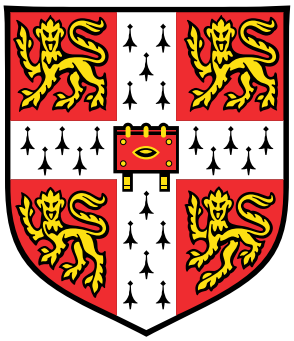
Genetic alterations defining human primary melanoma and mechanisms of immune evasion



Sofia Yixin Chen

Wellcome Sanger Institute

University of Cambridge

This dissertation is submitted for the degree of

Doctor of Philosophy

Newnham College

December 2019

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text with details in Appendix C. This dissertation contains less than 60,000 words and has less than 150 figures.

Sofia Yixin Chen
December 2019

Abstract

Genetic alterations defining human primary melanoma and mechanisms of immune evasion

The somatic mutations found in melanomas reflect the biological processes that govern tumour development. They also help shape how tumours evolve and escape immune regulation. Studying these changes can therefore help to refine our understanding of melanoma progression with the added potential to offer new perspectives for disease management.

Most prior melanoma sequencing studies have focused on advanced disease. Thus, somatic alterations that influence the behaviour of early-stage tumours have not been fully explored. Consequently, in this thesis I study a collection of 524 primary melanomas on which extensive clinical data have been collected for almost two decades. I describe the mutational landscape of these tumours including driver genes, new recurrent variants, mutually exclusive genetic interactions and copy number alterations. I discuss and associate these features with aspects of tumour pathology, sun exposure, immunogenicity and patient outcomes. To identify genes required for melanoma survival, I intersect my genomic analysis with a dataset of CRISPR-Cas9 dropout screens and discover a melanoma-associated genetic vulnerability mediated by *Interferon Regulatory Factor 4 (IRF4)*. I then begin to experimentally validate and explore the biological pathway by which *IRF4* may function in the context of melanoma.

Checkpoint inhibitors have revolutionised melanoma care, yet only a minority of patients respond to these treatments and our comprehension of the mechanisms governing PD-L1 expression on melanoma cells is still limited. In the second part of this thesis, I examine the regulation of the key checkpoint receptor PD-L1, which is often upregulated in melanoma to facilitate tumour escape. To improve our understanding of the processes controlling PD-L1 expression, and how the PD-1/PD-L1 axis can be targeted to overcome immune evasion, I employ a genome-wide CRISPR-Cas9 screening approach. I identify genes which elicit downregulation of PD-L1 when disrupted in melanoma cells, capturing several central processes including basal transcription, N-linked glycosylation and intracel-

lular transport. A second extensive screen in eight cancer cell lines of melanoma, bladder and lung cancer origin validate these findings and link novel candidate genes, including *Sphingolipid Transporter 1 (SPNS1*), to the control of PD-L1 cell surface expression.

Additional work is required to further validate and understand the regulation of PD-L1 through *SPNS1*, as well as its contribution to immune surveillance.

In summary, I present the first comprehensive evaluation into the somatic alteration landscape of primary melanomas, gaining insight into the molecular architecture of these tumours. Additionally, I introduce novel genes and processes regulating PD-L1 gene transcription, processing and presentation on the cell surface. Collectively, these results improve our understanding of the genetic processes that govern primary melanomas, as well as providing valuable insights into the mechanism of PD-L1 regulation.

Sofia Yixin Chen

Till min familj för all er kärlek, omtanke och stöd.
妈妈和爸爸, 你们是我一生中最大的灵感。

Acknowledgements

The process towards obtaining a doctorate degree has been both fulfilling and challenging, and I would like to express my sincere appreciation to everyone who has supported me throughout this journey.

First and foremost, I would like to thank my supervisor David Adams, for your guidance and endless support through all the challenges I faced during my PhD. In particular, thank you for believing in me despite failed projects in my first year, and having to start over shaping two new projects for my second year onwards.

To everyone in the "Adam's family", for your warm welcome which truly made Cambridge my second home. Special thanks to Aravind, Nicky, Marco, Katharina, Manu, Mamun, James and Daniela, for the laughter, culinary experiences, fika, travels, shenanigans, being my running companions and above all, for your friendship. Thanks to Marco, Vicky, Nicky, Gemma, Manu, Annie and Chi, for helpful project discussions and advice on experimental work. Mamun, Vivek and Aravind, I am grateful you convinced me to pursue bioinformatics, and for everything computational you've taught me.

I am eternally grateful to all my collaborators, in particular Julia Newton-Bishop, Tim Bishop and Mark Harland of the Leeds University group. Thank you for generously sharing the Leeds melanoma cohort dataset with me, as well as providing lots of helpful advice and discussions. My thesis would not have been possible without you.

I would like to acknowledge the Wellcome Sanger Institute and the MELGEN network for funding my PhD. I feel extremely privileged to have had the opportunity to be part of this fantastic research environment, to meet and collaborate with world-leading scientists, and travel across the world to present my work.

Finally, I would like to thank my family and friends, for your never-ending encouragement, support and love. Mom and dad, thank you for inspiring me to work hard, aim high and stay motivated. Gustaf, from the depth of my heart, I am forever grateful to you for your understanding, patience, and for keeping me sane over these past years. I could not have done this without you.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms / Abbreviations

AJCC American Joint Committee on Cancer

ALM Acral lentiginous melanoma

ASCAT Allele-specific copy number analysis of tumours

AUC Area under the curve

BAF B allele frequency

BET Bromodomain and extraterminal

Bp Base pairs

Cas CRISPR-associated

CGP Cancer Genome Project

COG Conserved oligomeric Golgi

CRISPR Clustered regularly interspaced short palindromic repeats

CSD Chronically sun-damaged

CTL Cytotoxic T lymphocyte

CTLA-4 Cytotoxic T-lymphocyte-associated protein 4

DSPc Dual specificity phosphatase catalytic domain

ER Endoplasmic reticulum

ETS E26 transformation-specific or E-twenty-six

ExAC The Exome Aggregation Consortium

FACS Fluorescence activated cell sorting

FFPE Formalin-fixed paraffin-embedded

FPKM Fragments per kilobase of transcript per million mapped reads

gRNA guide RNA

HDAC Histone deacetylase

HDR Homology-directed repair

HLA Human leukocyte antigen

HOPS Homotypic fusion and vacuole protein sorting

ICGC International Cancer Genome Consortium

IGV Integrative genomics viewer

LMM Lentigo maligna melanoma

LOH Loss of heterozygosity

MAF Minor allele frequency

MAPK Mitogen-activated protein kinase

MHC Major histocompatibility complex

MOI Multiplicity of infection

MTC Multi-subunit tethering complex

NHEJ Non-homologous end joining

NM Nodular melanoma

NSCLC Non-small cell lung cancer

OR Odds ratio

OST Oligosaccharyltransferase

PAM Protospacer adjacent motif

PD-1 Programmed cell death protein 1

PD-L1 Programmed death-ligand 1

PE Paired end

PI3K Phosphoinositide 3-kinase

PIP3 Phosphatidylinositol triphosphate

ROC Receiver operating characteristic

RRA Robust rank aggregation

RTK Receptor tyrosine kinase

SE Single end

SNP Single nucleotide polymorphism

SNV Single nucleotide variant

SSM Superficial spreading melanoma

STRING Search tool for the retrieval of interacting genes/proteins

TAF TBP-associated factors

TBP TATA box binding protein

TCGA The Cancer Genome Atlas

TFIID Transcription factor II D

TMA Tissue microarray

TRAPP Transport protein particle

UTR Untranslated region

VEP Variant effect predictor

WGS Whole genome sequencing

WT Wild-type

Chapter 1

Introduction

1.1 The origin and epidemiology of melanoma

Melanoma, the deadliest form of skin cancer, is the fifth most commonly diagnosed cancer in the UK [1] and US [2], and accounts for 1.6% of all newly diagnosed cancers worldwide [3]. The incidence of melanoma in fair-skinned individuals is among the most rapidly increasing cancer types (Fig. 1.1). However, the survival rates have almost doubled in the past 40 years [1], possibly owing to earlier detection and successful clinical interventions. Early stage melanomas often have a good prognosis, but the five-year survival rate declines rapidly if the tumour has metastasised [4, 5].

UV exposure is one of the most important environmental risk factors for the development of melanoma [6–10], where total sun exposure, sunburns and sun exposure patterns have been linked to melanoma susceptibility. Other established risk factors include phenotypic characteristics [11, 12], such as freckles, fair skin, pigmentation and naevi count, a history of other cancers [13] and genetic factors driving hereditary melanoma [14, 15].

Melanoma arises from melanocytes [16–18], a slow-proliferating, melanin pigment-producing cell present on body sites including the skin, eye, sinonasal, genital tract and mucosal surfaces. Melanin synthesis is responsible for pigmentation, a process where *TYR* and *MC1R* are critical components. However, it is also important in protecting against UV damage [19–22]. The melanocyte lineage is overseen by the master regulator *MITF*, with *SOX10*, *PAX3* and *KIT* being other key genes in melanocyte development [19]. Interestingly, the reactivation of such melanocyte-linage-associated genes have been found to increase the aggressiveness and metastatic propensity of melanoma cells [23, 24].

The transformation of a melanocyte into melanoma is proposed to follow one of the following two trajectories: either they arise *de novo* or from a pre-existing naevus [25–28].

