver genes whose function is under selective pressure in numerous malignancies. Therefore, it is important to establish if the *MTAP* gene, which is often co-deleted, is also under a selection pressure in CTCL as this would support a role for *MTAP* as a driver of CTCL. The main focus of this section is to determine if *MTAP* is lost independently of *CDKN2A* in CTCL. Independent loss would indicate a selective pressure against *MTAP*, supporting the hypothesis that *MTAP* is a driver gene in CTCL. Other mechanisms of inactivation will also be investigated such as the presence of SNPs using the NGS data generated from the other section of the project. Furthermore, the presence or absence of promoter hypermethylation will also be investigated. In addition to genetic and epigenetic characterization of the *MTAP* status in CTCL patient samples, a preliminary study of the functional effect of *MTAP* loss will be undertaken. Specifically, characterization of the effect of *MTAP* loss on protein-arginine methylation by the PRMT family of genes will be sought.

1.13.2 Identification of driver events in CTCL: A high throughput global approach

A major aim of this thesis aims to examine CTCL using a genomic approach to screen for driver events in CTCL. NGS is a maturing technology which enables the detection of a large number and variety of genetic perturbations, including SNVs and CNVs, simultaneously in one large experiment. As SS is the most severe form of the disease with the worst prognosis, this section

of the project will focus on SS. CTCLs are an incredibly heterogeneous malignancy and whilst similarities are likely to be detected at the level of single genes, it is important to consider that genetic perturbations may be more common at the level of gene families, gene regulatory networks and signalling pathways. Therefore, particular attention will be focused on identifying these higher level genomic patterns and features to identify common themes that appear in the disease. Validation will be sought for biological processes that are commonly perturbed and are likely to be key contributors to the disease such as mechanisms that effect genome stability. Work in this area has important implications from the perspective of translational medicine as it will aim to identify targets for possible therapeutic intervention and as a long term objective; foundations will be laid for personalized therapeutic strategies.

2 Materials and Methods

2.1 Collections and Processing of Samples

All patients fulfilled the WHO-EORTC diagnostic criteria for SS (4). A dominant clonal T-cell receptor (TCR) gene rearrangement was detected in all samples using PCR based methods (320). Furthermore, in matched samples (skin and PBMC) from SS patients an identical T-cell clone was demonstrated. Patient samples were obtained from the nationally approved CTCL research tissue bank (National Research Ethics Committee: 07/H10712/111+5); healthy control samples were obtained with the approval of the Guy's and St Thomas' Hospital Research Ethics Committee (EC01/301). Written and informed consent was obtained from all patients/volunteers.

Whole exome sequencing discovery samples: DNA was extracted from CD4+ enriched peripheral blood mononuclear cells (PBMCs) using RosetteSep (Stemcell Technologies, Cambridge, UK) and matched primary fibroblasts from skin explants obtained from 10 untreated SS patients at diagnosis. Targeted capture samples: DNA was extracted from PBMCs of 101 SS and 32 healthy control samples.

MTAP targeted study: Samples from 280 CTCL patients were used in this study (319 total samples), 234 of these patients were analyzed by Q-PCR (273 samples), 77 patients were analyzed by NGS, of which 31/77 were also analysed by QPCR. This cohort of 77 were a subset of the 101 prevalence screen patients that underwent targeted capture during the genomic screening study and were used alongside the 32 healthy controls (HCs) also used in this part of the study to call mutations as well as CNVs across the 9p21 region with the exome depth package (See section 2.28 for further detail of this). DNA samples from PBMCs of 84 HCs were obtained from the Human Random Control DNA Panel 4 (ECACC, Sailsbury, UK) for use as controls in the QPCR gene copy number analysis, a further cohort of.14 age matched HCs were recruited at King's College London to analyse gene expression of *MTAP, CDKN2A* and control

genes by QPCR in CD4+ T-cells for comparison with SS tumour CD4+ T-cells. The CTCL cell lines, HuT78 and SeAx were gifts from Dr S John (King's College London) and Professor M Vermeer (Leiden University Medical Centre) respectively. Jurkat and MyLa cells were obtained from the ECACC (Sailsbury, UK).

2.2 Isolation of PBMCs from Whole Blood

Whole blood samples were made up to 20ml with PBS and then layered on top of lymphoprep before centrifuging to separate PMBCs from eurythrocytes. PBMCs were then harvested, washed with PBS, RCLB (red cell lysis buffer) and RPMI medium (Roswell Park Memorial Institute) before re-suspending in 'RNAlater' (Qiagen, Manchester, UK).

2.3 Enrichment of CD4+ T-Cells from PBMCs

Whole blood samples were made up to 20ml with PBS (phosphate buffered saline) and incubated at room temperature with Rossettesep (1ml per 20ml blood) for 20 minutes. Samples were then layered on top of lymphoprep and CD4+ T-cells were isolated using the same method as for PBMCs.

2.4 Isolation of Fibroblasts

Fibroblasts were obtained from primary skin explants. Cultures were maintained in RPMI containing 10% fetal calf serum and 1% penicillin/streptomycin (Invitrogen, Paisley, UK)

2.5 Mycoplasma testing

Mycoplasma testing was carried out monthly. Cell media was collected from a minimum of 24 hours after passage. Testing was carried out using the 'Mycoplasma detection kit-quick test' (B39032, Biotool) according to manufacturer's instructions.

2.6 DNA Extraction

Frozen or fresh cell pellets were treated with nuclei lysis buffer (Promega, Southampton) and digested with proteinase K (Qiagen, Manchester, UK) for up to 24 hours before removing

proteins with protein precipitation solution (Promega, Southampton, UK). DNA was then precipitated in isopropanol and washed in 70% ethanol before re-suspension in MilliQ or low TE. Alternatively, where high yields were not necessary the DNeasy blood and tissue kit (Qiagen, Manchester, UK) was used according to manufacturer's instructions.

2.7 RNA Extraction

RNA was extracted from cell pellets using the RNeasy mini kit (Qiagen, Machester, UK) as instructed by the manufacturer's guidelines.

2.8 Synthesis of cDNA

High capacity cDNA reverse transcription kit (Applied Biosystems, Paisley, UK) was used to convert RNA into cDNA from 1000ng of RNA per 20ul reaction according to instructions supplied by the manufacturer.

2.9 PCR

PCR primers were designed using the Primer3 online software and assessed in-silico using the NCBI online primer design tools before optimization. PCR was performed in 25µl reactions containing 0.1-0.1-0.2mM of each primer, 0.06-0.08mM of each dNTP and 1.0-2.0 Units of AmpliTaq Gold DNA polymerase (Applied Biosystems, Paisley, UK). The concentration of MgCl2 was 1.5-2.5mM. Primer-pair annealing temperatures were optimised in parallel starting at the theoretical melting point and decreasing in 1°C decrements to 5°C below the theoretical melting point. PCR cycles were fixed at 35 cycles for end point experiments and ranged down to 25 cycles for semi-quantative experiments.

PCR products were electrophoresed through a 2% agarose gel supplemented with 0.005% ethidium bromide for visualization of PCR products. A negative control containing water was run with each master mix preparation. Appropriate cell line DNA was used as a positive control where possible.

2.10 QPCR

Real-time quantitative PCR (qPCR) reactions were carried out on an ABI 7900 (Applied Biosystems, Paisley, UK). Reactions continued for 40 cycles with an initial denaturation at 95°C for 10 minutes with a 60°C, 60 second annealing and extension phase. Probes/primer sets (**Table 1**) were combined with TaqMan gene expression mastermix (Applied Biosystems, Paisley, UK). Reaction volumes were 25ul and samples were analyzed in triplicate for both probes and with 3 wells of water per probe as a negative control. The ddCT method was used for the analysis with standard deviation of healthy samples used to calculate normal distribution. The ddCT method quantifies the amount of nucleic acid by measuring the number of cycles it takes for the reaction to progress to the log phase. This number of cycles (CT-value) can then be compared to housekeeping genes to quantify relative abundance of target nucleic acid sequences and then compared to healthy controls or other experimental samples to measure relative expression or copy number. In order to determine the normal range of ddCT for each experiment, the mean and standard deviation of the ddCT of the HC group was calculated and 3 standard deviations was set as the boundary for normal copy number or gene expression. Samples with ddCT lower than 3 standards deviations of the HC mean were said to have reduced gene copy number or gene expression for the respective experiments. ANOVA and Tukeys HSD was used to assess differences between groups used in QPCR for gene CN assays.

Target	Catalogue Number	Manufacturer	Experiment
CDKN2A	Hs00923894_m1	Applied Biosystems	Expression
PPIA	Hs99999904_m1	Applied Biosystems	Expression
MTAP	2.10.1.1 Hs00559618_m1	Applied Biosystems	Expression
RnaseP	4404631	Applied Biosystems	Copy number
TERT	4401663	Applied Biosystems	Copy number
MTAP P1	Custom (M'Soka et al 2000)	Applied Biosystems	Copy number
CDKN2A P1	Custom (M'Soka et al 2000)	Applied Biosystems	Copy number
MTAP P2	Hs06918311_cn	Applied Biosystems	Copy number
CDKN2A P2	Hs03721302_cn	Applied Biosystems	Copy number

Table 9 Taqman QPCR probes.

2.11 Whole Exome Capture and Sequencing

Whole exome capture was performed by 'in-solution hybridization' using the Agilent Sure Select XT target enrichment system for Illumina paired-end sequencing Libraries according to manufacturer's instructions (Agillent, Sureselect-XT). Subsequent massively parallel sequencing (MPS) was performed on an Illumina HiSeq 2000.

DNA was extracted from enriched CD4+ tumour cells of 10 SS patients. Libraries were prepared from patient CD4+ tumour cells and from corresponding patient fibroblasts individually prior to exome capture. Patient matched fibroblasts were chosen as a control source of DNA for subsequent MPS analysis in order to distinguish between germline mutations and somatic mutations detected in the CD4+T-cell population. Fibroblasts and CD4+T-cell reads were distinguished post-sequencing by the use of separate index-tags ligated during the post capture PCR amplification stage of the library preparation. A total of 10 PCR cycles was selected as the optimum cycle number used to increase the post capture library DNA concentration whilst maintaining fragment diversity in the library.

2.12 Targeted Capture and Sequencing

Capture probes were designed based on the exons of genes found to carry variants fitting three criteria: non-synonymous, novel, somatic tumour mutations from patients who had undergone whole exome sequencing. Novel was defined as a mutation not found in the 1000 genomes, dbSNP or in house exome databases' and non-synonymous also included all exonic stop, frameshift, INDELs and splice site variants. Somatic tumour mutations are variants that were found only in the tumour DNA and not in the fibroblast control DNA. Probes were also designed to cover recurrent regions of LOH (loss of heterozygousity). In addition, probes were designed to capture the JAK, STAT and TET family of genes as well as to cover all coding exons, all CpG islands and all multi-transcription factor ChIP sites from the encode project within 5kb of coding exons within the chromosomal region of 9p21. See Appendix 'Final targeted capture BED list.txt' for complete list of target regions.

Targeted capture library preparation was performed according to manufacturer's instructions (Agilent, Sureselect-XT) as for whole exome library preparation (**Figure 23**).Sequencing was performed on an Illumina Highseq-2000 in the BRC genomics core facility.

2.13 Preparation of Whole Cell Lysates

Fresh RIPA buffer was prepared prior to lysing (sc24948A, Santa-Cruz Biotech). Cells were initially washed in PBS to remove excess freezing medium or growth medium and cell pellets produced by centrifuging at 1500rpm for 5 minutes. The supernatant was then decanted and 1ml of complete RIPA buffer was added per $2x10^7$ cells. Cells were then disrupted by pippeting and incubated at 4˚C for 20-30 minutes. Supernatent was then decanted and protein concentration quantified by BCA (Thermoscientific, UK) and diluted to appropriate concentrations and denaturing at 95˚C with 4x Laemmli buffer (161-0747, Biorad, UK) for 5 minutes.

2.14 Immunoblotting

Whole cell lysates containing 15-40ug of total protein were loaded onto 5-15% PAGE gels (BioRad, UK). Following electrophoresis proteins were transferred to nitrocellulose or PVDF. Blocking was carried out in with 3-5% BSA or 3-5% fat-free milk in Tris-buffered saline and 1% tween (TBST) depending on antibody for 1 hour minimum. Antibodies were diluted to optimized concentrations in 3-5% BSA or 3-5% milk in TBST and incubated overnight under agitation at 4°C (primary antibody) (**Table 2**). Membranes were washed in TBST from between 20 minutes and 2 hours to remove residual primary antibody and blocked again before being incubated with secondary antibody for 2-4 hours.

Antibody	Manufacturer	Incubation time	Blocking time	Other Conditions	Dilution
Goat polyclonal HRP-conjugated anti-rabbit (secondary)	Abcam-ab97051	2-4 hours	1hour	5% milk/TBST	1:3000
Goat Anti-mouse (secondary)	Sigma - A2304	1 hour	1 hour	5% milk/TBST	1:8000
Dimethyle-Arginine asymetric	Active Motif SA - 39231	12-16 hours at 4˚C	2-4 hours	5% milk/TBST	1:1000
Mouse polyclonal MTAP	Abnova – H00004507-M01	12-16 hours at 4˚C	1 hour	5% milk/TBST	1:500
Anti-beta Actin	Abcam – ab8227	1 hour		5% milk/TBST	1:2000

Table 10 List of antibodies and optimized conditions for use.

2.15 Stripping Membranes for Re-probing

Stripping buffer was prepared using 20ml of 10% SDS, 12.5ml of 0.5M Tris HCL, 67.5ml of milli-Q water and 0.8ml of β-mercaptoethanol. Membranes were incubated in a fume hood at 50˚C for up to 45 minutes with stripping buffer. Membranes were then rinsed for 1-2 hours with running water to remove residual β-mercaptoethanol, then washed in TBST before blocking for re-probing.

2.16 Detection of Proteins

Detection was performed using the ECL plus Western blotting system (RPN2232, GE Healthcare UK Ltd) combined with exposure to X-ray film (12705325, FUJI, Fisher Scientific Ltd, UK).

2.17 Thawing Frozen Cells

Cells were thawed at 37°C for 2-3 minutes before re-suspending in fresh medium. Cells were then briefly centrifuged again before further re-suspending in fresh medium to remove excess DMSO. Growth media was replaced again after 24 hours and cells were assessed for growth and viability.

2.18 SeAx culture and passage

SeAx cells were maintained in RPMI medium (21875-034, Gibco) supplemented with 10% FBS (51800-500, Biosera, UK) and 1% penicillin/streptomycin (15140-122, Gibco, UK). Media was also supplemented with 25U/ml IL-2 (PHC0026, Gibco, UK). Cells were passaged at a confluency of 1 x10^6 cells/ml to 1/5 or 1/10.

2.19 HEK293 culture and passage

HEK293 cells were maintained in DMEM medium (61965-026, Gibco, UK) supplemented with 10% FBS (51800-500, Biosera, UK) and 1% penicillin/streptomycin (15140-122, Gibco, UK). Cells were washed in sterile PBS and then passaged at a confluency of 90-95% to 1/5 to 1/20 depending on requirements.

2.20 Freezing Cells

Log phase cells were prepared for long term storage by suspending up to 10 x 10^6 cells in 1ml of freshly prepared freezing medium (50% DMSO, 50% FBS) (sc358801, Santa-Cruz Biotechnology). Cryovials were stored at -80°C for 24 hours in a controlled rate freezing device before being transferred to liquid nitrogen.

2.21 Nucleofection

Transfection of SeAx cells was performed using the Amaxa human T-cell nucleofector kit (VPA-1002, Lonza, UK) optimized for unstimulated human T-cells according to manufacturer's instructions. Between 5-10x10^6 cells were transfected in each cuvette. One siRNA (**table 3**) or plasmid was transfected per cuvette unless specifically stated. 2µg of a GFP expressing control plasmid (pmaxGFP, supplied with the kit) was transfected in 1 cuvette as a positive control to ensure transfection took place successfully with each batch of nucleofection. Between 30-300nM of siRNA was transfected per 100µl cell suspension.

Transfected siRNA 5'->3'	Transfected siRNA 3'->5'	Target	Manufacturer and Catalogue Number
GCGAUAUUGUCAUUAUUGAtt	UCAAUAAUGACAAUAUCGCcg	MTAP	Ambion-s9021
CGAGAGAGGUUCUUAUAGAtt	UCUAUAAGAACCUCUCUCGtt	MTAP	Ambion-s9022
GCUGGAAUUUGUUACGCAAtt	UUGCGUAACAAAUUCCAGCct	MTAP	Ambion-s9023
Not provided	Not provided	Scrambled control	Ambion - AM4611

Table 11 List of siRNAs.

2.22 Transfection with Lipofectamine 2000

Transfection of Seax cells using Lipofectamine2000 (11668-030, Invitrogen) was attempted using 1-4x10^5 cells per ml in 24 well plates as suggested by the manufacturers but was ineffective at all concentrations of siRNA (5-300pmol, final concentration when combined with cells) and lipofectamine (0.5-3.55µl/50ml) attempted. Serum free and antibiotic free media were also attempted but no difference was observed. Transfection of pGFPmax plasmid was attempted (100-1500ng/ml final concentration) but observation using florescence microscope (Zeiss Axio Observer Z1) showed very few cells had undergone transfection (20-30 cells per 4x10^5 cells in any given well).

HEK293 cells were seeded in 24 well plates at 50% confluency at least 24 hours prior to transfection in antibiotic free media. Lipofectamine2000 was mixed with Optimem (61965-026, Gibco) in a range of concentrations between 0.5-2.5µl/50ml as suggested by manufacturers (Thermofisher Scientific). Media was mixed with siRNA (**table 3**) at concentrations between 5-50pmol (final concentration when combined with cells). Media/Lipofectamine and media/siRNA were incubated for 5 minutes before being gently combined and further incubated for 20 minutes to allow formation of complexes containing nucleic acid and Lipofectamine2000. After incubation, the solution was added dropwise to the

cell cultures and incubated for at least 24 hours before extraction of protein lysates and nucleic acids.

2.23 Micro-dissection

Sections were cut 9µm thick and mounted on glass slides. After drying for 24 hours the sections were dewaxed in Xylene and rehydrated. Sections were then stained with CD3+ and H/E. Target tissues were then removed and dissolved in lysis buffer. DNA was extracted using the QIAamp DNA micro kit (Qiagen, Manchester, UK) whilst the remaining areas of the sections were dehydrated and cover slipped.

2.24 Variant calling

Paired-end sequencing library preparation was performed according to manufacturer's instructions (Agilent, Sureselect-XT) and sequenced on an Illumina Hi-Seq2000 with reads aligned to Hg19 using Novoalign v2.07.11 and post-alignment processing performed by picard tools 1.114.

For WES, Varscan2 Somatic (v2.3.5) was used to separate tumour variants from patient matched fibroblasts. ANNOVAR was used for variant annotation (version available in January 2013) (321) (WES annotation was carried out with ANNOVAR in April 2013). Somatic and non-synonymous variants were selected based on exclusion of variants in dbSNP, the 1000 genomes project, exome variant server, in-house exome database and genes reported to be error prone in NGS analysis due to sequence repeats and high GC content (322).

For targeted capture Varscan2 and ANNOVAR were also used but the threshold on the minimum allele frequency for calling tumour variants was calibrated to account for the heterogeneity of tumour samples derived from PBMCs (see selection of minimum allele frequency). Mpileup2cns was used for SNV and INDEL identification with >=20x depth, >=15 phred score, >=6% minimum variant frequency and read frequency <=90% in either direction. Variants from 32 non-matched healthy controls were used to identify tumour specific variants

and exclude sequencing artefacts. Variants selected from WES and targeted capture data were validated by Sanger sequencing on original tumour and additional skin, lymph node and tumour derived cDNA samples from the same patients.

2.25 Selection of minimum allele frequency

A low allele frequency of 6% was chosen to apply for the targeted capture (TC) samples based on the empirical data obtained from 10 WES samples that were also included in TC. The list of high quality variants from 10 WES samples was used to identify the variants in each of the TC samples with WES. In the comparison of variants between WES and TC samples, variants that were identified from both WES and TC were considered to be True Positives (TP), while those identified in TC and not in WES as False Positives (FP). Variants identified in WES and not in TC were considered False Negatives (FN) and variants not present in both WES and targeted capture were considered True Negatives (TN). Information was obtained from TC samples at various thresholds starting from 1 to 60% and calculated Matthew's Correlation Coefficient score (MCC: 0 to 1) at each threshold. A threshold with highest MCC score was considered as optimal threshold for calling variants in the targeted capture data. Individual samples resulted in a range of thresholds (5% to 30%) optimal for each sample, although combined analysis of all the samples has resulted in a maximum MCC score of 0.271 at a threshold of 6%.

2.26 Mutational pattern analysis.

Several types of mutational pattern analysis were conducted, custom in-house Python scripts were used to assess SNV base changes and the proportions of different types of variant (Synonymous, Non-synonymous etc). Analysis of SNV base change patterns and mutation context (motif) analysis upstream and downstream at 3bp was assessed using the 'R' Bioconducter package 'Somatic Signatures' (time of analysis: September 2015).

2.27 Identification of SNV drivers

Several parallel criteria were used to identify genes affected by SNVs. These included MutSigCV version 1.3 (298), the 20/20 rule as described by Vogelstein et al (297) and simple frequency filtering of over 5% after removing genes previously identified as problematic (322). We also compared the list of candidate driver genes to those present in the network of cancer genes version 4.0 (291).

2.28 Gene copy number analysis

Tumour specific CNVs were identified through integrative analysis of discovery and targeted capture data generated using exome/targeted sequencing and SNP array technologies. Data from WES was analyzed by Excavator v2.2 in matched pairs. HumanOmni5Exome arrays were analyzed using OncoSNP v1.4. Raw data (BAF and LRR) required for OncoSNP was extracted using Illumina Genome Studio software v1.9.4. Data from WES and SNP array were combined for final analysis (n=16). Remaining prevalence samples (n=91) were analysed with ExomeDepth software (323) using the targeted capture data and 32 healthy controls. This analysis was restricted to targeted capture genes (n=549) but allowed deeper resolution.

2.29 Pathway analysis

To investigate significant perturbations at the pathway level, gene set enrichment analysis was performed on WES and TC SNVs using the MSigDB repositories (date of analysis was January 2015). MSigDB is a collection of a curated gene sets including gene sets from pathways in KEGG, REACTOME, BIOCARTA and PID. Initially the list of genes used for targeted capture was used to look for enriched pathways with a Fisher's exact test P.value <= 0.05.

Next, enriched pathways in each tumour sample were selected using all the Non-Synomymous variants from that sample. To perform this task, the list of genes from each tumour sample was searched against MSigDB for a significant overlap. All the samples with a significant enrichment of genes in a pathway were counted and the pathway(s) with highest number of tumour samples

were reported. In this analysis, the background and the size of the pathways in MSigDB used to calculate the Fisher's exact test P-value were corrected to represent the priory selected and small size of the genes in the targeted capture. A background of 598 genes were used and resulted in the decrease of number of genes in a pathway which corresponds to the genes seen in 598 genes background.

Pathway-level perturbations were quantified using two inter-related metrics. One metric, "fraction of pathway genes mutated", captures the proportion of pathway genes involved in non-synonymous SNV's or indels across all patients. The second metric, "pathway perturbation frequency score", captures how often each pathway is perturbed as a proportion of all samples (both uncorrected and corrected for pathway size), assuming perturbation occurs if at least one of the pathway's genes is mutated. As this analysis was predominantly qualitative, i.e.a hit was counted if any gene in the pathway was found to harbour a mutation, we therefore did not consider it appropriate to calculate a false discovery rate.

3 Independent loss of methylthioadenosine phosphorylase *(MTAP)* in Primary Cutaneous T-cell Lymphoma

3.1 Candidate contributions to chapter 3.

The chapter is a publication that several members of a research team contributed work towards. My personal contributions were as follows: initial draft and subsequent revisions to the manuscript with oversight of my supervisors, collating the data and producing final figures 1-3, producing the data for figures 1b, 2a, 2d, 3, producing all tables, design of targeted capture experiment used in the NGS work, library preparations for the NGS work, downstream analysis of mutation data, producing and calculating normalized depth plots from ExomeDepth R-objects.

3.2 Additional statistical tests for chapter 3.

Correlation between QPCR CN assay probes 1 and 2 (Figure 1e and 1f). Probes were compared for correlation between P1 and P2 using Pearson's correlation. Disease samples assayed using both *MTAP* probes were highly correlated (PCC=0.79, P<0.01). Disease samples assayed using both *CDKN2A* probes were also correlated, but with weaker correlation (PCC=0.37, P<0.01). As discussed in the published work, the reason for the discrepancy is likely to be due to the different locations of the probes corresponding to sites of focal deletion in some cases, and the other probe being located outside of the boundary of the focal deletion. Another source of variation could be differences in the efficiency of PCR reactions for each different probe.

***MTAP* expression correlated with loss of *MTAP* CN in SS samples.** To clarify this statement; 'concordance' is a more accurate description as these data were assigned into qualitative categories based on being within or below a threshold of 3 standard deviations of the healthy control mean. All samples where a loss of *MTAP* was detected also showed reduced expression that was <= -3 standard deviations of the mean expression in the healthy control group. Interestingly, 9 patient samples that did not demonstrate copy number changes in *MTAP* also showed expression lower than 3 standard deviations of the healthy control mean. There were no healthy samples that showed expression lower than 3 standard deviations. In total, 20

tumours were analysed for expression. To further test for concordance between *MTAP* CN loss and reduced *MTAP* expression, samples with reduced *MTAP* CN (*MTAP* CN less than 3 standard deviations below HC mean) were assigned a categorical variable 'CN loss'. Samples with *MTAP* expression below 3 standard deviations below the HC expression mean were assigned a categorical value 'reduced EXP'. These categorical variables were compared using a Chi-squared test which resulted in a $P<0.01$.

Independent Loss of Methylthioadenosine Phosphorylase *(MTAP)* in Primary Cutaneous T-Cell Lymphoma

Wesley J. Woollard[1], Nithyha P. Kalaivani[1], Christine L. Jones[1], Catherine Roper[1], Lam Tung[1], Jae Jin Lee[1], Bjorn R. Thomas[1], Isabella Tosi[1], Silvia Ferreira[2], Carl Z. Beyers[2], Robert C.T. McKenzie[1], Rosie M. Butler[1], Anna Lorenc[3], Sean J. Whittaker[1] and Tracey J. Mitchell[1]

Methylthioadenosine phosphorylase (*MTAP*) and the tumor suppressor genes *CDKN2A-CDKN2B* are frequently deleted in malignancies. The specific role of *MTAP* in cutaneous T-cell lymphoma subgroups, mycosis fungoides (MF) and Sézary syndrome (SS), is unknown. In 213 skin samples from patients with MF/SS, MTAP copy number loss (34%) was more frequent than CDKN2A (12%) in all cutaneous T-cell lymphoma stages using quantitative reverse transcription PCR. Importantly, in early stage MF, MTAP loss occurred independently of CDKN2A loss in 37% of samples. In peripheral blood mononuclear cells from patients with SS, codeletion with CDKN2A occurred in 18% of samples but loss of MTAP alone was uncommon. In $CD4^+$ cells from SS, reduced MTAP mRNA expression correlated with *MTAP* copy number loss ($P < 0.01$) but reduced *MTAP* expression was also detected in the absence of copy number loss. Deep sequencing of *MTAP/CDKN2A-CDKN2B* loci in 77 peripheral blood mononuclear cell DNA samples from patients with SS did not show any nonsynonymous mutations, but read-depth analysis suggested focal deletions consistent with MTAP and CDKN2A copy number loss detected with quantitative reverse transcription PCR. In a cutaneous T-cell lymphoma cell line, promoter hypermethylation was shown to downregulate MTAP expression and may represent a mechanism of MTAP inactivation. In conclusion, our findings suggest that there may be selection in early stages of MF for MTAP deletion within the cutaneous tumor microenvironment.

Journal of Investigative Dermatology (2016) **136,** 1238–1246; doi:10.1016/j.jid.2016.01.028

INTRODUCTION

Primary cutaneous T-cell lymphomas (CTCLs) are a heterogeneous group of extranodal lymphomas that arise from skin homing T cells. Mycosis fungoides (MF) accounts for more than 50% of CTCL cases (Agar et al., 2010; Willemze et al., 2005) and responds well to skin-directed therapy. In contrast, a related CTCL subtype Sézary syndrome (SS) is an aggressive leukemic variant with no treatment options proven to prolong overall survival.

In both MF and SS, no disease specific translocations have been identified although genomic instability, illustrated by recurrent "hotspots" of chromosomal structural and numerical abnormalities, is a consistent feature. Studies from our own group and others using multiple genomic platforms have identified chromosome 9p21 as a region of frequent deletion (Laharanne et al., 2010; Mao et al., 2003; Salgado et al., 2010; van Doorn et al., 2009). The 9p21 locus includes *CDKN2A/CDKN2B*, which encode the well-characterized tumor suppressor genes $p14^{ARF}$, $p16^{INK4A}$, and $p15^{INK4B}$.

The *MTAP* gene lies 110 kb telomeric to *CDKN2A* and encodes the metabolic enzyme 5′-methylthioadenosine phosphorylase. MTAP metabolizes 5′-methylthioadenosine, a byproduct of polyamine biosynthesis, to almost undetectable levels (Kamatani et al., 1981). This occurs in all normal cells as part of the adenine nucleotide salvage and methionine biosynthesis pathways and is considered a housekeeping function. Loss of *MTAP* expression has been described in a range of solid malignancies, including melanoma (Behrmann et al., 2003; Garcia-Castellano et al., 2002; Hellerbrand et al., 2006; Watanabe et al., 2009), and hematological malignancies (Batova et al., 1996). *MTAP*-deficient tumor cells exhibit specific alterations in polyamine metabolism induced by elevated 5′-methylthioadenosine levels that could drive a selective growth advantage (Kamatani and Carson, 1980). Melanoma cells with elevated 5′-methylthioadenosine levels show increased expression of matrix metalloproteinase and growth factor genes (Stevens et al., 2009).

Emerging evidence supports a role for *MTAP* as a tumor suppressor gene independent from *CDKN2A/CDKN2B*

[1]*St John's Institute of Dermatology, Division of Genetics and Molecular Medicine, Faculty of Life Sciences & Medicine, King's College London, London, UK;* [2]*Viapath, Skin Tumour Unit, St John's Institute of Dermatology, Guy's Hospital, London, UK; and* [3]*Transformational Bioinformatics, NIHR Research Biomedical Research Center at Guy's and St Thomas' Hospital Foundation Trust and Kings College London, London, UK*

Correspondence: Tracey J. Mitchell, St John's Institute of Dermatology, Division of Genetics and Molecular Medicine, Faculty of Life Sciences & Medicine, King's College London, 9th Floor, Tower Wing, Guy's Hospital, London SE1 9RT, UK. E-mail: tracey.mitchell@kcl.ac.uk

Abbreviations: CN, copy number; CTCL, cutaneous T-cell lymphoma; HC, healthy control; MF, mycosis fungoides; MTAP, methylthioadenosine phosphorylase; NGS, next generation sequencing; PBMC, peripheral blood mononuclear cells; Q-PCR, quantitative reverse transcriptase in real time; SS, Sézary syndrome

Received 5 January 2015; revised 7 December 2015; accepted 20 January 2016; accepted manuscript published online 9 February 2016; corrected proof published online 30 March 2016

(Christopher et al., 2002). In mice, homozygous deletion of *MTAP* is embryonically lethal, whereas mice heterozygous for germ-line *MTAP* knockout are viable but die prematurely of a mature T-cell lymphoma (Kadariya et al., 2009). A meta-analysis of genetic association studies in cutaneous melanoma highlights an association between single nucleotide polymorphisms in *MTAP* and cutaneous melanoma risk (Antonopoulou et al., 2014). In addition, copy number (CN) alterations in eight genes including *MTAP* differentiate superficial spreading melanoma from nodular melanoma (Rose et al., 2011).

In common with other genes (*CDKN2A/CDKN2B*) on chromosome 9p21, epigenetic silencing of *MTAP* via promoter methylation is also recognized as an alternative mechanism for *MTAP* inactivation in human malignancy. Studies of melanoma (Behrmann et al., 2003) and a range of other malignancies (Berasain et al., 2004; Leal et al., 2007; Watanabe et al., 2009) have all demonstrated loss of *MTAP* expression is associated with inactivation of *MTAP* by aberrant promoter hypermethylation.

Our aim was to investigate genetic aberrations of *MTAP* in CTCL. Analysis of tumor DNA derived from MF and SS skin and blood samples demonstrated that loss of *MTAP* CN is common and frequently occurs independently of *CDKN2A* deletion. Furthermore, in SS $CD4^+$-enriched tumor cells derived from PBMCs, reduced *MTAP* mRNA expression is associated with *MTAP* CN loss. Studies of CTCL cell lines identified aberrant *MTAP* promoter hypermethylation as an alternative mechanism of *MTAP* silencing and that methylation reversal can restore *MTAP* expression. In contrast, we demonstrate that mutational events in *MTAP, CDKN2A,* and *CDKN2B* are rare in SS and are unlikely to contribute significantly to the disease. In conclusion, our findings suggest that loss of *MTAP* is an early event in MF and further investigation is now justified to establish whether MTAP loss may have a role in disease initiation.

RESULTS

MTAP and *CDKN2A* gene copy number loss in mycosis fungoides and Sézary syndrome

It is well established that chromosome 9p21 is a region of frequent deletion in advanced stage CTCL (Laharanne et al., 2010; Scarisbrick et al., 2002; van Doorn et al., 2009). We sought to establish if *MTAP* loss occurs simultaneously with *CDKN2A/CDKN2B* as part of a larger region of deletion, or if loss of these genes occurs independently.

Previously published quantitative reverse transcriptase in real time (Q-PCR) CN assays for *MTAP* and *CDKN2A* were performed with the two control genes *PFKL* and *BCMA* (M'Soka et al., 2000). CN status for *MTAP* and *CDKN2A* was determined in the CTCL cell lines Hut78, MyLa, SeAx, and the leukemic T-cell line Jurkat. Homozygous loss of *MTAP* and *CDKN2A* was detected in both HuT78 and Jurkat, but not in MyLa or SeAx cells (data not shown).

MTAP and *CDKN2A* gene CN was then assessed in 84 healthy controls (HCs) and 273 tumor DNA samples from CTCL subgroups comprising MF IB-IIA, 101 samples; MF IIB-IV, 73 samples; and SS III-IV, 99 samples (including 78 samples from matched skin lesions and PBMCs from 39 patients, and 21 unmatched samples from PBMCs).

Loss of *MTAP* occurred in all disease stages and was statistically significant compared with HC samples ($P < 0.05$). Whereas loss of *CDKN2A* compared with HCs was only significant in advanced stages of patients with MF (IIB-IV) and SS ($P < 0.05$). Gain of *MTAP* or *CDKN2A* was an infrequent event across all of the patient samples. The distribution of CN variation for *MTAP* and *CDKN2A* is shown in Figure 1a and b, respectively, within patient and HC groups.

CN gains and losses were then compared between skin and blood and disease stage (Table 1). Strikingly, in the skin compartment, loss of *MTAP* was more frequent than *CDKN2A* alone or combined loss in all disease stages, this was particularly prominent in early stage MF (IB-IIA). In contrast, loss of *CDKN2A* was more frequent than *MTAP* in the tumor cells derived from the blood in advanced stage disease. Furthermore, data from matched skin and blood samples in the SS (III-IV) group (Table 2) also showed loss of *MTAP* CN in the skin but independent loss of *MTAP* in blood was infrequent.

To confirm the CN analysis, additional probes were used (probe set 2) for both genes on a subset of 94 DNA samples comprising 32 matched skin and PBMCs (64 samples) from stage SS III-IV; 14, MF IB-IIA skin; and 16, MF IIB-IV skin (Figure 1e and f). Overall concordance between CN probes was 75.0% and 58.3% for *MTAP* and *CDKN2A*, respectively, with some losses detected with each probe that were not detected on the other. Differences between probe sets are consistent with a pattern of focal deletions across the region rather than large structural variations.

To further investigate the pattern of CN variations across *MTAP* and *CDKN2A,* we applied next generation sequencing (NGS) depth analysis using previously published methods (Plagnol et al., 2012) to 31 SS PBMC-derived DNA samples that had also been analyzed by Q-PCR and 32 HCs (Figure 2). Of the 31 samples analyzed, the ExomeDepth package predicted CN variation with 90.3% concordance to at least one probe. Normalized exome depth was plotted as a heat map to highlight deviations in the depth of coverage consistent with focal deletions (Figure 2).