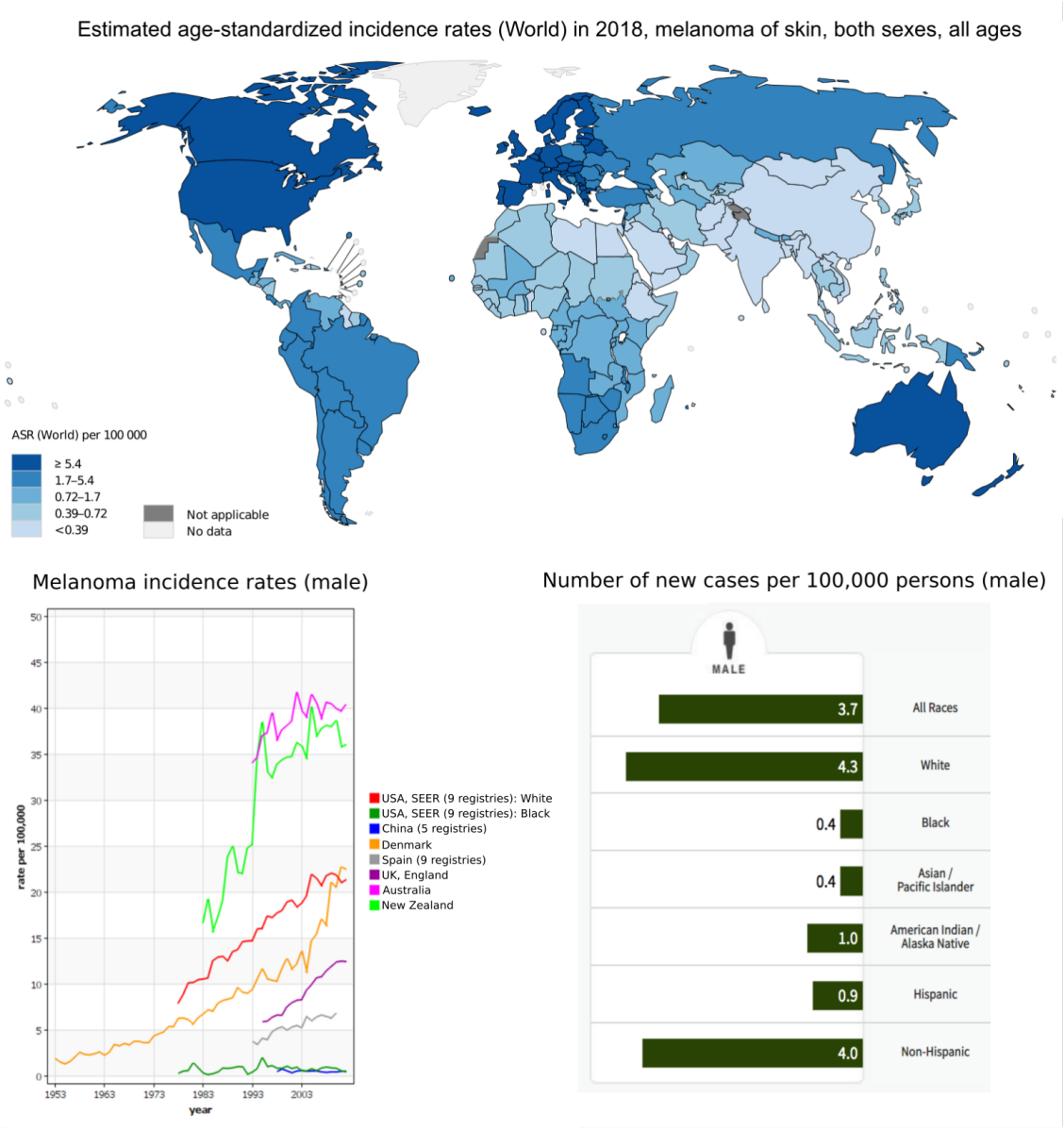


Figure 1.1: **Incidence rates of melanoma.** Top panel: Melanoma incidence rates, shown worldwide with 2018 statistics. Bottom panel, left figure: Melanoma incidence rates over time for selected countries (data for males shown, a similar trend was observed for females). Bottom panel, right figure: Number of melanoma cases per 100,000 persons by race/ethnicity (data for males shown, a similar trend was observed for females). All data shown are age-adjusted. Data extracted from GLOBOSCAN 2018 or SEER, graphs adapted from IARC (http://gco.iarc.fr/) and [2].

The early Clark model (Fig. 1.2), describes the process of malignant transformation through a series of events including the formation of a naevus, and has historically been seen as the standard model depicting melanoma development [29]. This model of melanoma development illustrates specific histopathological characteristics with each stage of progression. It starts with the proliferation of melanocytes to form a benign naevus lesion, followed by irregular growth causing structural abnormalities. In the following phases, cells begin to spread, first intraepidermally, and then vertically into the dermis, until finally acquiring metastatic properties to advance to other organs. In terms of the genetic components linking to the Clark model, the initiating event is credited to MAPK activation mainly through a *BRAF* V600E mutation, and a large proportion of benign moles also harbour activating *BRAF* mutations [30, 31]. The majority of moles do however not progress into melanoma, as oncogene-induced senescence needs to be overcome to facilitate melanoma development [32–34]. This can be achieved through inactivation of *CDKN2A* and associated cell cycle regulatory checkpoints. In the final stages, additional processes include overcoming replicative senescence through activation of the oncogene *TERT* [35], alteration of signalling pathways involving chromatin remodelling, cell adhesion or migration, or inactivation of other important cell growth and survival regulators may occur.

However, it is now known that the majority of melanomas form spontaneously rather than from a mole [26, 28, 37]. Both models share the fundamental basis that the acquisition of genetic alterations provide the tumour cells with favourable tumourigenic traits. Such traits include those associated with the hallmarks of cancer [38–40], such as increased cell proliferation and evasion of apoptosis. The proposed order of events, do however differ slightly. In support of the *de novo* hypothesis, the *BRAF* V600E mutation identified as a founder event in the Clark model does not show complete penetrance amongst melanocytic naevi [30, 41, 42], suggesting moles form through other processes. Furthermore, subclones within a naevus have been found to comprise of a mix of *BRAF*-mutant and wild-type (WT) cells [43], an observation which challenges the traditional Clark model which describes clonal expansion of *BRAF*-mutant cells. The two models do however agree in the necessity of losing apoptosis and cell cycle regulatory mechanisms, most notably through *CDKN2A* loss, as a step towards melanoma development. Whether this event would take place prior to, simultaneously or following Mitogen-activated protein kinase (MAPK) pathway activation remains a debate [44–46]. The final steps towards melanoma progression involving escape of the hayflick limit and acquisition of additional pro-tumourigenic properties remain concordant between both the Clark and the *de novo* models.

Recent work by Bastian and colleagues, support the sequential procurement of somatic

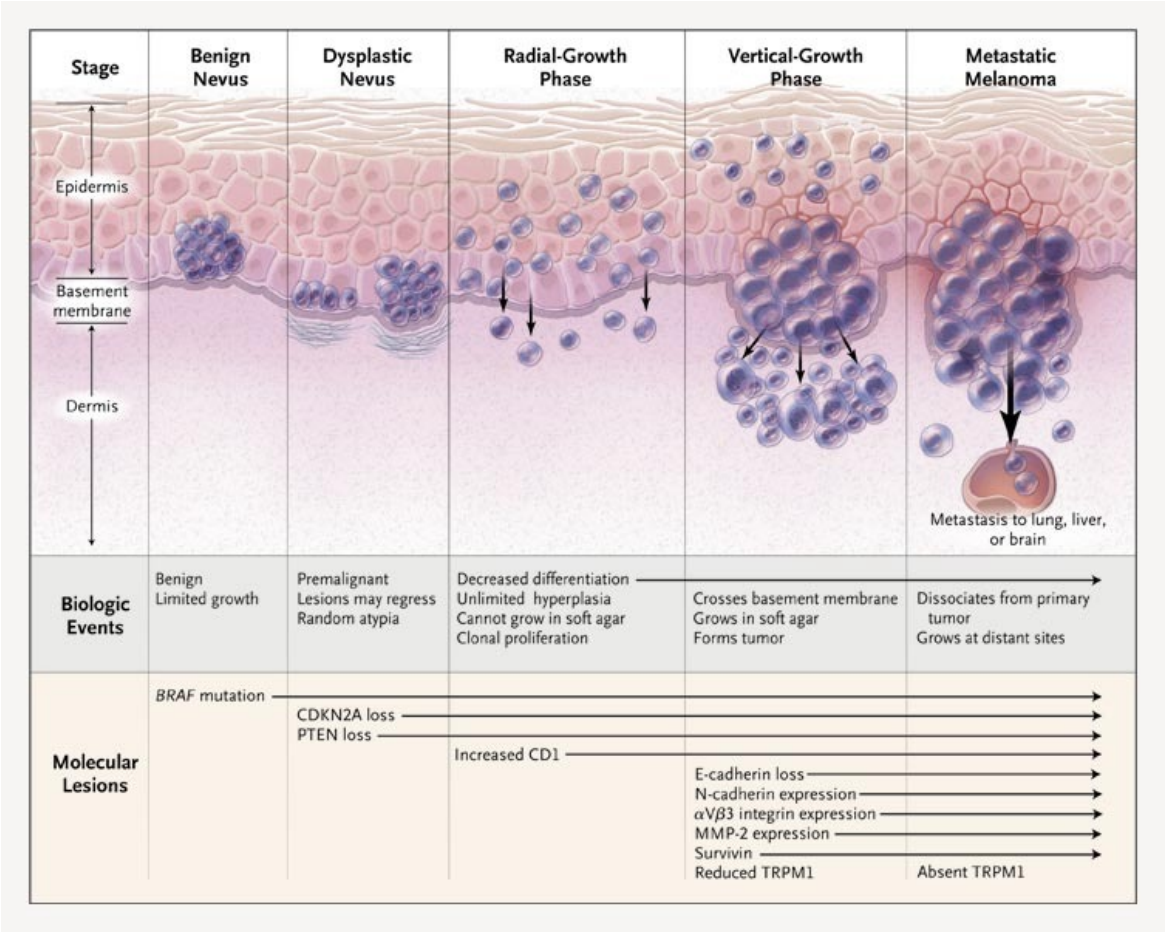


Figure 1.2: **The Clark model of melanoma development.** In this model, melanoma development begins with the formation of a benign naevus. This is followed by the dysplastic naevus stage, the radial-growth phase, the vertical-growth phase and finally the melanoma becoming metastatic. Each stage is represented by specific biological and molecular events. Reproduced with permission from [36], Copyright Massachusetts Medical Society.

alterations during melanoma evolution. This involves early activation of the MAPK pathway and upregulation of telomerase activity, followed by additional modifications targeting cell cycle regulation and epigenetic modifications in the transition to invasive melanoma [47, 48]. The later stages were also shown to be associated with inactivation of key tumour suppressors *PTEN* and *TP53,* as well as the presence of increased genomic instability with multiple activating mutations in key melanoma pathways.

The proposed models of melanoma development, progressing through an intermediate melanocytic naevus or not, might reflect heterogeneity in melanoma as a disease. Tumours harbouring *BRAF* V600E mutations are more likely to originate from a mole [18, 49]. Such tumours are also more frequently found on body sites with intermittent sun exposure and in younger patients [50–52]. In contrast, *de novo* arising melanomas are more frequently associated with tumours found on body sites heavily exposed to UV radiation [45, 53, 54]. This indicates that different progression models of melanoma could reflect distinct differences in melanoma biology [47]. These differences extend also to histopathological subtypes (discussed in Section 1.2), where the coexistence of nevoid cells is more common in the superficial spreading subtype of melanoma (although the majority still arise *de novo*) [28, 49, 55]. This is in contrast to much lower frequencies of nevoid cells found in for example lentigo maligna or acral lentiginous melanomas. Thus, melanomas cover a diverse range of characteristics, whereby the different roads leading to malignancy are shaped by genetic and environmental factors, which also contribute to its histopathological features.