Reduced *MTAP* mRNA expression in SS $CD4^+$-enriched tumor cell populations correlates with copy number loss

Next, we sought to determine if *MTAP* and *CDKN2A* CN loss was associated with reduced gene expression at the RNA level using cDNA derived from peripheral blood $CD4^+$-enriched tumor cell populations from 20 patients with SS.

A significant difference in *MTAP* mRNA expression was observed between SS and HC samples ($P < 0.01$). Furthermore, reduced *MTAP* expression correlated with loss of *MTAP* CN in all SS samples (Figure 3). In contrast, no decrease in *CDKN2A* expression was detected despite eight cases showing *CDKN2A* CN loss.

Aberrant *MTAP* promoter hypermethylation as an alternative mechanism of *MTAP* gene inactivation

We then examined promoter methylation as a possible regulatory mechanism of *MTAP* expression using the CTCL cell lines MyLa and SeAx. SeAx cells had normal *MTAP* CN and abundant expression of *MTAP*, whereas although MyLa cells had normal *MTAP* CN, *MTAP* expression was not detectable (Figure 4). The methylation status of the entire

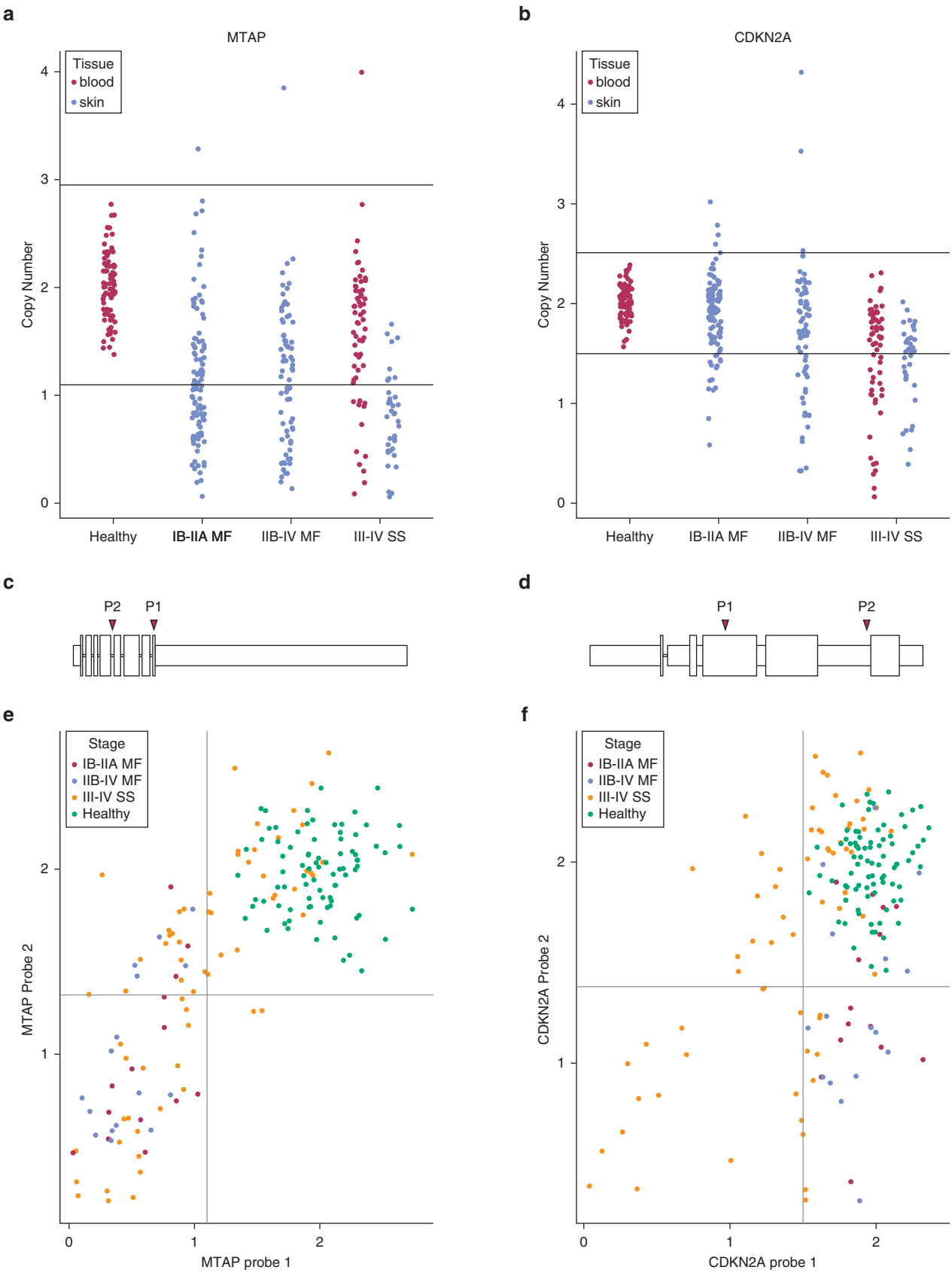


Figure 1. Gene CN analysis for *MTAP* and *CDKN2A* in MF and SS. (**a**) *MTAP* and (**b**) *CDKN2A* CN was determined by Q-PCR in 84 HCs, 52 MF IB-IIA, 73 MF IIB-IV, and 101 SS III-IV samples (64 PBMC samples and 39 skin samples). Horizontal lines represent 3 standard deviations from the HC mean. Schematic representations of (**c**) *MTAP* and (**d**) *CDKN2A* showing positions of Q-PCR probe sets 1 and 2. Gene CN of (**e**) *MTAP* and (**f**) *CDKN2A* determined by probe sets 1 (*y*-axis) and 2 (*x*-axis). Red lines represent 3 standard deviations from the HC mean for the corresponding probe set. Probe set 2 shows 75.0% and 58.3% concordance with probe set 1 for *MTAP* and *CDKN2A*, respectively. CN, copy number; MF, mycosis fungoides; *MTAP*, methylthioadenosine phosphorylase; SS, Sézary syndrome.

Table 1. Frequencies of copy number variation in *MTAP* and *CDKN2A* in CTCL

	MTAP loss only	*MTAP* and *CDKN2A* loss	*CDKN2A* loss only	*MTAP* gain only	*MTAP* and *CDKN2A* gain	*CDKN2A* gain only	No loss or gain of copy number
IB-IIA MF (101)	38S	13S	2S	0	1S	3S	44S
IIB-IV MF (73)	19S	13S	10S	0	1S	2S	28S
III-IV SS (39s/39b + 21b)	0B/16S	11B/15S	13B/1S	0	0	1B/0S	35B/7S
Total	73S	(52) 11B/41S	(26) 13B/13S	0	2S	(6) 1B/5S	(114) 35B/79S

Abbreviations: B/S, blood/skin; MF, mycosis fungoides; *MTAP*, methylthioadenosine phosphorylase; PBMCs, peripheral blood mononuclear cells.
Tumor DNA was analyzed in 273 samples comprising MF lesional skin samples from 101 stage IB-IIA and 73 IIB-IV; and in SS (III-IV), 78 matched skin and PBMCs (39 patients) and an additional 21 from PBMCs. The number of samples of each disease stage harboring each type of copy number variation is shown.

CpG island, containing three discrete regions of CpG sites, termed CG1, CG2, and CG3, was analyzed by direct sequencing of bisulfite-converted DNA. In SeAx cells complete methylation of all CpG sites in CG1 and CG2 was observed, whereas the CG3 region was completely unmethylated (Supplementary Figure S1a online). In contrast, MyLa cells were found to be completely methylated at all CpG sites within the three regions (Supplementary Figure S1b). This suggests that differential methylation of CG3 may be regulating *MTAP* expression.

Promoter hypermethylation and loss of *MTAP* expression is reversible

To determine if the observed differential methylation of the CG3 region regulates *MTAP* expression, methylation reversal experiments using the demethylating agent 5′-aza-2′-deoxycytidine were performed. MyLa and SeAx cells were treated with 5′-aza-2′-deoxycytidine for up to 72 hours. After treatment, *MTAP* expression levels were measured by reverse transcriptase-PCR (Figure 4a) and intracellular flow cytometry (Figure 4b) to assay cDNA and protein levels, respectively. These studies demonstrated a strong induction of *MTAP* expression in 5′aza-treated MyLa cells and consistent *MTAP* expression before and after treatment in SeAx. This supports a role for methylation of the CG3 region regulating *MTAP* expression.

Polymorphisms of *MTAP, CDKN2A,* and *CDKN2B* are rare in Sézary syndrome

Finally, we sought to determine if rare variants in the genes *MTAP, CDKN2A,* and *CDKN2B* are present in SS tumor DNA samples. We defined rare variants as those not observed in the "1,000 genomes" data or dbSNP (build 138). Only one rare nonsynonymous variant was discovered in a coding region and only very few variants were discovered in noncoding regions, none of which were predicted to affect transcription (Supplementary Table S1 online). Although functional analysis has not been performed, these results suggest that somatic mutational events in *MTAP, CDKN2A,* and *CDKN2B* rarely contribute to the pathogenesis of SS.

DISCUSSION

This study of a large cohort of 280 patients with MF/SS (319 samples) demonstrates that *MTAP* gene deletion is common and can occur independently of *CDKN2A* gene deletion in early stage MF and skin-derived tumor cells from patients with SS. Loss of *MTAP* CN also correlates with reduced *MTAP* expression, but significantly reduced *MTAP* expression can also occur in the absence of *MTAP* loss. In CTCL-derived cell lines, aberrant promoter hypermethylation of one of the *MTAP* CpG islands (CG3) leads to loss of *MTAP* expression. Epigenetic mechanisms such as promoter hypermethylation may therefore be an alternative mechanism of *MTAP* gene inactivation in CTCL as demonstrated in other malignancies (Behrmann et al., 2003; Hellerbrand et al., 2006; Ishii et al., 2005; Leal et al., 2007). In contrast to loss of CN, somatic mutations in *CDKN2A/B* or *MTAP* are rare in SS.

Loss of *MTAP* has been reported in association with deletion of *CDKN2A/CDKN2B* in advanced disease stages of CTCL (Laharanne et al., 2010; Salgado et al., 2010; Scarisbrick et al., 2002; van Doorn et al., 2009) and has been shown to be an independent prognostic marker for survival. However, recent studies using mouse models have begun to define a role for *MTAP* as an independent tumor supressor gene (Kadariya et al., 2009, 2013). *MTAP*-heterozygous mice show reduced *MTAP* expression levels compared with wild-type mice and die prematurely because of the development of mature T-cell lymphomas. Furthermore, reduced *MTAP* expression has been reported in several rare lymphomas (Bertino et al., 2012). These data suggest that *MTAP* loss alone may be a contributing factor in the development of T-cell lymphomas. Our data support a role for *MTAP* loss in CTCL pathogenesis as we show significant

Table 2. Concordance of gene copy number between matched skin and PBMCs in SS

	Copy number loss in skin but not blood	Normal copy number in skin and blood	Copy number loss in blood but not skin	Copy number loss in blood and skin
MTAP loss	21	8	0	10
CDKN2A loss	2	20	3	14

Abbreviations: *MTAP*, methylthioadenosine phosphorylase; PBMCs, peripheral blood mononuclear cells; SS, Sézary syndrome.
Gene copy number of *MTAP* and *CDKN2A* was analyzed in both blood and skin tissue in 39 patients with SS stage III-IV (78 samples) and concordance compared between the two tissues.

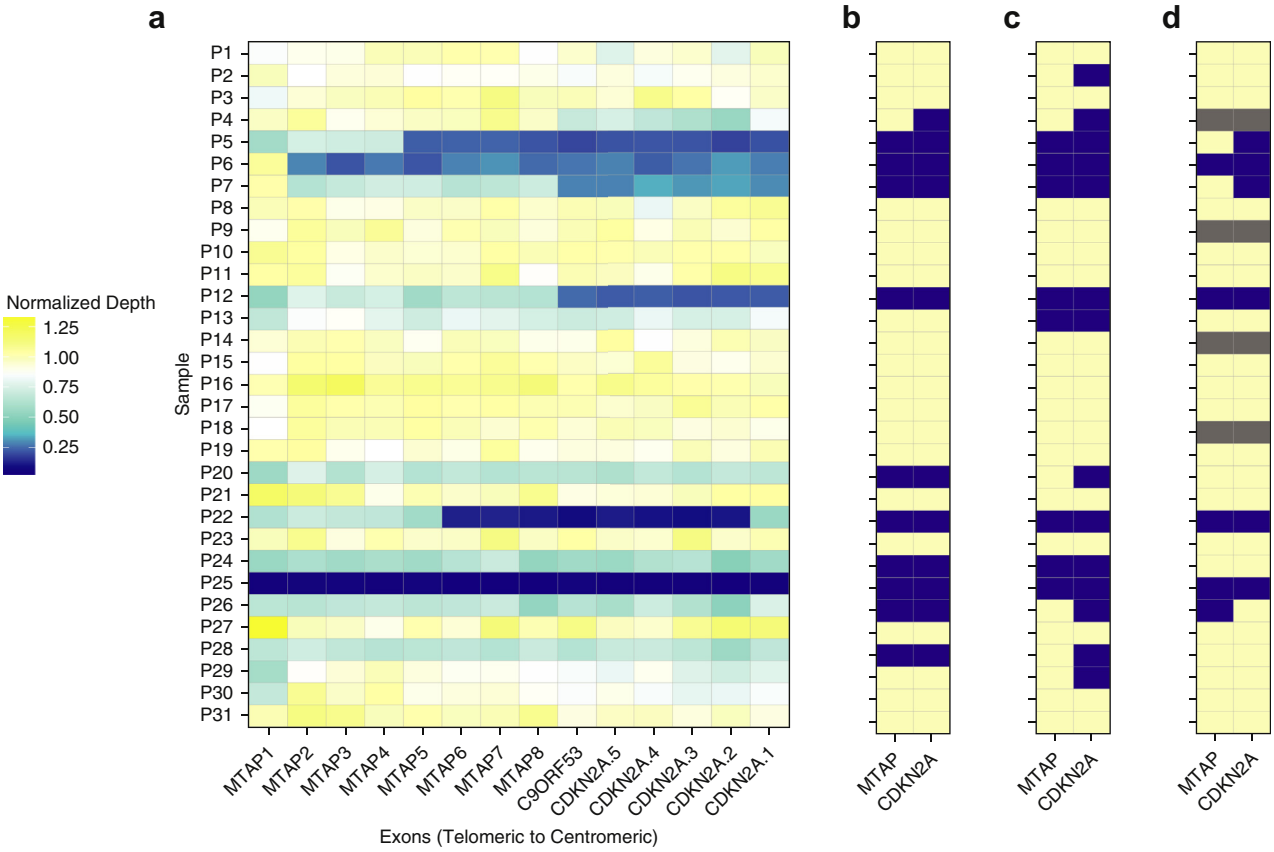


Figure 2. Comparison of methods to call gene copy number in SS. From left to right: (**a**) Normalized exome depth across each exon plotted as a heat map, see the Materials and Methods section for the description of normalized depth calculation. Exons are arranged telomeric to centromeric along the *x*-axis and SS samples that were assayed by Q-PCR are plotted on the *y*-axis. (**b**) Normal or loss (ivory/blue, respectively) as called by the "ExomeDepth" package. (**c**) Normal or loss as assayed by Q-PCR probe set 1, *MTAP*-P1 = exon 8, *CDKN2A*-P2 = exon 3. (**d**) Normal or loss as assayed by Q-PCR probe set 2, gray represents samples not assayed by probe set 2 *MTAP*-P2 = intron 4, *CDKN2A*-P1 = intron 1. *MTAP,* methylthioadenosine phosphorylase; Q-PCR, quantitative reverse transcriptase in real time; SS, Sézary syndrome.

MTAP CN loss in all disease stages. Importantly, in early stage of MF, *MTAP* CN loss is independent of *CDKN2A* suggesting *MTAP* loss precedes *CDKN2A* loss, which in contrast only becomes significant in advanced stages. Furthermore, our results suggest that *MTAP* CN loss is more frequent in lesional skin than in the peripheral blood of patients with SS. Discordant loss of *MTAP* and *CDKN2A* in different tumor compartments and stages likely reflects tumor heterogeneity and subclonal evolution. Although overall numbers studied are small, multilineage progression of genetically unstable tumor subclones has been reported in CTCL (Rubben et al., 2004) and may contribute to these. However, the apparent selection pressure for subclones with loss of *MTAP* in skin and retention of *MTAP* in the peripheral blood of patients with SS is intriguing and remains to be explained but could reflect a role for the skin microenvironment that is not present in the peripheral blood. Interestingly, a recent study suggested that MTAP may act as a tumor suppressor independently of its enzymatic function (Tang et al., 2014); however, the precise mechanism remains unclear.

Analysis using additional Q-PCR probes demonstrated 75% and 58.3% concordance for *MTAP* and CDKN2A, respectively, between probe sets 1 and 2. The probe sets were located in different gene regions that are likely to account for these differences and these data are consistent with a genomic landscape in SS and MF, which is shaped by a series of focal deletions. ExomeDepth analysis from the NGS data of 9p21 was 90.3% concordant with at least one probe set that confirms this interpretation. In addition, recent NGS studies examining late stage CTCL (Choi et al., 2015; da Silva Almeida et al., 2015; Kiel et al., 2015; Ungewickell et al., 2015; Wang et al., 2015) indicate that structural alterations and CN variations frequently involve 9p21. Losses commonly affect 9p21 and include *MTAP* and *CDKN2A* (Choi et al., 2015; da Silva Almeida et al., 2015; Kiel et al., 2015; McGirt et al., 2015; Wang et al., 2015) in many cases. However, the size and position of the deletions show considerable variation. Our data relating to advanced stage disease are consistent with these reports but currently no NGS study has addressed early stage CTCL.

Reduced *MTAP* expression has previously been reported in SS and tumor stage MF and shown to correlate with gene CN loss and reduced survival (van Doorn et al., 2009). Our findings confirm SS tumor samples with *MTAP* CN loss show a significant reduction in *MTAP* mRNA expression; however, reduced *MTAP* expression was also observed in the absence of a detectable loss in *MTAP* CN. This indicates that other mechanisms are disrupting expression such as mutations, polymorphisms, or epigenetic events. We did not detect any rare mutations that are likely to affect *MTAP* transcripts in a

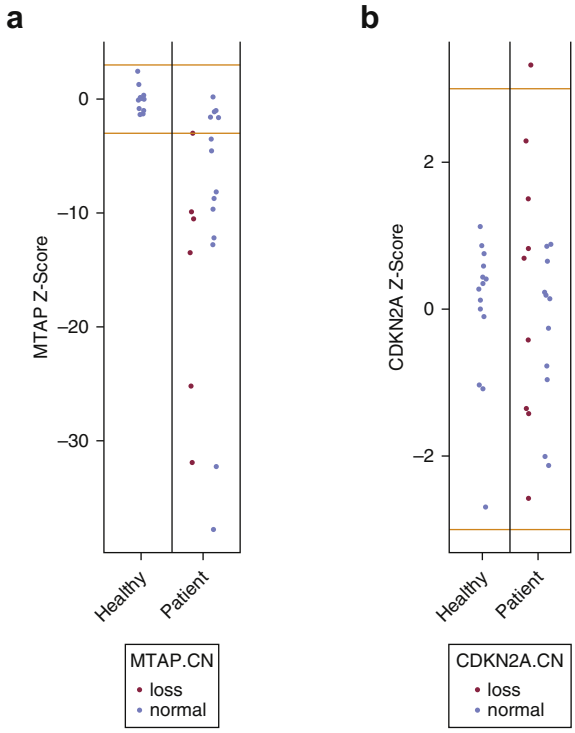


Figure 3. Expression of *CDKN2A* (p16) and *MTAP* in CD4$^+$ lymphocytes of patients with SS. Relative expression of (**a**) *MTAP* and (**b**) *CDKN2A* (p16) compared with cyclophilin (*PPIA*) in CD4$^+$ lymphocytes of HCs and 20 patients. The *y*-axis shows standard deviations from the HC mean ΔΔCT. Points represent *Z*-scores and red lines represent 3 standard deviations of the HC ΔΔCT. Samples from tumors where copy number changes have been identified are highlighted with red dots. Samples with no identified copy number change are represented by blue dots. HC, healthy control; *MTAP*, methylthioadenosine phosphorylase; SS, Sézary syndrome.

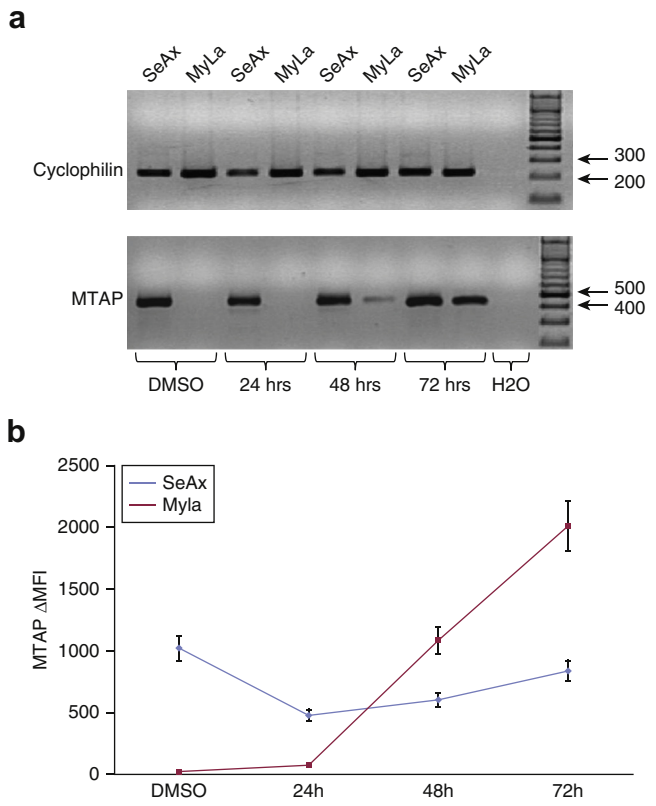


Figure 4. Functional evidence linking *MTAP* promoter hypermethylation and expression. (**a**) Demethylation restores MTAP RNA expression in MyLa cells. Lanes (1) SeAx and (2) MyLa cells treated with DMSO alone. Lanes 3–8 cell lines treated with 5 μM 5′-azadeoxycytidine (3) SeAx (4) MyLa, 24 hours; (5) SeAx (6) MyLa, 48 hours; (7) SeAx (8) MyLa, 72 hours. Lane (9) water control. RNA was analyzed by RT-PCR for (A) cyclophilin expression as an endogenous control and (B) MTAP expression. (**b**) Intracellular flow cytometry confirms demethylation of MyLa cells induces MTAP protein expression. SeAx and MyLa cells treated as described and MTAP protein expression determined by flow cytometry, 10,000 events were acquired per sample. Data expressed as delta mean fluorescent intensity (ΔMFI) calculated using Flowjo software. *MTAP*, methylthioadenosine phosphorylase; RT-PCR, reverse transcriptase-PCR.

large cohort of 77 patients with SS, which is consistent with other reports (Choi et al., 2015; da Silva Almeida et al., 2015; Kiel et al., 2015; Ungewickell et al., 2015; Wang et al., 2015). In contrast, epigenetic silencing of *MTAP* via promoter hypermethylation is known to produce *MTAP* inactivation in other tumors (Behrmann et al., 2003; Hellerbrand et al., 2006; Ishii et al., 2005; Leal et al., 2007). This is supported by the demonstration that in the CTCL cell line MyLa, *MTAP* expression is induced by treatment with the demethylating agent 5′-aza-2′-deoxycytidine. Reversible promoter hypermethylation of *MTAP* has also been shown in primary tissue and cell lines of other malignancies including histiocytic lymphoma (Ishii et al., 2005), hepatocellular carcinoma (Hellerbrand et al., 2006), gastric adenocarcinoma (Leal et al., 2007), and malignant melanoma (Behrmann et al., 2003).

Studies demonstrating reduced expression of *MTAP* in melanoma support a causal link with tumor cell proliferation, tumor progression, and metastasis (Behrmann et al., 2003; Bertino et al., 2012; Meyer et al., 2010; Wild et al., 2006). Specifically, reexpression of *MTAP* led to a significant reduction in the invasive potential of melanoma cells (Behrmann et al., 2003). These studies in melanoma have also identified *MTAP* expression as an independent positive prognostic marker for remission free and overall survival. Loss of *MTAP* expression has also been proposed as a predictive marker of interferon therapy resistance (Wild et al., 2006).

In contrast to *MTAP*, loss of *CDKN2A* CN did not result in loss of expression. This is consistent with previous reports of heterozygous *CDKN2A* deletion producing mixed expression patterns (Zhao et al., 2012). In addition, two reports suggest that mRNA expression does not correlate with protein expression in *CDKN2A* (Brownhill et al., 2007; Takasaki et al., 2003). Together these reports present a complex pattern of regulatory control for CDKN2A.

In conclusion, our findings indicate *MTAP* CN loss occurs more frequently than *CDKN2A* in early stage MF and skin-derived tumor cells from SS. Interestingly, our data indicate that reduced *MTAP* expression is a prominent feature of SS and can be attributed to more than one mechanism. This adds weight to the mounting evidence that *MTAP* acts as an independent tumor suppressor gene. The functional impact of *MTAP* gene inactivation in CTCL has yet to be defined, but may yield potential novel therapeutic possibilities.

MATERIALS AND METHODS

Patient samples and cell lines

Samples from 280 patients with CTCL were used in this study (319 total samples); 234 of these patients were analyzed by Q-PCR (273 samples), 77 patients by NGS, and 31 patients by both methods. Patients were classified according to the TNM staging system. Of the 280 patients, 174 were MF including 101 stage 1B-IIA and 73 stage IIB-IV and 106 stage III-IV SS. All patients fulfilled the WHO-EORTC diagnostic criteria for either MF or SS (Willemze et al., 2005). A dominant clonal TCR gene rearrangement was detected in all samples using PCR-based methods (Brüggemann et al., 2007). Furthermore, in matched samples (skin and PBMC) from patients with SS, an identical T-cell clone was demonstrated. Supplementary Tables S2 and S3 online detail the breakdown of samples used in different study components.

DNA samples from lesional skin and PBMCs were obtained from the nationally approved cutaneous lymphoma research tissue bank (National Research Ethics Committee: 07/H10712/111+5). Written and informed consent was obtained from all patients/volunteers in this study. DNA samples from PBMCs of 84 HCs were obtained from the Human Random Control DNA Panel 4 (ECACC, Sailsbury, UK) for use as controls.

The CTCL cell lines, HuT78 and SeAx were gifts from Dr S. John (King's College London) and Professor M. Vermeer (Leiden University Medical Centre), respectively. Jurkat and MyLa cells were obtained from the ECACC. All cell lines were maintained in the RPMI medium containing 10% fetal calf serum and 1% penicillin/streptomycin (Invitrogen, Paisley, UK). SeAx cells were supplemented with IL-2 (25 U/ml).

Assessment of gene copy number by Q-PCR

Probe set 1. Q-PCR was performed on 273 patient samples and 84 HCs. Reactions were performed in triplicate on the ABI Prism 7000 Sequence Detection System (Applied Biosystems, Cheshire, UK) using previously published primer/probe sequences for *CDKN2A* (M'Soka et al., 2000), *MTAP* (M'Soka et al., 2000), *BCMA* (De Preter et al., 2002), and *PFKL* (De Preter et al., 2002). *BCMA* and *PFKL* control genes were used to control the amount of input DNA and selected to detect diploid genes in chromosomal regions not previously reported to be affected by structural or numerical changes in CTCL. Mean ΔΔCT was calculated from the ΔΔCT of both control genes and patient ΔΔCT was compared with the mean and standard deviation of the HC group to infer gene CN. Further statistical analyses are described below.

Probe set 2. Q-PCR was performed on 94 of 273 samples with additional probes to validate findings from the probe set 1. Probes for *MTAP* (Hs06918311_cn), *CDKN2A* (Hs03721302_cn), and the control gene TERT (probes purchased from Life Technologies, Paisley, UK) were run in triplicate. Analysis was performed using the ΔΔCT method and the mean and standard deviation of 84 HCs were used to determine the normal ΔΔCT for the probes. Samples below 3 standard deviations of the HC mean were called CN loss.

Assessment of gene expression

RNA was extracted from $CD4^+$-enriched PBMCs of 20 patients with SS and 14 HCs and converted to cDNA. Previously published primer/probe sequences for *cyclophilin-A/PPIA* (Riemer et al., 2012), *MTAP* (Hs00559618_m1), and *CDKN2A* (Hs00923894_m1) were used to perform RT-Q-PCR on the ABI 7900HT high-throughput real-time PCR system (Applied Biosystems); reactions were performed in triplicate. *Cyclophilin-A/PPIA* was used as an endogenous control. The ΔΔCT thresholds of *MTAP* and *CDKN2A* were compared between patient cDNA and HC cDNA. The normal range of *MTAP* and *CDKN2A* expression was defined as within 3 standard deviations of the HC mean ΔΔCT ($P < 0.01$). A sample with ΔΔCT below this range was considered to have reduced expression.

Statistical analysis

All statistical analysis was performed in R (www.r-project.org/). A CN of 2 was determined to be the mean ± 3 standard deviations of the HCs with 99% confidence. Any patient values that fell outside of these limits were considered to have CN loss or gain. For Q-PCR probe set 1 analysis of variance was used to test for significant differences between healthy and patient subgroups followed by Tukey's HSD test for pairwise comparisons. A Wilcoxon-rank-sum test was used to test for a significant difference in expression between patient and healthy groups.

Bisulfite conversion of DNA

One microgram of DNA was bisulfite converted using the Epitect Bisulfite kit (Qiagen, Manchester, UK) according to the manufacturer's instructions.

Bisulfite sequencing PCR

A set of PCR primers specific to bisulfite-treated DNA were designed for each discrete region CG1, CG2, and CG3 of the *MTAP* CpG island using Methprimer software (Li and Dahiya, 2002). DNA was then subjected to PCR amplification with primers specific to:

CG1 (F-ATTGGATTATTTTAGTAGGGAAGGG, R-AAAACCTTAC ACAACTCCCAATCTA)

CG2 (F-AGTTTTGGGTTAAGTTTATTTTAGT, R-AAAAAAAACAA CACTCCCTACTTAAC)

CG3 (F-GGGGAGATTTTATATAAGTAGTTAATT, R-TAATACCAA AAACCATATCTACAC)

PCR was carried out using AmpliTaq Gold (Applied Biosystems) with 2 mM $MgCl_2$ and an annealing temperature of 59 °C for CG1, 61 °C for CG2, and 58 °C for CG3.

Reverse transcriptase-PCR

RNA was extracted fresh from cell lines. The following primers were used for PCR amplification of cDNA:

MTAP (F-CTCCCGCGCAGTGAGGTTGG, R-CGCCGGGCTGAA ATCTCCTCC)

CDKN2A (F-GCTGCCCAACGCACCGA, R-GCGCTGCCCATCA TCATG)

Cyclophillin (F-AAAGCATACGGGTCCTGGCATC, R-CGAGTTG TCCACAGTCAGCAATG)

PCR was carried out using AmpliTaq Gold (Applied Biosystems) with an annealing temperature and $MgCl_2$ concentration of 55 °C, 1.5 mM for cyclophillin and 60 °C, 1.5 mM for CDKN2A. A touchdown PCR approach using 2.0 mM $MgCl_2$ was employed for MTAP.

Intracellular flow cytometry

The cell lines SeAx and MyLa were fixed in 1% formaldehyde and then permeabilized in 90% methanol at 1×10^6 cells/ml. Cells were incubated with an MTAP monoclonal antibody (ab55517, AbCam, Cambridge, UK) followed by treatment with a fluorescently labeled

secondary antibody (Alexa Fluor 555 goat antimouse, A-21422, Life Technologies). To detect background fluorescence, samples were incubated with the mouse isotype control (Life Technologies). Flow cytometry analysis was performed on a FACSAriaII (BD Biosciences, Oxford, UK); machine settings were standardized and retained throughout the study, 10,000 events were acquired per sample, and delta mean fluorescent intensity was calculated by subtracting the mean fluorescent intensity of the isotype control from the mean fluorescent intensity of the specific antibody using Flowjo software (Tree Star, Ashland, OR).

Capture library design

A targeted capture was designed using the "Agilent SureDesign" probe selection software (Agilent, Cheshire, UK). Probes were designed to cover all exons and UTRs of *MTAP*, *CDKN2A,* and *CDKN2B* (Supplementary Table S4 online). Additional probes were added to the 3′ and 5′ ends of each exon to ensure efficient capture of splice site sequences.

Library preparation, target capture, and sequencing

Tumor DNA was derived from PBMCs of 77 patients with SS and 32 HCs. Libraries were prepared according to the manufacturer's instructions using the SureSelect XT library preparation kit (Agilent). Libraries were incubated for 24 hours with biotinylated custom capture probes designed against the target region. Streptavidin-coated magnetic beads were utilized to isolate the probe-bound target DNA and any unbound DNA was washed off. A low cycle number PCR was then used to amplify the bound DNA but minimize PCR duplicates. The target enriched DNA was finally sequenced with 100 basepairs paired-end reads on an Illumina (San Diego, CA) Hi-Seq2000.

Next generation sequencing analysis

Reads were aligned to the reference genome (hg19, NCBI build) using Novoalign (Novocraft Technologies). SAMtools was used to call variants and remove PCR duplicates. ANNOVAR (Wang et al., 2010) was used to annotate variants with respect to coding genes. In-house scripts and the BEDTools package were used to calculate the depth of sequencing and overall coverage. A total of 32 HCs underwent NGS and were used as controls to exclude sequencing artifacts from the capture.

For 31 patient with SS samples that were also included in the Q-PCR study, CN analysis was performed on NGS data using the "ExomeDepth" R-package (Plagnol et al., 2012). Briefly, this package uses normalized exome depth, which is the ratio of depth at any given base to the total number of reads for that sample. This metric is considered across each exon in comparison to expected values based on HCs included in our NGS study. The final output is based on a hidden markov model that outputs either deletion, normal, or duplication for each exon; see Plagnol et al (2012) for more detailed description.

CONFLICTS OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

This research was supported by grants from the British Skin Foundation and the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. We thank Michael Simpson for his assistance with the analysis of next generation sequencing data and critical appraisal of the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at http://dx.doi.org/10.1016/j.jid.2016.01.028

REFERENCES

Agar NS, Wedgeworth E, Crichton S, Mitchell TJ, Cox M, Ferreira S, et al. Survival outcomes and prognostic factors in mycosis fungoides/Sézary syndrome: validation of the revised International Society for Cutaneous Lymphomas/European Organisation for Research and Treatment of Cancer staging proposal. J Clin Oncol 2010;28:4730–9.

Antonopoulou K, Stefanaki I, Lill CM, Chatzinasiou F, Kypreou K, Karagianni F, et al. Updated field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma: The MelGene Database. J Invest Dermatol 2014;135:1074–9.

Batova A, Diccianni MB, Nobori T, Vu T, Yu J, Bridgeman L, et al. Frequent deletion in the methylthioadenosine phosphorylase gene in T-cell acute lymphoblastic leukemia: strategies for enzyme-targeted therapy. Blood 1996;88:3083–90.

Behrmann I, Wallner S, Komyod W, Heinrich PC, Schuierer M, Buettner R, et al. Characterization of methylthioadenosin phosphorylase (MTAP) expression in malignant melanoma. Am J Pathol 2003;163:683–90.

Berasain C, Hevia H, Fernandez-Irigoyen J, Larrea E, Caballeria J, Mato JM, et al. Methylthioadenosine phosphorylase gene expression is impaired in human liver cirrhosis and hepatocarcinoma. Biochim Biophys Acta 2004;1690:276–84.

Bertino JR, Lubin M, Johnson-Farley N, Chan WC, Goodell L, Bhagavathi S. Lack of expression of MTAP in uncommon T-cell lymphomas. Clin Lymphoma Myeloma Leuk 2012;5:306–9.

Brownhill SC, Taylor C, Burchill SA. Chromosome 9p21 gene copy number and prognostic significance of p16 in ESFT. Br J Cancer 2007;96:1914–23.