1.2 Melanoma classification

1.2.1 Histopathological subtypes

Cutaneous melanomas, originating from the skin, are commonly classified into subtypes based on histopathological features [56–60]. The major four subtypes comprise superficial spreading melanoma (SSM), nodular melanoma (NM), lentigo maligna melanoma (LMM) and acral lentiginous melanoma (ALM) (Fig. 1.3). Approximately 70% of all melanomas in patients of European descent are diagnosed as SSM, making it the most prevalent subtype. Most often, it arises on intermittently sun-exposed skin, and is characterised by a growth pattern of horizontal spread [61–63]. The more rapidly growing NM subtype, which makes up 5-25% of reported melanoma cases, spreads vertically rather than horizontally, and often show a raised and symmetric appearance [56, 61]. LMM (10%) originates from a slow-growing lentigo maligna precursor lesion, is often flat and gradually spreads horizontally

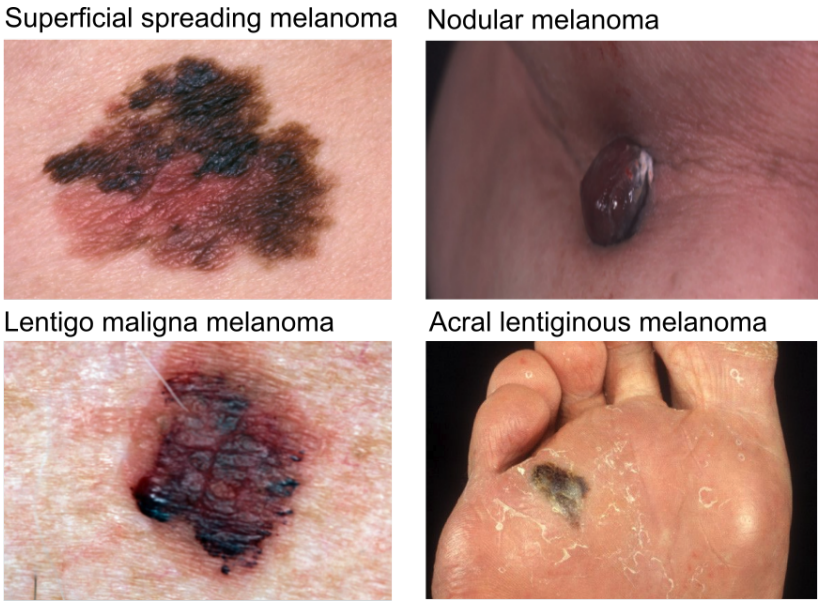


Figure 1.3: **The major histopathological melanoma subtypes.** Superficial spreading melanoma (SSM) is characterised by a growth pattern of horizontal spread. Nodular melanoma (NM) often show a raised and symmetric appearance. Lentigo maligna melanoma (LMM) is often flat and gradually spreads intraepidermally. Acral lentiginous melanoma (ALM) is commonly found on palms, soles and under the nails. SSM, NM and LMM figures reprinted with permission from [67], under © Crown copyright, and ALM from [68], under the Creative Commons Attribution License.

before becoming invasive [64]. This type of melanoma is more commonly associated with older age and high UV exposure. ALM is characteristically found on palms, soles and under the nails, and also shows an initial horizontal growth, similar to the SSM and LMM subtypes [65]. This subtype is rare in people of European descent; however, it is the most common type of melanoma affecting people with darker skin [66].

Besides these four major subtypes, a whole range of rarer histopathological subtypes exist including amelanotic, desmoplastic and spitzoid tumours, each with unique histopathological characteristics [60, 62]. Occasionally, melanoma can also arise from melanocytes residing on surfaces other than the skin, such as the eye (ocular/uveal melanoma) or mucosal membranes (mucosal melanoma) [69–72]. Recently, adaptations to these histopathological subtypes have been proposed, as several components including naevi count, UV exposure and genetic alterations have been shown to provide distinct characteristics to melanomas [54, 63, 73]. This has resulted in the World Health Organisation (WHO) incorporating a new classification system in 2018 taking into account epidemiological, clinical, histopathological and genomic attributes for the stratification of melanoma [62].

1.2.2 Prognostic classification

An alternative way of classifying melanomas, which is routinely used in clinical practice, is by prognostic factors which can be assessed from the primary tumour. Breslow thickness is an estimate of the tumour depth of invasion, measured vertically from the skin surface down to the deepest invasive cell [74]. A thicker tumour is associated with worse prognosis. A tumour showing signs of ulceration, where the epidermis covering the melanoma has been compromised, is another important factor contributing to a worse patient outlook [75, 76].

Breslow thickness and ulceration status makes up the T category of the widely adopted American Joint Committee on Cancer (AJCC) staging system [77]. In the previous edition, mitotic rate, defined as the number of mitoses per mm^2, was also part of the T stage classification [76], but in the newest 8th edition, mitotic rate was excluded due to lack of prognostic power in their multivariate model [5]. However, they argue it might still be an important factor in specific cases such as thin melanomas, supported also by other studies [78–80], and should therefore not be completely disregarded. Furthermore, the AJCC staging system also incorporates measures of metastasis, where the N stage corresponds to the presence of regional lymph node spread, and the M category illustrates if distant metastasis has occurred. Together, the T, N and M stages are jointly combined into main prognostic stages ranging from I-IV, with associated substages ranging from A-D (Table 1.1).

1.3 Genetics of melanoma

Genetic factors influence the susceptibility and pathogenicity of all cancers [81]. Cutaneous melanomas are particularly affected by genetic aberrations as a result of UV exposure, leading to excessively high mutation burdens often dominated by C>T transitions [82–86]. Together with the dysregulation of DNA damage repair genes [87, 88], these changes contribute to the genomic instability often found in melanoma and other cancers [89, 90]. These instabilities further provoke the acquisition of additional changes, paving way for diverse and opportunistic clones to manifest. This heterogeneity and multitude of genetic alterations will pose a great challenge in our attempts to unravel the important biological changes in melanoma as a disease.

As melanomas progress, they adapt and modulate key biological pathways in their favour (Fig. 1.4). This has resulted in the implementation of molecular subtypes in melanoma, categorising patients based on tumour alterations in key driver genes [73, 85, 91–93]. These subtypes describe the key genetic events driving melanoma progression in different tumours,

Table 1.1: **AJCC 8th edition pathologic staging groups for melanoma of the skin.** The prognostic AJCC staging system according to respective T, N and M stages. 5-year survival rates refer to melanoma-specific survival. Adapted and reprinted with permission from [57] and [5].

AJCC stage	T category[1]	N category[2]	M category[3]	5-year survival
0	Tis	N0	M0	>98%
IA	T1a-T1b	N0	M0	99%
IB	T2a	N0	M0	97%
IIA	T2b/T3a	N0	M0	94%
IIB	T3b/T4a	N0	M0	87%
IIC	T4b	N0	M0	82%
IIIA	T1a-2a	N1a/N2a	M0	93%
IIIB	T0-2a	N1b-c	M0	83%
	T1a-2a	N1b-c/N2b	M0	
	T2b-3a	N1a-2b	M0	
IIIC	T0	N2b-c/N3b-c	M0	69%
	T1a-3a	N2c-3c	M0	
	T3b-4a	Any N ≥ N1	M0	
	T4b	N1a-2c	M0	
IIID	T4b	N3a-c	M0	32%
IV	Any T	Any N	M1	15-20%

[1]Primary tumour definition
Tis: Melanoma *in situ*
T1a: <0.8mm and no ulceration
T1b: ≥0.8 mm or <0.8 mm with ulceration
T2a: >1.0–2.0 mm without ulceration
T3a: >2.0–4.0 mm without ulceration
T3b: >2.0–4.0 mm with ulceration
T4a: >4.0 mm without ulceration
T4b: >4.0 mm with ulceration

[3]Broad M category definition
M0: No evidence of distant metastasis
M1: Evidence of distant metastasis

[2]Number of tumour-involved regional lymph nodes
N0: No regional metastases detected
N1a: 1 clinically occult (i.e. detected by SLN biopsy)
N1b: 1 clinically detected
N1c: Presence of in-transit, satellite, and/or microsatellite metastases
N2a: 2–3 clinically occult (i.e. detected by SLN biopsy)
N2b: 2–3, at least 1 clinically detected
N2c: 1 clinically occult or detected, with in-transit, satellite, and/or microsatellite metastases
N3a: 4≤ clinically occult (i.e. detected by SLN biopsy)
N3b: 4≤ at least 1 of which clinically detected, or the presence of any number of matted nodes
N3c: 2≤ clinically occult or clinically detected and/or presence of any number of matted nodes or in-transit, satellite, and/or microsatellite metastases
N3b: 4≤ at least 1 of which clinically detected, or the presence of any number of matted nodes
N3c: 2≤ clinically occult or clinically detected and/or presence of any number of matted nodes or in-transit, satellite, and/or microsatellite metastases

highlighting the importance of genetics. Since the first discovery of the highly oncogenic *BRAF* V600E mutation [94], this event has emerged as one of the most important genetic alterations in melanoma, causing up to 700x increase in MAPK pathway activity. The *BRAF* gene is mutated in approximately half of all patients with sporadic melanoma [73], and is the predominant *RAF* isoform altered in melanoma and other cancers. This could be because it is the most efficient isoform in activating downstream MAPK components, while *ARAF* and *CRAF* are only found mutated in rare cases [95].

RAS genes are also frequently mutated in sporadic cases of melanoma and other cancers [96], where amino acid substitutions in codon 12, 13 and 61 are the most common oncogenic events resulting in a constitutively active form of the protein [97]. *NRAS* is altered in approximately 30% of melanoma cases [73, 85], whereas *HRAS* and *KRAS* mutations are less common. RAS proteins are critical for signal transduction through the MAPK pathway; however, they also control the Phosphoinositide 3-kinase (PI3K)/AKT pathway, another essential pathway in melanoma progression [98–100].

1.3.1 The MAPK pathway

The MAPK pathways are integral in signal transduction control, regulating essential biological functions including cellular proliferation, differentiation and apoptosis [101–103]. Multiple MAPK families exist: in mammalian cells the most well-characterised involve the ERK (also known as the classical MAPK), JNK and p38 pathways [102]. Throughout this thesis, the definition of the MAPK pathway will refer to the ERK-associated MAPK pathway. The MAPK cascade is initiated through the binding of external stimuli to receptor tyrosine kinases (RTKs), causing receptor dimerisation and intracellular autophosphorylation events [101, 104]. This in turn leads to the recruitment of GRB2 and SOS1, causing activation of small GTPase RAS proteins (HRAS, KRAS or NRAS) through the conversion of GDP to GTP. The active GTP-bound form of RAS is then capable of binding to RAF kinases (ARAF, BRAF or CRAF), causing conformational changes to occur which stimulate its activity [105]. Activated RAF then exerts its kinase function by phosphorylating MEK, which in turn phosphorylates ERK, leading to its translocation to the nucleus and the subsequent activation of various transcription factors [106]. The MAPK pathway is upregulated in up to 90% of all melanomas [92, 107], emphasising the central role of this signalling pathway in melanoma pathogenesis.

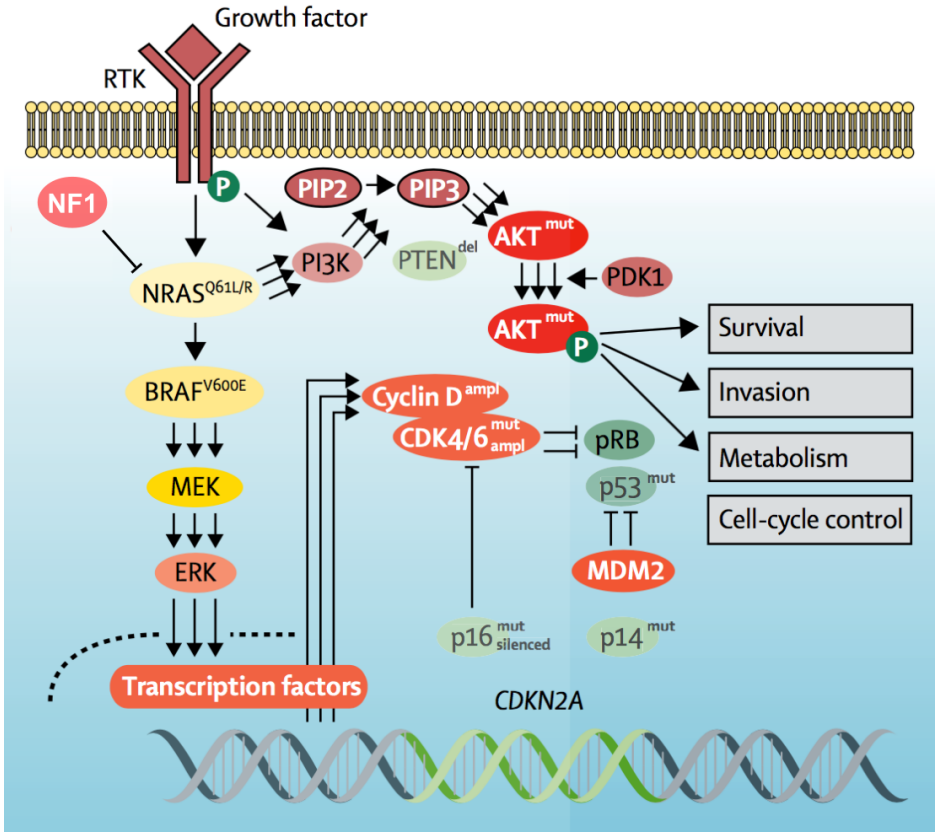


Figure 1.4: **Key signalling pathways in melanoma.** An overview of the three main signalling pathways dysregulated in melanomas: the Mitogen-activated protein kinase (MAPK) pathway, the Phosphoinositide 3-kinase (PI3K)/AKT pathway and the *CDKN2A*-associated pathways. Figure adapted and reprinted with permission from the publisher [92].