Brüggemann M, White H, Gaulard P, Garcia-Sanz R, Gameiro P, Oeschger S, et al. Powerful strategy for polymerase chain reaction-based clonality assessment in T-cell malignancies Report of the BIOMED-2 Concerted Action BHM4 CT98-3936. Leukemia 2007;21:215–21.

Choi J, Goh G, Walradt T, Hong B, Bunick C, Chen K, et al. Genomic landscape of cutaneous T cell lymphoma. Nat Genet 2015;47:1011–9.

Christopher SA, Diegelman P, Porter CW, Kruger WD. Methylthioadenosine phosphorylase, a gene frequently codeleted with p16(cdkN2a/ARF), acts as a tumor suppressor in a breast cancer cell line. Cancer Res 2002;62:6639–44.

da Silva Almeida A, Abate F, Khiabanian H, Martinez-Escala E, Guitart J, Tensen C, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. Nat Genet 2015;47:1465–70.

De Preter K, Speleman F, Combaret V, Lunec J, Laureys G, Eussen BH, et al. Quantification of MYCN, DDX1, and NAG gene copy number in neuroblastoma using a real-time quantitative PCR assay. Mod Pathol 2002;15:159–66.

Garcia-Castellano JM, Villanueva A, Healey JH, Sowers R, Cordon-Cardo C, Huvos A, et al. Methylthioadenosine phosphorylase gene deletions are common in osteosarcoma. Clin Cancer Res 2002;8:782–7.

Hellerbrand C, Muhlbauer M, Wallner S, Schuierer M, Behrmann I, Bataille F, et al. Promoter-hypermethylation is causing functional relevant down-regulation of methylthioadenosine phosphorylase (MTAP) expression in hepatocellular carcinoma. Carcinogenesis 2006;27:64–72.

Ishii M, Nakazawa K, Wada H, Nishioka J, Nakatani K, Yamada Y, et al. Methylthioadenosine phosphorylase gene is silenced by promoter hypermethylation in human lymphoma cell line DHL-9: another mechanism of enzyme deficiency. Int J Oncol 2005;26:985–91.

Kadariya Y, Tang B, Wang L, Al-Saleem T, Hayakawa K, Slifker MJ, et al. Germline mutations in Mtap cooperate with Myc to accelerate tumorigenesis in mice. PLOS One 2013;8:e67635.

Kadariya Y, Yin B, Tang B, Shinton SA, Quinlivan EP, Hua X, et al. Mice heterozygous for germ-line mutations in methylthioadenosine phosphorylase (MTAP) die prematurely of T-cell lymphoma. Cancer Res 2009;69:5961–9.

Kamatani N, Carson DA. Abnormal regulation of methylthioadenosine and polyamine metabolism in methylthioadenosine phosphorylase-deficient human leukemic cell lines. Cancer Res 1980;40:4178–82.

Kamatani N, Nelson-Rees WA, Carson DA. Selective killing of human malignant cell lines deficient in methylthioadenosine phosphorylase, a purine metabolic enzyme. Proc Natl Acad Sci USA 1981;78:1219–23.

Kiel M, Sahasrabuddhe A, Rolland D, Velusamy T, Chung F, Schaller M, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sézary syndrome. Nat Commun 2015;6:8470.

Laharanne E, Chevret E, Idrissi Y, Gentil C, Longy M, Ferrer J, et al. CDKN2A-CDKN2B deletion defines an aggressive subset of cutaneous T-cell lymphoma. Mod Pathol 2010;23:547–58.

Leal MF, Lima EM, Silva PN, Assumpcao PP, Calcagno DQ, Payao SL, et al. Promoter hypermethylation of CDH1, FHIT, MTAP and PLAGL1 in gastric adenocarcinoma in individuals from Northern Brazil. World J Gastroenterol 2007;13:2568–74.

Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics 2002;18:1427–31.

Mao X, Lillington DM, Czepulkowski B, Russell-Jones R, Young BD, Whittaker S. Molecular cytogenetic characterization of Sézary syndrome. Genes Chromosomes Cancer 2003;36:250–60.

M'Soka TJ, Nishioka J, Taga A, Kato K, Kawasaki H, Yamada Y, et al. Detection of methylthioadenosine phosphorylase (MTAP) and p16 gene deletion in T cell acute lymphoblastic leukemia by real-time quantitative PCR assay. Leuk Lymphoma 2000;14:935–40.

McGirt L, Jia P, Baerenwald D, Duszynski R, Dahlman K, Zic J, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. Blood 2015;126:508–19.

Meyer S, Wild PJ, Vogt T, Bataille F, Ehret C, Gantner S, et al. Methylthioadenosine phosphorylase represents a predictive marker for response to adjuvant interferon therapy in patients with malignant melanoma. Exp Dermatol 2010;19:e251–7.

Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics 2012;28:2747–54.

Riemer AB, Keskin DB, Reinherz EL. Identification and validation of reference genes for expression studies in human keratinocyte cell lines treated with and without interferon-γ—a method for qRT-PCR reference gene determination. Exp Dermatol 2012;21:625–9.

Rose AE, Poliseno L, Wang J, Clark M, Pearlman A, Wang G, et al. Integrative genomics identifies molecular alterations that challenge the linear model of melanoma progression. Cancer Res 2011;71:2561–71.

Rubben A, Kempf W, Kadin ME, Zimmermann DR, Burg G. Multilineage progression of genetically unstable tumor subclones in cutaneous T-cell lymphomas. J Exp Dermatol 2004;13:472–83.

Salgado R, Servitje O, Gallardo F, Vermeer MH, Ortiz-Romero PL, Karpova MB, et al. Oligonucleotide array-CGH identifies genomic subgroups and prognostic markers for tumor stage mycosis fungoides. J Invest Dermatol 2010;130:1126–35.

Scarisbrick JJ, Woolford AJ, Calonje E, Photiou A, Ferreira S, Orchard G, et al. Frequent abnormalities of the p15 and p16 genes in mycosis fungoides and Sézary syndrome. J Invest Dermatol 2002;118:493–9.

Stevens AP, Spangler B, Wallner S, Kreutz M, Dettmer K, Oefner PJ, et al. Direct and tumor microenvironment mediated influences of 5′-deoxy-5′-(methylthio)adenosine on tumor progression of malignant melanoma. J Cell Biochem 2009;106:210–9.

Takasaki Y, Yamada Y, Sugahara K, Hayashi T, Dateki N, Harasawa H, et al. Interruption of p16 gene expression in adult T-cell leukaemia/lymphoma: clinical correlation. Br J Haematol 2003;122:253–9.

Tang B, Kadariya Y, Chen Y, Slifker M, Kruger W. Expression of MTAP inhibits tumor-related phenotypes in HT1080 cells via a mechanism unrelated to its enzymatic function. G3 (Bethesda) 2014;5:35–44.

Ungewickell A, Bhaduri A, Rios E, Reuter J, Lee C, Mah A, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. Nat Genet 2015;47:1056–60.

van Doorn R, van Kester MS, Dijkman R, Vermeer MH, Mulder AA, Szuhai K, et al. Oncogenomic analysis of mycosis fungoides reveals major differences with Sézary syndrome. Blood 2009;113:127–36.

Wang K, Li M, Hakonarson H. ANNOVAR: Functional annotation of genetic variants from next-generation sequencing data. Nucleic Acids Res 2010;38:e164.

Wang L, Ni X, Covington K, Yang B, Shiu J, Zhang X, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. Nat Genet 2015;47:1426–34.

Watanabe F, Takao M, Inoue K, Nishioka J, Nobori T, Shiraishi T, et al. Immunohistochemical diagnosis of methylthioadenosine phosphorylase (MTAP) deficiency in non-small cell lung carcinoma. Lung Cancer 2009;63:39–44.

Wild PJ, Meyer S, Bataille F, Woenckhaus M, Ameres M, Vogt T, et al. Tissue microarray analysis of methylthioadenosine phosphorylase protein expression in melanocytic skin tumors. Arch Dermatol 2006;142:471–6.

Willemze R, Jaffe ES, Burg G, Cerroni L, Berti E, Swerdlow SH, et al. WHO-EORTC classification for cutaneous lymphomas. Blood 2005;105:3768–85.

Zhao W, Huang CC, Otterson GA, Leon ME, Tang Y, Shilo K, et al. Altered P16 INK4 and RB1 expressions are associated with poor prognosis in patients with nonsmall cell lung cancer. J Oncol 2012;2012:957437.

3.3 Additional figures and data for chapter 3

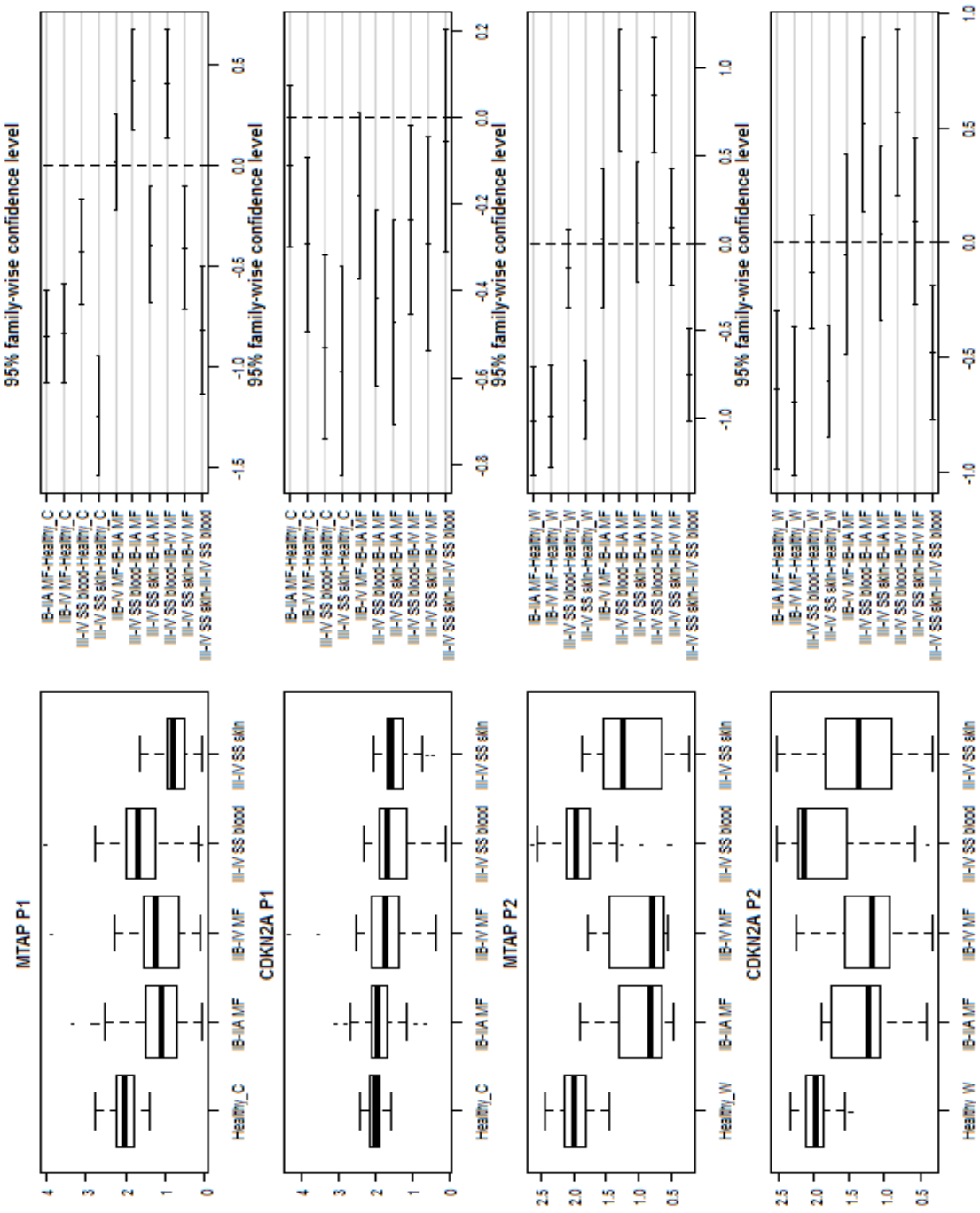


Figure 3.1. Tukey's HSD for probe sets 1 and 2 for gene copy number analysis. ANOVA, followed by post hoc testing was carried out for each probe for each gene. Box plots show the distribution of samples for each probe/gene combination and Tukeys test results are shown adjacent. Results for probe set 2 should be interpreted with caution as samples included in this set were selected from probe set 1 to validate original data.

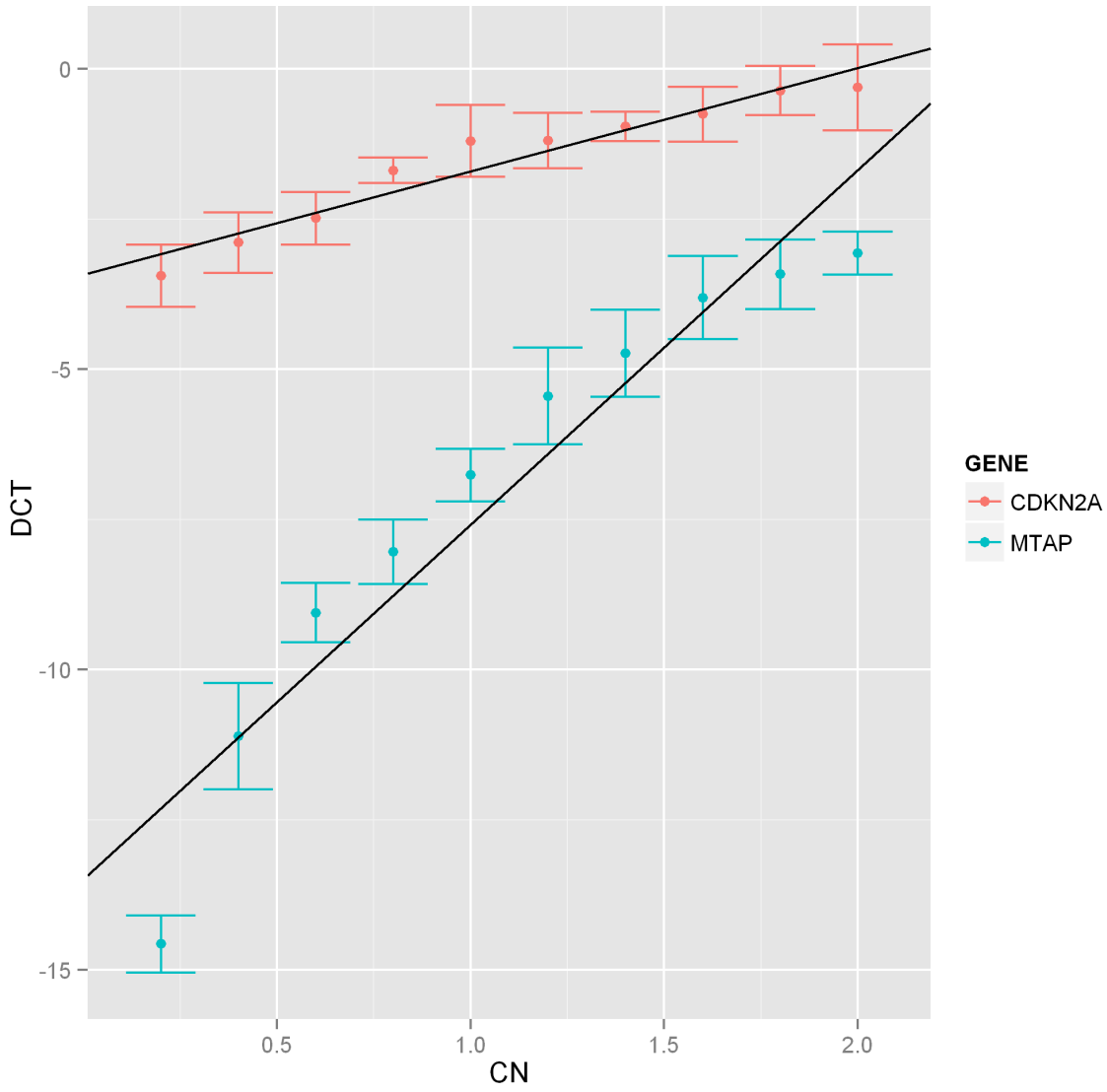
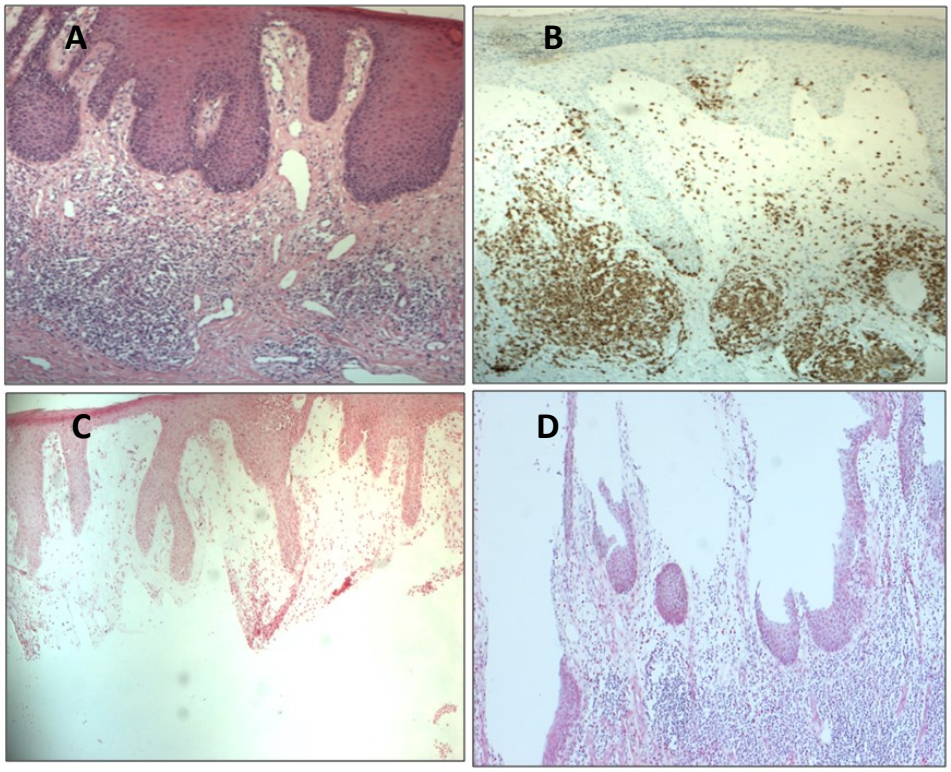


Figure 3.2. Linear Regression analysis to determine *MTAP/CDKN2A* Copy Number

The MyLa cell line has normal gene copy number of *MTAP* and *CDKN2A* whereas the Jurkat cell line has homozygous loss of both genes. Mixed DNA samples from the cell lines were assayed by Q-PCR for *MTAP* and *CDKN2A* using *TERT* as a control gene (See materials and methods section). The gene copy number for *MTAP* and *CDKN2A* in each mixed sample can be directly inferred from the ratios of MyLa and SeAx in the samples and is plotted on the x-axis. From left to right (x-axis) samples comprising 10%- MyLa /90% -Jurkat up to 100% MyLa/0% -Jurkat mixed in 10% intervals. Delta-CT was calculated and plotted on the y-axis. Linear regression is a method for modelling a scalar, dependent variables on the y-axis with an independent variable on the x-axis. Linear regression was used to derive a line of best fit demonstrating a clear positive trend for both *MTAP* and *CDKN2A* with increasing target DNA relative to *TERT*. Regression lines were extrapolated to infer copy number in subsequent microdissected DNA samples from patient tumour and epidermis (DCT=Delta-CT, CN=gene copy number).

Case 1



Case 2

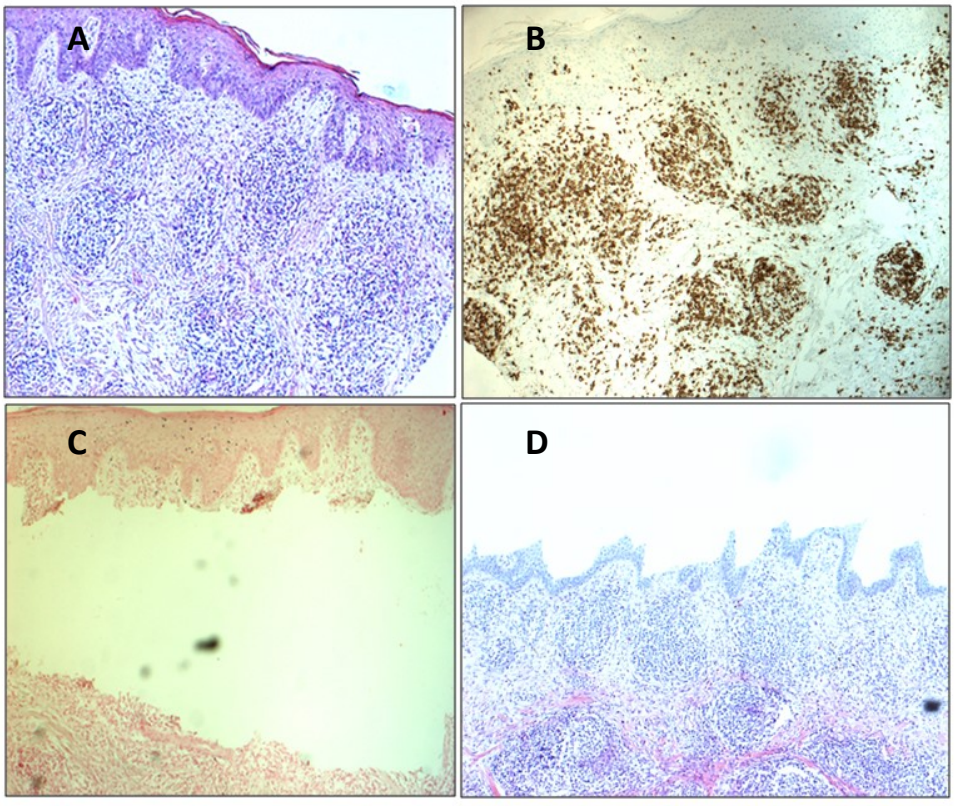
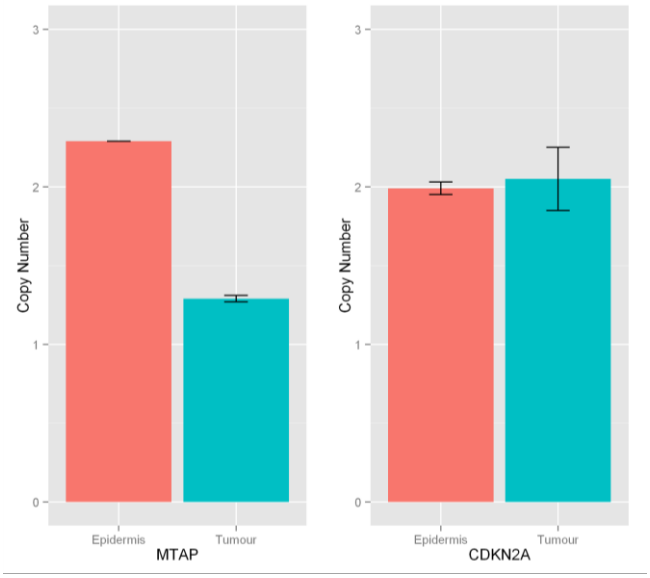


Figure 3.3. Micro-dissection of tumour rich tissue and epidermal regions from FFPE sections

Micro-dissection was performed on 2 MF cases for which paraffin embedded material of the original lesion (analysed by probe sets 1 and 2) was available. Slides were assessed and marked by an expert dermatopathologist. DNA was extracted from microdissected FFPE tissue from the epidermis and focal areas of tumour cells. Images were obtained of sections for each case: A = H/E stain of MF plaque showing the epidermis and lymphocyte infiltrates. B = CD3+ stain of the same tumour, lymphocytes can be seen clearly in large numbers. C = Microdissected area containing tumour cells which was used for enriched tumour DNA extraction. D = Microdissected area showing regions where epidermis was removed for DNA extraction. An additional H&E was taken after cutting the sections (not shown).

Case 1



Case 2

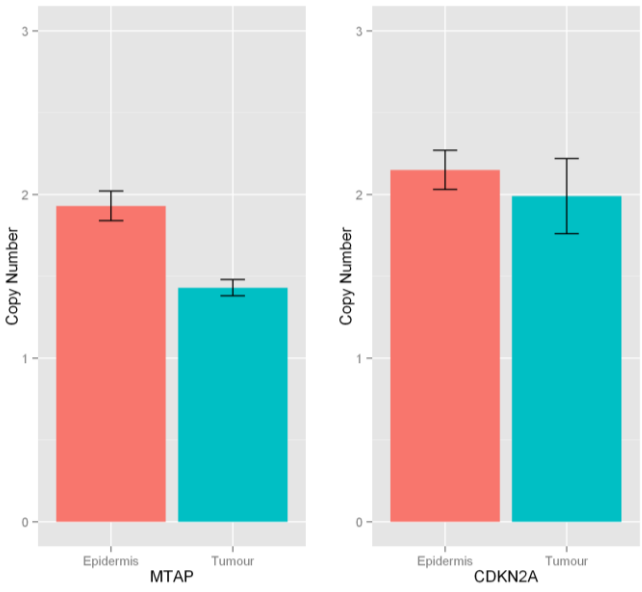


Figure 3.4. Determination of *CDKN2A* and *MTAP* Copy Number in DNA from micro-dissected tissue.

Q-PCR was performed on DNA extracted from tumour enriched and epidermal FFPE tissue (shown in figure 1A). Delta CT values were calculated for *MTAP* and *CDKN2A* relative to the housekeeping gene TERT. Gene copy number was calculated from the regression analysis of Delta-CT values (shown in figure 1). The data demonstrates reduced *MTAP* gene copy number and no loss of *CDKN2A* in the tumour compared to the epidermis in both cases. Error bars show standard error of gene CN. T-tests were carried out for *MTAP* and *CDKN2A* comparing to a gene copy number of 2 which was previously determined by the cell line MyLa. In both cases; for *MTAP* a significant difference was detected between tumour and MyLa ($P<0.05$), there was no significant difference between the *MTAP* CN in epidermis and MyLa, or *CDKN2A* from the tumour or epidermis compared to MyLa in either case.

4 The effect of MTAP loss on type I PRMT activity

4.1 Backgound

Loss of *MTAP* is known to occur in a range of haematological malignancies (246-252) and several studies suggest MTAP has a role as a tumour suppressor. The most prominent lines of evidence include a mouse model where loss of *MTAP* leads to mature T-cell malignancies (253) and studies involving xenografted tumours where tumour growth is inhibited upon reintroduction of *MTAP* (141).

The only known function of MTAP is to metabolize MTA (241, 242) and the most prominent hypothesis of MTAPs tumour suppressing function is that it prevents the accumulation of MTA and thus prevents MTA from inhibiting other cellular processes (144, 242).

One class of enzymes thought to be inhibited by the accumulation of MTA is the PRMT family. Increased MTA has been shown to inhibit the activity of PRMT1 and prevent the deactivation of STAT1 leading to constitutive STAT1 activation (267). Levels of MTA have also been shown to correlate with MAPK/ERK activity, this is a well-known oncogenic pathway known to be modulated by PRMT5 (271). Restoring MTAP function has been shown to lead to the restoration or PRMT function which correlates with a reduction in overactive ERK (270).

To date, the effect of loss of *MTAP* on the PRMT family has not been investigated in CTCL. The aim of this investigation is as a preliminary study to first determine if PRMTs are expressed in CTCL and if loss of MTAP has an effect on their activity.

4.2 Expression of PRMTs in SeAx and Jurkat cells

In order to determine that PRMT family members are expressed in model cell lines, qualitative expression of the PRMT gene family was assessed in SeAx and Jurkat CTCL cell lines using RT-PCR. Expression of all but one PRMT family member was confirmed, with the only exception being PRMT8 which showed no detectable expression (**Figure 4.1**).

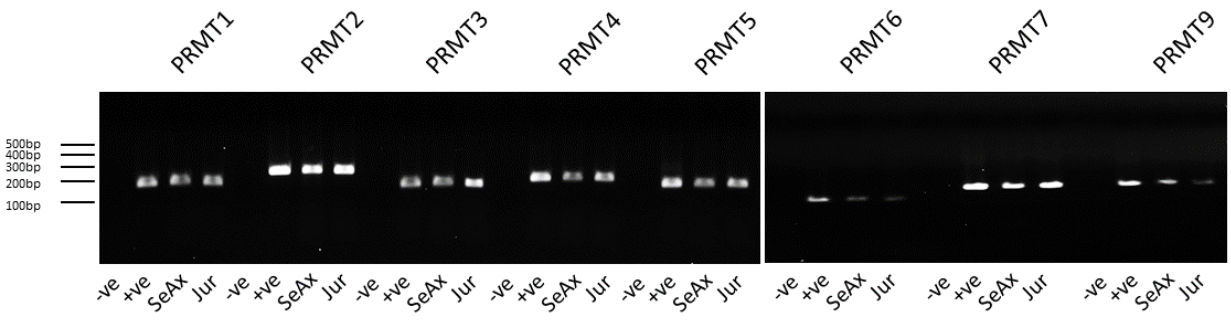


Figure 4.1 Expression of PRMTs in SeAx and Jurkat.
Qualitative expression of PRMTs is shown in the positive control (Healthy CD4+ T-cells) and the cell lines SeAx and Jurkat. Single bands of the expected length were shown in each case.

4.3 Transient knock-down of MTAP in SeAx cells

Seax cells were chosen as a model to assess in vitro effects of *MTAP* loss. Other lines reviewed as candidate models included the cell lines Hut78 and Jurkat, however both copies of *MTAP* are deleted in these lines. Also, the cell line MyLa was considered but *MTAP* is silenced by promotor methylation. Therefore, SeAx was considered the most suitable T-cell model.

Transient knockdown of *MTAP* was attempted in SeAx using a siRNA construct complimentary to *MTAP* and a scrambled siRNA control. Initial attempts by nucleofection yielded a high rate of cell transformation (**Figure 4.2**). However, this was accompanied by high rates of cell death and aggregation in the following 24 hours so that lysates and RNA could not be extracted consistently.

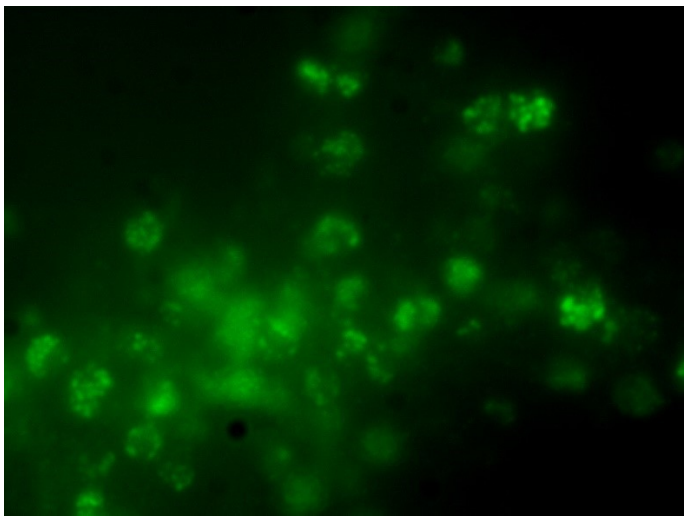


Figure 4.2 Nucleofection of SeAx cells.
SeAx cells underwent nucleofection using GFP expression vector as a control. Qualitative assessment of nucleofection was performed using fluorescence microscopy; the image shows the typical appearance of cells 24 hours after transfection.

Subsequent transfection of SeAx was attempted by Lipofectamine 2000. This was shown to have low to no toxicity to SeAx cells however the rate of transformation was considered too low (<0.1% transformation) to detect a signal with all recommended experimental conditions (**Figure 4.3**). Transfection was also attempted at concentrations of both DNA and Lipofectamine2000 beyond those recommended by the manufacturer (Invitrogen, UK) and showed no significant improvement. Transient transfection of SeAx was considered unpractical at this point for the planned investigation.

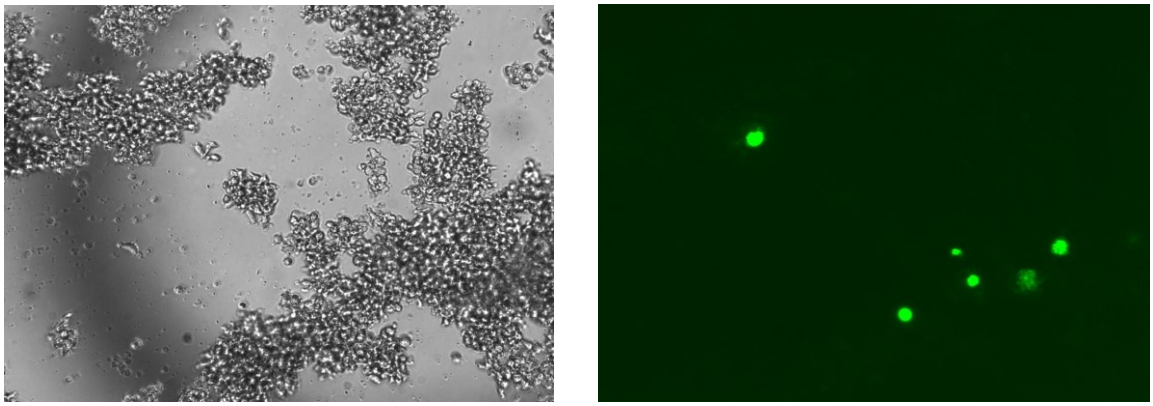


Figure 4.3 Transfection of SeAx by Lipofectamine2000.
SeAx cells transfected with 1000ng GFP expression vector using 2.5µl Lipofectamine2000 in 50µl of antibiotic free Optimem media per 4x105 SeAx cells. Limited transfection occurs.

4.4 Proof of principle experiments in HEK293 cells

The human embryonic kidney cell line HEK293 was considered for use in a proof of principle experiment to assess the in-vitro effects of MTAP knock-down. This cell line is known to be easily transfected and so signals from transient knock-down of MTAP should be detectable in downstream experiments.

4.5 Expression of PRMTs in HEK293 cells

Qualitative expression of the PRMT gene family was assessed in HEK293 cells using RT-PCR in the first instance. Expression of most (8/9) PRMT family members was confirmed, similar to SeAx and Jurkat with the only exception being PRMT8 which showed no detectable expression (**Figure 4.4**).

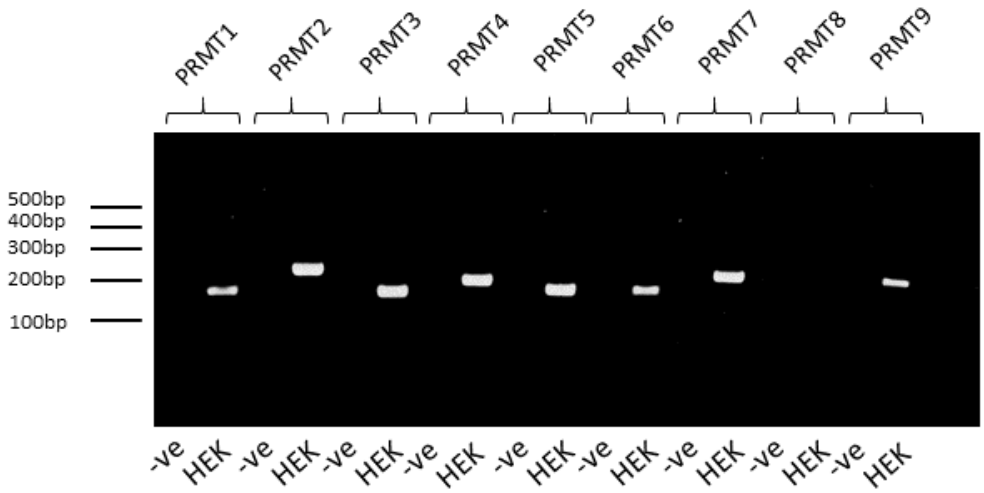


Figure 4.4 Expression of PRMTs in HEK293 cells.
All PRMTs except PRMT8 show bands of the expected size demonstrating expression in HEK293 cells. A negative control containing no target DNA was run for each reaction.