1.3.2 The PI3K/AKT pathway

Similar to the MAPK pathway, the PI3K/AKT pathway can be activated through stimulation of RTKs and activated RAS proteins. Different classes and isoforms of PI3Ks exist and might play different roles depending on tissue- and disease context [108].

Most relevant for human cancers are the heterodimeric class I PI3Ks for which *PIK3CA*, *PIK3CB*, *PIK3CD* or *PIK3CG* genes encode the catalytic domain (p110), whilst *PIK3R1*, *PIK3R2* and *PIK3R3* translate into the regulatory subunit (p85). When activated, PI3Ks phosphorylate phosphatidylinositols (PIPs, also sometimes referred to as PtdIns) into phosphatidylinositol triphosphate (PIP3). PIP3 in turn regulates various signalling molecules, such as AKTs, which are activated through phosphorylation by PDK1 (sometimes referred to as PDPK1) [109] and mTORC2 [110]. AKT is a protein kinase (also called Protein Kinase B), encoded by *AKT1*, *AKT2* and *AKT3*, which has a wide range of targets includ-

ing TSC2 [111–113], BAD [114] and FOXO proteins [115, 116]. It is therefore involved in regulating a range of cellular processes that are hijacked to promote tumour development [112, 117, 118]. Other targets of AKT phosphorylation have been identified which impact important hallmark pathways in promoting cancer progression. Such targets include MDM2, which has a direct impact on p53 degradation [119, 120], and GSK3 [121]. When phosphorylated by AKT, GSK3 returns to its inactive form, which promotes activation of downstream effectors part of the WNT/β-catenin pathway [122], such as β-catenin [123], c-myc [124] and cyclin D1 [125].

As the PI3K/AKT signalling pathway has such important biological functions, there are also several control mechanisms in place, which could partly explain the great difficulties encountered in the development of single agent drugs to inhibit this signalling cascade [126, 127]. In the context of melanoma, the most important negative regulator of the PI3K/AKT pathway is PTEN [128–132]. The role of PTEN is to convert PIP3 back to PIP2, thereby terminating the signalling cascade. However, as *PTEN* is a tumour suppressor gene which is commonly silenced, this alteration would lead to increased activity through the PI3K/AKT pathway.

1.3.3 *CDKN2A*-associated regulatory pathways

Inherited germline variants in genes such as *CDKN2A* and *CDK4* have been found to play a significant role in the development of familial melanoma [133–135]. However, somatic alterations in these genes are also important contributors to sporadic melanoma [73, 136], where *CDKN2A* is one of the most frequently silenced genes. *CDKN2A* encodes two protein products: p14ARF and p16INK4A. p16INK4A is involved in the regulation of cell cycle progression and replicative senescence [137], through negative regulation of CDK4 and CDK6, thus preventing complex formation with Cyclin D (*CCND1*) and subsequent phosphorylation of retinoblastoma protein (*RB1*) [138]. p14ARF, the alternative protein produced by *CDKN2A,* has tumour-suppressive properties through its activation of *TP53* directly or through inhibition of *MDM2* [139–143]. Direct silencing of *TP53* is less common in melanoma compared to other cancers [144–146], suggesting compensatory mechanisms such as the ones mentioned above are substituting for the lack of such changes.

1.3.4 Additional important pathways

TERT is essential for telomere maintenance, and is another frequently altered gene found both in hereditary and sporadic cases of melanoma [147, 148]. Oncogenic activation of

TERT has been indicated as a critical event in the cellular progression towards malignancy, by mediating escape of the hayflick limit [35]. Frequencies of *TERT* mutations are higher in metastatic melanoma compared to primaries [73, 148–150], and mostly clusters in the promoter region where the mutation creates an E26 transformation-specific or E-twenty-six (ETS) transcription factor binding motif, resulting in increased gene expression [148, 151].

Finally, genetic alterations targeting other biological processes and pathways are also found in melanoma. Wnt signalling might play a role in melanoma progression [152–154], epigenetic modulators such as *ARID2* have been implicated as driver genes [85, 155], and the impact of genetic alterations on controlling immune regulation is emerging as an interesting area to study [156, 157].

1.3.5 Advantage of large-scale genome profiling

With the rapid advancement of next-generation sequencing technologies over the past two decades, it quickly became feasible to utilise this powerful method to aid cancer research. It is now possible to rapidly and comprehensively analyse DNA from tiny tumours, at an affordable price. This impressive development has led to the establishment of large collaborations such as the cancer genome atlas (TCGA) [96, 158], the cancer genome project (CGP) [159] and the international cancer genome consortium (ICGC) [160], which are continuously providing the research community with invaluable knowledge through characterising the genomic landscape of various human cancers. These and other large-scale genome profiling studies have made essential findings in the field of melanoma research, with the discovery of novel melanoma genes, identification of genomic subclasses of melanoma with potential prognostic value, and further insights into the molecular processes dysregulated in melanoma [73, 85, 91, 161]. Yet, there is still much to learn, especially in the context of primary melanoma where only smaller cohorts have been studied to date.

1.4 Melanoma management

1.4.1 Diagnosis

Suspicious melanoma lesions are often initially evaluated based on the ABCDE rule (Fig. 1.5), where A stands for asymmetry, B for border irregularity, C for colour variegation, D for diameter >6mm and E for evolving (Fig. 1.5).

A combination of these markers [162–164] and other features such as the “ugly duckling” sign [165], could indicate the presence of an early melanoma. However, a biopsy

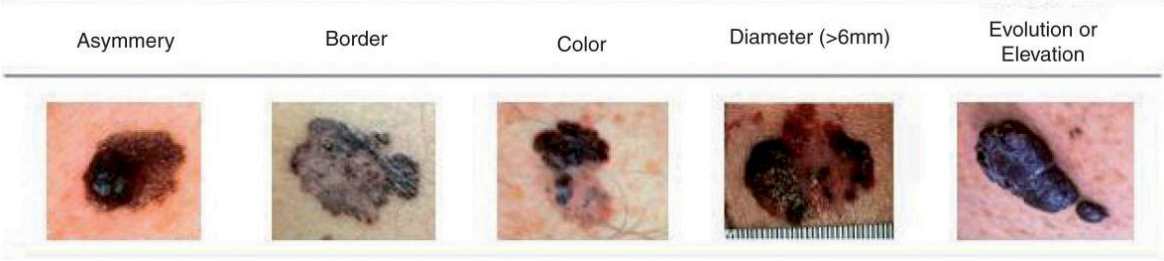


Figure 1.5: **The ABCDE rule of melanoma diagnosis.** A: Asymmetry, B: Border, C: Colour, D: Diameter >6 mm, E: Evolving. Reproduced from [57], under a Creative Commons CC BY-NC 4.0 license.

is needed for in-depth pathological assessment for accurate diagnostic and prognostic purposes [166]. The early detection of melanoma is crucial as early stage melanomas have an exceptionally good prognosis, while the survival rate declines rapidly with advancement of tumour stage [2, 77]. As such, many efforts to raise awareness, educate the general public on skin self-examination, as well as the initiation of mass screenings have been undertaken over the past decades [167–175].

Many newly diagnosed melanoma patients with intermediate to thick primaries also undergo a sentinel lymph node biopsy to exclude the possibility of nodal spread [176–178]. Additionally, if the presence of tumour cells is found in the biopsy, complete lymph node dissection usually follows. However, recent data suggest the excision of all remaining regional lymph nodes might be excessive, as melanoma-specific survival is not improved with this strategy [179].

1.4.2 Therapy for melanoma

Tumour characteristics from biopsy of the primary melanoma are used to classify tumours according to prognostic factors such as AJCC stage, and are informative in therapy decisions. Patients with early stage melanomas, where the tumour is confined to a localised region only, are mostly treated surgically, where complete excision of the primary can be curative [25, 166, 180]. In cases where removal of the primary tumour is considered insufficient or the tumour is deemed high-risk, such as when loco-regional spread has occurred, adjuvant therapy can be offered to lower the risk of recurrence [25, 181–183]. However, limited effects on survival have been observed and there is no established standard of care [184, 185]. For stage IV melanomas, where distant metastases are found, surgical resection is performed where possible [186]. However, chemotherapy, radiotherapy, targeted therapy and immunotherapy have historically also been considered treatment options [182, 187].

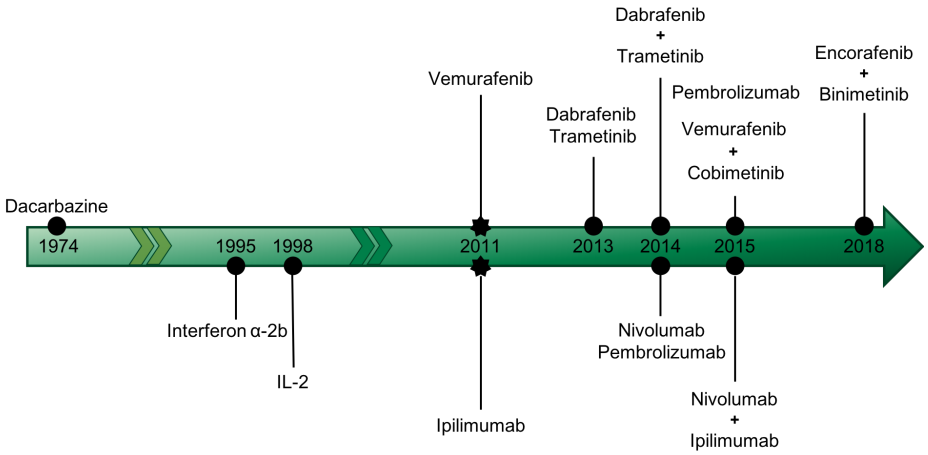


Figure 1.6: **Timeline showing selected key melanoma FDA approvals since 1974.** FDA-approved drugs from 1974 to 2018.

There have been dramatic improvements in non-surgical interventions for melanoma treatment over the past decades (Fig. 1.6). For a long time, palliative care using Dacarbazine, was the predominant option to treat inoperable advanced melanoma [188, 189]. Complete response on Dacarbazine was rare (5%) and the median response rate was limited to 5-6 months [190–192]. Therefore, several Dacarbazine-combination regimes have been attempted in clinical trials, unfortunately with limited improvement of survival rates [193–199].