4.6 Transient knock-down of MTAP in HEK293 cells

Transient knockdown of *MTAP* was attempted in HEK293 cells using the siRNA constructs that were used for the previous experiment in SeAx. The Lipofectamine2000 chemical transfection kit was used according to manufacturer's recommendations (Invitogen, UK). Qualitative assessment of transfection using a GFP expression vector and fluorescence microscope indicated that transfection was efficient (**Figure 4.5**).

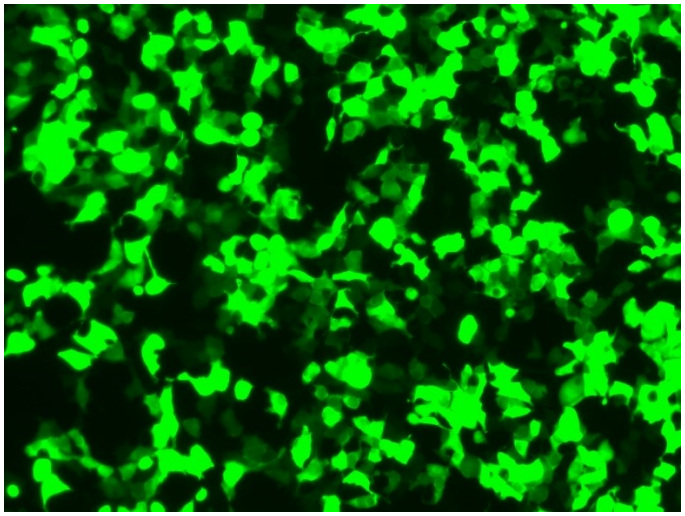


Figure 4.5 HEK293 cells transfected with pmaxGFP.
High rates of transfection of the GFP expressing vector pmaxGFP was confirmed qualitatively by fluorescence microscopy. The image shows a typical transfection after 24 hours.

4.6.1 Assessment of *MTAP* expression by RT-PCR

In order to determine that the *MTAP* siRNAs can effectively target *MTAP*, expression of *MTAP* in HEK293 cells was analysed in untransfected cells, cells that were transfected with two different siRNA constructs complimentary to *MTAP* and cells were transfected with a scrambled siRNA control. Cells were collected prior to transfection (0h) and at 24, 48 and 72 hours post transfection to extract RNA and cell lysates. RT-PCR of *MTAP* and PPIA was performed for all samples and products were run on a 2% agarose gel. The cells transfected with *MTAP* siRNA show reduced expression of *MTAP* compared to untransfected cells and cells transfected with scrambled siRNA suggesting that transfection with *MTAP* siRNA can reduce *MTAP* mRNA (**Figure 4.6**).

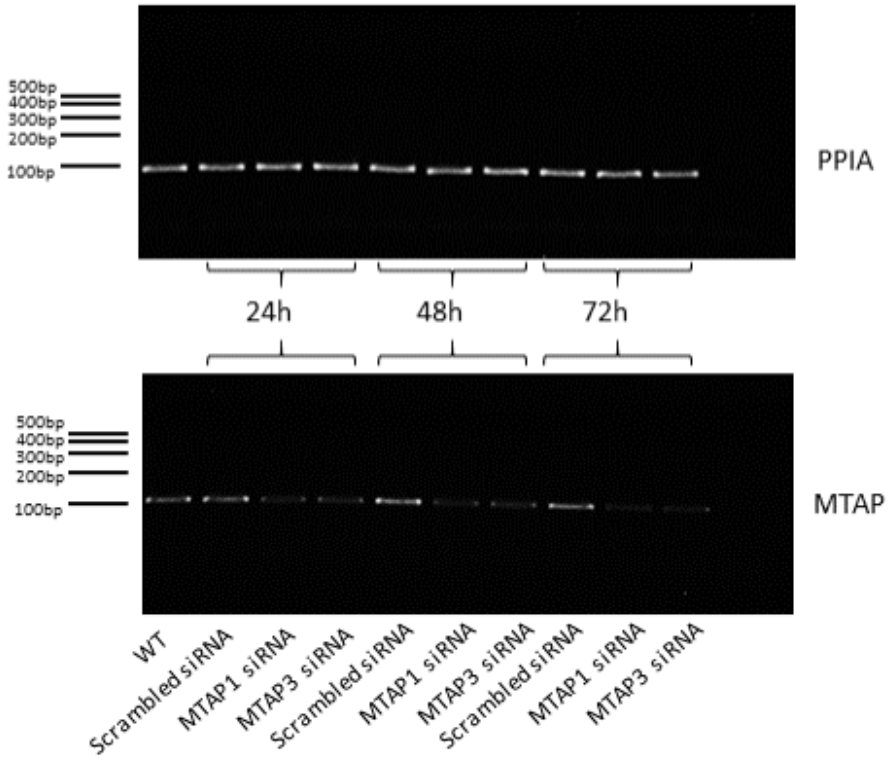


Figure 4.6 Expression of *MTAP* and PPIA by RT-PCR in HEK293.
Expression of *MTAP* and PPIA was analysed in HEK293 cells. From left to right; untransfected, scrambled siRNA, *MTAP* siRNA1*, MTAP* siRNA 3 after 24 hours, scrambled siRNA, *MTAP* siRNA1, *MTAP* siRNA 3 after 48 hours, scrambled siRNA, *MTAP* siRNA1, *MTAP* siRNA 3 after 72 hours.

4.6.2 Assessment of MTAP expression by RT-QPCR

MTAP expression was further analysed by RT-QPCR which confirmed with greater accuracy reduced *MTAP* expression in cells transfected with *MTAP* siRNA compared to cells transfected with scrambled siRNA or untransfected cells (**Figure 4.7**). Untransfected HEK293 cells had comparable expression to cells prior to transfection (data not shown).

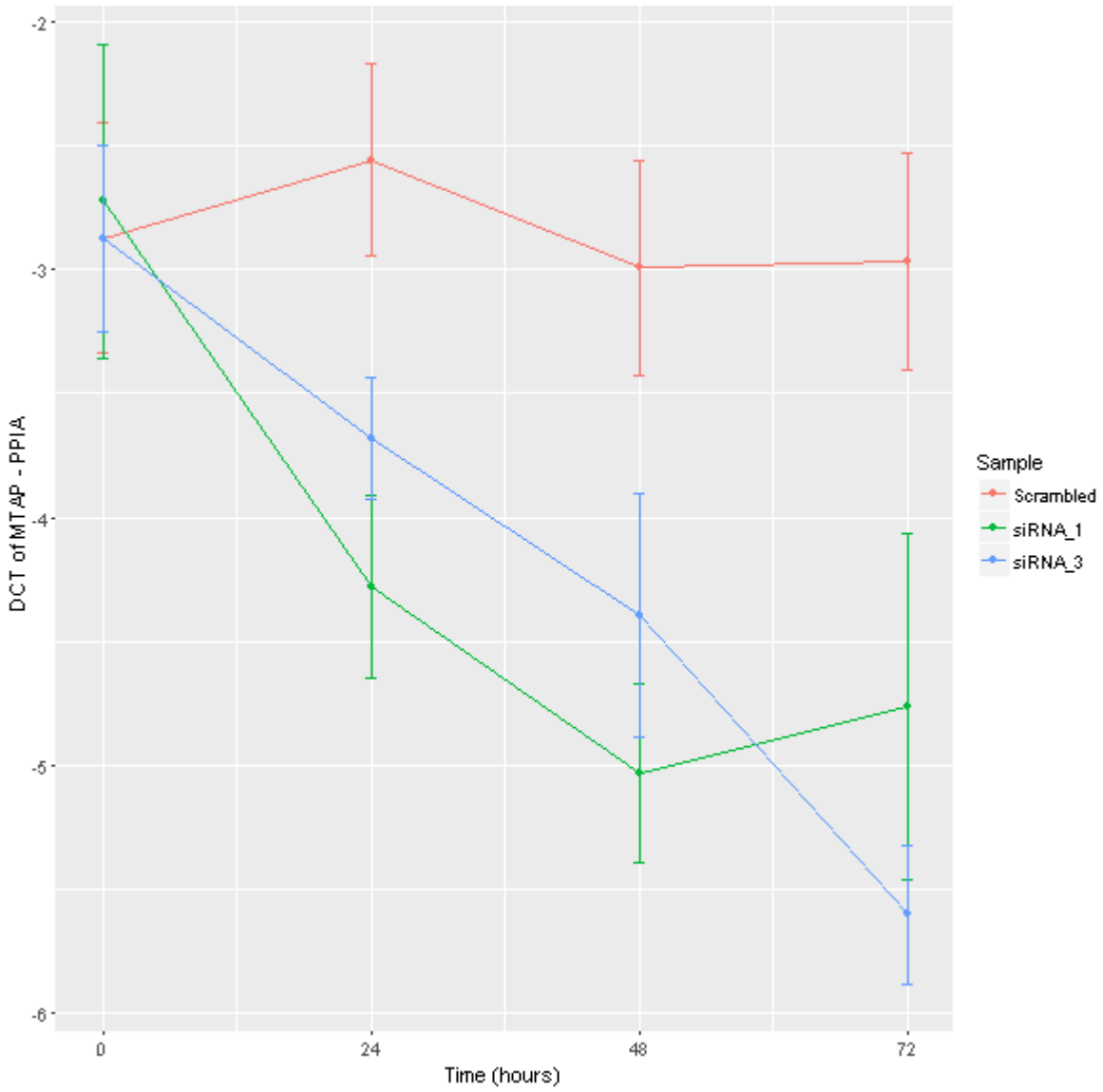


Figure 4.7 *MTAP* expression in HEK293 cells analysed by RT-QPCR.
Expression of *MTAP* relative to PPIA in HEK293 cells transfected with scrambled siRNA, *MTAP* siRNA1 and *MTAP* siRNA3. Expression was analysed before transfection and at 24, 48 and 72 hours after transfection. All samples were run in triplicate. Error bars show standard error.

4.7 Assessment of MTAP protein expression

To assess if loss of *MTAP* mRNA correlates with loss of MTAP protein, whole cell protein lysates were prepared from; untransfected HEK293 cells, cells transfected with a scrambled siRNA control, cells transfected with *MTAP* siRNA1 and cells transfected with *MTAP* siRNA3. Lysates were extracted from transfected cells at 24, 48 and 72 hours post-transfection and run on a 5-15% SDS-PAGE gel.

Immunoblot analysis of MTAP shows similar expression of untransfected HEK293s compared to cells transfected with scrambled siRNA at all timepoints (**Figure 4.8**). Cells transfected with *MTAP* siRNA appear to show reduced MTAP protein expression suggesting that *MTAP* siRNAs can reduce MTAP protein. However, this assay was only performed a single time and further confirmation would be required to validate this claim.

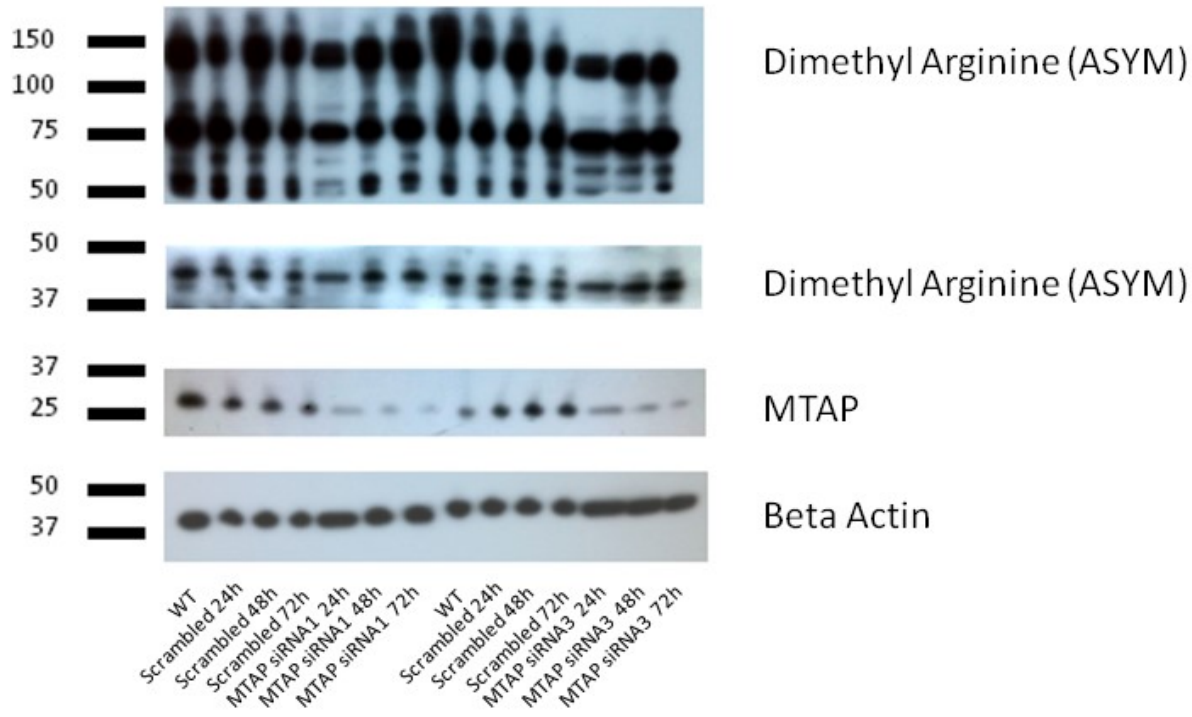


Figure 4.8 Immunoblot analysis of MTAP, Asymetric dimethyl arginine methylation and β-actin.
Whole cell lysates from untransfected HEK293s, HEK293s transfected with scrambled siRNA, *MTAP* siRNA 1, and *MTAP* siRNA 3 were run on a 5-15% SDS-PAGE gel. Lysates were extracted from transfected cells at 24, 48 and 72 hours after transfection and probed with the indicated antibodies.

4.8 Assessment of pan asymmetric-dimethyle-arginine methylation in MTAP knock-down HEK293 cells

To determine if reduced MTAP affects the activity of the PRMT family the blot was re-probed to assess asymmetric dimethyl arginine methylation. This is an assay to detect the activity of type I PRMT family members which include PRMTs 1, 2, 3, 6 and 8.

The inferences reported in this section relating to the Western blot analysis are suggestive hints derived from a single experiment which require further confirmation rather than conclusive results. Untransfected HEK293s and HEK293s transfected with scrambled siRNA show similar patterns of dimethyl arginine methylation at all timepoints (**Figure 4.8**). HEK293 cells transfected with *MTAP* siRNA appear to show slightly reduced dimethyl arginine methylation on several proteins. This can be inferred from the slightly reduced intensity of bands at ~150KDa, ~80KDa, ~50KDa, ~45KDa and ~40KDa. Cells transfected with *MTAP* siRNA1 also appear to show fainter bands at ~60KDa. The reduction in dimethyl arginine methylation peaks at 24 hours in the majority of cases.

The apparent reduction in asymmetric dimethyl-arginine methylation correlates with reduced MTAP expression. This assay suggests the possibility that loss of *MTAP* could reduce pan-dimethyle-arginine methylation, likely via reducing the activity of type I PRMT family members. However, this assay has not been repeated and the result is very weak at this stage so would require further optimisation to draw any strong conclusion.

5 Candidate driver genes in Sézary syndrome: frequent perturbations of genes involved in genome maintenance and DNA repair

5.1 Candidate contributions to chapter 5

This chapter is work that has been contributed to and published by a team of researchers. My personal contributions are as follows: library preparation of all tumour and germline WES samples, downstream filtering of somatic WES variants, design and library preparation of all targeted capture samples including patient samples and HC's, Sanger sequencing validation of numerous WES and targeted capture variants, comparison of genes in our dataset to the network of cancer genes, comparison of variants in our dataset to identical variants in other similar NGS studies, Figure 1, all tables, all supplementary tables and heatmaps, selection of genes to be included in submatrices of figure 4, I also produced alternate versions of figure 2a, 2b and matrices from figure 4 though versions produced by the BRC were selected for the final manuscript, furthermore I produced the original draft and subsequent revisions of the manuscript with oversight from my PhD supervisors.

5.2 Introductory remarks for chapter 5

In this study, patients from the cohort used in the *MTAP* study (Chapter 3) were shared with this study. In total 31/101 patients that underwent NGS in this study were used in the *MTAP* study and overlapped with the QPCR gene CN study in chapter 3. A total of 77/101 were used for mutational analysis across 9p21 in the chapter 3.

Gene copy number variations and mutations across the chromosome 9p21 region have not been analysed in this study as this was looked at more specifically in chapter 3 and the data had been published in the Journal of Investigative Dermatology. Data from this chapter was published in the journal 'Blood' and the studies are bound by publishing rules of both journals which do not allow data to be duplicated across separate studies. For this reason data relating to 9p21 is not analysed as part of chapter 5.

Regular Article

blood

LYMPHOID NEOPLASIA

Candidate driver genes involved in genome maintenance and DNA repair in Sézary syndrome

Wesley J. Woollard,[1] Venu Pullabhatla,[2] Anna Lorenc,[2] Varsha M. Patel,[1] Rosie M. Butler,[1] Anthony Bayega,[1] Nelema Begum,[1] Farrah Bakr,[1] Kiran Dedhia,[1] Joshua Fisher,[1] Silvia Aguilar-Duran,[1] Charlotte Flanagan,[1] Aria A. Ghasemi,[1] Ricarda M. Hoffmann,[1] Nubia Castillo-Mosquera,[1] Elisabeth A. Nuttall,[1] Arisa Paul,[1] Ceri A. Roberts,[1] Emmanouil G. Solomonidis,[1] Rebecca Tarrant,[1] Antoinette Yoxall,[1] Carl Z. Beyers,[1] Silvia Ferreira,[1] Isabella Tosi,[1] Michael A. Simpson,[3] Emanuele de Rinaldis,[2] Tracey J. Mitchell,[1,*] and Sean J. Whittaker[1,*]

[1]St John's Institute of Dermatology, Division of Genetics and Molecular Medicine, Tower Wing, King's College London, London, United Kingdom; [2]BRC Bioinformatics Core, Tower Wing, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom; and [3]Department of Genetics, Division of Genetics and Molecular Medicine, Tower Wing, King's College London, London, United Kingdom

Key Points

- Aberrations in genome maintenance and DNA repair genes including *POT1* occur at a high frequency in Sézary syndrome.
- Candidate driver genes and affected pathways in Sézary syndrome show extensive heterogeneity but overlap with other mature T-cell lymphomas.

Sézary syndrome (SS) is a leukemic variant of cutaneous T-cell lymphoma (CTCL) and represents an ideal model for study of T-cell transformation. We describe whole-exome and single-nucleotide polymorphism array–based copy number analyses of CD4$^+$ tumor cells from untreated patients at diagnosis and targeted resequencing of 101 SS cases. A total of 824 somatic nonsynonymous gene variants were identified including indels, stop–gain/loss, splice variants, and recurrent gene variants indicative of considerable molecular heterogeneity. Driver genes identified using MutSigCV include *POT1,* which has not been previously reported in CTCL; and *TP53* and *DNMT3A*, which were also identified consistent with previous reports. Mutations in *PLCG1* were detected in 11% of tumors including novel variants not previously described in SS. This study is also the first to show *BRCA2* defects in a significant proportion (14%) of SS tumors. Aberrations in *PRKCQ* were found to occur in 20% of tumors highlighting selection for activation of T-cell receptor/NF-κB signaling. A complex but consistent pattern of copy number variants (CNVs) was detected and many CNVs involved genes identified as putative drivers. Frequent defects involving the *POT1* and *ATM* genes responsible for telomere maintenance were detected and may contribute to genomic instability in SS. Genomic aberrations identified were enriched for genes implicated in cell survival and fate, specifically PDGFR, ERK, JAK STAT, MAPK, and TCR/NF-κB signaling; epigenetic regulation (*DNMT3A, ASLX3, TET1-3*); and homologous recombination (*RAD51C, BRCA2, POLD1*). This study now provides the basis for a detailed functional analysis of malignant transformation of mature T cells and improved patient stratification and treatment. (*Blood.* 2016; 127(26):3387-3397)

Introduction

Primary cutaneous T-cell lymphomas (CTCL) represent a heterogeneous group of mature T-cell lymphomas. Targeted treatment options for advanced stages of CTCL are limited and associated with modest and short-lived responses.[1,2] Sézary syndrome (SS) is a leukemic variant of CTCL and represents an ideal model for defining the molecular pathways involved in the malignant transformation of mature T cells.

Recent studies[3-9] have revealed marked genomic heterogeneity in SS illustrated by extensive copy number variants (CNVs) and single-nucleotide variants (SNVs) affecting many genes, including known cancer genes, and selection for genes involved in T-cell receptor (TCR), JAK-STAT, and NF-κB signaling[3,5,6,9] and epigenetic regulation.[3,7,8]

Furthermore, nodal T-cell lymphomas (TCL) show considerable genomic overlap with CTCL.[10-13] Although adult T-cell leukemia lymphoma (ATLL) is associated with HTLV-1 transformation, both ATLL and CTCL are mature T-cell lymphomas of skin-homing memory CD4$^+$ T cells with marked clinical and phenotypic overlap. A recent comprehensive genomic study of ATLL has also shown striking similarities at the genomic level with high rates of CNV.[13] However the underlying basis for genomic instability, reflected in the high prevalence of CNVs detected in mature T-cell lymphomas, including CTCL, has yet to be clarified.

We have performed a discovery screen using next-generation sequencing (NGS) to analyze enriched tumor cell populations and matched normal DNA from samples obtained at diagnosis from untreated patients using whole-exome sequencing and SNP arrays. This was followed with a prevalence screen in a large cohort of SS samples using targeted resequencing.

Submitted February 11, 2016; accepted April 13, 2016. Prepublished online as *Blood* First Edition paper, April 27, 2016; DOI 10.1182/blood-2016-02-699843.

*T.J.M. and S.J.W. contributed equally to this study.

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

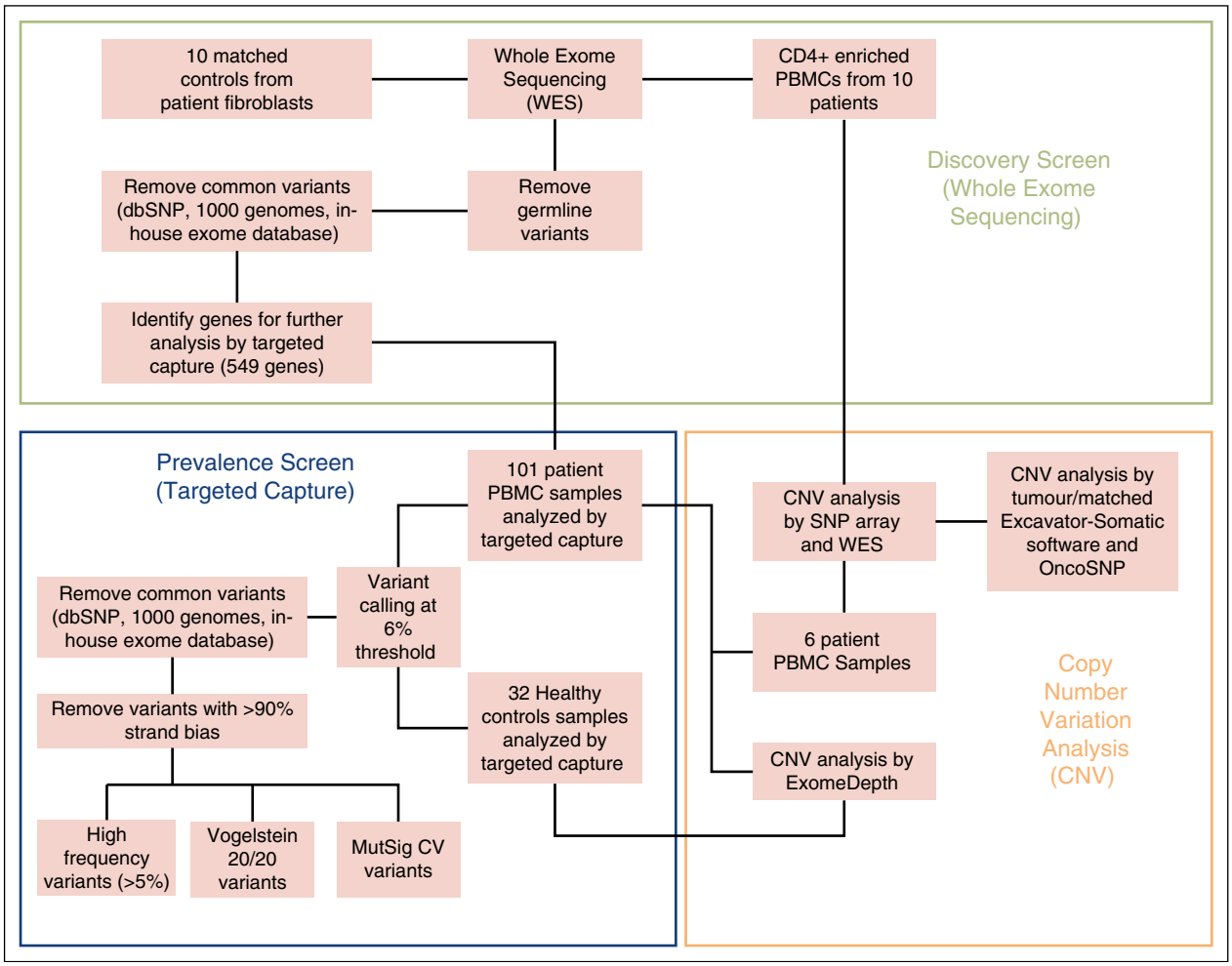


Figure 1. Workflow overview of experimental methods. Discovery screen (whole-exome sequencing), prevalence screen (targeted capture), and CNV analysis.

Patients, materials, and methods

Samples

All patients fulfilled the WHO-EORTC diagnostic criteria for SS.[2] Patient samples were obtained from the nationally approved CTCL research tissue bank (National Research Ethics Committee: 07/H10712/111+5); healthy control samples were obtained with the approval of the Guy's and St Thomas' Hospital Research Ethics Committee (EC01/301). Written and informed consent were obtained from all patients and volunteers. Discovery samples: DNA was extracted from $CD4^+$-enriched peripheral blood mononuclear cells (PBMCs) using RosetteSep (Stemcell Technologies, Cambridge, UK) and matched primary fibroblasts from skin explants obtained from 10 untreated patients with SS at diagnosis. Targeted capture samples: DNA was extracted from PBMCs of 101 SS and 32 healthy control samples (supplemental Table 1, available on the *Blood* Web site).

Whole-exome sequencing (WES) and targeted capture

The workflow overview is summarized in Figure 1. Paired-end sequencing library preparation was performed according to manufacturer's instructions and sequenced on an Illumina Hi-Seq2000 with reads aligned to Hg19 using Novoalign v2.07.11 and postalignment processing performed by picard tools.

For WES, Varscan2 Somatic was used to separate tumor variants from patient-matched fibroblasts. ANNOVAR was used for variant annotation.[14] Somatic and nonsynonymous variants were selected based on exclusion of variants in dbSNP, the 1000-genomes project, exome variant server, in-house exome database, and genes reported to be error prone in NGS analysis because of sequence repeats and high GC content.[15]

For targeted capture, Varscan2 and ANNOVAR were also used but the threshold on the minimum allele frequency for calling tumor variants was calibrated to account for the heterogeneity of tumor samples derived from PBMCs (supplemental Methods). Mpileup2cns was used for SNV and INDEL identification with $\geq$20$\times$ depth, $\geq$15 phred score, $\geq$6% minimum variant frequency, and read frequency $\leq$90% in either direction. Variants from 32 nonmatched healthy controls were used to identify tumor-specific variants and exclude sequencing artifacts. Variants selected from WES and targeted capture data were validated by Sanger sequencing on original tumor and additional skin, lymph node, and tumor-derived cDNA samples from the same patients.

Mutational pattern analysis

Several types of mutational pattern analysis were conducted using custom in-house Perl scripts. These included proportions of different types of variant (synonymous, nonsynonymous), SNV base change patterns, and mutation context (motif) analysis upstream and downstream at 3 bp.

Identification of SNV drivers

Several parallel criteria were used to identify genes affected by SNVs. These included MutSigCV,[16] the 20/20 rule[17] (see supplemental Methods for details), and simple frequency filtering of >5% after removing genes previously identified as problematic.[15] We also compared the list of candidate driver genes to those present in the network of cancer genes.[18]

Gene copy number analysis

Tumor-specific CNVs were identified through integrative analysis of discovery and targeted capture data generated using exome/targeted sequencing and SNP array technologies. Data from WES were analyzed by Excavator v2.2 in matched pairs. HumanOmni5Exome arrays were analyzed using OncoSNP, v1.4. Raw data (BAF and LRR) required for OncoSNP was extracted using Illumina Genome Studio software. Data from WES and SNP array were combined for final analysis (n = 16). Remaining prevalence samples (n = 91) were analyzed with ExomeDepth software[19] using the targeted capture data and 32 healthy controls (Figure 1). This analysis was restricted to targeted capture genes (n = 549) but allowed deeper resolution. The genotype array data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE80650 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE80650).

Pathway analysis

To investigate for significant perturbations at the pathway level, we performed a gene set enrichment analysis on WES and TC SNVs (supplemental Methods) using the MSigDB repositories. Pathway-level perturbations were quantified using 2 inter-related metrics. One metric, "fraction of pathway genes mutated," captures the proportion of pathway genes involved in nonsynonymous SNVs or indels across all patients. The second metric, "pathway perturbation frequency score," captures how often each pathway is perturbed as a proportion of all samples (both uncorrected and corrected for pathway size), assuming perturbation occurs if at least one of the pathway's genes is mutated.

Results

Whole-exome sequencing of CD4$^+$-enriched cells and matched controls (discovery screen)

For 10 CD4$^+$-enriched/control-matched DNA samples, we obtained depth >20 reads covering 82% to 95% of the target region across all samples, with a median of 91.12%. The most frequent type of variant effect (Figure 2A) was nonsynonymous (63%) followed by synonymous (26%) and stop-gain (3.5%). After filtering, we identified 824 somatic, nonsynonymous variants (750 genes; supplemental Table 2) from which we selected 549 genes for targeted capture analysis. Overall mutation rates for filtered somatic tumor variants were between 0.54 and 4.2 mutations per megabase with total nonsynonymous variants per tumor from 23 to 182, with a median of 98 comparable with other NGS studies of CTCL.[3,4,7] Furthermore this rate is similar to rates reported for other non-Hodgkin lymphomas[13] and distinct from tumors with a higher median (130-160) associated with specific carcinogens such as lung cancer and melanoma.[16] Two samples showed low levels of nonsynonymous variants (23 and 41 SNVs) with the youngest patient having the lowest number (supplemental Table 2). The most common type of nucleotide change (Figure 2B) was C>T and G>A (61%), in keeping with what has been observed already in many types of cancers.[20] Specifically, 42% of the C>T variants occurred at NpCpG sites, reflecting age-related spontaneous deamination at methylated CpG sites, and 27.5% occurred at NpCpC sites, but there were <1% CC>TT mutations (14 of 1520 SNVs affecting 6% of samples). Although interpretation is limited by our sample size, mutation context analysis shows consistency with several trinucleotide signatures (Figure 2C) identified in a recent study.[20]

Copy number analysis of WES sequence data from the discovery panel (Figure 3, outer track) revealed recurrent (>1) tumor-specific CNVs consisting of large (>1 MB) and focal (<1 MB) regions of amplification and deletion and confirmed using SNP array (Figure 3, inner track). Noticeably, 2 deleted regions on chromosomes 7 and 14 containing the γ and α TCR genes occurred in all samples, reflecting clonal TCR gene rearrangements.

Analysis of SNP array and WES CNV data confirms and extends previous findings in SS using array CGH and cytogenetic techniques,[21-25] including large complex chromosomal abnormalities such as isochromosome 17q (loss of 17p >70% and gain of 17q >50% of tumor samples),[26-28] and recurrent focal CNVs often affecting individual genes already subject to SNVs. Other frequently observed large CNVs were losses on 1p, 2p, 13p, and 10q, and gains on 8, consistent with previous reports.[3,21-25] However, array data showed gains on chromosome 4 that were not observed in the WES data, although these have been previously reported in SS.[21]

Targeted capture sequencing of 101 SS samples from patients (prevalence screen)

In the targeted custom capture of 549 genes—depth was between 149 and 848 reads starting from the list of all variants in tumors—we filtered out those also present in the healthy control panel. To enrich for somatic variants, we further filtered out variants present in dbSNP, the 1000 Genomes Project, Exome Variant Server (National Heart, Lung, and Blood Institute Exome Sequencing Project), and our in-house exome database. We also filtered out variants in genes reported to be error-prone in NGS analysis because of sequence repeats and high GC content.[15] From this list we focused on the final subgroup of nonsynonymous variants including Indels, stop gain/loss, splice variants, and indels, for a total of 1520 variants. There were between 2 and 93 variants per tumor, with a median of 13 variants per tumor (supplemental Table 3).

Analysis of CNVs in the targeted capture samples revealed a similar distribution of CNVs to the discovery WES samples. Overall, 453 of 549 genes in the targeted capture were affected by CNVs in at least 1 tumor (supplemental Table 4). Illustration of CNVs and mutations occurring in each gene are reported in Figure 4 and supplemental Figure 2.

Identification of potential driver genes

In view of the marked genomic heterogeneity, we applied different parallel criteria (supplemental Methods) to identify 21 potential driver genes based on SNVs and 42 genes based on high rates of CNVs (Figure 4; supplemental Tables 4 and 5).

We identified 51 genes with SNVs occurring in >5% of tumors. Analysis of all SNVs using MutSigCV[16] confirmed that 5 of these 51 genes (*DNMT3A, FAM47A, POT1, CADPS, TP53)* were mutated more often than expected by chance ($q < 0.1$) and all except *FAM47A* have been implicated as driver gene mutations in other cancer types. Two genes (*PREX2* and *PCLO*) had MutSigCV values close to significance ($q = 0.11$). In this data set, *ATM* and both *TP53* and *DNMT3A* identified with MutSigCV, have been previously defined as driver genes based on specific criteria applied to COSMIC and the Cancer Genome Atlas and functional validation.[17,29] In contrast *CSMD1, CSMD3, PCLO*, and *CNTNAP5*[16] have been identified as likely false-positive cancer genes, although overexpression of *CSMD3* has been associated with growth advantage in epithelial cells.[30] Finally, correlation of these putative driver genes with data sets annotating candidate driver genes[18] identified 21 genes as potential or established driver genes. Analysis of SNVs using the "20/20" rule[17] identified 16 of these 21 genes (supplemental Table 5) as either potential oncogenes (5 genes including *FAM47A, PLCG1*, and *GPR158*) or tumor-suppressor genes (11 genes including *POT1, ANK3, UNC13C*, *ATM, DNMT3A*, and *TP53).*

The overall frequency of SNVs affecting these 21 potential driver genes in our prevalence data set ranged from 5.5% to 19% (6-21), with 7 mutated genes affecting >10% of tumors. These consisted of known

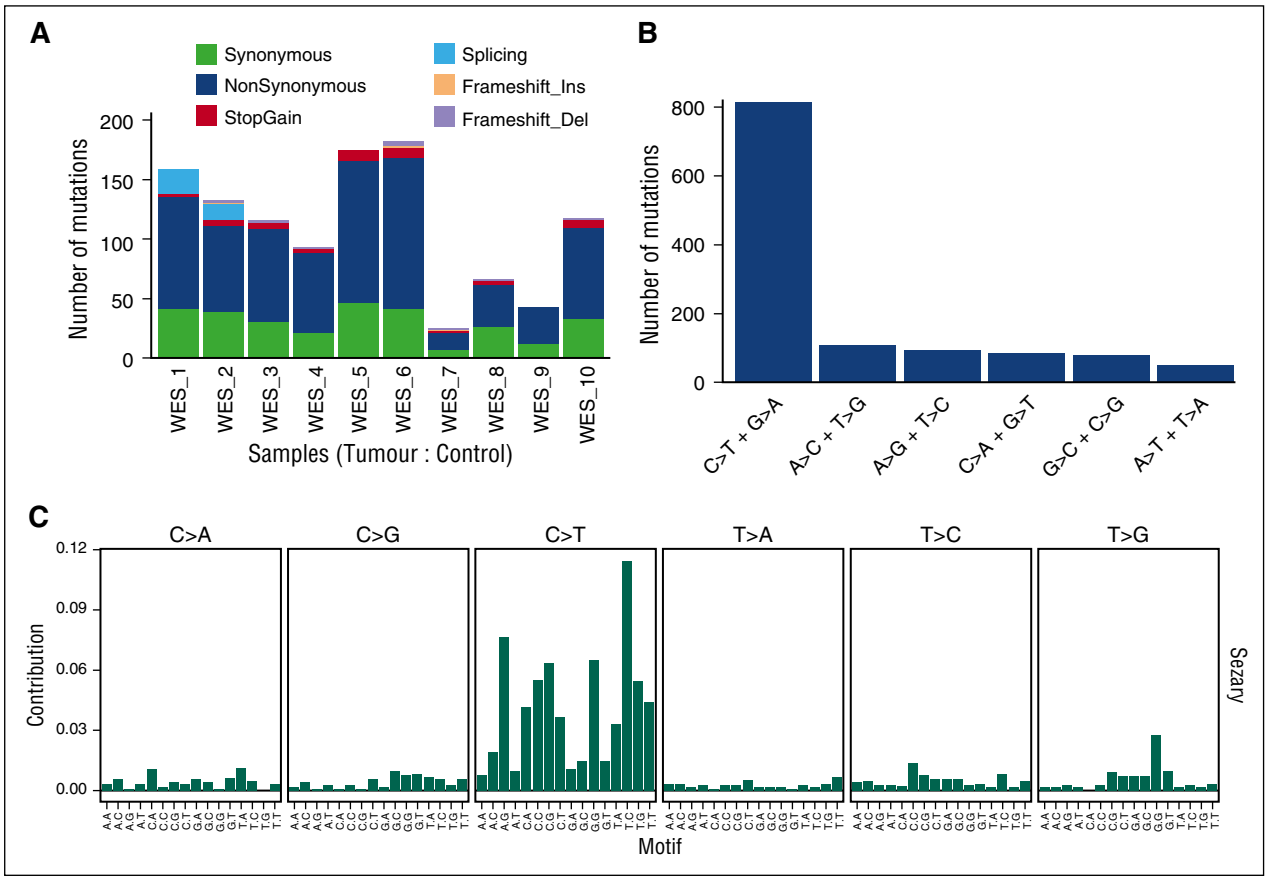


Figure 2. Summary of somatic tumor mutations in 10 whole-exome sequences. (A) Types and ratios of all somatic mutations detected in the discovery panel. (B) Distribution of all SNV base changes from the panel. (C) Trinucleotide analysis showing frequency of bases immediately 5′ and 3′ of the mutated bases.

tumor suppressor genes *TP53* and *FAT3/FAT4* (upstream regulators of the Hpo pathway[31]) and putative oncogenes such as *PLCG1*,[6,32] as well as *GPR98, CADPS*, and *CACNAIE*,[33] whose potential functional role in cancer has yet to be established. Of these potential driver genes, 15 had identical recurrent variants, with 2 genes having more than one recurrent variant, namely *GPR98* (2) and *PLCG1* (5).