In the late 1990s, three early immunotherapy drugs (interferon α-2b, interleukin-2, and ontak) gained FDA approval, showing slight survival benefit in subgroups of melanoma patients [200–202], such as those with ulcerated tumours [203]. The first breakthrough however, came almost a decade later, when the *BRAF* V600E mutation was discovered as a key genetic alteration in melanoma patients [94]. This finding led to the development of selective BRAF inhibitors, with Vemurafenib showing significant improvement compared to standard of care in terms of response and survival [204–207]. In 2011, it became the first approved mutation-dependent targeted therapy in melanoma, with many others to follow in the subsequent years [208–210]. In conjunction with this progress, the genetic component of melanoma became a hot topic, and molecular testing to determine the presence of not only *BRAF*-mutations are now routine in clinical practice [182, 211]. Unfortunately, despite an initial tumour regression, resistance to targeted therapy tend to occur. This is likely due to reactivation of the MAPK pathway and other key cell proliferation and survival-regulatory pathways [212, 213]. Intensive research into combating resistance is ongoing, where dual targeting of MAPK pathway components is a promising approach [185, 214–217].

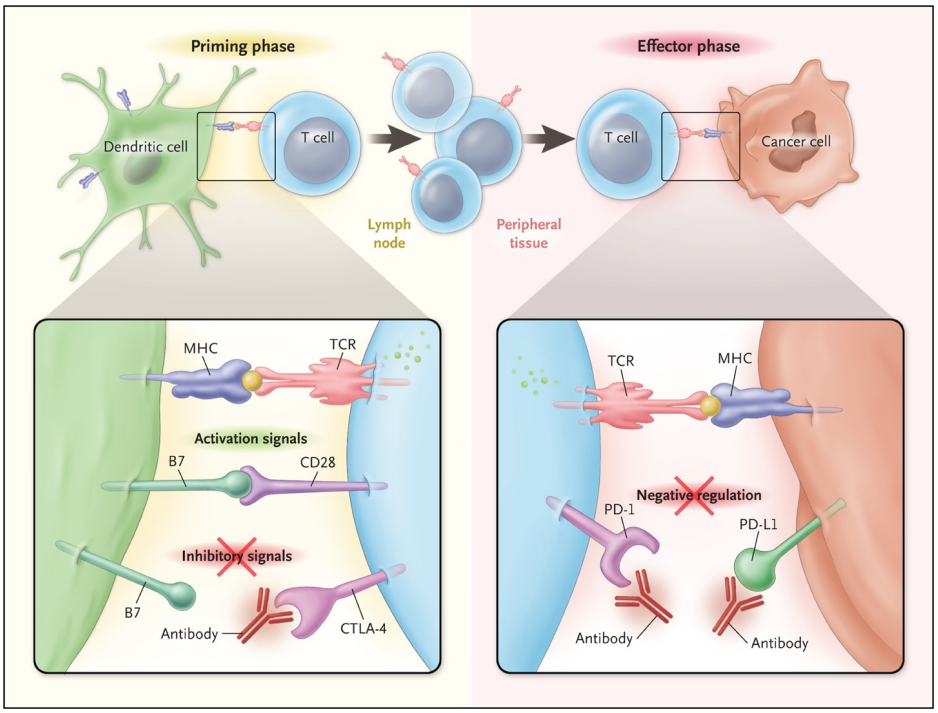


Figure 1.7: **Immune checkpoint blockade targeting PD-1/PD-L1 or CTLA-4 signalling.** In the priming phase, blockade of CTLA-4 restores T cell activation. In the effector phase, blockade of the PD-1/PD-L1 interaction restores effector T cell function. Reproduced with permission from [227], Copyright Massachusetts Medical Society.

2011 was a landmark year for melanoma, but not only due to the FDA approval of the first BRAF inhibitor. The same year, Ipilimumab, an immune checkpoint blockade antibody targeting Cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), was also approved for use in melanoma [218]. Additionally, evasion of the immune system was added as a hallmark of cancer, recognising the importance of immune surveillance in mounting an anti-tumour response [38]. As cancer cells evolve, the genetic alterations they acquire distinguishes them from a normal cell. Each mutation increases the risk of neoantigen presentation on the cell surface to signal to immune cells for destruction [219, 220]. With a functional immune system, the tumour should be identified as foreign and destroyed. Therefore, the development of a tumour indicates impairment of host protection capabilities [221–224]. A common mechanism involves the dominance of inhibitory signals in the tumour microenvironment, leading to immune cell exhaustion and reduced anti-tumour activity [225, 226]. Immunotherapy using immune checkpoint inhibitors work by blocking these inhibitory signals sent to immune cells, thereby awakening the force of previously tumour-suppressed cells to attack the tumour (Fig. 1.7).

To activate a naive T cell, a major histocompatibility complex (MHC)-bound peptide must interact with the T cell receptor, followed by a second T cell activation signal provided by the binding of CD28 to its ligands CD80 (also known as B7.1) and CD86 (also known as B7.2) [228]. In an activated T cell, the expression of CTLA-4 is upregulated, a feedback mechanism to control the immune response. The inhibitory receptor CTLA-4 then competes with the co-stimulatory receptor CD28 for binding to CD80 and CD86, but CTLA-4 has a higher affinity for its ligand, causing a dampened T cell response [228–230]. High expression of inhibitory receptors such as CTLA-4 is a marker of T cell exhaustion, indicating hampered T cell function, commonly found in a suppressive tumour microenvironment [231, 232]. By blocking CTLA-4 using antibodies like Ipilimumab, this suppressive signal is removed, and T cell activity is restored. Additional anti-tumour mechanisms linked to anti-CTLA-4 antibodies have also been proposed, including depletion of intratumoural regulatory T cells [233–235] and blocking of trans-endocytosis of B7 molecules on dendritic cells [236].

Another major inhibitory signal in the tumour microenvironment results from the interaction between the programmed cell death protein 1 (PD-1) receptor on immune cells and its ligand programmed death-ligand 1 (PD-L1) (Fig. 1.7) [231, 237]. PD-1 is mainly expressed on activated T cells, B cells and myeloid cells [228, 238, 239], and the expression is maintained as a consequence of continuous antigen exposure, causing immune cell exhaustion. PD-1 binds PD-L1 and PD-L2. PD-L1 expression is more widespread than PD-L2 and has been shown to be upregulated in melanoma and other cancers [226, 240–242]. While blocking CTLA-4 impacts T cell activation, inhibition of the PD-1/PD-L1 axis modulates T cell effector functions (Fig. 1.7). Blocking the interaction between PD-1 and PD-L1 therefore reverses the exhausted state of effector immune cells, facilitating T cell mediated tumour killing. Just a few years after Ipilimumab was approved, Pembrolizumab and Nivolumab, two anti-PD-1 antibodies were also granted FDA approval [243–246]. Immune checkpoint inhibitors have since been quickly making their way into standard clinical practice [185, 247, 248]. Furthermore, several anti-PD-L1 antibodies have now been approved [208, 249–251], and combination strategies targeting both CTLA-4 and PD-1/PD-L1 show success in melanoma and other cancers [252–254].

1.4.3 Future outlook

Targeted therapy and the newer immunotherapies have revolutionised melanoma care. While the standard treatment a decade ago did not improve patient survival [195], current phase

3 clinical trial data of combined Nivolumab (anti-PD-1 treatment) and Ipilimumab (anti-CTLA-4 treatment) show an impressive 58% overall survival after 3 years [253, 254]. Additionally, reports indicate these newer strategies demonstrate durable responses, with the potential of achieving long-term remission in a subset of patients [214, 255, 256]. This emphasises the power of firstly utilising genetic information to target important signalling pathways, and secondly modulating immunological mechanisms to achieve an effective host anti-tumour response. But unfortunately these newer therapies have not been completely successful, as complications including resistance to targeted therapy [257–260] and severe adverse events with immunotherapy [252, 253, 261] can occur. Therefore, immunotherapy might not be a good option in particular for frail, elderly or patients with an abnormal immune function. In addition, targeted therapy can only be administered to a subgroup of patients showing a certain mutation profile. Consequently, much work is still needed to address the limitations with current treatment regimens. These include understanding intrinsic and acquired resistance, research into designing effective treatments to patients currently not eligible for targeted or immunotherapy, as well as understanding how to improve outcome by patient stratification. In late 2017, the MSK-IMPACT panel used to profile tumours to identify genetic alterations was approved by the FDA [262], and will likely pave the way for personalised discovery of clinically actionable targets. Additionally, recent work has suggested biomarkers such as PD-L1 tumour expression, tumour mutational load and increased markers of cytolytic activity might help identify patients more likely to respond to immunotherapy [220, 263–270]. The important research leading to the development of immune checkpoint inhibitors was recognised by the 2018 Nobel prize in medicine which was awarded to the two immunologists James Allison and Tasuku Honjo for their discovery of CTLA-4 and PD-1 [271]. We are just on the forefront of exploring genetics and immunology to combat melanoma, with many interesting and important discoveries lying ahead of us.

1.5 Regulation of PD-L1 expression to treat melanoma

PD-L1, also known as B7-H1, encoded by *CD274*, was discovered in 1999 and is a member of the B7 family of immune-regulatory cell surface membrane receptors [272]. It is constitutively expressed on a mRNA level by a wide range of tissues [272]; however, protein expression is mostly confined to the eyes, placenta, endothelial cells and activated immune cells [226, 238, 273]. Additionally, PD-L1 expression can be induced by pro-inflammatory molecules, most importantly IFN-γ [274].

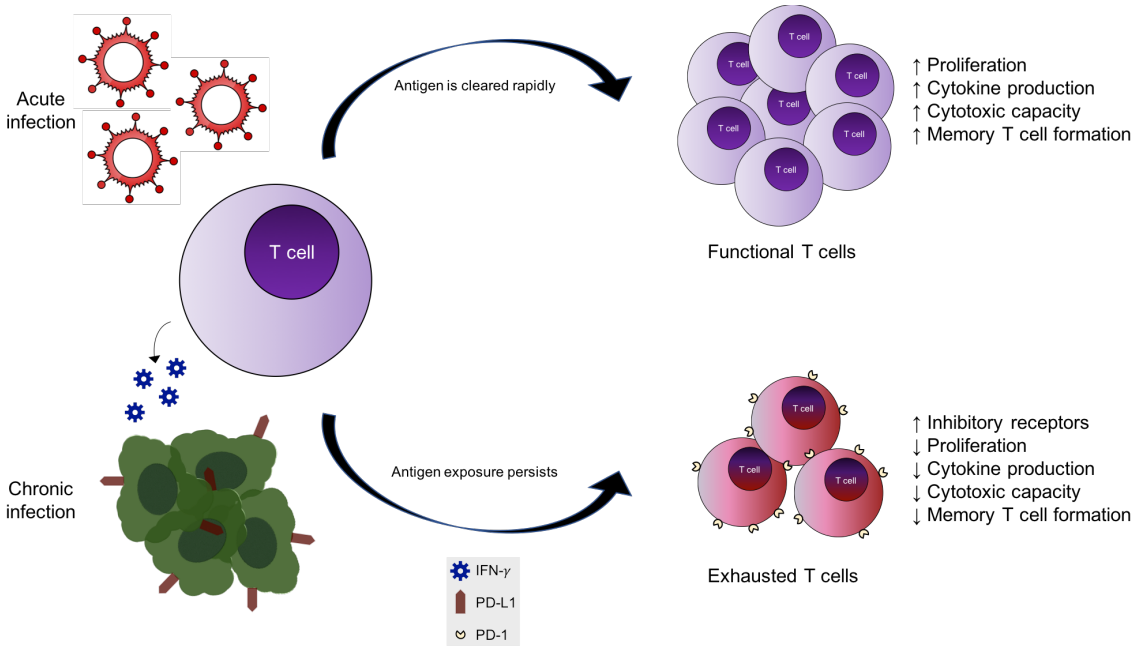


Figure 1.8: **T cell response to acute or chronic antigen exposure.** During an acute infection, antigen is cleared rapidly, therefore leading to a functional T cell response. These T cells have high proliferative capability, effector functions and are capable of forming memory T cells. During a chronic infection, the antigen exposure persists leading to T cell exhaustion. These T cells have an increased expression of inhibitory receptors such as PD-1, decreased effector functions and impaired capability of forming functional memory T cells. PD-L1 expression on tumours during a chronic infection can be triggered by IFN-γ in the tumour microenvironment.

1.5.1 PD-L1 biology

PD-L1 can bind to PD-1, which is induced on activated T cells, and this interaction mediates an inhibitory signal to suppress T cell proliferation, cytokine production and cytotoxicity [273, 275]. PD-L1 therefore functions as a control mechanism to regulate excessive immune responses, and in maintaining tolerance against self-antigens. Its expression patterns indicate it might play an important role in foetal-maternal tolerance and in restricting T cell reactivity to maintain the immune privileged status of the eye [276–278]. However, in the context of chronic infection or cancer, persistent antigen exposure and inflammation causes a state of T cell exhaustion [237]. This constant pro-inflammatory condition causes failure to generate memory T cells, loss of proliferative capability, impaired effector functions and an abundant expression of inhibitory receptors (Fig. 1.8).

In cancer, PD-L1 expression is often upregulated in the tumour microenvironment by both cancer cells and immune cells [226], which promotes tumourigenesis and facilitates tu-

mour escape [279]. Physical blockade of the PD-1/PD-L1 interaction has proven successful in reversing T cell exhaustion and restoring normal immune cell function [280, 281]. These therapies are therefore now routinely used to treat patients with melanoma, non-small cell lung cancer, bladder cancer and many other cancer types [282, 283].

In addition to PD-1, PD-L1 can also interact with CD80, which produces another inhibitory signal to T cells, stalling their activation, proliferation and cytokine production [284, 285]. This interaction has been proposed to function in the induction and maintenance of T cell tolerance [286]. Therefore, blocking PD-L1 rather than PD-1 might provide an additional advantage. Animal studies have shown that blockade with a dual-specific anti-PD-L1 antibody creates superior T cell responses in chronically infected mice compared to anti-PD-1 blockade [280]. Furthermore, PD-1 also has a second interaction partner: PD-L2, which has a more conserved expression pattern than PD-L1. PD-L2 is only inducibly expressed on dendritic cells, macrophages and bone-derived mast cells [284], and its expression is regulated by different factors compared to PD-L1 [287]. Contrary to PD-L1, PD-L2 expression on melanoma cells is only found in a small subset of samples, and its expression in the tumour microenvironment is mostly confined to stromal or immune cell subsets [288, 289]. Moreover, some studies suggest PD-L2 might also act as a co-stimulatory ligand [290], possibly through a PD-1-independent mechanism [291, 292]. This suggests a context-dependent role of PD-L2 interactions [293, 294]. Taken together, blocking PD-L1 expression rather than PD-1 might therefore be advantageous.

1.5.2 Processes controlling PD-L1 expression

The production of IFN-γ in the tumour microenvironment has multiple consequences on immune cell activity. On one hand, secretion of IFN-γ promotes a pro-inflammatory immune response by activating macrophages, T cells, B cells and NK cells, upregulating expression of MHC molecules, as well as mediating direct anti-tumour effects [228, 295–299]. On the other hand, through the process of adaptive immune resistance, melanoma cells escape immune destruction by upregulating expression of PD-L1 as a response to IFN-γ produced by tumour-infiltrating immune cells [248, 300]. Because of this, high expression of PD-L1 on tumours could both indicate an active ongoing immune response [301] and a dysfunctional state of immune cell exhaustion [237]. It is therefore not surprising to find contradictory reports on how PD-L1 tumour expression influences patient prognosis [300, 302–308].

Although IFN-γ is the strongest inducer of PD-L1 expression, other cytokines and extrinsic factors can also alter PD-L1 expression (Fig. 1.9) [309–312]. PD-L1 expression can

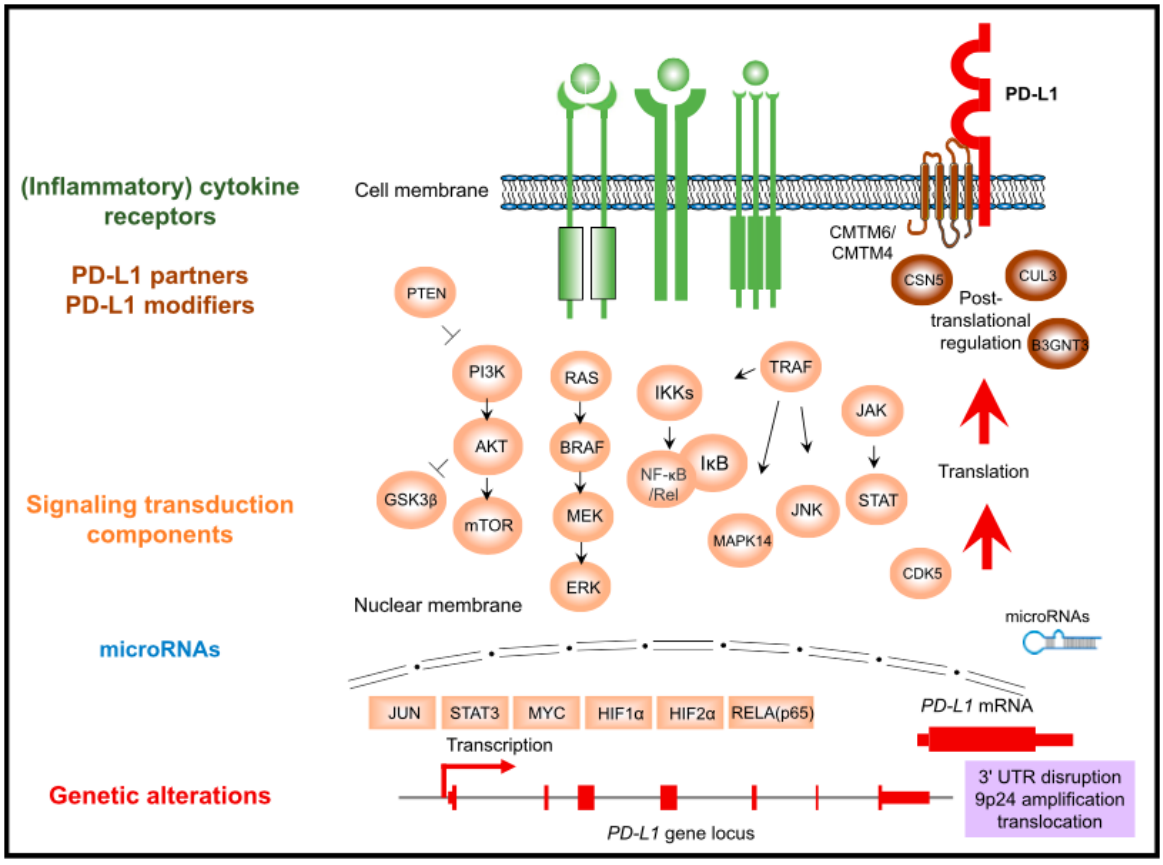


Figure 1.9: **Overview of mechanisms controlling PD-L1 expression.** In a simple model, PD-L1 expression can be regulated on multiple levels: through stimulation with cytokines, intrinsic signal transduction pathways, microRNAs and genetic alterations. Reprinted with permission from [323].

be modulated on an epigenetic level, by promoter methylation [313, 314], hypoxia [315], bromodomain or histone deacetylase regulation [316–319], or by microRNAs [313, 320–322]. Furthermore, tumour PD-L1 expression can also be induced by intrinsic factors as well as external stimuli.

By acquiring genetic alterations that would render tumours less immunogenic or confer resistance to immunotherapy, cancer cells learn to adapt and survive. Genetic alterations targeting *CD274* (PD-L1) have been found in various cancers, with cases of mutations [324], amplifications [325–329], 3'-UTR disruption [330] and genomic rearrangements [331, 332] being reported. Several oncogenic pathways can play a role in modulating the anti-tumour immune response [156, 248, 332–334], with some exerting their effect through altering PD-L1 tumour expression [335, 336].

The MAPK pathway is overactive in the vast majority of melanomas [73, 92, 107]. Sev-

eral reports have shown that oncogenic activation of this pathway contributes to immune escape [337]. Cells which acquire resistance to BRAF inhibitors show upregulation of PD-L1, possibly through transcriptional regulation by c-Jun and STAT3 [338]. Subsequently, using knock-down experiments or drugs targeting components of the MAPK pathway, PD-L1 expression could be reduced [312, 339–341]. Furthermore, patients treated with BRAF inhibitors have been shown to have an increase in immune cell infiltrate to the tumour [342, 343], improved T cell recognition [344, 345] and effector T cell functions [345]. Co-treatment using BRAF inhibitors with immunomodulatory regimens show a favourable anti-tumour response in several preclinical models [337, 346–348], and similar combinations are now being tested in clinical trials [349, 350].

Oncogenic RAS signalling, targeting both the MAPK and PI3K/AKT pathways, have also been shown to upregulate PD-L1 expression, through a mechanism of reduced tristetraprolin activity causing stabilisation of PD-L1 mRNA transcripts [351]. Loss of *PTEN* is a frequent event activating the PI3K/AKT pathway in melanoma, and this event has been linked to increased PD-L1 expression, hampered immune responses and resistance to immunotherapy [156, 335, 352]. Aberrant PD-L1 expression as a result of regulating other members of the PI3K/AKT pathway has also been reported [338, 353–355]. Furthermore, additional genetic alterations could also impact PD-L1 expression, as exemplified by *EGFR* mutations [356], CDK5 disruption [357], altered Hippo pathway activity mediated by YAP1 or TAZ activation [358, 359], and transcriptional regulation by amplified *MYC* oncogene expression [360]. It is important to note, many studies have mainly looked at the effect on PD-L1 expression on a gene level, and such changes might not always translate to a similar change in protein expression [226, 361, 362].

Research on PD-L1 is a rapidly expanding field, with emerging data from just the past few years uncovering important mechanisms of PD-L1 post-translational regulation. Some of these mechanisms causing abundant PD-L1 protein expression include targeting glycosylation through GSK3β [363], deubiquitination-associated stabilisation via CSN5 [364], cyclin D/CDK4-dependent proteasomal degradation [365] and protection versus induction of lysosomal degradation by CMTM6 [366, 367] and HIP1R [368], respectively. Regulation of PD-L1 expression on a protein level could provide valuable insights into the mechanisms that may constitute future therapeutic targets.

1.6 CRISPR-Cas9 screening approaches to identify regulators of cell surface proteins

The concept of genomic engineering, enabling precise modifications of DNA in living organisms, reached a new era with the discovery of the clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated (Cas) system [369–371]. Originally found functioning as part of the adaptive immune system in prokaryotes [371–373], researchers have over the past 20 years learned to utilise this process in creating targeted alterations in the genome of living cells [374–376]. The CRISPR locus is defined by short conserved DNA repeats scattered across the genome, separated by non-repetitive and unique sequences of a similar length [370]. Three major CRISPR-Cas systems exist, where the type II CRISPR-Cas9 system is the most widely utilised today [377].