Two tumor samples were notable for having relatively few SNVs detected in the whole-exome study (supplemental Table 2). Both tumors had aberrations affecting only a few of our potential driver genes including identical variants reported in COSMICv71 namely $POT1^{R117C}$, $JAK3^{A573V}$, and *PREX2* in one tumor and *PREX2, DNMT3A*, $STAT5B^{N642H}$, $FAT4^{R3615W}$, and $GPCR158^{R757C}$ in the other tumor, suggesting that these gene mutations could be sufficient for tumor development.

We identified 42 genes with at least 1 SNV and CNVs affecting >10% of tumors as additional potential driver genes. CNVs affected all 21 potential driver genes identified based on analysis of SNVs. We also identified a higher prevalence of aberrations (14%-48% of tumor samples) for other putative driver genes such as *DNAH9, ENPP2, ELAVL2, RFX6, PDCD11, GPR158*, *PTPRK*, *PRKCQ*, *BRAC2*, *TET1*, *RAD51C*, and *PREX2*. Notably few tumor samples had SNVs affecting individual *JAK* and *STAT* genes, but overall 55% of tumors had combined SNVs and CNVs affecting these genes including regulators of *STAT3* such as *SOCS7*.

Overall, 510 of 549 genes in our targeted capture had SNVs and/or CNVs reported in recent studies of SS and mycosis funcgoides (MFs) (supplemental Table 6). Specifically, SNVs affecting 15 of our 21 potential driver genes have been detected in recent studies of SS and/or MF.[3-9] Identical gene variants have also been functionally validated in these and other studies, namely *TP53*,[34], *POT1*,[35] *PLCG1*,[3,6,36-38] *ATM*,[39] *JAK3*,[40,41] *STAT3*,[42] and *STAT5B*[43-46] (Table 1). In addition, identical variants without functional validation have been reported in COSMICv71 for 4 other genes from our 21 potential drivers, namely $FAT3^{R4213C}$, $FAT4^{R3615W}$, $GPR158^{R757C}$, and $UNC13C^{R2037H/G2150R}$.[17]

Analysis of signaling pathways affected by SNVs

Our gene set enrichment analysis highlighted aberrations affecting numerous pathways involved in cell fate, cell survival, genome maintenance, and immune-related functions in both the TC and WES data sets (Figure 5; supplemental Figure 3). Notably, several pathways have SNVs affecting the same gene(s), specifically *JAKs, STATs, PLCG1*, and *TP53*. This analysis showed enrichment for genetic aberrations involving many of the putative driver genes affecting pathways including homologous recombination (*RAD51C, BRAC2, POLD1*: 45%) and DNA repair (*ATM, TP53, BRAC2*: 32%).

Gene perturbations (SNVs and CNVs; supplemental Figure 4) were also grouped into families with related functions including DNA repair (at least 1 perturbation 64%; >1 perturbation 36%), global epigenetic regulation (at least 1 perturbation 42%; >1 perturbation 14%), and programmed cell death (at least 1 perturbation 64%; >1 perturbation 37%) in line with well-known hallmarks of cancer.[47]

Lack of correlation with clinical outcome

A pairwise analysis of gene mutations using Bonferroni adjustment failed to identify any SNVs and CNVs, which either occurred together or were mutually exclusive. In addition, we did not detect any correlation

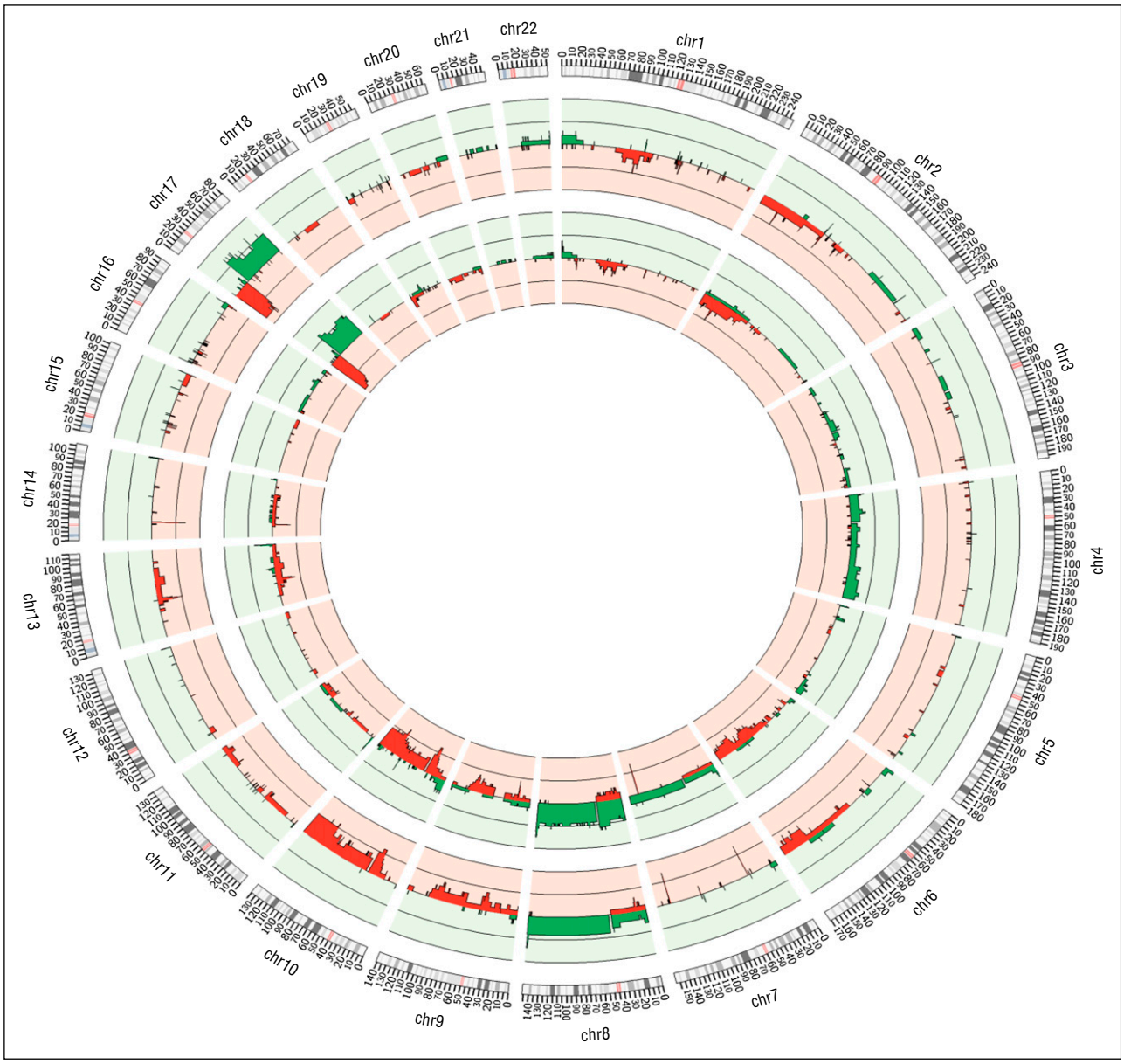


Figure 3. Summary of CNVs identified genome-wide for 10 WES samples. The outermost track represents results from WES samples anaylzed using Excavator; the middle track represents 16 SNP array samples analyzed with OncoSNP. In all the tracks, red denotes losses and green denotes gains.

between mutational load (including total SNVs and CNVs) and overall survival. Analysis of individual genes in the targeted capture data set identified 10 genes affected by SNV, CNV, or a combination of both, which were associated with a worse overall survival (supplemental Table 7). Only one of these 10 genes (*RELN*) was identified as a potential driver gene. However in view of lack of power, these results should be interpreted cautiously because the likelihood of chance occurrence is 1 in 5.

Validation of potential driver gene mutations

Sanger resequencing was performed on variants from 55 genes from the Discovery Panel (supplemental Table 8). A total of 97 of 101 variants were validated, consistent with other NGS cancer studies.[48-52] Similar proportions of variants were successfully validated on the prevalence screen data (134/139). Highly recurrent gene mutations in the targeted capture analysis were also validated in multiple and different tissue samples (blood, lesional skin, and lymph node) from the same patients at diagnosis and at disease progression, further supporting their role as candidate drivers (supplemental Table 8). In contrast for those patients who achieved a complete clinical remission after reduced intensity allogeneic transplantation, we could not detect specific gene variants, identified in the diagnostic samples, in the post-transplant tissue samples consistent with the absence of the original T-cell clone and a complete molecular remission.

The presence of *POT1*^{R117C} and *ATM*^{G2863V} variants were confirmed in additional blood and skin samples and at the transcriptional level in mRNA from enriched CD4^{+} tumor cells (Figure 6 and supplemental Table 8). Interestingly, for *POT1*^{R117C}, predominant expression of the mutant was detected over the wild-type allele. This is likely attributable to LOH affecting the wild-type allele. This was confirmed by sequencing genomic DNA from the same CD4^{+}-enriched tumor cells, in which *POT*^{R117C} was also predominantly detected (data not shown).

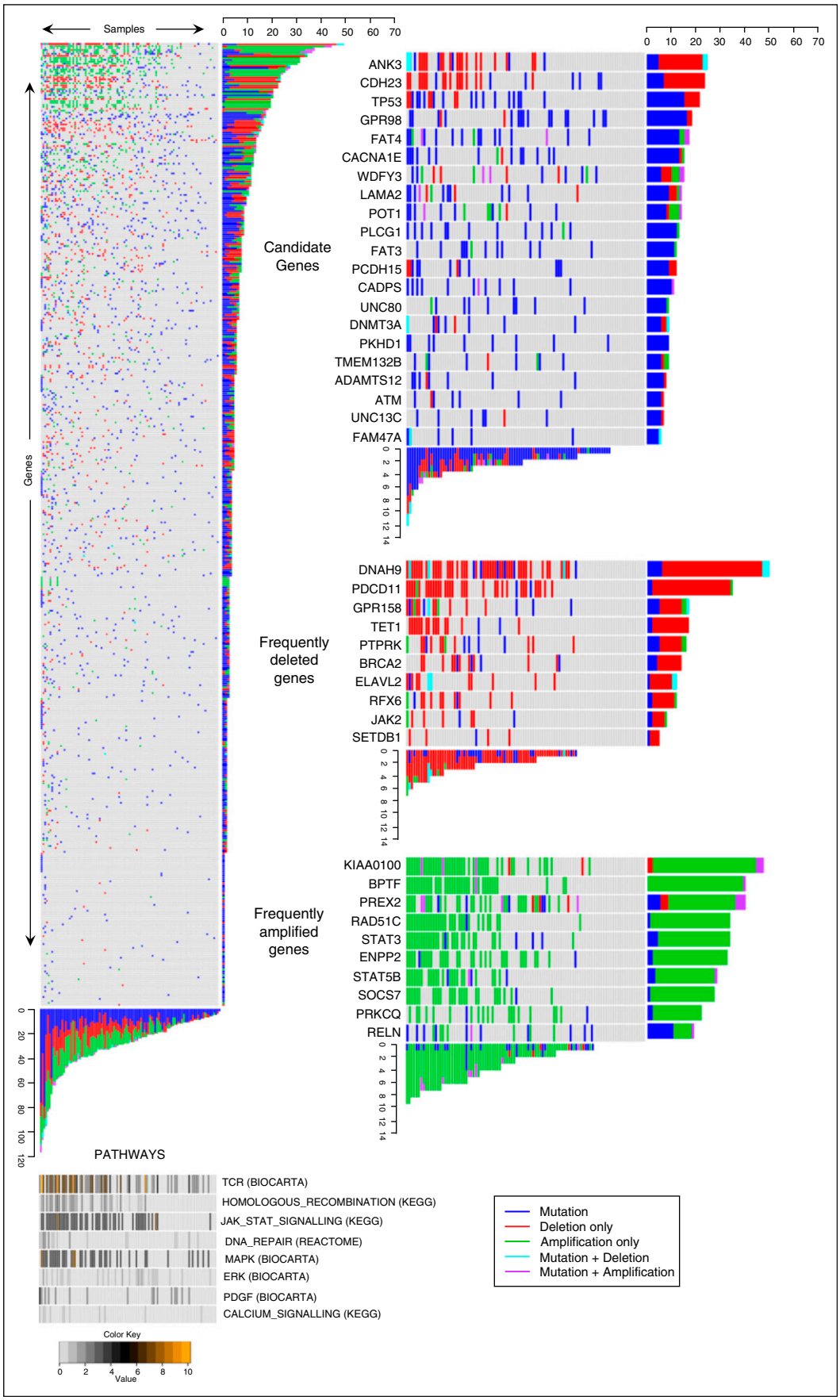


Figure 4.

Table 1. Identical gene variants reported in other studies

Gene	Variant	Functional validation	References
JAK3	A573V	Yes	40, 41
STAT3	Y640F	Yes	42
STAT5B	Y665F	Yes	43, 44
STAT5B	N642H	Yes	43, 45, 46
STAT5B	E150Q	Yes	46
PLCG1	S345F	Yes	3-6, 8, 9, 13, 36, 37
PLCG1	S520F	Yes	4, 6, 8
PLCG1	D342N	Yes	3, 37, 38
PLCG1	R48W	Not done	3, 4, 7, 13
PLCG1	E1163K	Not done	3, 4, 7, 8, 13
PLCG1	D1165H	Not done	4, 13
POT1	R117C	Yes	35
TP53	S127F	Yes	>10 papers
TP53	H20R	Not done	5
TP53	R37X	Not done	5
ATM	E2423K	Yes	39
CSMD1	A408V	Not done	4
ENPEP	V97L	Not done	7
LRP1B	R790Q	Not done	4

Several specific variants have been reported previously in CTCL and other malignancies. Functionally validated specific variants are indicated.

Finally, we sought to identify variants from our study that had been detected and/or functionally validated in previous studies. Several such variants (Table 1) were identified including *PLCG1*$^{(S345F, S520F, D342N)}$, 3,6,3[6]-3[8] *JAK3*^{A573V},4[0],4[1] *STAT3*^{Y640F},4[2] *STAT5B*$^{(Y665F\ N642H)}$,4[3]-4[5] *ATM*^{E2423K},[39] and *POT1*^{R117C},3[5]. Other candidate driver genes from our study including *PTPRK* and *FAT3* (as opposed to proven drivers such as *TP53*) have also been functionally validated based on analysis of different variants.[18]

Discussion

Our analysis of a large series of SS patient samples has identified novel variants and CNVs predicted to be potential drivers. Specifically, we found a high frequency of perturbations in *POT1*, *ATM*, and *BRCA2,* which are involved in genome maintenance. Dysregulation of genome maintenance processes may contribute to the high prevalence of structural variation observed in CTCL. We also detected mutations of *JAK-STAT*, *DNMT3A*, *TP53*, and *PLCG1*, genes previously reported as likely drivers in other lymphomas[17,53] and CTCL.[3-5,7-9] These putative driver gene mutations were present in diagnostic blood, skin, and node samples and samples collected at disease progression time points, but that were absent in samples from those patients who achieved a complete remission after stem cell transplant.

Genomic instability is a feature of SS with complex copy number variation reported using different techniques.[3,21,22,24] More than 50% of tumors had SNVs and/or CNVs affecting genes involved in DNA repair and telomere maintenance. Notably, a significant number of tumors (23%) had mutations and/or loss of genes involved in telomere maintenance such as *POT1* and *ATM*. *POT1* is part of the multiprotein Shelterin complex responsible for telomere length and loss of *POT1* function increases chromosomal instability.[54] All the *POT1* variants occurred in the oligonucleotide/oligosaccharide binding (OB) domains and loss of OB function has been shown to cause extensive telomere elongation[55] and frequent telomere fusions.[56] Recurrent *POT1* mutations have also been detected in a subset of patients with CLL (5%)[57] and ATLL.[13] In addition, the *POT1*^{R117C} variant has recently been identified as the cause of an inherited cancer syndrome in which loss of function causes an age-related increase in telomere length and genomic instability, contributing to the development of malignancies including lymphomas.[35] *ATM* is a PI3 kinase involved in the recognition of DNA double-strand breaks and recruitment of telomerase, and copy number losses have recently been reported in SS.[3,58] The *ATM*^{E2423K} variant is associated with loss of function in non–small-cell lung cancer.[39] *POT1* also represses the *ATM* damage response checkpoint.[56] Other genes involved in telomere maintenance include *ATRX* and *TEP1*, both with somatic mutations. We also detected frequent SNVs/CNVs affecting genes involved in homologous recombination such as *RAD51C*, *BRCA2*, and *POLD1*, a component of the DNA polymerase δ complex,[59] and loss of function *TP53* mutations, which are described in CTCL.[60,61] Previous mouse models showed that combined defects of telomerase and cell-cycle genes are associated with the development of mature T-cell lymphomas.[62] Loss of cell-cycle control (*TP53*), telomere maintenance (*POT1/ATM*), and DNA repair initiation (*BRCA2*) could contribute to the genomic instability, which is a consistent feature of SS.

Overall, 40% of tumors had somatic mutations affecting genes involved in TCR/NF-κB signaling. We detected recurrent *PLCG1* gene variants in 11 patients including several variants reported previously[3-9,37] in MF and SS, as well as PTCL, AITL, and ATLL.[13,36] The *PLCG1*^{S345F} variant has been shown to induce expression of both NFAT via IP_3 activation and NF-κB via DAG activation of PKC signaling. This mutation is predicted to impair the auto-inhibitory function of PLCG1, which limits TCR signaling downstream of receptor ligation.[6,63] In addition, a further recurrent variant (*PLCG1*^{D342N}) has been shown to increase inositol phosphate production in COS-7 cells.[38] It is not yet clear whether the other recurrent *PLCG1* variants identified affect this same catalytic function and whether these variants are sufficient alone to enable constitutive TCR signaling without costimulatory signals, but recent studies in SS and ATLL have detected activating *CD28* mutations and CTLA4-CD28 and ICOS-CD28 gene fusions.[3,13] Although we did not detect abnormalities of *CD28*, key findings in our study included mutations of other TCR/NF-κB signaling genes, notably *PRKCQ* (20% of cases) as well as *NFATC2*, *NFκB1*, and *PAK7*. *PRKCQ* belongs to the PKC family of serine/threonine kinases, is highly expressed in T cells, and has a pivotal role downstream of *PLCG1* in transducing TCR and costimulatory CD28 signals.[13] In all but 2 cases, *PRKCQ* aberrations were independent of *PLCG1* mutations. In ATLL, studies have identified gain-of-function *PRKCB* mutations and associated downstream activating mutations of *CARD11*, leading to enhanced NF-κB activation.[13] *CARD11*[3-5,7,8] activating mutations and *PRKCQ*[3,4] SNVs and CNVs have also been detected recently in SS.

Constitutive activation of NF-κB is described in CTCL,[64,65] and recurrent gain-of-function mutations affecting the *TNFRSF1B* gene in MF/SS have been shown to enhance noncanonical NF-κB signaling.[5] In PTCL, the t(5;9)(q33;22) results in an *ITK-SYK* fusion kinase, which induces constitutive TCR activation,[66] and *LCK* mutations have been documented in lymphoma.[67] These findings now provide compelling support for the hypothesis that the survival of malignant T cells in

Figure 4. Genomic data of 549 genes from 101 Sézary tumors identifies candidate driver genes. Heat map showing all genes (y-axis) and all tumors (x-axis) (left); pathways identified as frequently perturbed are aligned below the main panel (bottom). The color code represents the percentage of mutationally perturbed genes in each tumor sample for each pathway. Candidate driver genes showing high frequencies of SNVs (top right), frequently deleted genes (middle right), and frequently amplified genes (bottom right) are subsetted from the main panel.

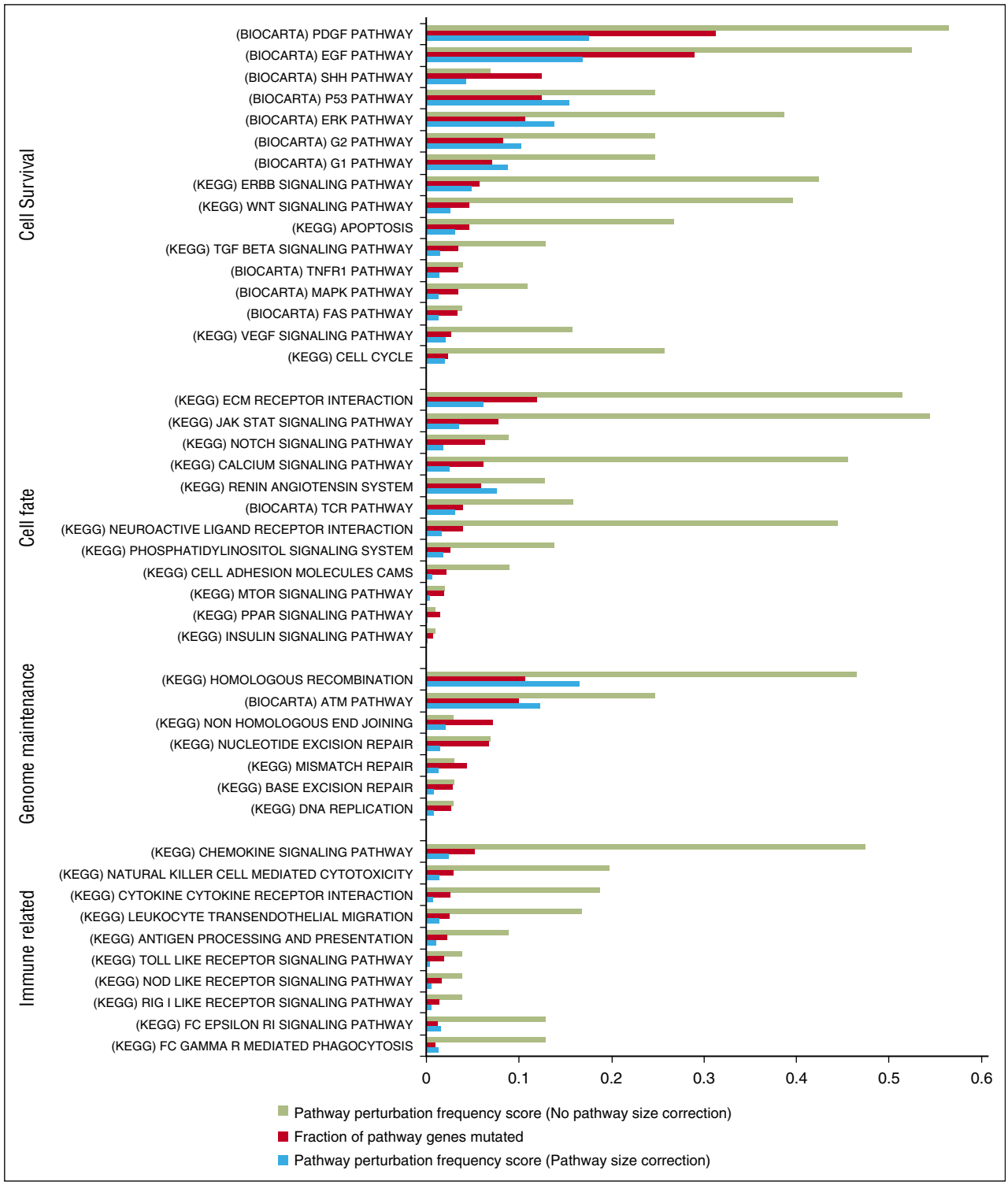


Figure 5. Gene-set enrichment analysis identifies key pathways. Highly perturbed pathways from the KEGG and BIOCARTA repositories were grouped according to related functions. The proportion of tumors showing perturbations in a given pathway is shown in green. Proportion of genes in each pathway with a perturbation across all affected tumors is shown in red bar. Pathway perturbation frequency score (blue bar) shows how often a pathway is perturbed but taken in consideration of the pathway size.

SS and other mature T-cell lymphomas is at least partly dependent on TCR and NF-κB signaling: In mature T-cell lymphomas such as SS, selection for activating mutations of *PLCG1*, *PRKCQ*, *PREX2*, and *CARD11* is likely to enhance cell survival if accompanied by appropriate costimulatory signals and resistance to *TNFRSF*-mediated apoptosis.[68-72]

Although we did not detect a high frequency of SNVs affecting individual *JAK STAT* genes, the presence of activating *JAK1/3*, *STAT3*, and *STAT5A/B* mutations and copy number gains of 17q including *STAT3* and *STAT5* could explain constitutive STAT3 activation in some cases of SS.[63,65,73-75] *STAT5B* mutations have recently been described in γδ T-cell lymphomas[10] and *JAK/STAT* mutations have

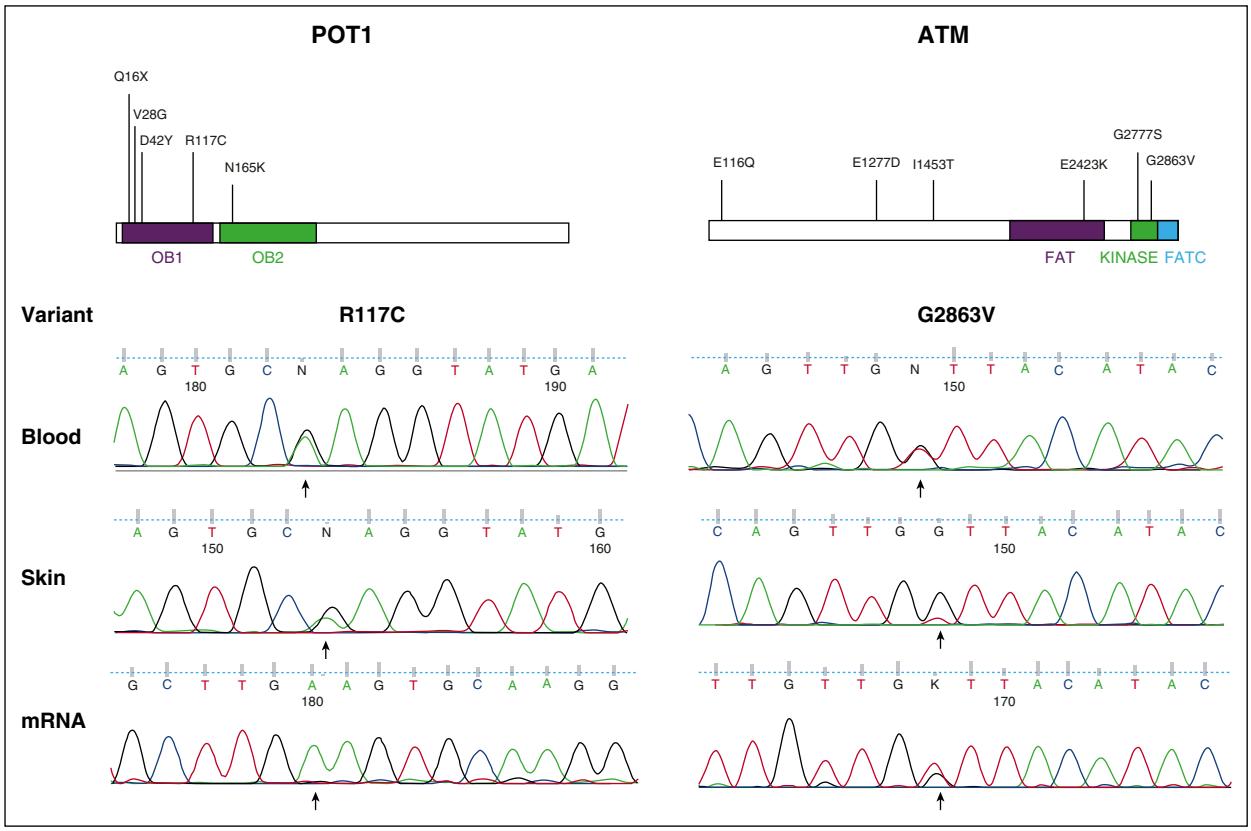


Figure 6. Validation of candidate driver genes *POT1* and *ATM*. Schematics of each protein are shown with the locations of mutations identified (top). Sanger validation of SNVs in lesional skin, diagnostic blood samples, and mRNA isolated from $CD4^{+}$-enriched tumor cells (bottom).

now been documented in other extranodal and nodal T-cell lymphomas including SS.[8,42,44,76] Recent studies have also shown that constitutive STAT3 expression in ALCL can be the result of gene fusions.[77] Previously we reported constitutive STAT3 protein expression in 10 patients with SS, all of whom were included in this study.[76] Overall, 6 of these cases showed copy number gains but no SNVs of *JAKs, STAT3*, or *SOCS7*, suggesting that other upstream events can lead to aberrant STAT3 signaling in SS. There is growing interest in the role of GPCRs in malignancy, and one, *SIPR1*, is involved in noncanonical activation of JAK STAT signaling in lymphoid cells via PI3K signaling.[78] Whether other GPCRs are implicated in JAK STAT signaling is unclear, but SNVs affecting 2 GPCRs (98, 158) were detected in 30% of tumors. One variant ($GPR158^{R757C}$) has been reported in COSMIC, and recent studies in ATLL also detected a high prevalence of *GPCR* aberrations including SNVs affecting *GPR183*.[13]

Overall, 40% of tumors had either SNVs and/or CNVs affecting genes involved in epigenetic regulation including *ASLX3, TET1*, *TET2*, and *DNMT3A*, which have been described in myeloid malignancies and lymphomas.[79,80] These include inactivating mutations of *TET1/2*, *ASLX3*, and *DNMT3A*, which have a role in DNA methylation, and *IDH2* mutations affecting histone methylation in AITL.[81] Both histone acetylation and methylation are known to be critical for T-cell differentiation and memory. Loss of epigenetic regulation in SS is reflected by promoter hypermethylation of multiple genes[82] and clinical responses to HDAC inhibitors such as Romidepsin.[83] In addition, recent studies in both CTCL and PTCL have shown mutations of *ARID1A/B* involved in chromatin remodeling.[3,84] These findings suggest that chromatin modification plays a key role in malignant transformation of mature T cells as recently described for B-cell non-Hodgkin lymphoma.[85]

Analysis of our data sets revealed that 42% of the C>T variants occurred at NpCpG sites, which could be consistent with at least 5 of 21 recently described signatures including age-related deamination of methylated cytosines.[21,86,87] Although UV-specific TP53 mutations (CC>TT transversions at pyrimidine sites) have been described previously in MF,[61] we only detected very rare CC-TT transversions in SS. MF is considered to be derived from skin resident memory T cells,[88] which may be exposed to environmental UV, and MF patients are often treated with phototherapy. In contrast, SS is thought to derive from central memory T cells. Further studies of larger data sets are required to define the mutational signatures associated with SS and other CTCL variants including MF.

In conclusion, our findings illustrate that the genomic landscape of SS is markedly heterogeneous. We suggest that the high prevalence of perturbations in genes maintaining genome integrity is a likely cause of the loss of genome stability in SS. Furthermore, there is selection for gene mutations/structural variation contributing to deregulation of key pathways regulating T-cell homeostasis, cell survival, and global epigenetic processes. These findings provide the basis for detailed functional analyses to define novel therapeutic targets for CTCL.

Acknowledgments

This research was supported by grants from The British Skin Foundation, Skin Tumour Unit Fund 461, Guy's and St Thomas' Charity, and the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. C.F., R.M.B., and R.T. were supported by the King's Bioscience Institute and the Guy's and St Thomas' Charity Prize Programme in Biomedical and Translational Science.

Authorship

Contribution: W.J.W., V.M.P., R.M.B., A.B., N.B., F.B., K.D., J.F., S.A.D., C.C.C., A.A.G., R.M.H., N.M., E.A.N., A.P., C.A.R., E.G.S., R.T., A.Y., C.Z.B., S.F., and I.T. performed experiments and data analysis; W.J.W., V.P., A.L., M.A.S., and E.d.R. performed bioinformatic analysis; W.J.W., V.P., A.L., F.B., and E.d.R. produced figures and tables; W.J.W., V.P., A.L., E.d.R., T.J.M., and S.J.W. wrote the manuscript; and T.J.M. and S.J.W. led the project.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Tracey J. Mitchell, St John's Institute of Dermatology, Division of Genetics and Molecular Medicine, Tower Wing, King's College London, London SE1 9RT, United Kingdom; e-mail: tracey.mitchell@kcl.ac.uk.