1.6.1 The CRISPR-Cas9 mechanism

The CRISPR-Cas9 system works in three steps (Fig. 1.10). In the first adaptation (also known as spacer acquisition) stage, a short segment of foreign DNA is integrated into the CRISPR locus [373, 378]. This is followed by the interference (also known as the CRISPR RNA (crRNA) biogenesis) phase where the CRISPR locus is transcribed and processed to generate the crRNA [379]. The crRNA functions as a targeting motif, and is comprised of the unique foreign sequence, followed by a repeat protospacer adjacent motif (PAM). In the final interference step, crRNA forms a complex with trans-encoded RNA (tracrRNA), a highly abundant endogenous non-coding sequence showing part complementarity to the repeat region of the crRNA [380]. This fusion generates the guide RNA (gRNA) which then directs Cas9 proteins to the DNA site showing complementarity to the gRNA sequence. This is finally followed by the Cas9 protein exerting its function by recognising and cleaving the double-stranded DNA at a site-specific position upstream of the PAM motif [376].

In the context of prokaryotic immunity, this process leads to degradation of the foreign DNA and also provides the host with a genetic memory, protecting against future encounters with viruses containing the same recognised sequence. However, this mechanism can be designed to target any position in the human genome located three base pairs adjacent to a PAM sequence using a customised synthetic gRNA together with expression of Cas9 [376, 381, 382]. In mammalian cells, the double-stranded DNA breaks introduced by the CRISPR-Cas9 machinery will subsequently be repaired through two main processes: non-homologous end joining (NHEJ) or homology-directed repair (HDR) [383]. NHEJ is an

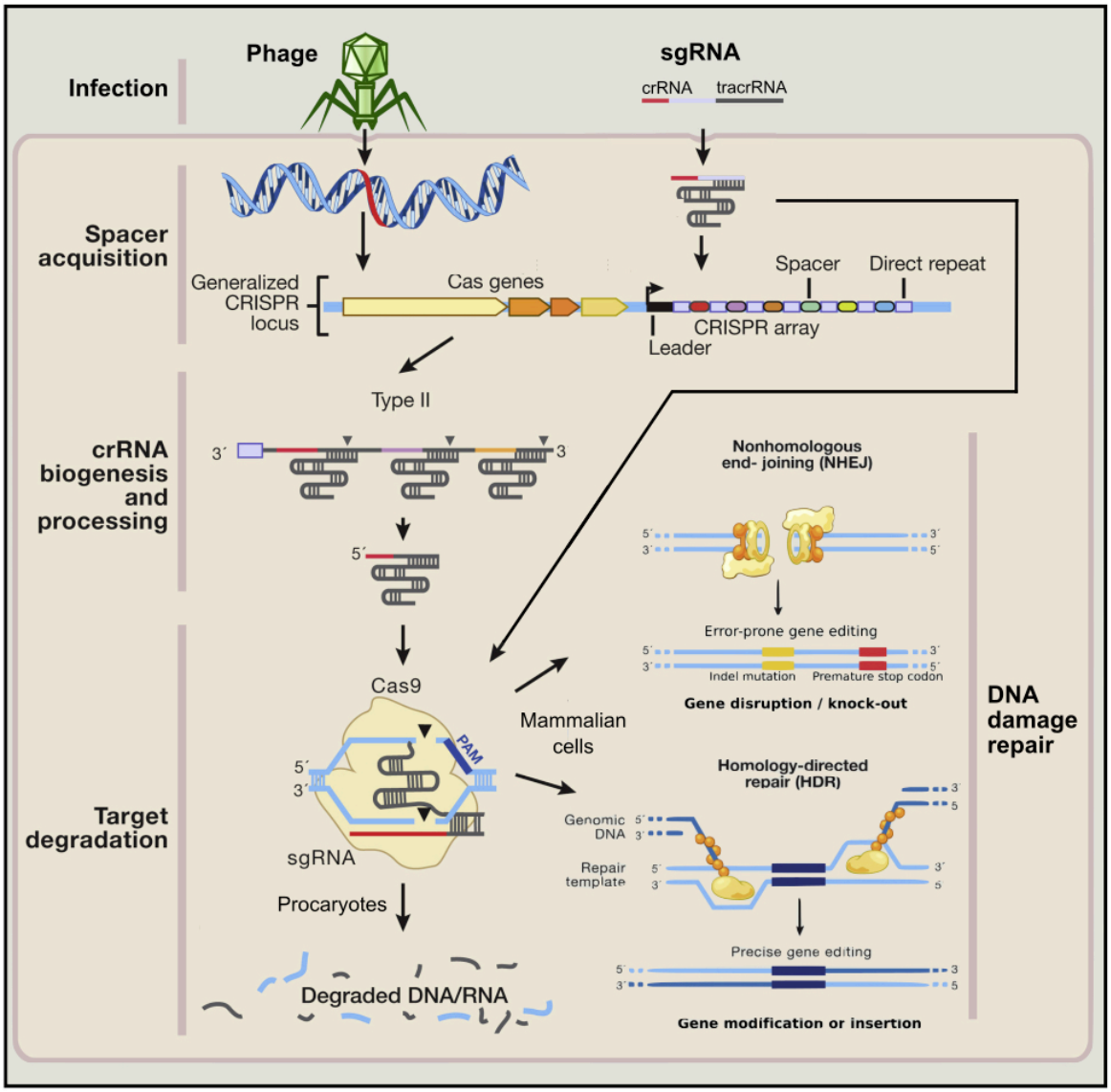


Figure 1.10: **An overview of the CRISPR-Cas9 mechanism.** DNA is delivered through viral infection, whereby the DNA is integrated into the CRISPR locus (adaptation or spacer acquisition phase). This is followed by the interference (or crRNA biogenesis) phase where the CRISPR locus is transcribed and processed. Finally, in the target degradation stage, the target is degraded by Cas9 proteins. In mammalian cells, DNA damage repair pathways generate gene modifications through non-homologous end joining or homology-directed repair. Adapted from [375] and printed with permission.

error-prone mechanism which generates insertions or deletions when repairing the cut, and is the dominating pathway when a repair template is absent. The mutations created with NHEJ could render the targeted gene dysfunctional, thereby creating a gene knock-out. HDR is much more precise and could be used to generate targeted modifications by using a customised repair template. The repair template is specifically designed to contain the desired alteration such as point mutations or sequence insertions, which would be integrated into the genome through this process. The CRISPR-Cas9 technology can therefore be used to create specific modifications to alter gene function in cells or living animals.

The field of CRISPR-Cas9 research is rapidly advancing [384], with continuous improvements including increased target specificity or efficacy [385–387], new applications such as transcriptional activation, manipulation of epigenetic marks or chromatin, inducible *in vivo* CRISPR systems and combined approaches with single-cell sequencing technology [388–392].

1.6.2 The pooled CRISPR-Cas9 screening approach

The discovery that a synthetic gRNA could be designed to alter almost any gene in the genome, led to the rapid development of optimised CRISPR components and computational tools to enhance specificity and efficacy of genome engineering [376, 381, 393–396]. Furthermore, libraries containing a pool of gRNAs with different specificities, were found successful in the application of forward genetic screens [381, 397, 398]. The first implementations of genome-wide screens utilising the CRISPR-Cas9 approach in human cells were published five years ago [399, 400], and is now widely used to study pan-genome loss of function perturbations. This approach is superior to older techniques such as RNA interference because it is capable of stable and permanent gene silencing, in addition to being highly versatile and specific [374, 401]. In a pooled genome-wide CRISPR-Cas9 screen, first a library of gRNAs needs to be designed, synthesised and cloned into plasmids (Fig. 1.11). These libraries can be custom-made using various gRNA design tools [402, 403]; however, several well-designed and frequently used gRNA libraries are also available for purchase [403–408]. The plasmid library can be packaged into lentivirus, to allow delivery of the gRNA library into the Cas9-expressing target cells. Finally, a selection pressure is applied whereby the desired phenotype can be studied by sampling the evolving cell population and comparing the gRNAs present at each stage.

Three main types of screening approaches are commonly used to identify the genes responding to the perturbation of interest [409, 410]. In a dropout screen (also known as neg-

ative selection), the majority of cells are expected to survive the selection pressure, whereas the minority of cells showing increased cell death upon treatment could provide clues as to what genes cause the decreased viability [411–413]. In an enrichment screen (also known as positive selection), the increased survival is instead measured, where the vast majority of cells are expected to die upon treatment. In the few surviving cells, this allows for identification of the genes overcoming the effects of the selection pressure such as drug resistance [400, 406, 414]. The third screening approach, is not viability or proliferation based, but instead looks at a phenotypic signal such as the expression of cell surface markers [415–417]. A potential readout would then be the consequence of gene knock-out on the expression of the marker of interest, which can be measured by comparing a control cell population with one subjected to fluorescence activated cell sorting (FACS).

Essential for all screening approaches is the subsequent sequencing of the various cell populations, to identify the gRNAs giving rise to the desired phenotype upon treatment with the selection pressure. Finally, several bioinformatic packages exist to aid the selection and ranking of hits from the screen [418, 419].

A phenotypic marker-based screening approach could be particularly interesting in understanding how important cell-surface proteins are regulated, and will be covered in Chapter 5 of my thesis to study regulation of tumour PD-L1 expression. The target of interest could be a fluorescently labelled protein or an endogenous cell surface protein which could be labelled with antibodies coupled to a fluorescent tag. The change in expression level of this target could then be measured by FACS or related strategies, whereby the genetic perturbations causing this alteration could be examined. It is possible to study either or both genetic perturbations causing a downregulation or upregulation of the target expression. These types of screens have proven successful in identifying regulators of key signalling processes, including genes controlling the host response to pathogens through assessing the LPS-mediated induction of TNF [415], or regulators of Hedgehog [420] and Wnt signalling pathways [421]. A caveat with this approach however, is the time-consuming nature of the FACS methodology, limiting the possible size of the screen. Yet, it holds a major advantage compared to the viability or proliferation-associated screens, in the possibility to examine a non-binary readout. Rather than studying death or survival, varying degrees of marker expression could be assessed separately, as well as simultaneously identifying regulators which mediate an increase or decrease in the expression of the target of interest. In conclusion, this marker-based CRISPR-Cas9 screening approach has been and will continue to be a useful strategy in identifying regulators of cell-surface proteins.

An interesting function-based approach of CRISPR-Cas9 screens has also emerged re-

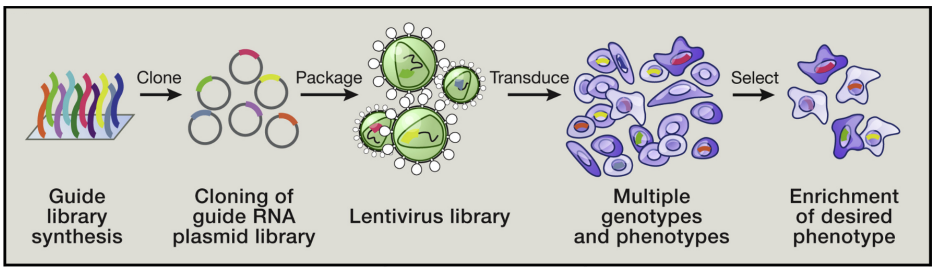


Figure 1.11: **A simple overview of a pooled CRISPR-Cas9 screening approach.** The first step is to create a library of gRNAs, which is cloned into plasmids and packaged into lentivirus. The targets cells are then transduced with this library, generating a heterogenous population with varying genetic perturbations. A selection pressure is applied and afterwards the desired phenotype can be studied. Printed with permission from [375].

cently, which has proven particularly useful in the field of immuno-oncology. These approaches typically use *in vivo* or co-culture systems of tumour and immune cells, whereby genes or mechanisms enhancing tumour cell killing [422], conferring protection to immune cell cytotoxicity [423–425], or facilitating tumour escape [333, 426] could be identified. The advent of CRISPR-Cas9 technology has revolutionised the field of genome engineering, and its rapid advancement will certainly continue to transform the way research is being conducted in ways beyond imagination.