References

1. Agar NS, Wedgeworth E, Crichton S, et al. Survival outcomes and prognostic factors in mycosis fungoides/Sézary syndrome: validation of the revised International Society for Cutaneous Lymphomas/European Organisation for Research and Treatment of Cancer staging proposal. *J Clin Oncol.* 2010;28(31):4730-4739.
2. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood.* 2005;105(10):3768-3785.
3. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat Genet.* 2015;47(9):1011-1019.
4. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat Genet.* 2015;47(12):1426-1434.
5. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. *Nat Genet.* 2015;47(9):1056-1060.
6. Vaqué JP, Gómez-López G, Monsálvez V, et al. PLCG1 mutations in cutaneous T-cell lymphomas. *Blood.* 2014;123(13):2034-2043.
7. da Silva Almeida AC, Abate F, Khiabanian H, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat Genet.* 2015;47(12):1465-1470.
8. Kiel MJ, Sahasrabuddhe AA, Rolland DC, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sézary syndrome. *Nat Commun.* 2015;6(8470):8470.
9. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood.* 2015;126(4):508-519.
10. Küçük C, Jiang B, Hu X, et al. Activating mutations of STAT5B and STAT3 in lymphomas derived from γδ-T or NK cells. *Nat Commun.* 2015;6:6025.
11. Sakata-Yanagimoto M, Enami T, Yokoyama Y, Chiba S. Disease-specific mutations in mature lymphoid neoplasms: recent advances. *Cancer Sci.* 2014;105(6):623-629.
12. Palomero T, Couronné L, Khiabanian H, et al. Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. *Nat Genet.* 2014;46(2):166-170.
13. Kataoka K, Nagata Y, Kitanaka A, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. *Nat Genet.* 2015;47(11):1304-1315.
14. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
15. Fuentes Fajardo KV, Adams D, Mason CE, et al; NISC Comparative Sequencing Program. Detecting false-positive signals in exome sequencing. *Hum Mutat.* 2012;33(4):609-613.
16. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013;499(7457):214-218.
17. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LAJ Jr, Kinzler KW. Cancer genome landscapes. *Science.* 2013;339(6127):1546-1558.
18. An O, Dall'Olio GM, Mourikis TP, Ciccarelli FD. NCG 5.0: updates of a manually curated repository of cancer genes and associated properties from cancer mutational screenings. *Nucleic Acids Res.* 2016;44(D1):D992-D999.
19. Plagnol V, Curtis J, Epstein M, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics.* 2012;28(21):2747-2754.
20. Alexandrov LB, Nik-Zainal S, Wedge DC, et al; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain. Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415-421.
21. Mao X, Lillington D, Scarisbrick JJ, et al. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: identification of common genetic alterations in Sézary syndrome and mycosis fungoides. *Br J Dermatol.* 2002;147(3):464-475.
22. Caprini E, Cristofoletti C, Arcelli D, et al. Identification of key regions and genes important in the pathogenesis of sezary syndrome by combining genomic and expression microarrays. *Cancer Res.* 2009;69(21):8438-8446.
23. Salgado R, Servitje O, Gallardo F, et al. Oligonucleotide array-CGH identifies genomic subgroups and prognostic markers for tumor stage mycosis fungoides. *J Invest Dermatol.* 2010;130(4):1126-1135.
24. Iżykowska K, Przybylski GK. Genetic alterations in Sezary syndrome. *Leuk Lymphoma.* 2011;52(5):745-753.
25. Carbone A, Bernardini L, Valenzano F, et al. Array-based comparative genomic hybridization in early-stage mycosis fungoides: recurrent deletion of tumor suppressor genes BCL7A, SMAC/DIABLO, and RHOF. *Genes Chromosomes Cancer.* 2008;47(12):1067-1075.
26. Berger R, Baranger L, Bernheim A, Valensi F, Flandrin G. Cytogenetics of T-cell malignant lymphoma. Report of 17 cases and review of the chromosomal breakpoints. *Cancer Genet Cytogenet.* 1988;36(1):123-130.
27. Nowell PC, Finan JB, Vonderheid EC. Clonal characteristics of cutaneous T cell lymphomas: cytogenetic evidence from blood, lymph nodes, and skin. *J Invest Dermatol.* 1982;78(1):69-75.
28. Thangavelu M, Finn WG, Yelavarthi KK, et al. Recurring structural chromosome abnormalities in peripheral blood lymphocytes of patients with mycosis fungoides/Sézary syndrome. *Blood.* 1997;89(9):3371-3377.
29. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature.* 2013;502(7471):333-339.
30. Liu P, Morrison C, Wang L, et al. Identification of somatic mutations in non-small cell lung carcinomas using whole-exome sequencing. *Carcinogenesis.* 2012;33(7):1270-1276.
31. Katoh M. Function and cancer genomics of FAT family genes (review). *Int J Oncol.* 2012;41(6):1913-1918.
32. Kunze K, Spieker T, Gamerdinger U, et al. A recurrent activating PLCG1 mutation in cardiac angiosarcomas increases apoptosis resistance and invasiveness of endothelial cells. *Cancer Res.* 2014;74(21):6173-6183.
33. Natrajan R, Little SE, Reis-Filho JS, et al. Amplification and overexpression of CACNA1E correlates with relapse in favorable histology Wilms' tumors. *Clin Cancer Res.* 2006;12(24):7284-7293.
34. Tassi E, Zanon M, Vegetti C, et al. Role of Apollon in human melanoma resistance to antitumor agents that activate the intrinsic or the extrinsic apoptosis pathways. *Clin Cancer Res.* 2012;18(12):3316-3327.
35. Calvete O, Martinez P, Garcia-Pavia P, et al. A mutation in the POT1 gene is responsible for cardiac angiosarcoma in TP53-negative Li-Fraumeni-like families. *Nat Commun.* 2015;6(8383):8383.
36. Manso R, Rodríguez-Pinilla SM, González-Rincón J, et al. Recurrent presence of the PLCG1 S345F mutation in nodal peripheral T-cell lymphomas. *Haematologica.* 2015;100(1):e25-e27.
37. Caumont C, Gros A, Boucher C, et al. PLCG1 Gene Mutations Are Uncommon in Cutaneous T-Cell Lymphomas. *J Invest Dermatol.* 2015;135(9):2334-2337.
38. Everett KL, Bunney TD, Yoon Y, et al. Characterization of phospholipase C gamma enzymes with gain-of-function mutations. *J Biol Chem.* 2009;284(34):23083-23093.
39. Dhanasekaran SM, Balbin OA, Chen G, et al. Transcriptome meta-analysis of lung cancer reveals recurrent aberrations in NRG1 and Hippo pathway genes. *Nat Commun.* 2014;5(5893):5893.
40. Degryse S, de Bock CE, Cox L, et al. JAK3 mutants transform hematopoietic cells through JAK1 activation, causing T-cell acute lymphoblastic leukemia in a mouse model. *Blood.* 2014;124(20):3092-3100.
41. Koo GC, Tan SY, Tang T, et al. Janus kinase 3-activating mutations identified in natural killer/T-cell lymphoma. *Cancer Discov.* 2012;2(7):591-597.

42. Koskela HL, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. *N Engl J Med.* 2012;366(20):1905-1913.
43. Kontro M, Kuusanmäki H, Eldfors S, et al. Novel activating STAT5B mutations as putative drivers of T-cell acute lymphoblastic leukemia. *Leukemia.* 2014;28(8):1738-1742.
44. Rajala HL, Eldfors S, Kuusanmäki H, et al. Discovery of somatic STAT5b mutations in large granular lymphocytic leukemia. *Blood.* 2013;121(22):4541-4550.
45. Bandapalli OR, Schuessele S, Kunz JB, et al. The activating STAT5B N642H mutation is a common abnormality in pediatric T-cell acute lymphoblastic leukemia and confers a higher risk of relapse. *Haematologica.* 2014;99(10):e188-e192.
46. Yamada K, Ariyoshi K, Onishi M, et al. Constitutively active STAT5A and STAT5B in vitro and in vivo: mutation of STAT5 is not a frequent cause of leukemogenesis. *Int J Hematol.* 2000;71(1):46-54.
47. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011;144(5):646-674.
48. Nikolaev SI, Santoni F, Vannier A, et al. Exome sequencing identifies putative drivers of progression of transient myeloproliferative disorder to AMKL in infants with Down syndrome. *Blood.* 2013;122(4):554-561.
49. Pleasance ED, Stephens PJ, O'Meara S, et al. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature.* 2010;463(7278):184-190.
50. Turajlic S, Furney SJ, Lambros MB, et al. Whole genome sequencing of matched primary and metastatic acral melanomas. *Genome Res.* 2012;22(2):196-207.
51. Sutton LA, Ljungström V, Mansouri L, et al. Targeted next-generation sequencing in chronic lymphocytic leukemia: a high-throughput yet tailored approach will facilitate implementation in a clinical setting. *Haematologica.* 2015;100(3):370-376.
52. Vater I, Montesinos-Rongen M, Schlesner M, et al. The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing. *Leukemia.* 2015;29(3):677-685.
53. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med.* 2014;20(12):1472-1478.
54. Robles-Espinoza CD, Harland M, Ramsay AJ, et al. POT1 loss-of-function variants predispose to familial melanoma. *Nat Genet.* 2014;46(5):478-481.
55. Loayza D, De Lange T. POT1 as a terminal transducer of TRF1 telomere length control. *Nature.* 2003;423(6943):1013-1018.
56. Lazzerini Denchi E, Celli G, de Lange T. Hepatocytes with extensive telomere deprotection and fusion remain viable and regenerate liver mass through endoreduplication. *Genes Dev.* 2006;20(19):2648-2653.
57. Ramsay AJ, Quesada V, Foronda M, et al. POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. *Nat Genet.* 2013;45(5):526-530.
58. Ambrose M, Gatti RA. Pathogenesis of ataxia-telangiectasia: the next generation of ATM functions. *Blood.* 2013;121(20):4036-4045.
59. Valle L, Hernández-Illán E, Bellido F, et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. *Hum Mol Genet.* 2014;23(13):3506-3512.
60. Marrogi AJ, Khan MA, Vonderheid EC, Wood GS, McBurney E. p53 tumor suppressor gene mutations in transformed cutaneous T-cell lymphoma: a study of 12 cases. *J Cutan Pathol.* 1999;26(8):369-378.
61. McGregor JM, Crook T, Fraser-Andrews EA, et al. Spectrum of p53 gene mutations suggests a possible role for ultraviolet radiation in the pathogenesis of advanced cutaneous lymphomas. *J Invest Dermatol.* 1999;112(3):317-321.
62. Canela A, Martín-Caballero J, Flores JM, Blasco MA. Constitutive expression of tert in thymocytes leads to increased incidence and dissemination of T-cell lymphoma in Lck-Tert mice. *Mol Cell Biol.* 2004;24(10):4275-4293.
63. Koss H, Bunney TD, Behjati S, Katan M. Dysfunction of phospholipase Cγ in immune disorders and cancer. *Trends Biochem Sci.* 2014;39(12):603-611.
64. Sors A, Jean-Louis F, Pellet C, et al. Down-regulating constitutive activation of the NF-kappaB canonical pathway overcomes the resistance of cutaneous T-cell lymphoma to apoptosis. *Blood.* 2006;107(6):2354-2363.
65. Izban KF, Ergin M, Qin JZ, et al. Constitutive expression of NF-kappa B is a characteristic feature of mycosis fungoides: implications for apoptosis resistance and pathogenesis. *Hum Pathol.* 2000;31(12):1482-1490.
66. Pechloff K, Holch J, Ferch U, et al. The fusion kinase ITK-SYK mimics a T cell receptor signal and drives oncogenesis in conditional mouse models of peripheral T cell lymphoma. *J Exp Med.* 2010;207(5):1031-1044.
67. Wright DD, Sefton BM, Kamps MP. Oncogenic activation of the Lck protein accompanies translocation of the LCK gene in the human HSB2 T-cell leukemia. *Mol Cell Biol.* 1994;14(4):2429-2437.
68. Contassot E, French LE. Targeting apoptosis defects in cutaneous T-cell lymphoma. *J Invest Dermatol.* 2009;129(5):1059-1061.
69. Wu J, Nihal M, Siddiqui J, Vonderheid EC, Wood GS. Low FAS/CD95 expression by CTCL correlates with reduced sensitivity to apoptosis that can be restored by FAS upregulation. *J Invest Dermatol.* 2009;129(5):1165-1173.
70. Zoi-Toli O, Vermeer MH, De Vries E, Van Beek P, Meijer CJ, Willemze R. Expression of Fas and Fas-ligand in primary cutaneous T-cell lymphoma (CTCL): association between lack of Fas expression and aggressive types of CTCL. *Br J Dermatol.* 2000;143(2):313-319.
71. Jones CL, Wain EM, Chu CC, et al. Downregulation of Fas gene expression in Sézary syndrome is associated with promoter hypermethylation. *J Invest Dermatol.* 2010;130(4):1116-1125.
72. Serwold T, Hochedlinger K, Swindle J, Hedgpeth J, Jaenisch R, Weissman IL. T-cell receptor-driven lymphomagenesis in mice derived from a reprogrammed T cell. *Proc Natl Acad Sci USA.* 2010;107(44):18939-18943.
73. Netchiporouk E, Litvinov IV, Moreau L, Gilbert M, Sasseville D, Duvic M. Deregulation in STAT signaling is important for cutaneous T-cell lymphoma (CTCL) pathogenesis and cancer progression. *Cell Cycle.* 2014;13(21):3331-3335.
74. Nielsen M, Nissen MH, Gerwien J, et al. Spontaneous interleukin-5 production in cutaneous T-cell lymphoma lines is mediated by constitutively activated Stat3. *Blood.* 2002;99(3):973-977.
75. McKenzie RC, Jones CL, Tosi I, Caesar JA, Whittaker SJ, Mitchell TJ. Constitutive activation of STAT3 in Sézary syndrome is independent of SHP-1. *Leukemia.* 2012;26(2):323-331.
76. Jerez A, Clemente MJ, Makishima H, et al. STAT3 mutations unify the pathogenesis of chronic lymphoproliferative disorders of NK cells and T-cell large granular lymphocyte leukemia. *Blood.* 2012;120(15):3048-3057.
77. Crescenzo R, Abate F, Lasorsa E, et al; European T-Cell Lymphoma Study Group, T-Cell Project: Prospective Collection of Data in Patients with Peripheral T-Cell Lymphoma and the AIRC 5xMille Consortium "Genetics-Driven Targeted Management of Lymphoid Malignancies". Convergent mutations and kinase fusions lead to oncogenic STAT3 activation in anaplastic large cell lymphoma. *Cancer Cell.* 2015;27(4):516-532.
78. Yu H, Lee H, Herrmann A, Buettner R, Jove R. Revisiting STAT3 signalling in cancer: new and unexpected biological functions. *Nat Rev Cancer.* 2014;14(11):736-746.
79. Lemonnier F, Couronné L, Parrens M, et al. Recurrent TET2 mutations in peripheral T-cell lymphomas correlate with TFH-like features and adverse clinical parameters. *Blood.* 2012;120(7):1466-1469.
80. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med.* 2010;363(25):2424-2433.
81. Cairns RA, Iqbal J, Lemonnier F, et al. IDH2 mutations are frequent in angioimmunoblastic T-cell lymphoma. *Blood.* 2012;119(8):1901-1903.
82. van Doorn R, Zoutman WH, Dijkman R, et al. Epigenetic profiling of cutaneous T-cell lymphoma: promoter hypermethylation of multiple tumor suppressor genes including BCL7a, PTPRG, and p73. *J Clin Oncol.* 2005;23(17):3886-3896.
83. Whittaker SJ, Demierre MF, Kim EJ, et al. Final results from a multicenter, international, pivotal study of romidepsin in refractory cutaneous T-cell lymphoma. *J Clin Oncol.* 2010;28(29):4485-4491.
84. Schatz JH, Horwitz SM, Teruya-Feldstein J, et al. Targeted mutational profiling of peripheral T-cell lymphoma not otherwise specified highlights new mechanisms in a heterogeneous pathogenesis. *Leukemia.* 2015;29(1):237-241.
85. Hopp L, Löffler-Wirth H, Binder H. Epigenetic Heterogeneity of B-Cell Lymphoma: DNA Methylation, Gene Expression and Chromatin States. *Genes (Basel).* 2015;6(3):812-840.
86. Fischer A, Vázquez-García I, Illingworth CJ, Mustonen V. High-definition reconstruction of clonal composition in cancer. *Cell Reports.* 2014;7(5):1740-1752.
87. Lee DH, Pfeifer GP. Deamination of 5-methylcytosines within cyclobutane pyrimidine dimers is an important component of UVB mutagenesis. *J Biol Chem.* 2003;278(12):10314-10321.
88. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood.* 2010;116(5):767-771.

2016 127: 3387-3397
doi:10.1182/blood-2016-02-699843 originally published online April 27, 2016

Candidate driver genes involved in genome maintenance and DNA repair in Sézary syndrome

Wesley J. Woollard, Venu Pullabhatla, Anna Lorenc, Varsha M. Patel, Rosie M. Butler, Anthony Bayega, Nelema Begum, Farrah Bakr, Kiran Dedhia, Joshua Fisher, Silvia Aguilar-Duran, Charlotte Flanagan, Aria A. Ghasemi, Ricarda M. Hoffmann, Nubia Castillo-Mosquera, Elisabeth A. Nuttall, Arisa Paul, Ceri A. Roberts, Emmanouil G. Solomonidis, Rebecca Tarrant, Antoinette Yoxall, Carl Z. Beyers, Silvia Ferreira, Isabella Tosi, Michael A. Simpson, Emanuele de Rinaldis, Tracey J. Mitchell and Sean J. Whittaker

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

5.3 Additional figures for chapter 5

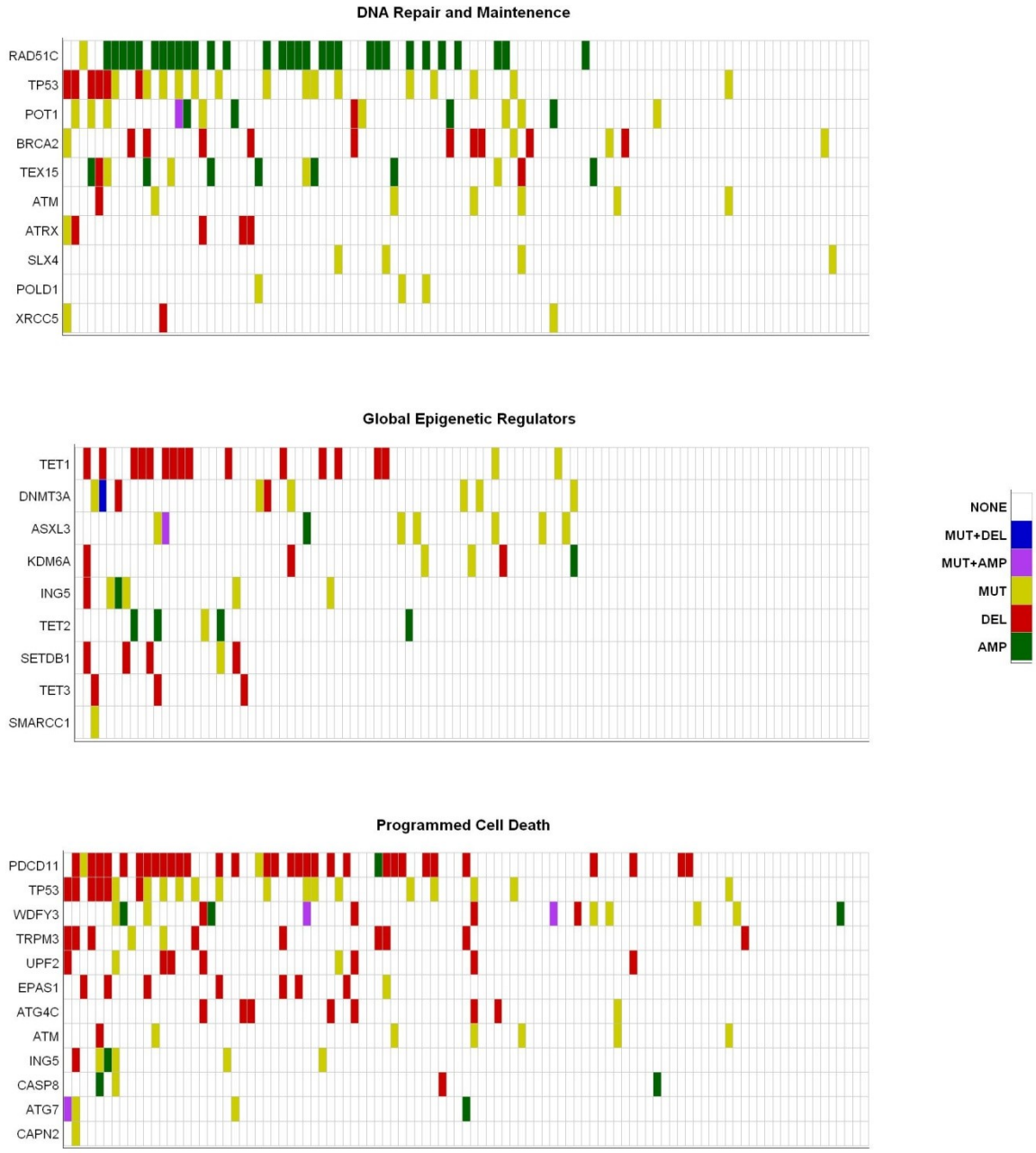


Figure 5.1. Heatmap of tumours affected by perturbations in gene families.

DNA Repair and Maintenance (Top): 64% of Tumours are affected by a perturbation in at least 1 of these genes and 36% have perturbations in more than 1 of these genes. **Global Epigenetic Regulators (Middle):** 42% of Tumours are affected by a perturbation in at least 1 of these genes and 14% have perturbations in more than 1 of these genes. **Programmed Cell Death (Bottom):** 64% of Tumours are affected by a perturbation in at least 1 of these genes and 37% have perturbations in more than 1 of these genes.

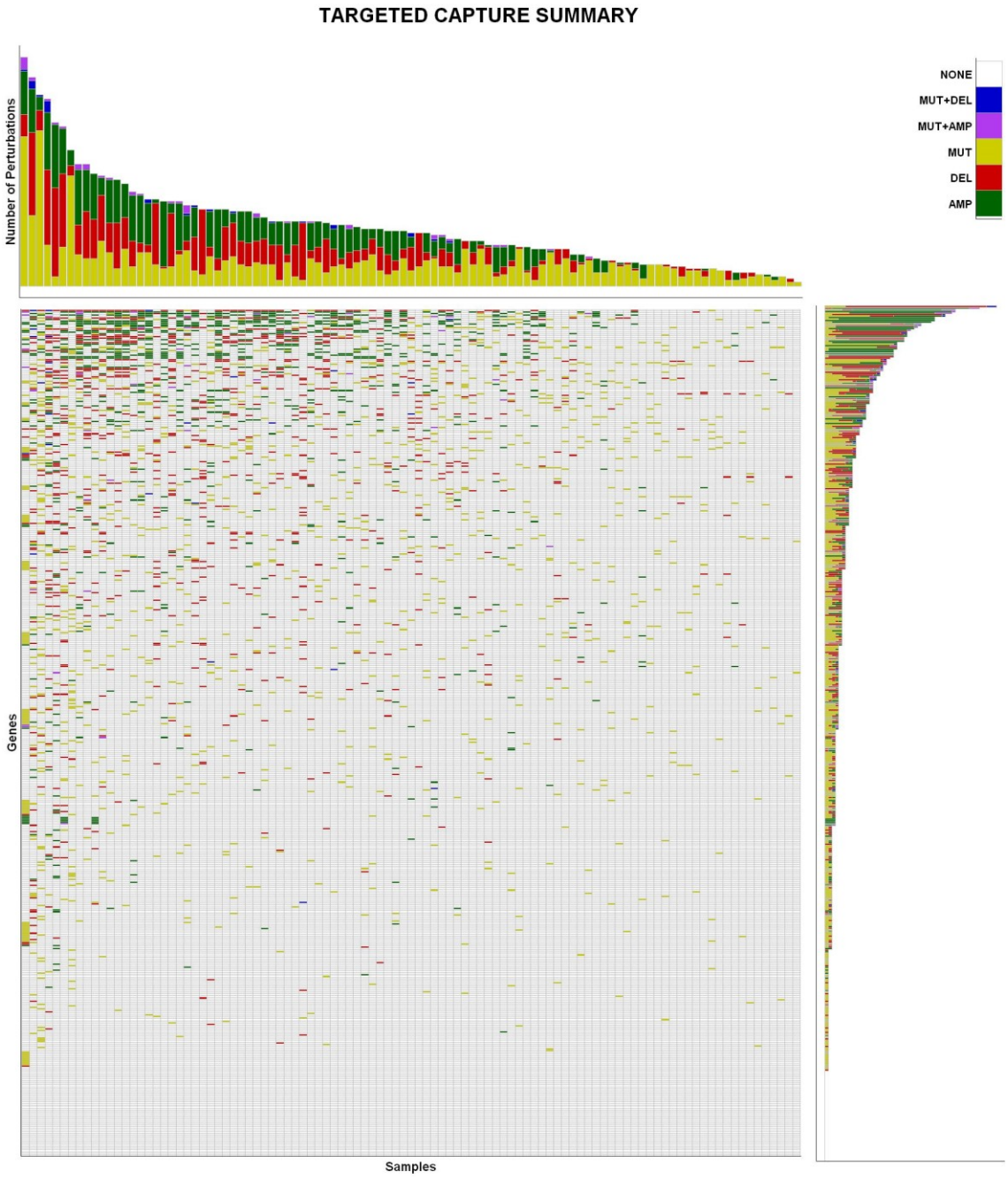


Figure 5.2. Summary heatmap showing targeted capture gene alterations.

This was a preliminary alternative heatmap figure that was produced as a summary of the data produced in this chapter. An alternative version produced by the BRC was chosen for inclusion in the manuscript.

6 Discussion

6.1 Focal deletions of 9p21 in CTCL

In CTCL several previous studies have demonstrated loss of the broader region of 9p21 including *CDKN2A*, *CDKN2B* and *MTAP* (22, 23, 27, 127). In addition recent NGS studies have confirmed with higher resolution that focal deletions in this region are very common (16, 49), but whilst the majority of losses cover *MTAP* and *CDKN2A* and *CDKN2B* the lost regions can vary considerably in size and position. Two important aspects of these NGS studies is that (a) they focus on advanced stage disease and (b) derive tumour samples exclusively from circulating blood making it currently unfeasible to contrast differences in tumour sub compartments such as skin. A novel finding of this study is that it shows that *MTAP* is often lost as an earlier event in MF, preceding the loss of *CDKN2A*. The SS skin tumour sub compartment also shows a propensity for independent *MTAP* loss on 9p21, although the reason for this remains unclear. Four independent QPCR probes (two on *MTAP* and two on *CDKN2A*) were used to identify copy number variations across 9p21 and the pattern of loss was in many cases consistent with the presence of several small focal deletions being present across the region. Data from the next generation sequencing study was able to validate the QPCR methodology as 31 overlapping samples were shared by both studies. These samples were used to reassess copy number variation and the synteny between both datasets was over 90%. When this study was designed the available DNA samples of early MF tumours were not of sufficient purity or quantity to undergo NGS, it would have been of tremendous benefit to validate *MTAP* loss as an early event in MF. Two cases were selected for microdissection and subsequent QPCR of the higher purity micro-dissected tumour DNA which allowed more specific interrogation of MF tumour tissue. This experiment served as a further validation to the original QPCR results. Recent advancements in laser capture microdissection and advancements in NGS library preparation would now make NGS studies of early MF very feasible.

One curious outcome of the *MTAP* CN study was the apparent difference in independent *MTAP* CN loss between the SS blood and SS skin compartments. It is conceivable that the reason for this may be related to tumour subclone selection pressure being different in the tumour microenviroment of the skin compared to the microenviroment of the circulating blood.

Gene expression analysis of *MTAP* was concordant with *MTAP* gene loss but loss of *MTAP* expression also seemed to be occurring in samples where loss of *MTAP CN* was not detected.

One limitation of this study is that we were not able to validate the promoter hypermethylation hypothesis in patient samples where *MTAP* gene CN was normal but *MTAP* expression was reduced. The reason for this was that the DNA samples available which were used in the gene CN study were predominantly legacy samples extracted from PBMCs and therefore were not pure tumour DNA. This would have made it challenging to detect *MTAP* promoter methylation differences between patient and healthy control samples using bisulphite sequencing as PBMCs are a mixed population of cells. Any differences would not be able to robustly attributed to tumour cells and any signal that may have been present would have been weakened by the presence of other (non-tumour) cell types in the DNA sample. It would be interesting to investigate *MTAP* promoter hypermethylation using bisulphite sequencing in a new cohort with DNA and RNA extracted from CD4+ enriched tumour samples.

6.2 The effects of MTAP loss on type I PRMT activity

Part of the work contributing to this thesis has shown that *MTAP* is lost as an early event in CTCL. Several previous studies suggest that *MTAP* may act as a tumour suppressor (141, 253, 324). However, questions remain about the possible mechanisms by which MTAP may be contributing to tumourigenesis in CTCL.

Evidence largely points to a build-up of MTAP's known substrate, MTA, as playing the main role (141, 144, 242). There has been a study suggesting there is at least one other as yet unknown mechanism that MTAP may act via which differs to its enzymatic function (272).

However, the majority of evidence hints that MTA accumulation is the primary tumourigenic driver. Downstream of MTA, processes that may be disrupted include the conversion of peutricine into spermidine (141). Spermidine itself promotes autophagy (136) by inhibiting acetyltransferase EP300 (325). Autophagy is an important part of cellular homeostasis and under certain circumstances acts as a backup to apoptosis (113, 133, 136). The long term effects of reduced or partially compromised homeostatic functions such as autophagy are not well understood but there is suggestion that it may create an environment that is conducive to the acquisition of further tumorigenic advantages (326, 327). It is also conceivable that loss of *MTAP* may indirectly prevent autophagy induced programmed cell death, which alone may be a minor contribution to tumourigenesis but amongst a plethora of other damaging events could prove significant.

The accumulation of MTA is also known to inhibit PRMT enzymes such as PRMT1 (267). This has been shown to prevent the deactivation of STAT1 and lead to overactive JAK/STAT signalling (267). Furthermore, in melanoma asymmetric and symmetric dimethyl-argnine methylation was shown to correlate with loss of MTAP and increased MTA (270). This justifies similar investigation in other cancers including CTCL.

As part of this thesis it has been demonstrated that all PRMT family members except PRMT8 are expressed in the CTCL cell line SeAx. However, this cell line was difficult to transfect making downstream experiments challenging. As proof of principle the HEK293 cell line was chosen to assay the effects of *MTAP* loss on asymmetrically dimethyled-arginine. This cell line was also shown to express all PRMTs except PRMT8. Reduced MTAP expression was successfully demonstrated in HEK293s transfected with *MTAP* siRNAs at the mRNA level compared to untransfected HEK293s or HEK293s transfected with scrambled siRNAs. One preliminary experiment also showed reduced *MTAP* expression at the protein level though this would require further demonstration to confirm. The same western blot assay also seemed to show subtle suggestions of reduced asymmetrically dimethylated-arginine levels on at least 5

proteins ranging in size from 40KDa to 150KDa, though again this experiment was only performed once and lacked a clear signal. This would need to be repeated to draw any conclusion from. This indicates that reduced *MTAP* expression may affect type I PRMT activity although which of the type I PRMTs are affected is not yet known. It is thought that PRMT1 has the highest activity of the PRMT type I family (266) but the other type I PRMTs; 3, 4 and 6 are also plausible candidates for inhibition by MTA and their inhibition may also be contributing to reduced asymmetrically dimethylated arginine. The PRMT family have already been shown to regulate STAT1 (267) but other STATs such as STAT3 may be affected by PRMT1 or other PRMTs. The STAT family have been shown to have a selective pressure targeting them in SS by mutation and copy number variation. This has been demonstrated as part of the NGS study contributing to this thesis and by others previously (50, 90). Therefore, establishing other mechanisms by which abrogated STAT signalling could occur in CTCL is of significant interest. This work has suggested that several type I PRMT targets may be affected by reduced MTAP expression so further experiments may be able to confirm this. If successful, then identifying them could help identify further pathways where loss of *MTAP* may have functional consequences.

6.3 High throughput screening for putative driver genes in CTCL

It is well established that high genomic instability is a consistent feature of CTCL. Recent NGS studies (16, 28, 29, 49-51) strongly support and extend the previous lower resolution and array based studies however the underlying cause of genomic instability associated with this malignancy has remained elusive. Establishing the cause of genomic instability is likely to pay dividends for both scientific and clinical purposes. A major novel finding of this study is that the screen highlighted a large proportion of tumours as having perturbations in genes involved in genome maintenance which now warrant functional investigation.

Global epigenetic regulators were also significantly affected by aberrations and this study confirms several genes of this class previously reported in CTCL including *DNMT3A*, *SETDB1*

and the *TET* family (50). Global epigenetic regulators not previously reported in CTCL but highlighted by this study include *ASXL3* and *KDM6A,* which were perturbed in 8% and 6% of tumours respectively. The ASXL family are polycomb group proteins involved in maintaining transcriptional repression of developmental genes. They have been reported in other cancers including *ASXL3* in melanoma and *ASXL2* in ATLL (328). The histone modifier *KDM6A* catalyses the demethylation of tri/dimethylated histone H3 and has been reported as altered by several mutation types in 24% of bladder cancers (329).

Loss of apoptotic regulatory machinery is a hallmark feature of all cancers so it is not unexpected that SS tumours share this feature. This study confirms loss of several previously reported apoptotic regulators including *ATM* and *TP53* (16, 49, 51). In addition, several further candidates that may impact programmed cell death processes were identified including; *PDCD11* (also known as *ALG4*) in 35% of tumours and *TRPM3* in 11% of tumours. The majority of defects in *PDCD11* identified in this study were loss or deletion. The FAS cell death pathway is known to be activated by upregulation of *PDCD11* (330) so loss is likely to have a significant effect on apoptotic activation via this pathway. *TRPM3* is known to play a major role in the regulation of oncogenic autophagy in clear cell renal cell carcinoma (331). Whilst a specific role for autophagy in CTCL is currently unclear, defects in autophagy are emerging to have key roles in cancers and are intricately linked to programmed cell death (138). In addition the histone deacetylase inhibitor vorinostat, is thought to partially act by attenuating oncogenic autophagy (139, 140) offering the possibility of more targeted therapy to patients identified as having defective autophagy components.

Another signal transduction pathway related to proliferation that showed numerous mutations in this study is the MAP-Kinase network, with 53% of tumours showing defects in at least one of 14 genes listed in the KEGG MAPK gene list. Other NGS studies have also reported defects in MAP-kinase downstream components including *MAPK1* (51) and upstream components such

as PI3K members (29). The MAPK pathways are arguably the most well studied pathways in cancer biology so it perhaps comes as no surprise to find they are disrupted in CTCL.

Given the diversity of SNVs and CNVs identified by this study and others and the insight gained from these studies that CTCLs are largely driven by alterations at the pathway level; a potential improvement to this study could have been to increase the size of the discovery cohort. If the discovery cohort were larger, then the number of true driver genes included in the targeted capture would also have been larger and we may have covered a higher percentage of genes in the key pathways identified.

As we move into an era of personalized genomic medicine we will begin to diagnose and treat malignancies based more on genetic events and their consequences for specific molecular pathways, rather than traditional criteria such as tissue of origin and morphological or histological features. Therefore, characterising the many facets of genomic variation in CTCL and other cancers will become increasingly important and continue to grow over the coming years. Extensive genomic screens are the first important milestones that must be undertaken for cancer studies in the genomic era.

7 Conclusions

The work produced through the course of this thesis has identified putative driver genes of CTCL using two different approaches. Firstly, the 9p21 candidate region was analysed in detail using several methods. This work highlights that *MTAP* loss often occurs as an early event prior to the loss of the more established driver gene CDKN2A, on 9p21. This suggests that *MTAP* is under selective loss and may therefore act as a tumour suppressor gene. Secondly, a larger number of putative driver genes were identified using a high throughput next generation sequencing approach. Many candidates identified by this approach are known to play important roles in other malignancies but have been highlighted for the first time as potential

driver genes of CTCL. The identification of multiple candidate driver genes supports the hypothesis that a variety of different driver events underlie CTCL.

Both approaches used in this thesis are in concordance with the established knowledge that genetic events in CTCLs are highly variable and support the hypothesis of this thesis. However, the high throughput screen has hinted at a possible explanation as to why. The screen identified two groups of genes which may drive this variability. The first is the identification of genetic events occurring in genes that maintain genome integrity, particularly genes that facilitate double stranded break repair. Without adequate maintenance of the genome it is easily envisaged that mutations, copy number changes and structural variants such as 9p21 loss and isochromosome 17 would accumulate over several cell divisions. Those alterations that convey an advantage to the tumour, namely those that contribute to the hallmarks of cancer (68, 69) are maintained in the lineage in a manor akin to Darwinian natural selection. The 9p21 region has been highlighted by numerous previous studies as one such region (16, 22, 23, 27, 49, 127), and is thought to be particularly susceptible to deleterious DSB repair events (332) which may explain why 9p21 often undergoes loss in a malignancy highly affected by loss of DSB repair genes.

An interesting feature of SS shown by the whole exome data is the existence of significant parts of the genome where loss of heterozygousity (LOH) events have occurred. The existence of numerous LOH regions supports the model that SS tumours lose the ability to accurately repair DSB events. Moreover, the presence of numerous LOH regions could also indicate that DSB events are taking place in SS tumours more regularly than in other cells. It is conceivable that this is linked to a high rate of cell division as is the case in many cancers but the presence of LOH events is intriguing and justifies further research into DSB events and the related genes in SS and other CTCLs.

The INK and ARF cell cycle regulators transcribed by *CDKN2A* and *CDKN2B* are by far the most obvious tumour suppressors on 9p21 and this has perhaps overshadowed the potential role of

MTAP. As a metabolic gene in close proximity on the loci it could easily be viewed as coincidentally deleted as part of a larger deletion (243, 333). Contrary to expectations, *MTAP* appears to be deleted as an early stage event in CTCL and at a higher frequency than *CDKN2A* in MF and SS skin. This implication of these findings is that there is selective pressure for CTCL tumours to lose *MTAP* at an early stage in development prior to a selective loss of *CDKN2A*. One further insightful perspective is that selective loss of *MTAP* may have been overlooked by a broad genomic screen. This demonstrates a niche for candidate gene based studies to build on initial screens.

A common thread that emerges through this thesis is the selection for epigenetic mechanisms of gene silencing by the tumour. The high throughput screen identified numerous genes that had undergone CNVs and/or mutations with key roles in altering the epigenetic landscape of the cell. Novel genes identified include *ASXL3* and *KDM6A*, whilst others previously reported include DNMT3A and the TET family. On 9p21 the *MTAP* gene is silenced by several mechanisms; whilst CNVs are very common, repression also occurs by other mechanisms. No mutations were detected in *MTAP* implicating epigenetic mechanisms as a possible cause. The cell line work performed as part of this study indicates that DNA methylation could be a possible mechanism of *MTAP* silencing in CTCL tumours and the identification of loss of DNMT3A in several NGS studies including this one could possibly facilitate hypermethlation of the *MTAP* promoter. Additional promotor methylation studies of *MTAP* would also be required, ideally in enriched CD4+ tumour samples from patient samples to confirm this hypothesis and support the cell line studies conducted in chapter 3. However, the possibility of other epigenetic silencing mechanisms such as histone modifications driving *MTAP* repression cannot be excluded at this stage. Changes to global histone modification genes *ASXL3* and *KDM6A*, identified during the course of this thesis, open the possibility that changes to local histone structure may underlie repression of MTAP as well as numerous other genes that may not be detected in any genomic screen. This perspective gives insight into a whole alternative class of driver genes known as epigenetic driver genes (334) which may be of particular

importance to CTCL given that we detected 42% of tumours with aberrations in global epigenetic regulatory genes, and similarly high proportions were also detected by others (50). The identification of multiple candidate driver genes affecting global epigenetic regulation supports the hypothesis that different driver genes may contribute to disruption in the same gene regulatory networks.

7.1 Future work

The findings of this thesis have identified several areas that could be further investigated. The first study has identified *MTAP* as a candidate tumour suppressor and it would be highly pertinent to identify possible mechanisms of how *MTAP* may act as a tumour suppressor. The key aim here would be to identify further downstream targets that may be therapeutically exploitable. Preliminary work has shown that loss of MTAP may affect type I PRMT activity. Which type I PRMTs specifically may be affected by loss of MTAP has yet to be confirmed. Furthermore, an evaluation of type II PRMTs would be of great interest, as this subgroup may also be affected by loss of MTAP in CTCL. It would also be of significant interest to identify specific targets downstream of PRMTs that are affected by reduced MTAP expression. These studies could be extended to more appropriate models such as cells from primary tumours that show loss of *MTAP* copy number or gene expression. This would help to establish if the pattern of reduced PRMT activity occurs in primary tumour cells, rather than relying on in-vitro manipulated cell lines and potentially identify therapeutic targets which are currently undiscovered.

One important general consideration for future NGS studies is that gene expression data is an invaluable accompaniment to genomic data and would ideally accompany genomic datasets wherever possible in the form of transcriptomic data. This adds a level of robustness to any genomic screen as it would serve to validate copy number changes, small indels and stopgains in genes that may be associated with changes in gene expression and thus identify genes that are haplosufficient or haploinsufficient. In the absence of readily obtainable sample material,

as can be the case with many diseases, candidate gene based RT-QPCR is the next best alternative.

The other major arm of investigation in this thesis surrounds the considerable amount of genomic data that has now been collected on late stage CTCL. Future studies should focus on defining the roles of pathways and individual candidate driver genes using loss and gain of function experiments. Future NGS studies should also attempt to focus on early stage disease as the technical challenges associated with it are now surmountable and early putative drivers have not yet been identified. A comprehensive understanding the role of these will likely form a crucial part of early intervention strategies. Histological sections of early stage CTCL could undergo laser capture microdissection to extract DNA from enriched early stage tumour cells. Antibodies that target CD4+ cells could be used and stained cells from histological features such as Pautrier microabsesses could be laser captured to yield enriched DNA from early stage tumours. Library prep methods for NGS have developed considerably since the initial studies in this thesis were undertaken and sequencing can now be efficiently performed on samples with as little as 20ng of DNA.

One group of genes identified as putative drivers in this thesis and by others (50) is the master epigenetic regulators. This group of genes is likely to affect transcription across significant parts of the genome by targeting the regulation of genes which are not affected by SNVs or CNVs. This subgroup of downstream targets are known as epigenetic driver genes (159, 335) and cannot be identified by genomic screening. Further work addressing their roles in mature T-cells could highlight important targets for therapeutic intervention. One approach could be to use a Chip-Seq approach in T-cells to identify regions targeted by master epigenomic regulators, specifically to identify regions that are differentially regulated in CTCL tumour cells compared to healthy control T-cells. The first step could involve checking transcript levels of master epigenomic regulators using simple QPCR, ideally followed by assays at the protein level. The next step would be to Chip-seq candidates that show the highest expression change.

This approach could be used to identify genes, that may be therapeutically targetable, that would not be identified by genomic screening. In addition, this approach could be combined to identify genes that are targeted by both epigenetic changes; and genes that had been previously identified as putative drivers by NGS genomic screens. Identifying genes that are targeted by two (or more) different selective pressures such as CNVs and epigenetic alterations strongly increases the likelihood of that gene being a bona fide driver gene. This has been shown to be the case for *MTAP* in this thesis,

7.2 Concluding remarks

Before this thesis *MTAP* was known to be lost in CTCL but it was not known that *MTAP* loss frequently occurred independently of *CDKN2A*. It is also a novel finding that additional mechanisms of gene silencing act to supress *MTAP* expression in CTCL. A selective loss of *MTAP*, by copy number loss and as well as by epigenetic mechanisms, supports the hypothesis that MTAP acts as a tumour suppressor in CTCL. Evidence from other tumour types regarding *MTAP* loss adds further weight to this view. Prior to the NGS study produced as part of this thesis it was widely known that genomic instability was a key feature of CTCLs. However, the underlying cause of this instability remains elusive. This study is the first to propose that genomic instability is acquired in CTCL tumours because genes involved in genome maintenance are highly perturbed. This study is also the first to highlight this in CTCL at the level of gene families and networks rather at the level of individual driver genes.

A gene network orientated view must be taken to comprehend the constellation of SNVs and CNVs and integrate our understanding of focal deletions on 9p21 and other perturbation in SS. Genomic perturbations in CTCL, as in many cancers, undoubtedly encompasses numerous passengers as well as driver events. But despite the range of genetic events and lack of consistency at the level of individual genes, patterns begin to emerge when a molecular pathway or network is viewed as a functional unit rather than genes in isolation. This is highly

evident in this thesis as multiple components involved in genome maintenance, global epigenetic regulators and pathways that regulate growth and cell fate are perturbed in significant proportions of SS tumours. Broad challenges remain in cancer genomics with regards to standardising; (i) sample collection and processing, (ii) data generation, and (iii) analysis techniques. Applying a consistent approach as well as integrating genomic, transcriptomic and proteomic data in a systems biology approach will likely deliver the most efficient benefit in the long term to both patients and the research community.

8 References

1. Bradford PT, Devesa SS, Anderson WF, Toro JR. Cutaneous lymphoma incidence patterns in the United States: a population-based study of 3884 cases. Blood. 2009;113(21):5064-73.
2. Kim E, Hess S, Richardson S, Newton S, Showe L, Benoit B, et al. Immunopathogenesis and therapy of cutaneous T cell lymphoma. Journal of Clinical Investigation. 2005;115(4):798-812.
3. Agar N, Wedgeworth E, Crichton S, Mitchell T, Cox M, Ferreira S, et al. Survival outcomes and prognostic factors in mycosis fungoides/Sézary syndrome: validation of the revised International Society for Cutaneous Lymphomas/European Organisation for Research and Treatment of Cancer staging proposal. Journal of Clincal Oncolgy. 2010;28(31):4730-9.
4. Willemze R, Jaffe ES, Burg G, Cerroni L, Berti E, Swerdlow SH, et al. WHO-EORTC classification for cutaneous lymphomas. Blood. 2005;105(10):3768-85.
5. Weinstock MA, Gardstein B. Twenty-year trends in the reported incidence of mycosis fungoides and associated mortality. American Journal of Public Health. 1999;89(8):1240-4244.
6. Chuang TY, Su WP, Muller SA. Incidence of cutaneous T cell lymphoma and other rare skin cancers in a defined population. Journal of American Academy of Dermatology. 1990;23(2):254-6.
7. Kim Y, Liu H, Mraz-Gernhard S, Varghese A, Hoppe R. Long-term outcome of 525 patients with mycosis fungoides and Sezary syndrome: clinical prognostic factors and risk for disease progression. Archives of Dermatology. 2003;139(7):857-66.
8. Kazakov DV, Burg G, Kempf W. Clinicopathological spectrum of mycosis fungoides. Journal of the European Academy of Dermatology and Venereology. 2004;184(397-415).
9. Quaglino P, Pimpinelli N, Berti E, Calzavara-Pinton P, Alfonso Lombardo G, Rupoli S, et al. Time course, clinical pathways, and long-term hazards risk trends of disease progression in patients with classic mycosis fungoides: a multicenter, retrospective follow-up study from the Italian Group of Cutaneous Lymphomas. Cancer. 2012;118(23):5830-9.
10. Smoller BR SM, Wood GS, Whittaker SJ. Histopathology and Genetics of cutaneous T-cell lymphoma. Hematol Oncol Clin North AM. 2003(17):1277-311.
11. Nickoloff BJ. Light-microscopic assessment of 100 patients with patch/plaque stage mycosis fungoides. American Journal of Dermatopathology. 1988(10):469-77.
12. Diamandidou E, Colome-Grimmer M, Fayad L, Duvic M, Kurzrock R. Transformation of mycosis-fungoides/Sezary syndrome: clinical characteristics and prognosis. Blood. 1998(92):1150-9.
13. Zinzani PL, Ferreri AJM, Cerroni L. Mycosis fungoides. Critical Reviews in Oncology/Hematology. 2008;65(2):172-82.
14. Bernengo M, Novelli M, Quaglino P, Lisa F, De Matteis A, Savoia P, et al. The relevance of the CD4+ CD26- subset in the identification of circulating Sézary cells. British Journal of Dermatology. 2001;144(1):125-35.
15. Wood G, Hong S, Sasaki D, Abel E, Hoppe R, Warnke R, et al. Leu-8/CD7 antigen expression by CD3+ T cells: comparative analysis of skin and blood in mycosis fungoides/Sézary syndrome relative to normal blood values. Journal of American Academy of Dermatology. 1990;22(4):602-7.
16. Wang L, Ni X, Covington K, Yang B, Shiu J, Zhang X, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. Nature Genetics. 2015;47(12):1426-34.
17. Olsen E, Vonderheid E, Pimpinelli N, Willemze R, Kim Y, Knobler R, et al. Revisions to the staging and classification of mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Cancer (EORTC). Blood. 2007;110(6):1713-22.

18. Rook A, Gottlieb S, Wolfe J, Vowels B, Sood S, Niu Z, et al. Pathogenesis of cutaneous T-cell lymphoma: implications for the use of recombinant cytokines and photopheresis. Clinical and Experimental Immunology. 1997;107(Supplement 1):16-20.
19. Campbell S, Peters S, Zirwas M, Wong H. Immunophenotypic diagnosis of primary cutaneous lymphomas: a review for the practicing dermatologist. Journal of Clinical and Aesthetic Dermatology. 2010;3(10):21-5.
20. Russell-Jones R. Diagnosing erythrodermic cutaneous T-cell lymphoma. British Journal of Dermatology. 2005;153(1):1-5.
21. Vonderheid EC. On the diagnosis of erythrodermic cutaneous T-cell lymphoma. Journal of Cutaneous Pathology. 2006;33(1):27-42.
22. Salgado R, Servitje O, Gallardo F, Vermeer M, Ortiz-Romero P, Karpova M, et al. Oligonucleotide array-CGH identifies genomic subgroups and prognostic markers for tumor stage mycosis fungoides. Journal of Investigative Dermatology. 2010;130(4).
23. van Doorn R, van Kester MS, Dijkman R, Vermeer MH, Mulder AA, Szuhai K, et al. Oncogenomic analysis of mycosis fungoides reveals major differences with Sezary syndrome. Blood. 2009;113(1):127-36.
24. Carbone A, Bernardini L, Valenzano F, Bottillo I, De Simone C, Capizzi R, et al. Array-based comparative genomic hybridization in early-stage mycosis fungoides: recurrent deletion of tumor suppressor genes BCL7A, SMAC/DIABLO, and RHOF. Genomes Chromosomes Cancer. 2008;47(12):1067-75.
25. Laharanne E ON, Bonnet F. Genome-wide analysis of cutaneous T-cell lymphomas identifies three clinically relevant classes. J Invest Dermatol. 2010;130(1707-1708).
26. Mao X, Lillington D, Scarisbrick J, Mitchell T, Czepulkowski B, Russell-Jones R, et al. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: identification of common genetic alterations in Sézary syndrome and mycosis fungoides. British Journal of Dermatology. 2002;147(3):464-75.
27. Laharanne E, Chevret E, Idrissi Y, Gentil C, Longy M, Ferrer J, et al. CDKN2A-CDKN2B deletion defines an aggressive subset of cutaneous T-cell lymphoma. Mod Pathol. 2010;23(4):547-58.
28. McGirt L, Jia P, Baerenwald D, Duszynski R, Dahlman K, Zic J, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. Blood. 2015;126(4):508-19.
29. Ungewickell A, Bhaduri A, Rios E, Reuter J, Lee C, Mah A, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. Nature Genetics. 2015.
30. Vaqué J, Gómez-López G, Monsálvez V, Varela I, Martínez N, Pérez C, et al. PLCG1 mutations in cutaneous T-cell lymphomas. Blood. 2014;123(13):2034-43.
31. Iżykowska K, Przybylski G. Genetic alterations in Sézary syndrome. Leukemia & Lymphoma. 2011;52(5):745-53.
32. Cuneo A, Castoldi G. Mycosis funcgoides/Sezary's syndrome. Atlas of Genetics and Cytogenetics in Oncology and Haematology. 2005;9(3):242-3.
33. Wieselthier JS, Koh HK. Sézary syndrome: diagnosis, prognosis, and critical review of treatment options. . 22. 1990;3(381-401).
34. Sentis HJ, Willemze R, Scheffer E. Histopathologic studies in Sézary syndrome and erythrodermic mycosis fungoides: a comparison with benign forms of erythroderma. J Am Acad Dermatol. 1986;15(6):1217-26.
35. Trotter MJ, Whittaker SJ, Orchard GE, Smith NP. Cutaneous histopathology of Sézary syndrome: a study of 41 cases with a proven circulating T-cell clone. J Cutan Pathol 1997;24(5):286-91.
36. Scheffer E, Meijer C, van Vloten W, Willemze R. A histologic study of lymph nodes from patients with the Sézary syndrome. Cancer. 1986;57(12):2375-80.
37. Sibaud V, Beylot-Barry M, Thiébaut R, Parrens M, Vergier B, Delaunay M, et al. Bone marrow histopathologic and molecular staging in epidermotropic T-cell lymphomas. American Journal of Clinical Pathology. 2003;119(3):414-23.

38. Vonderheid EC, Bernengo MG, Burg G, Duvic M, Heald P, Laroche L, et al. Update on erythrodermic cutaneous T-cell lymphoma: report of the International Society for Cutaneous Lymphomas. J Am Acad Dermatol. 2002;46(1):95-106.
39. Jones C, Ferreira S, McKenzie R, Tosi I, Caesar J, Bagot M, et al. Regulation of T-plastin expression by promoter hypomethylation in primary cutaneous T-cell lymphoma. Journal of Investigative Dermatology. 2012;132(8):2042-9.
40. Karenko L, Kähkönen M, Hyytinen ER, Lindlof M, Ranki A. Notable losses at specific regions of chromosomes 10q and 13q in the Sézary syndrome detected by comparative genomic hybridization. 112. 1999;3(392-395).
41. Vermeer M, van Doorn R, Dijkman R, Mao X, Whittaker S, van Voorst Vader P, et al. Novel and highly recurrent chromosomal alterations in Sézary syndrome. Cancer Research. 2008;68(8):2689-98.
42. Fischer TC, Gellrich S, Muche JM, Sherev T, Audring H, Neitzel H, et al. Genomic aberrations and survival in cutaneous T cell lymphomas. Journal of Investigative Dermatology. 2004;122(3):579-86.
43. Karenko L, Hahtola S, Ranki A. Molecular cytogenetics in the study of cutaneous T-cell lymphomas (CTCL). Cytogenetic and Genome Research. 2007;118((2-4)):353-61.
44. Scarisbrick JJ, Woolford AJ, Calonje E, Photiou A, Ferreira S, Orchard G, et al. Frequent abnormalities of the p15 and p16 genes in mycosis fungoides and sezary syndrome. Journal of Investigative Dermatol 2002;118(3):493-9.
45. Barba G, Matteucci C, Girolomoni G, Brandimarte L, Varasano E, Martelli MF, et al. Comparative genomic hybridization identifies 17q11.2 approximately q12 duplication as an early event in cutaneous T-cell lymphomas. Cancer Genetics and Cytogenetics. 2008;184(1):48-51.
46. Scarisbrick JJ WA, Russell-Jones R, Whittaker SJ. Loss of heterozygosity on 10q and microsatellite instability in advanced stages of primary cutaneous T-cell lymphoma and possible association with homozygous deletion of PTEN. Blood. 2000;95(9):2937-42.
47. Limon J, Nedoszytko B, Brożek I, Hellmann A, Zajączek S, Lubiński J, et al. Chromosome aberrations, spontaneous SCE, and growth kinetics in PHA-stimulated lymphocytes of five cases with Sézary syndrome. Cancer Genetics and Cytogenetics. 1995;83(1):75-81.
48. Caprini L, Cristofoletti C, Arcelli D, Fadda P, Citterich M, Sampogna F, et al. Identification of key regions and genes important in the pathogenesis of sézary syndrome by combining genomic and expression microarrays. Cancer Research. 2009;69(21):8438-46.
49. Choi J, Goh G, Walradt T, Hong B, Bunick C, Chen K, et al. Genomic landscape of cutaneous T cell lymphoma. Nature Genetics. 2015.
50. Kiel M, Sahasrabuddhe A, Rolland D, Velusamy T, Chung F, Schaller M, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sézary syndrome. Nature Communications. 2015;6(8470).
51. da Silva Almeida A, Abate F, Khiabanian H, Martinez-Escala E, Guitart J, Tensen C, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. Nature Genetics. 2015;47(12):1465-70.
52. Burg G, Kempf W, Cozzio A, Feit J, Willemze R, S Jaffe E, et al. WHO/EORTC classification of cutaneous lymphomas 2005: histological and molecular aspects. Journal of Cutaneous Pathology. 2005;32(10):647-74.
53. Diwan A, Prieto V, Herling M, Duvic M, Jone D. Primary Sézary syndrome commonly shows low-grade cytologic atypia and an absence of epidermotropism. American Journal of Clinical Pathology. 2005;123(4):510-5.
54. Campbell J, Clark R, Watanabe R, Kupper T. Sézary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. Blood. 2010;116(5):767-71.
55. Whittaker S, Marsden J, Spittle M, R. RJ, Dermatologists BAo, Group UKCL. Joint British Association of Dermatologists and U.K. Cutaneous Lymphoma Group guidelines for the

management of primary cutaneous T-cell lymphomas. British Journal of Dermatology. 2003;149(6):1095-107.
56. Jones G, Kacinski B, Wilson L, Willemze R, Spittle M, Hohenberg G, et al. Total skin electron radiation in the management of mycosis fungoides: Consensus of the European Organization for Research and Treatment of Cancer (EORTC) Cutaneous Lymphoma Project Group. Journal of American Academy of Dermatology. 2002;47(3):364-70.
57. Olsen EA, Bunn PA. Interferon in the treatment of cutaneous T-cell lymphoma. Hematology/Oncology Clinics of North America. 1995;9(5):1089-107.
58. Duvic M, Cather JC. Emerging new therapies for cutaneous T-cell lymphoma. Dermatologic Clinics. 2000;18(1):147-56.
59. Kaye F, Bunn PA J, Steinberg S, Stocker J, Ihde D, Fischmann A, et al. A randomized trial comparing combination electron-beam radiation and chemotherapy with topical therapy in the initial treatment of mycosis fungoides. New England Journal of Medicine. 1989;321(26):1784-90.
60. Edelson R, Berger C, Gasparro F, Jegasothy B, Heald P, Wintroub B, et al. Treatment of cutaneous T-cell lymphoma by extracorporeal photochemotherapy. Preliminary results. New England Journal of Medicine. 1987;316(6):297-303.
61. Russell-Jones R. Extracorporeal photopheresis in cutaneous T-cell lymphoma. Inconsistent data underline the need for randomized studies. British Journal of Dermatology. 2000;142(1):16-21.
62. Lundin J, Hagberg H, Repp R, Cavallin-Ståhl E, Fredén S, Juliusson G, et al. Phase 2 study of alemtuzumab (anti-CD52 monoclonal antibody) in patients with advanced mycosis fungoides/Sezary syndrome. Blood. 2003;101(11):4267-76.
63. Grant C, Rahman F, Piekarz R, Peer C, Frye R, Robey R, et al. Romidepsin: a new therapy for cutaneous T-cell lymphoma and a potential therapy for solid tumors. Experimental review of anticancer therapy. 2010;10(7):997-1008.
64. Kavanaugh SM, White LA, Kolesar JM. Vorinostat: A novel therapy for the treatment of cutaneous T-cell lymphoma. Am J Health Syst Pharm. 2010;67(10):793-7.
65. Guitart J, Wickless S, Oyama Y, Kuzel T, Rosen S, Traynor A, et al. Long-term remission after allogeneic hematopoietic stem cell transplantation for refractory cutaneous T-cell lymphoma. Archives of Dermatology. 2002;138(10):1359-65.
66. Prochazkova M CE, Mainhaguiet G, Sobotka J, Vergier B, Belaud-Rotureau MA, Beylot-Barry M, Merlio JP. Common chromosomal abnormalities in mycosis fungoides transformation. Genes Chromosomes Cancer. 2007;46(9):828-38.
67. Renan MJ. How many mutations are required for tumorigenesis? Implications from human cancer data. Mol Carcinog. 1993;7(3):139-46.
68. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57-70.
69. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.
70. Chang TP, Vancurova I. NFκB function and regulation in cutaneous T-cell lymphoma. American Journal of Cancer Research. 2013;3(5):433-45.
71. DiDonato JA, Mercurio F, Karin M. NF-κB and the link between inflammation and cancer. Immunological reviews. 2012;246(1):379-400.
72. Izban K, Ergin M, Qin J, Martinez R, Pooley RJ J, Saeed S, et al. Constitutive expression of NF-kappa B is a characteristic feature of mycosis fungoides: implications for apoptosis resistance and pathogenesis. Human Pathology. 2000;31(12):1482-90.
73. Juvekar A, Manna S, Ramaswami S, Chang T, Vu H, Ghosh C, et al. Bortezomib induces nuclear translocation of IκBα resulting in gene-specific suppression of NF-κB--dependent transcription and induction of apoptosis in CTCL. Molecular Cancer Research. 2011;9(2):183-94.
74. Oeckinghaus A, Ghosh S. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harbour perspectives in biology. 2009;1(14):a000034.

75. Schuster M, Annemann M, Plaza-Sirvent C, Schmitz I. Atypical IκB proteins - nuclear modulators of NF-κB signaling. Cell Communication and Signalling. 2013;11(1):23.
76. Karin M, Ben-Neriah Y. Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. Annual Review of Immunology. 2000;18(621):663.
77. Dobrzanski P, Ryseck RP, Bravo R. Both N- and C-terminal domains of RelB are required for full transactivation: role of the N-terminal leucine zipper-like motif. Molecular Cell Biology. 1993;13(3):1572-82.
78. Huang B, Yang XD, Lamb A, Chen LF. Posttranslational modifications of NF-kappaB: another layer of regulation for NF-kappaB signaling pathway. Cell Signalling. 2010;22(9):1281-90.
79. Sors A, Jean-Louis F, Pellet C, Laroche L, Dubertret L, Courtois G, et al. Down-regulating constitutive activation of the NF-kappaB canonical pathway overcomes the resistance of cutaneous T-cell lymphoma to apoptosis. Blood. 2006;107(6):2354-63.
80. Dutta P, Willis LX. Role of the JAK-STAT Signalling Pathway in Cancer. John Wiley & Sons, Ltd. 2013.
81. Thomas S, Snowden J, Zeidler M, Danson S. The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours. British Journal of Cancer. 2015;113(3):365-71.
82. Furqan M, Mukhi N, Lee B, Liu D. Dysregulation of JAK-STAT pathway in hematological malignancies and JAK inhibitors for clinical application. Biomarker Research. 2013;1(1):5.
83. Brooks A, Dai W, O'Mara M, Abankwa D, Chhabra Y, Pelekanos R, et al. Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. Science. 2014;344(6185):1249783.
84. Li WX. Canonical and non-canonical JAK-STAT signaling. Trends in Cell Biology. 2008;18(11):545-51.
85. Shi S, Larson K, Guo D, Lim S, Dutta P, Yan S, et al. Drosophila STAT is required for directly maintaining HP1 localization and heterochromatin stability. Nature Cell Biology. 2008;10(4):489-96.
86. Koo GC, Tan SY, Tang T, Poon SL, Allen GE, Tan L, et al. Janus kinase 3-activating mutations identified in natural killer/T-cell lymphoma. Cancer discovery. 2012;2(7):591-7.
87. Koskela H, Eldfors S, Ellonen P, van Adrichem A, Kuusanmäki H, Andersson E, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. New England Medical Journal. 2012;366(20):1905-13.
88. Kontro M, Kuusanmäki H, Eldfors S, Burmeister T, Andersson E, Bruserud O, et al. Novel activating STAT5B mutations as putative drivers of T-cell acute lymphoblastic leukemia. Leukemia. 2014;28(8):1738-42.
89. Netchiporouk E, Litvinov I, Moreau L, Gilbert M, Sasseville D, Duvic M. Deregulation in STAT signaling is important for cutaneous T-cell lymphoma (CTCL) pathogenesis and cancer progression. Cell Cycle. 2014;13(21):3331-5.
90. McKenzie RCT, Jones CL, Tosi I, Caesar JA, Whittaker SJ, Mitchell TJ. Constitutive activation of STAT3 in Sézary syndrome is independent of SHP-1. Leukemia. 2012;26(2):323-31.
91. Mitchell T, Whittaker S, John S. Dysregulated expression of COOH-terminally truncated Stat5 and loss of IL2-inducible Stat5-dependent gene expression in Sézary Syndrome. Cancer Research. 2003;63(24):9048-54.
92. Engelman JA, Luo J, Cantley LC. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nature Reviews Genetics. 2006;7(8):606-19.
93. Hemmings BA, Restuccia DF. PI3K-PKB/Akt pathway. Cold Spring Harbour perspectives in biology. 2012;4(9):a011189.
94. Burke JE, Williams RL. Synergy in activating class I PI3Ks. Trends in Biochemical Science. 2015;40(2):88-100.
95. Chalhoub N, Baker SJ. PTEN and the PI3-kinase pathway in cancer. Annual Review of Pathology. 2009;4:127-50.

96. Wullschleger S, Loewith R, Hall MN. TOR signaling in growth and metabolism. Cell. 2006;124(3):471-84.
97. Cristofoletti C, Picchio M, Lazzeri C, Tocco V, Pagani E, Bresin A, et al. Comprehensive analysis of PTEN status in Sezary syndrome. Blood. 2013;122(20):3511-20.
98. Lee C, Ungewickell A, Bhaduri A, Qu K, Webster D, Armstrong R, et al. Transcriptome sequencing in Sézary syndrome identifies Sézary cell and mycosis fungoides-associated lncRNAs and novel transcripts. Blood. 2012;120(16):3288-97.
99. Steelman L, Franklin R, Abrams S, Chappell W, Kempf C, Bäsecke J, et al. Roles of the Ras/Raf/MEK/ERK pathway in leukemia therapy. Leukemia. 2011;25(7):1080-94.
100. Downward J. Targeting RAS signalling pathways in cancer therapy. Nature Reviews Cancer. 2003;3(1):11-22.
101. Regad T. Targeting RTK Signaling Pathways in Cancer. Cancers. 2015;7(3):1758-84.
102. Logue JS, Morrison DK. Complexity in the signaling network: insights from the use of targeted inhibitors in cancer therapy. Genes and Development. 2012;26(7):641-50.
103. Marais R, Light Y, Paterson H, Mason C, Marshall C. Differential regulation of Raf-1, A-Raf, and B-Raf by oncogenic ras and tyrosine kinases. Journal of Biological Chemistry. 1997;272(7):4378-83.
104. Mason C, Springer C, Cooper R, Superti-Furga G, Marshall C, Marais R. Serine and tyrosine phosphorylations cooperate in Raf-1, but not B-Raf activation. The EMBO Journal. 1999;18(8):2137-48.
105. Dhillon A, Hagan S, Rath O, Kolch W. MAP kinase signalling pathways in cancer. Oncogene. 2007;26(22):3279-90.
106. Kiessling M, Oberholzer P, Mondal C, Karpova M, Zipser M, Lin W, et al. High-throughput mutation profiling of CTCL samples reveals KRAS and NRAS mutations sensitizing tumors toward inhibition of the RAS/RAF/MEK signaling cascade. Blood. 2011;117(8):2433-40.
107. Bliss-Moreau M, Coarfa C, Gunaratne P, Guitart J, Krett N, Rosen S. Identification of p38β as a therapeutic target for the treatment of Sézary syndrome. Journal of Investigative Dermatology. 2014;135(2):599-608.
108. Lattanzio R, Piantelli M, Falasca M. Role of phospholipase C in cell invasion and metastasis. Advances in biological regulation. 2013;53(3):309-18.
109. Elmore S. Apoptosis: a review of programmed cell death. Toxicologic Pathology. 2007;35(4):495-516.
110. Ouyang L, Shi Z, Zhao S, Wang F, Zhou T, Liu B, et al. Programmed cell death pathways in cancer: a review of apoptosis, autophagy and programmed necrosis. Cell Proliferation. 2012;45(6):487-98.
111. Portt L, Norman G, Clapp C, Greenwood M, Greenwood M. Anti-apoptosis and cell survival: a review. Biochimica and Biophysica Acta. 2011;1813(1):238-59.
112. Tan M, Ooi J, Ismail N, Moad A, Muhammad T. Programmed cell death pathways and current antitumor targets. Pharmaceutical Research. 2009;26(7):1547-60.
113. Bialik S, Zalckvar E, Ber Y, Rubinstein A, Kimchi A. Systems biology analysis of programmed cell death. Trends in Biochemical Science. 2010;35(10):556-64.
114. Kerr JF, Wyllie AH, Currie AR. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. British Journal of Cancer. 1972;26(4):239-57.
115. Nishida K, Yamaguchi O, Otsu K. Crosstalk between autophagy and apoptosis in heart disease. Circulation Research. 2008;103(4):343-51.
116. Al-Yacoub N, Fecker L, Möbs M, Plötz M, Braun F, Sterry W, et al. Apoptosis induction by SAHA in cutaneous T-cell lymphoma cells is related to downregulation of c-FLIP and enhanced TRAIL signaling. Journal of Investigative Dermatology. 2012;132(9):2263-74.
117. Hsu H, Xiong J, Goeddel DV. The TNF receptor 1-associated protein TRADD signals cell death and NF-kappa B activation. Cell. 1995;81(4):495-504.
118. Wajant H. The Fas signaling pathway: more than a paradigm. Science. 2002;296(5573):1635-6.

119. Waring P, Müllbacher A. Cell death induced by the Fas/Fas ligand pathway and its role in pathology. Immunology and Cell Biology. 1999;77(4):312-7.
120. Brunelle JK, Letai A. Control of mitochondrial apoptosis by the Bcl-2 family. journal of Cell Science. 2009;122(Pt 4):437-41.
121. Yacoubian TA, Standaert DG. Targets for neuroprotection in Parkinson's disease. Biochim Biophys Acta. 2009;1792(7):676-87.
122. Ghobrial IM, Witzig TE, Adjei AA. Targeting apoptosis pathways in cancer therapy. CA: A cancer journal for clinicians. 2005;55(3):178-94.
123. Zoi-Toli O, Vermeer M, De Vries E, Van Beek P, Meijer C, Willemze R. Expression of Fas and Fas-ligand in primary cutaneous T-cell lymphoma (CTCL): association between lack of Fas expression and aggressive types of CTCL. British Journal of Dermatology. 2000;143(2):313-9.
124. Dereure O, Portales P, Clot J, Guilhou JJ. Decreased expression of fas (APO-1/CD95) on lesional CD4+ T lymphocytes in cutaneous T cell lymphomas: correlations with blood data. British Journal of Dermatology. 2001;145(6):1031-2.
125. Stutz N, Johnson RD, Wood GS. The Fas apoptotic pathway in cutaneous T-cell lymphomas: frequent expression of phenotypes associated with resistance to apoptosis. Journal of American Academy of Dermatology. 2012;67(6):e1-10.
126. Wu J, Nihal M, Siddiqui J, Vonderheid E, Wood G. Low FAS/CD95 expression by CTCL correlates with reduced sensitivity to apoptosis that can be restored by FAS upregulation. Journal of Investigative Dermatology. 2009;129(5):1165-73.
127. Mao X, Lillington DM, Czepulkowski B, Russell-Jones R, Young BD, Whittaker S. Molecular cytogenetic characterization of Sezary syndrome. Genes Chromosomes Cancer. 2003;36(3):250-60.
128. Nagasawa T, Takakuwa T, Takayama H, Dong Z, Miyagawa S, Itami S, et al. Fas gene mutations in mycosis fungoides: analysis of laser capture-microdissected specimens from cutaneous lesions. Oncology 2004;67(2):130-4.
129. Dereure O, Levi E, Vonderheid EC, Kadin ME. Infrequent Fas mutations but no Bax or p53 mutations in early mycosis fungoides: a possible mechanism for the accumulation of malignant T lymphocytes in the skin. J Invest Dermatol 2002;118(6):949-56.
130. van Doorn R, Dijkman R, Vermeer M, Starink T, Willemze R, Tensen C. A novel splice variant of the Fas gene in patients with cutaneous T-cell lymphoma. Cancer Research. 2002.
131. Jones C, Wain E, Chu C, Tosi I, Foster R, McKenzie R, et al. Downregulation of Fas gene expression in Sézary syndrome is associated with promoter hypermethylation. Journal of Investigative Dermatology. 2010;130(4):1116-25.
132. Kourtis N, Tavernarakis N. Autophagy and cell death in model organisms. Cell Death and Differentiation. 2009;16(1):21-30.
133. He C, Klionsky DJ. Regulation mechanisms and signaling pathways of autophagy. Annual Review of Genetics. 2009;43:67-93.
134. Gottlieb RA, Mentzer RM. Autophagy during cardiac stress: joys and frustrations of autophagy. Annual Review of Physiology. 2010;72:45-59.
135. Kroemer G, Levine B. Autophagic cell death: the story of a misnomer. National Review of Molecular Cell Biology. 2008;9(12):1004-10.
136. Wang SY, Yu QJ, Zhang RD, Liu B. Core signaling pathways of survival/death in autophagy-related cancer networks. The international journal of biochemistry and cell biology. 2011;43(9):1263-6.
137. Eisenberg-Lerner A, Bialik S, Simon HU, Kimchi A. Life and death partners: apoptosis, autophagy and the cross-talk between them. Cell Death and Differentiation. 2009;16(7):966-75.
138. Kundu M, Thompson CB. Autophagy: basic principles and relevance to disease. Annual Review of Pathology. 2008;3:427-55.
139. Duvic M, Vu J. Update on the treatment of cutaneous T-cell lymphoma (CTCL): Focus on vorinostat. Biologics: Targets & Therapy. 2007;1(4):377-92.

140. Shao Y, Gao Z, Marks PA, Jiang X. Apoptotic and autophagic cell death induced by histone deacetylase inhibitors. Proceedings of the National Academy of Sciences of the United States of Americal. 2004;101(52):18030-5.
141. Christopher SA, Diegelman P, Porter CW, Kruger WD. Methylthioadenosine phosphorylase, a gene frequently codeleted with p16(cdkN2a/ARF), acts as a tumor suppressor in a breast cancer cell line. Cancer Res. 2002;62(22):6639-44.
142. Eisenberg T, Knauer H, Schauer A, Büttner S, Ruckenstuhl C, Carmona-Gutierrez D, et al. Induction of autophagy by spermidine promotes longevity. Nature Cell Biology. 2009;11(11):1305-14.
143. Marchitto KS, Ferro AJ. The metabolism of 5'-methylthioadenosine and 5-methylthioribose 1-phosphate in Saccharomyces cerevisiae. Journal of General Microbiology. 1985;131(9):2153-64.
144. Avila MA, García-Trevijano ER, Lu SC, Corrales FJ, Mato JM. Methylthioadenosine. Int J Biochem Cell Biol 2004;36(11):2125-30.
145. Proskuryakov SY, Gabai VL. Mechanisms of tumor cell necrosis. Current Pharmaceutical Design. 2010;16(1):56-68.
146. Festjens N, Vanden Berghe T, Cornelis S, Vandenabeele P. RIP1, a kinase on the crossroads of a cell's decision to live or die. Cell Death and Differentiation. 2007;14(3):400-10.
147. Temkin V, Karin M. From death receptor to reactive oxygen species and c-Jun N-terminal protein kinase: the receptor-interacting protein 1 odyssey. Immunological reviews. 2007(220):8-21.
148. Meurette O, Rebillard A, Huc L, Le Moigne G, Merino D, Micheau O, et al. TRAIL induces receptor-interacting protein 1-dependent and caspase-dependent necrosis-like cell death under acidic extracellular conditions. Cancer Research. 2007;67(1):218-26.
149. Vakkila J, Lotze MT. Inflammation and necrosis promote tumour growth. 4. 2004;8(641-648).
150. Gloeckler Ries L, Reichman M, Lewis D, Hankey B, Edwards B. Cancer survival and incidence from the Surveillance, Epidemiology, and End Results (SEER) program. Oncologist. 2003;8(6):541-52.
151. Krejsgaard T, Litvinov I, Wang Y, Xia L, Willerslev-Olsen A, Koralov S, et al. Elucidating the role of interleukin-17F in cutaneous T-cell lymphoma. Blood. 2013;122(6):943-50.
152. Wu X, Hwang ST. Cutaneous T-Cell Lymphoma: The Yin and Yang of Inflammation and Neoplasia. Journal of Investigative Dermatology Symposium proceedings. 2015;17(1):34-5.
153. Colotta F, Allavena P, Sica A, Garlanda C, Mantovani A. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. Carcinogenesis. 2009;30(7):1073-81.
154. Rakoff-Nahoum S. Why cancer and inflammation. The Yale journal of biology and medicine. 2006;79(3-4):123-30.
155. Rokavec M, Öner MG, Hermeking H. lnflammation-induced epigenetic switches in cancer. Cell and molecular life sciences. 2015.
156. Aggarwal BB, Vijayalekshmi RV, Sung B. Targeting inflammatory pathways for prevention and therapy of cancer: short-term friend, long-term foe. Clinical Cancer Research. 2009;15(2):425-30.
157. Dawson MA, Kouzarides T. Cancer epigenetics: from mechanism to therapy. Cell. 2012;150(1):12-27.
158. Bird A. DNA methylation patterns and epigenetic memory. Genes and Development. 2002;16(1):6-21.
159. Baylin SB. DNA methylation and gene silencing in cancer. Nature clinical practice Oncology. 2005;2(1):S4-11.
160. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. New England Journal of Medicine. 2003;349(21):2042-54.
161. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. Nature Reviews Cancer. 2011;11(10):726-34.

162. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992;69(6):915-26.
163. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999;99(3):247-57.
164. Goll M, Kirpekar F, Maggert K, Yoder J, Hsieh C, Zhang X, et al. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science. 2006;311(5759):395-8.
165. Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. Genes and Development. 2011;25(23):2436-52.
166. Lee JS, Smith E, Shilatifard A. The language of histone crosstalk. Cell. 2010;142(5):682-5.
167. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. Molecular Biology of the Cell. Garland Science. 2002;4th edition.
168. Neganova I, Lako M. G1 to S phase cell cycle transition in somatic and embryonic stem cells. Journal of Anatomy. 2008;213(1):30-44.
169. Hannon GJ, Beach D. p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest. Nature. 1994;371(6494):257-61.
170. Serrano M, Hannon GJ, Beach D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature. 1993;366(6456):704-7.
171. Roussel MF. The INK4 family of cell cycle inhibitors in cancer. Oncogene. 1999;18(38):5311-7.
172. Gartel AL, Radhakrishnan SK. Lost in transcription: p21 repression, mechanisms, and consequences. Cancer Research. 2005;65(10):3980-5.
173. Alarcon-Vargas D, Ronai Z. p53–Mdm2—the affair that never ends. Carcinogenesis. 2002;23(4):541-7.
174. Toledo L, Murga M, Gutierrez-Martinez P, Soria R, Fernandez-Capetillo O. ATR signaling can drive cells into senescence in the absence of DNA breaks. Genes and Development. 2008;22(3):297-302.
175. Taylor WR, Stark GR. Regulation of the G2/M transition by p53. Oncogene. 2001;20(15):1803-15.
176. Stark GR, Taylor WR. Control of the G2/M transition. Molecular biotechnology. 2006;32(3):227-48.
177. Boutros R, Lobjois V, Ducommun B. CDC25 phosphatases in cancer cells: key players? Good targets? Nature Reviews Cancer. 2007;7(7):495-507.
178. Malumbres M, Barbacid M. Mammalian cyclin-dependent kinases. Trends in Biochemical Sciences. 2005;30(11):630-41.
179. Curiel-Lewandrowski C, Yamasaki H, Si C, Jin X, Zhang Y, Richmond J, et al. Loss of nuclear pro-IL-16 facilitates cell cycle progression in human cutaneous T cell lymphoma. Journal of Clinical Investigation. 2011;121(12):4838-49.
180. Biskup E, Manfé V, Kamstrup M, Gniadecki R. Growth dynamics and cyclin expression in cutaneous T-cell lymphoma cell lines. Dermatology reports. 2010;2(1):e8.
181. Nihal M, Stutz N, Schmit T, Ahmad N, Wood G. Polo-like kinase 1 (Plk1) is expressed by cutaneous T-cell lymphomas (CTCLs), and its downregulation promotes cell cycle arrest and apoptosis. Cell cycle. 2011;10(8):1303-11.
182. Wu S. Targeting expression or function of Plk1 in CTCL, that is a question. Cell Cycle. 2011;10(10):1526.
183. Mathews LA, Dexheimer TS. DNA repair of cancer stem cells, Chapter 2: DNA repair pathways and mechanisms. Springer Science. 2013:19-32.
184. Branzei D, Foiani M. Regulation of DNA repair throughout the cell cycle. Nature reviews molecular cell biology. 2008;9(4):297-308.
185. Kitagawa R, Kastan MB. The ATM-dependent DNA damage signaling pathway. Cold Spring Harbour Symposia on quantitative biology. 2005;70:99-109.

186. Cimprich KA, Shin TB, Keith CT, Schreiber SL. cDNA cloning and gene mapping of a candidate human cell cycle checkpoint protein. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(7):2850-5.
187. Matsuoka S, Ballif B, Smogorzewska A, McDonald ER r, Hurov K, Luo J, et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science. 2007;316(5828):1160-6.
188. Meek DW. Tumour suppression by p53: a role for the DNA damage response? Nature Reviews Cancer. 2009;9(10):714-23.
189. Canman C, Lim D, Cimprich K, Taya Y, Tamai K, Sakaguchi K, et al. Activation of the ATM kinase by ionizing radiation and phosphorylation of p53. Science. 1998;281(5383):1677-9.
190. Ahn J, Schwarz J, Piwnica-Worms H, Canman C. Threonine 68 phosphorylation by ataxia telangiectasia mutated is required for efficient activation of Chk2 in response to ionizing radiation. Cancer Research. 2000;60(21):5934-6.
191. Matsuoka S, Rotman G, Ogawa A, Shiloh Y, Tamai K, Elledge S. Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. Proceedings of the National Academy of Sciences of the United States of Americal. 2000;97(19):10389-94.
192. Maya R, Balass M, Kim S, Shkedy D, Leal J, Shifman O, et al. ATM-dependent phosphorylation of Mdm2 on serine 395: role in p53 activation by DNA damage. Genes and Development. 2001;15(9):1067-77.
193. Lim D, Kim S, Xu B, Maser R, Lin J, Petrini J, et al. ATM phosphorylates p95/nbs1 in an S-phase checkpoint pathway. Nature. 2000;404(6778):613-7.
194. Taniguchi T, Garcia-Higuera I, Xu B, Andreassen P, Gregory R, Kim S, et al. Convergence of the fanconi anemia and ataxia telangiectasia signaling pathways. Cell. 2002;107(4):459-72.
195. Xu B, O'Donnell A, Kim S, Kastan M. Phosphorylation of serine 1387 in Brca1 is specifically required for the Atm-mediated S-phase checkpoint after ionizing irradiation. Cancer Research. 2002;62(16):4588-91.
196. Bao S, Tibbetts R, Brumbaugh K, Fang Y, Richardson D, Ali A, et al. ATR/ATM-mediated phosphorylation of human Rad17 is required for genotoxic stress responses. Nature. 2001;411(6840):969-74.
197. Riley T, Sontag E, Chen P, Levine A. Transcriptional control of human p53-regulated genes. Nature reviews molecular cell biology. 2008;9(5):402-12.
198. Lowe SW, Sherr CJ. Tumor suppression by Ink4a-Arf: progress and puzzles. Current opinion in genetics and development. 2003;13(1):77-83.
199. Brooks CL, Gu W. New insights into p53 activation. Cell research. 2010;20(6):614-21.
200. McGregor J, Crook T, Fraser-Andrews E, Rozycka M, Crossland S, Brooks L, et al. Spectrum of p53 gene mutations suggests a possible role for ultraviolet radiation in the pathogenesis of advanced cutaneous lymphomas. Journal of Investigative Dermatology. 1999;112(3):317-21.
201. Kim YJ, Wilson DM. Overview of base excision repair biochemistry. Current molecular pharmaclogy. 2012;5(1):3-13.
202. Zharkov DO. Base excision DNA repair. Cell and molecular life sciences. 2008;65(10):1544-65.
203. Kunkel TA, Erie DA. DNA missmatch repair. Annual Review of Biochemistry. 2005;74:681-710.
204. Modrich P. Mechanisms in eukaryotic mismatch repair. The journal of biological chemistry. 2006;281(41):30305-9.
205. Li GM. Mechanisms and functions of DNA mismatch repair. Cell Research. 2008;18(1):85-98.
206. Fukui K. DNA mismatch repair in eukaryotes and bacteria. Journal of nucleic acids. 2010;27.
207. Larrea AA, Lujan SA, Kunkel TA. SnapShot: DNA mismatch repair. Cell. 2010;141(4):730e.

208. Peltomäki P. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Human Molecular Genetics. 2001;10(7):735-40.
209. Portis T, Harding JC, Ratner L. The contribution of NF-kappa B activity to spontaneous proliferation and resistance to apoptosis in human T-cell leukemia virus type 1 Tax-induced tumors. Blood. 2001;98(4):1200-8.
210. Scarisbrick J, Mitchell T, Calonje E, Orchard G, Russell-Jones R, Whittaker S. Microsatellite instability is associated with hypermethylation of the hMLH1 gene and reduced gene expression in mycosis fungoides. Journal of Investigative Dermatology. 2003;121(4):894-901.
211. Shuck SC, Short EA, Turchi JJ. Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. Cell Research. 2008;18(1):64-72.
212. Costa R, Chigançcas V, Galhardo Rda S, Carvalho H, Menck C. The eukaryotic nucleotide excision repair pathway. Biochimmie. 2003;85(11):1083-99.
213. Nouspikel T. Nucleotide excision repair and neurological diseases. DNA Repair. 2008;7(7):1155-67.
214. Li X, Heyer WD. Homologous recombination in DNA repair and DNA damage tolerance. Cell Research. 2008;18(1):99-113.
215. Mao Z, Bozzella M, Seluanov A, Gorbunova V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. Cell Cycle. 2008;7(18):2902-6.
216. Rodgers K, McVey M. Error-Prone Repair of DNA Double-Strand Breaks. Journal of cellular physiology. 2016;231(1):15-24.
217. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annual Review of Biochemistry. 2010;79(181-211).
218. DeFazio LG, Stansel RM, Griffith JD, Chu G. Synapsis of DNA ends by DNA-dependent protein kinase. The EMBO journal. 2002;21(12):3192-31200.
219. Chappell C, Hanakahi L, Karimi-Busheri F, Weinfeld M, West S. Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. The EMBO Journal. 2002;21(11):2827-32.
220. Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell. 2006;124(2):301-13.
221. Ferenczi K, Ohtola J, Aubert P, Kessler M, Sugiyama H, Somani A, et al. Malignant T cells in cutaneous T-cell lymphoma lesions contain decreased levels of the antiapoptotic protein Ku70. British Journal of Dermatology. 2010;163(3):564-71.
222. Gomez D, Armando R, Farina H, Menna P, Cerrudo C, Ghiringhelli P, et al. Telomere structure and telomerase in health and disease. International Journal of Oncology. 2012;41(5):1561-9.
223. O'Sullivan RJ, Karlseder J. Telomeres: protecting chromosomes against genome instability. Nature reviews molecular cell biology. 2010;11(3):171-81.
224. Hänsel R, Löhr F, Foldynová-Trantírková S, Bamberg E, Trantírek L, Dötsch V. The parallel G-quadruplex structure of vertebrate telomeric repeat sequences is not the preferred folding topology under physiological conditions. Nucleic Acids Research. 2011;39(13):5768-75.
225. Palm W, de Lange T. How shelterin protects mammalian telomeres. Annual Review of Genetics. 2008;42:301-34.
226. Cooper GM. DNA replication. The Cell: A molecular approach 2000;2nd Edition. Sunderland (MA), Sinauer Assocaites.
227. Hayflick L, Moorhead PS. The serial cultivation of human diploid cell strains. Experimental cell research. 1961;25:585-621.
228. Campisi J. The biology of replicative senescence. European Journal of Cancer. 1997;33(5):703-9.
229. Hastie N, Dempster M, Dunlop M, Thompson A, Green D, Allshire R. Telomere reduction in human colorectal carcinoma and with ageing. Nature. 1990;346(6287):866-8.

230. Bodnar A, Kim N, Effros R, Chiu C. Mechanism of telomerase induction during T cell activation. Experimental cell research. 1996;228(1):58-64.
231. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. Nature. 1990;345(6274):458-60.
232. Chevret E, Andrique L, Prochazkova-Carlotti M, Ferrer J, Cappellen D, Laharanne E, et al. Telomerase functions beyond telomere maintenance in primary cutaneous T-cell lymphoma. Blood. 2014;123(12):1850-9.
233. Wu K, Lund M, Bang K, Thestrup-Pedersen K. Telomerase activity and telomere length in lymphocytes from patients with cutaneous T-cell lymphoma. Cancer. 1999;86(6):1056-63.
234. Mao X MS. Functional copy number changes in Sézary syndrome: toward an integrated molecular cytogenetic map III. Cancer Genet Cytogenet. 2008;185(2):86-94.
235. Senff N, Zoutman W, Vermeer M, Assaf C, Berti E, Cerroni L, et al. Fine-mapping chromosomal loss at 9p21: correlation with prognosis in primary cutaneous diffuse large B-cell lymphoma, leg type. Journal of Investigative Dermatology. 2009;129(5):1149-55.
236. Brown V, Harwood C, Crook T, Cronin J, Kelsell D, Proby C. p16INK4a and p14ARF tumor suppressor genes are commonly inactivated in cutaneous squamous cell carcinoma. Journal of Investigative Dermatology. 2004;122(5):1284-92.
237. Fountain J, Karayiorgou M, Ernstoff M, Kirkwood J, Vlock D, Titus-Ernstoff L, et al. Homozygous deletions within human chromosome band 9p21 in melanoma. Proceedings of the National Academy of Sciences of the United States of Americal. 1992;89(21):10557-61.
238. Behrmann I, Wallner S, Komyod W, Heinrich PC, Schuierer M, Buettner R, et al. Characterization of methylthioadenosin phosphorylase (MTAP) expression in malignant melanoma. Am J Pathol. 2003;163(2):683-90.
239. Jonsson A, Tuominen R, Grafström E, Hansson J, Egyhaz S. High Frequency of p16INK4A Promoter Methylation in NRAS-Mutated Cutaneous Melanoma. Journal of Investigative Dermatology. 2010;130:2809-17.
240. Vo QN, Geradts J, Boudreau DA, Bravo JC, Schneider BG. CDKN2A promoter methylation in gastric adenocarcinomas: clinical variables. Hum Pathol. 2002;33(12):1200-4.
241. Kamatani N, Nelson-Rees WA, Carson DA. Selective killing of human malignant cell lines deficient in methylthioadenosine phosphorylase, a purine metabolic enzyme. Proc Natl Acad Sci U S A. 1981;78(2):1219-23.
242. Bertino JR, Waud WR, Parker WB, Lubin M. Targeting tumors that lack methylthioadenosine phosphorylase (MTAP) activity: current strategies. Cancer Biology and Therapy. 2011;11(7):627-32.
243. Illei PB, Rusch VW, Zakowski M, Ladanyi M. Homozygous deletion of CDKN2A and codeletion of the methylthioadenosine phosphorylase gene in the majority of pleural mesotheliomas. Clinical Cancer Research. 2003;9(6):2108-13.
244. Hustinx S, Hruban R, Leoni L, Iacobuzio-Donahue C, Cameron J, Yeo C, et al. Homozygous deletion of the MTAP gene in invasive adenocarcinoma of the pancreas and in periampullary cancer: a potential new target for therapy. Cancer Biology and Therapy. 2005;4(1):83-6.
245. Nobori T, Karras J, Della Ragione F, Waltz T, Chen P, Carson D. Absence of methylthioadenosine phosphorylase in human gliomas. Cancer Research. 1991;51(12):3193-7.
246. Harasawa H, Yamada Y, Kudoh M, Sugahara K, Soda H, Hirakata Y, et al. Chemotherapy targeting methylthioadenosine phosphorylase (MTAP) deficiency in adult T cell leukemia (ATL). Leukemia. 2002;16(9):1799-807.
247. M'Soka TJ, Nishioka J, Taga A, Kato K, Kawasaki H, Yamada Y, et al. Detection of methylthioadenosine phosphorylase (MTAP) and p16 gene deletion in T cell acute lymphoblastic leukemia by real-time quantitative PCR assay. Leukemia & Lymphoma. 2000;14(5):935-40.
248. Traweek S, Riscoe M, Ferro A, Braziel R, Magenis R, Fitchen J. Methylthioadenosine phosphorylase deficiency in acute leukemia: pathologic, cytogenetic, and clinical features. Blood. 1988;71(6):1568-73.

249. Hori Y, Hori H, Yamada Y, Carrera C, Tomonaga M, Kamihira S, et al. The methylthioadenosine phosphorylase gene is frequently co-deleted with the p16INK4a gene in acute type adult T-cell leukemia. International Journal of Cancer. 1998;75(1):51-6.
250. Dreyling M, Roulston D, Bohlander S, Vardiman J, Olopade O. Codeletion of CDKN2 and MTAP genes in a subset of non-Hodgkin's lymphoma may be associated with histologic transformation from low-grade to diffuse large-cell lymphoma. Genes Chromosomes Cancer. 1998;22(1):72-8.
251. Mirebeau D, Acquaviva C, Suciu S, Bertin R, Dastugue N, Robert A, et al. The prognostic significance of CDKN2A, CDKN2B and MTAP inactivation in B-lineage acute lymphoblastic leukemia of childhood. Results of the EORTC studies 58881 and 58951. Haematologica. 2006;91(7):881-5.
252. Marcé S, Balagué O, Colomo L, Martinez A, Höller S, Villamor N, et al. Lack of methylthioadenosine phosphorylase expression in mantle cell lymphoma is associated with shorter survival: implications for a potential targeted therapy. Clinical Cancer Research. 2006;12(12):3754-61.
253. Kadariya Y, Yin B, Tang B, Shinton SA, Quinlivan EP, Hua X, et al. Mice heterozygous for germ-line mutations in methylthioadenosine phosphorylase (MTAP) die prematurely of T-cell lymphoma. Cancer Res. 2009;69(14):5961-9.
254. Schmid M, Malicki D, Nobori T, Rosenbach MD, Campbell K, Carson DA, et al. Homozygous deletions of methylthioadenosine phosphorylase (MTAP) are more frequent than p16INK4A (CDKN2) homozygous deletions in primary non-small cell lung cancers (NSCLC). Oncogene 1998;17(20):2669-75.
255. Rose AE, Poliseno L, Wang J, Clark M, Pearlman A, Wang G, et al. Integrative genomics identifies molecular alterations that challenge the linear model of melanoma progression. Cancer Res. 2011;71(7):2561-71.
256. Wild PJ, Meyer S, Bataille F, Woenckhaus M, Ameres M, Vogt T, et al. Tissue microarray analysis of methylthioadenosine phosphorylase protein expression in melanocytic skin tumors. Arch Dermatol. 2006;142(4):471-6.
257. Meyer S, Wild PJ, Vogt T, Bataille F, Ehret C, Gantner S, et al. Methylthioadenosine phosphorylase represents a predictive marker for response to adjuvant interferon therapy in patients with malignant melanoma. Exp Dermatol. 2010;19(8):e251-7.
258. Yang XR, Liang X, Pfeiffer RM, Wheeler W, Maeder D, Burdette L, et al. Associations of 9p21 variants with cutaneous malignant melanoma, nevi, and pigmentation phenotypes in melanoma-prone families with and without CDKN2A mutations. Fam Cancer 2010;9(4):625-33.
259. Conway C, Beswick S, Elliott F, Chang Y, Randerson-Moor J, Harland M, et al. Deletion at chromosome arm 9p in relation to BRAF/NRAS mutations and prognostic significance for primary melanoma. Genes Chromosomes Cancer. 2010;49(5):425-38.
260. Leal MF, Lima EM, Silva PN, Assumpcao PP, Calcagno DQ, Payao SL, et al. Promoter hypermethylation of CDH1, FHIT, MTAP and PLAGL1 in gastric adenocarcinoma in individuals from Northern Brazil. World J Gastroenterol. 2007;13(18):2568-74.
261. Watanabe F, Takao M, Inoue K, Nishioka J, Nobori T, Shiraishi T, et al. Immunohistochemical diagnosis of methylthioadenosine phosphorylase (MTAP) deficiency in non-small cell lung carcinoma. Lung Cancer. 2009;63(1):39-44.
262. Berasain C, Hevia H, Fernandez-Irigoyen J, Larrea E, Caballeria J, Mato JM, et al. Methylthioadenosine phosphorylase gene expression is impaired in human liver cirrhosis and hepatocarcinoma. Biochim Biophys Acta. 2004;1690(3):276-84.
263. Hellerbrand C, Muhlbauer M, Wallner S, Schuierer M, Behrmann I, Bataille F, et al. Promoter-hypermethylation is causing functional relevant downregulation of methylthioadenosine phosphorylase (MTAP) expression in hepatocellular carcinoma. Carcinogenesis. 2006;27(1):64-72.
264. Stevens AP, Spangler B, Wallner S, Kreutz M, Dettmer K, Oefner PJ, et al. Direct and tumor microenvironment mediated influences of 5'-deoxy-5'-(methylthio)adenosine on tumor progression of malignant melanoma. Journal of Cell Biochemistry. 2009;106(2):210-9.

265. Basu I, Cordovano G, Das I, Belbin T, Guha C, Schramm V. A transition state analogue of 5'-methylthioadenosine phosphorylase induces apoptosis in head and neck cancers. Journal of Biological Chemistry. 2007;282(29):21477-86.
266. Bedford MT, Clarke SG. Protein arginine methylation in mammals: who, what, and why. Molecular Cell. 2009;33(1):1-13.
267. Mowen KA, Tang J, Zhu W, Schurter BT, Shuai K, Herschman HR, et al. Arginine methylation of STAT1 modulates IFNalpha/beta-induced transcription. Cell. 2001;104(5):731-41.
268. Chen W, Daines MO, Hershey GK. Methylation of STAT6 modulates STAT6 phosphorylation, nuclear translocation, and DNA-binding activity. J Immunol 2004;172(11):6744-50.
269. Iwasaki H, Kovacic JC, Olive M, Beers JK, T. Y, Crook MF, et al. Disruption of protein arginine N-methyltransferase 2 regulates leptin signaling and produces leanness in vivo through loss of STAT3 methylation. Circ Res. 2010;107(8):992-1001.
270. Limm K, Ott C, Wallner S, Mueller DW, Oefner P, Hellerbrand C, et al. Deregulation of protein methylation in melanoma. European Journal of Cancer. 2013;49(6):1305-013.
271. Andreu-Pérez P, Esteve-Puig R, de Torre-Minguela C, López-Fauqued M, Bech-Serra J, Tenbaum S, et al. Protein arginine methyltransferase 5 regulates ERK1/2 signal transduction amplitude and cell fate through CRAF. Science Signalling. 2011;4(190):ra58.
272. Tang B, Kadariya Y, Chen Y, Slifker M, Kruger W. Expression of MTAP inhibits tumor-related phenotypes in HT1080 cells via a mechanism unrelated to its enzymatic function. G3 (Bethesda). 2014;5(1):35-44.
273. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology. 1975;94(3):441-8.
274. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977;74(12):5463-7.
275. Margulies M, Egholm M, Altman W, Attiya S, Bader J, Bemben L, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005;437(7057):376-80.
276. Health NIo. The Human Genome Project Completion: Frequently Asked Questions. National Human Genome Research Institute. 2010.
277. Hayden EC. Technology: The $1,000 genome. Nature. 2014;507(7492):294-5.
278. Bahassi eM, Stambrook PJ. Next-generation sequencing technologies: breaking the sound barrier of human genetics. Mutagenesis. 2014;29(5):303-10.
279. Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. Nature Genetics. 2015;47(11):1304-15.
280. Black JS, Salto-Tellez M, Mills KI, Catherwood MA. The impact of next generation sequencing technologies on haematological research – A review. Pathogenesis. 2015;2(1):9-16.
281. Grossmann V, Kohlmann A, Klein H, Schindela S, Schnittger S, Dicker F, et al. Targeted next-generation sequencing detects point mutations, insertions, deletions and balanced chromosomal rearrangements as well as identifies novel leukemia-specific fusion genes in a single procedure. Leukemia. 2011;24(4):671-80.
282. Network CGA. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490(7418):61-70.
283. Phillips H, Kharbanda S, Chen R, Forrest W, Soriano R, Wu T, et al. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. Cancer Cell. 2006;9(3):157-73.
284. Verhaak R, Hoadley K, Purdom E, Wang V, Qi Y, Wilkerson M, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell. 2010;17(1):98-110.
285. Brennan C, Verhaak R, McKenna A, Campos B, Noushmehr H, Salama S, et al. The somatic genomic landscape of glioblastoma. Cell. 2013;155(2):462-77.
286. Network. CGAR. Comprehensive genomic characterization of squamous cell lung cancers. Nature. 2012;489(7417):519-25.

287. Desai AN, Jere A. Next-generation sequencing: ready for the clinics? Clinical Genetics. 2012;81(6):503-10.
288. Schatz MC, Langmead B, Salzberg SL. Cloud computing and the DNA data race. Nature Biotechnology. 2010;28(7):691-3.
289. Reddy EP, Reynolds RK, Santos E, Barbacid M. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature. 1982;300(5888):149-52.
290. Tabin C, Bradley S, Bargmann C, Weinberg R, Papageorge A, Scolnick E, et al. Mechanism of activation of a human oncogene. Nature. 1982;300(5888):143-9.
291. An O, Dall'Olio G, Mourikis T, Ciccarelli F. NCG 5.0: updates of a manually curated repository of cancer genes and associated properties from cancer mutational screenings. Nucleic Acids Research. 2016;44(D1):D992-9.
292. Warr A, Robert C, Hume D, Archibald A, Deeb N, Watson M. Exome Sequencing: Current and Future Perspectives. G3 (Bethesda). 2015;5(8):1543-50.
293. Ruffalo M, LaFramboise T, Koyutürk M. Comparative analysis of algorithms for next-generation sequencing read alignment. Bioinformatics. 2011;27(20):2790-6.
294. Hwang S, Kim E, Lee I, Marcotte EM. Systematic comparison of variant calling pipelines using gold standard personal exome variants. Scientific Reports. 2015;5:17875.
295. Tattini L, D'Aurizio R, Magi A. Detection of Genomic Structural Variants from Next-Generation Sequencing Data. Frontiers in bioengineering and biotechnology. 2015;3:92.
296. Ng P, Levy S, Huang J, Stockwell T, Walenz B, Li K, et al. Genetic variation in an individual human exome. PLOS Genetics. 2008;4(8):e1000160.
297. Vogelstein B, Papadopoulos N, Velculescu V, Zhou S, Diaz LA J, Kinzler K. Cancer genome landscapes. Science. 2013;339(6127):1546-58.
298. Lawrence M, Stojanov P, Polak P, Kryukov G, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013;499(7457):214-8.
299. Wood L, Parsons D, Jones S, Lin J, Sjöblom T, Leary R, et al. The genomic landscapes of human breast and colorectal cancers. Science. 2007;318(5853):1108-13.
300. Greenman C, Stephens P, Smith R, Dalgliesh G, Hunter C, Bignell G, et al. Patterns of somatic mutation in human cancer genomes. Nature. 2007;446(7132):153-8.
301. Goldman N, Yang Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. Molecular biology and evolution. 1994;11(5):725-36.
302. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nature Genetics. 1999;22(3):231-8.
303. Halushka M, Fan J, Bentley K, Hsie L, Shen N, Weder A, et al. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. Nature Genetics. 1999;22(3):239-47.
304. Weng A, Ferrando A, Lee W, Morris JP t, Silverman L, Sanchez-Irizarry C, et al. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. Science. 2004;306(5694):269-71.
305. Agrawal N, Frederick M, Pickering C, Bettegowda C, Chang K, Li R, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science. 2011;333(6046):1154-7.
306. Agrawal N, Jiao Y, Bettegowda C, Hutfless S, Wang Y, David S, et al. Comparative genomic analysis of esophageal adenocarcinoma and squamous cell carcinoma. Cancer discovery. 2012;2(10):899-905.
307. Van den Eynden J, Fierro AC, Verbeke LP, Marchal K. SomInaClust: detection of cancer genes based on somatic mutation patterns of inactivation and clustering. BMC Bioinformatics. 2015;16:125.

308. Tamborero D, Gonzalez-Perez A, Lopez-Bigas N. OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes. Bioinformatics. 2013;29(18):2238-44.
309. Forbes S, Bindal N, Bamford S, Cole C, Kok C, Beare D, et al. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. Nucleic Acids Research. 2011;39(Database Issue):D945-D50.
310. Forbes S, Tang G, Bindal N, Bamford S, Dawson E, Cole C, et al. COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. Nucleic Acids Research. 2010;38(Database Issue):D652-7.
311. Youn A, Simon R. Identifying cancer driver genes in tumor genome sequencing studies. Bioinformatics. 2011;27(2):175-81.
312. Fousteri M, Mullenders LH. Transcription-coupled nucleotide excision repair in mammalian cells: molecular mechanisms and biological effects. Cell Research. 2008;18(1):73-84.
313. Stamatoyannopoulos J, Adzhubei I, Thurman R, Kryukov G, Mirkin S, Sunyaev S. Human mutation rate associated with DNA replication timing. Nature Genetics. 2009;41(4):393-5.
314. Chen C, Rappailles A, Duquenne L, Huvet M, Guilbaud G, Farinelli L, et al. Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. Genome Research. 2010;20(4):447-57.
315. Koren A, Polak P, Nemesh J, Michaelson J, Sebat J, Sunyaev S, et al. Differential relationship of DNA replication timing to different forms of human mutation and variation. American Journal of Human Genetics. 2012;91(6):1033-40.
316. Dees N, Zhang Q, Kandoth C, Wendl M, Schierding W, Koboldt D, et al. MuSiC: identifying mutational significance in cancer genomes. Genome Research. 2012;22(8):1589-98.
317. Zhang J, Liu J, Sun J, Chen C, Foltz G, Lin B. Identifying driver mutations from sequencing data of heterogeneous tumors in the era of personalized genome sequencing. Briefings in Bioinformatics. 2014;15(2):244-55.
318. Tan H, Bao J, Zhou X. A novel missense-mutation-related feature extraction scheme for 'driver' mutation identification. Bioinformatics. 2012;28(22):2948-55.
319. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. eLife. 2013;2:e00471.
320. Brüggemann M, White H, Gaulard P, Garcia-Sanz R, Gameiro P, Oeschger S, et al. Powerful strategy for polymerase chain reaction-based clonality assessment in T-cell malignancies Report of the BIOMED-2 Concerted Action BHM4 CT98-3936. Leukemia. 2007;21(2):215-21.
321. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research. 2010;38(16):e164.
322. Fuentes Fajardo K, Adams D; NISC Comparative Sequencing P, Mason C, Sincan M, Tifft C, Toro C, et al. Detecting false-positive signals in exome sequencing. Human Mutation. 2012;33(4):609-13.
323. Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics. 2012;28(21):2747-54.
324. Kadariya Y, Tang B, Wang L, Al-Saleem T, Hayakawa K, Slifker MJ, et al. Germline Mutations in Mtap Cooperate with Myc to Accelerate Tumorigenesis in Mice. PLOS one. 2013.
325. Pietrocola F, Lachkar S, Enot D, Niso-Santano M, Bravo-San Pedro J, Sica V, et al. Spermidine induces autophagy by inhibiting the acetyltransferase EP300. Cell Death and Differentiation. 2015;22(3):509-16.
326. Gewirtz DA. The four faces of autophagy: implications for cancer therapy. Cancer Research. 2014;74(3):647-51.
327. Tong Y, Liu YY, You LS, Qian WB. Perifosine induces protective autophagy and upregulation of ATG5 in human chronic myelogenous leukemia cells in vitro. Acta pharmacologica sinica. 2012;33(4):542-50.

328. Katoh M. Functional and cancer genomics of ASXL family members. British Journal of Cancer. 2013;109(2):299-306.
329. Nickerson M, Dancik G, Im K, Edwards M, Turan S, Brown J, et al. Concurrent alterations in TERT, KDM6A, and the BRCA pathway in bladder cancer. Clinical Cancer Research. 2014;20(18):4935-48.
330. Lacana E, D'Adamio L. Regulation of Fas ligand expression and cell death by apoptosis-linked gene 4. Nature Medicine. 1999;5(5):542-7.
331. Hall D, Cost N, Hegde S, Kellner E, Mikhaylova O, Stratton Y, et al. TRPM3 and miR-204 establish a regulatory circuit that controls oncogenic autophagy in clear cell renal cell carcinoma. Cancer Cell. 2014;26(5):738-53.
332. Sasaki S, Kitagawa Y, Sekido Y, Minna J, Kuwano H, Yokota J, et al. Molecular processes of chromosome 9p21 deletions in human cancers. Oncogene. 2003;22(24):3792-8.
333. Powell EL, Leoni LM, Canto MI, Forastiere AA, Iocobuzio-Donahue CA, Wang JS, et al. Concordant loss of MTAP and p16/CDKN2A expression in gastroesophageal carcinogenesis: evidence of homozygous deletion in esophageal noninvasive precursor lesions and therapeutic implications. The American Journal of Surgical Pathology. 2005;29(11):1497-504.
334. De Carvalho D, Sharma S, You J, Su S, Taberlay P, Kelly T, et al. DNA methylation screening identifies driver epigenetic events of cancer cell survival. Cancer Cell. 2012;21(5):655-67.
335. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nature Reviews Genetics. 2002;3(6):415-28.