1.7 Outline of my thesis

My thesis comprises two separate projects, both united under the scope of melanoma research, with one focusing on melanoma genetics and the other on immune evasion. Therefore, my thesis will be structured into two main parts reflecting the respective projects I have worked on. Part I will contain the three chapters that constitute my melanoma genetics project: the genetic landscape of human primary melanoma. The first results chapter will describe the sequencing methods and quality control of the data, followed by two results chapters where I describe the key genetic alterations and dysregulated biological pathways in human primary melanoma. Part II, will outline through one extensive chapter: the design, set up, performance and initial validation of a CRISPR-Cas9 screen identifying regulators of PD-L1 tumour expression. Finally, the last chapter will bring together and conclude my thesis as a whole.

Part I

The genetic landscape of human primary melanoma

The overarching goal of Part I is to understand how genetic alterations shape melanoma as a disease. The genetic composition of melanomas holds the key to understanding melanoma development, such as the genes or signalling pathways which might play a large role in tumourigenesis, and the interplay between these components. Due to the significant heterogeneity found across melanomas, it is also important to outline the genetic differences between different subgroups of melanomas, which can be distinguished by histopathological subtype, sun exposure, mutational or pathway alterations. Ultimately, genetic changes could also have a prognostic impact, which is important to explore. Overall, with this project I hope to visualise how the successful application of large-scale sequencing of human primary melanoma can provide valuable insights into melanoma development, progression and patient prognosis.

Chapter 2

Sequencing methods and QC

2.1 Introduction

In this chapter, I will introduce the human primary melanoma dataset (Leeds melanoma cohort) I have generated sequencing data for and worked on for part I of my thesis (Section C.2). I describe the cohort characteristics and sequencing methods, including the key parameters that makes this dataset remarkable and unique compared to all previous studies. Like any project dealing with extensive sequencing data, it is essential to perform comprehensive quality control prior to analysis. Therefore, I have utilised a broad range of tools to assess and ensure the trustworthiness of the data used for this project.

Details of any software tools including specific parameters can be found in Section C.1.

2.1.1 Chapter aims

The aim of Chapter 2 is to describe the sequencing methods and confirm that the quality of the sequencing data is sufficient for downstream analyses by:

- Ensuring the low amounts of DNA from primary melanomas will provide enough coverage to assess the driver gene landscape
- Assessing the consequences of FFPE preservation on the data quality
- Ensuring the choice of method used for variant calling will provide the highest accuracy and is best suited for my dataset

2.2 Sequencing of human primary melanomas

2.2.1 Cohort description

In an effort to thoroughly examine the genetic landscape of primary melanoma, I have utilised one of the largest cohorts of primary melanomas to date, the Leeds melanoma cohort (Section. C.2). This cohort consists of over 2000 patients from the Northern UK region [427], recruited shortly after melanoma diagnosis. These patients have thereafter been followed-up in the clinic for over 15 years. A major advantage of this cohort is not only the length of follow-up, but also the additional extensive clinical, pathological and lifestyle data that have been gathered, see Table 2.1 for some key characteristics. Further variables such as evidence of regression in the primary tumour, smoking, sun exposure, tanning and sun sensitivity scores, history of other illnesses such as autoimmune disorders, diabetes or other cancers are also possible to study in conjunction with the sequencing data. The patients' primary tumours were biopsied shortly after recruitment, and where sufficient DNA could be obtained, data generated from these samples were used in my project. Targeted therapy and immunotherapy were not developed until the later stages of this study, therefore 97% of the patients in this cohort can be considered treatment-naive, as they have not received any such treatments. As such, the genetic alterations studied with this project reflect the innate genetic state of each patient, a trait most other studies does not benefit from.

2.2.2 Targeted capture bait design

To study genetic alterations in the primary melanoma cohort, I custom-designed an Agilent SureSelectXT library, summarised in Table 2.2. More details can be found in Table A.2 and Section B.1. Rather than using exome sequencing, I elected to tailor the design for my specific project. This not only cuts down sequencing costs by limiting the sequencing to only capture genes of particular interest, but also allowed me to capture non-coding regions and positions important for copy number calling and human leukocyte antigen (HLA) typing.

The bait library was designed to capture genes or genomic locations potentially altered in melanoma. Probe group 1 consists of probes targeting coding regions and untranslated regions (UTRs) of 551 genes. This list includes melanoma driver genes [73, 91], significantly mutated genes from TCGA [428], melanoma driver genes identified using a dN/dS model [429, 430] and genes in the Intogen database mutated in more than 3% of melanoma samples. Additionally, 254 genes associated with solid cancers, designed by Dr. Ultan McDermott (Section C.1), were also included in this group to allow discovery of new genes

Table 2.1: **Leeds melanoma cohort patient clinical characteristics**. The key clinical characteristics of patients in the Leeds melanoma cohort, which had their primary tumours sequenced as part of my project. The patients are split by the overall cohort or the current survival status of patients. Data originated from patient clinical records.

			Death	
	Overall (n=524)	Alive (n=351)	Melanoma (n=141)	Non-melanoma (n=32)
Sex				
Female	263 (50 %)	193 (55 %)	59 (42 %)	11 (34 %)
Male	261 (50 %)	158 (45 %)	82 (58 %)	21 (66 %)
Age (years)				
Mean (SD)	57 (± 12)	55 (± 12)	60 (± 12)	66 (± 8.1)
Stage[a]				
I	167 (32 %)	133 (38 %)	25 (18 %)	9 (28 %)
II	253 (48 %)	168 (48 %)	67 (48 %)	18 (56 %)
III	97 (19 %)	47 (13 %)	46 (33 %)	4 (12 %)
Breslow thickness[a] **(mm)**				
Mean (SD)	3.0 (± 2.4)	2.6 (± 1.9)	4.1 (± 3.2)	3.3 (± 2.5)
Ulceration[a]				
No	289 (55 %)	215 (61 %)	61 (43 %)	13 (41 %)
Yes	169 (32 %)	89 (25 %)	67 (48 %)	13 (41 %)
Mitotic rate[a] **(mitoses/ per mm2)**				
<1	66 (13 %)	52 (15 %)	11 (8 %)	3 (9 %)
>=1	402 (77 %)	264 (75 %)	116 (82 %)	22 (69 %)
Tumour-infiltrating lymphocytes[a]				
Absent	83 (16 %)	57 (16 %)	24 (17 %)	2 (6 %)
Yes (Unclassified)	47 (9 %)	36 (10 %)	6 (4 %)	5 (16 %)
Non-brisk	215 (41 %)	132 (38 %)	72 (51 %)	11 (34 %)
Brisk	77 (15 %)	61 (17 %)	13 (9 %)	3 (9 %)
Mutational subtype				
BRAF	205 (39 %)	138 (39 %)	57 (40 %)	10 (31 %)
NRAS	148 (28 %)	97 (28 %)	38 (27 %)	13 (41 %)
NF1	32 (6 %)	22 (6 %)	8 (6 %)	2 (6 %)
WT	139 (27 %)	94 (27 %)	38 (27 %)	7 (22 %)
Nonsynonymous mutation load per MB				
Mean (SD)	5.1 (± 7.2)	5.2 (± 6.7)	4.5 (± 7.6)	6.9 (± 10)
Relapse				
No	333 (64 %)	304 (87 %)	0 (0 %)	29 (91 %)
Yes	191 (36 %)	47 (13 %)	141 (100 %)	3 (9 %)
Immunotherapy (post biopsy)				
No	507 (97 %)	346 (99 %)	129 (91 %)	32 (100 %)
Yes	17 (3 %)	5 (1 %)	12 (9 %)	0 (0 %)

[a]Missing values have been excluded from table output.

Table 2.2: **Targeted capture bait design.** The bait design for the probe groups included in my targeted capture panel. More details can be found in Table A.2 and B.1.

Probe group	Description	Total probes	Region size (kbp)	Probe size (kbp)	Target regions	Tiling density	Masking	Boosting
1	Melanoma genes	203294	4481	5234	11404	3x	No masking	Max Performance
2	Copy number genes	762	44.9	49.9	111	2x	Moderately stringent	Balanced
3	Broad copy number probes	1934	115.2	172.0	960	2x	Most stringent	Balanced
4	Specific copy number probes	5227	5.0	409.6	2521	2x	No masking	Balanced
5	Promoter mutation genes	1018	66.9	62	23	4x	Least stringent	Balanced
6	Fusion genes	1018	66.9	62.0	23	2x	Least stringent	Balanced
7	HLA typing panel	16351	215.5	215.5				
8	Spiradeno-carcinoma SNPs	1116	0.3	28.1	134	4x	Least stringent	Max Performance

potentially important in melanoma development. Finally, this group also includes genes of the IFN-γ pathway [431], to make it possible to study genetic alterations influencing immune regulation. My baits were shared with an adnexal cancer (henceforth referred to as spiradenocarcinoma) project, hence I included 39 genes in this probe group because of their potential role in the development of this cancer type.

Probe group 2 was designed to extract coding regions of 6 additional genes: *BNC2*, *CASP1*, *NUP107*, *PSIP1*, *SETDB1* and *TERT*. These genes or locations of these genes have been found to be amplified or deleted in melanoma [432].

Probe group 3 was designed by Dr. Ultan McDermott and has 1934 probes against 960 reference Single Nucleotide Polymorphisms (SNPs), allowing a low resolution genome-wide analysis of copy number alterations.

Probe group 4 allows more detailed copy number estimation of 28 genes, where probes were tiled spanning 80-100 heterozygous SNPs per gene (designed by Dr. Vivek Iyer, Section C.1). Positions with SNPs with a Minor Allele Frequency (MAF) between 0.45-0.55 were selected. Where the number of SNPs found in these regions were below 80, the region flanking the gene was expanded until sufficient number of SNPs were captured.

Probe group 5 was designated to promoter mutations, with a total of 330 probes spanning 5 genomic locations for *TERT*, 3 genomic locations each for *SDHD* and *DPH3* and 1

genomic location for *NDUFB9* and *NFKBIE* each.

Probe group 6 focuses on gene fusions, designed by Dr. Ultan McDermott (Section C.1), where introns adjacent to exons known to be part of a gene fusion were targeted.

Probe group 7 captures HLA genes [433], to allow haplotyping for subsequent neoantigen prediction.

Probe group 8 was designed to capture SNPs for the spiradenocarcinoma project.

2.2.3 Sample preparation and sequencing

All sample preparation from sampling to DNA extraction were done by collaborators in Prof. Julia Newton-Bishop's group at the University of Leeds (Section C.1). Informed consent was obtained under the Multicentre Research Ethics Committee (UK): 01/3/57. Following recruitment, patient primary tumours were excised and Formalin-Fixed Paraffin-Embedded (FFPE). These blocks were then sampled using a Tissue Microarray (TMA) needle to allow 0.8 x 0.2 mm cores to be extracted. Horizontal sections were taken from the part of the tumour deemed to contain the least stroma and adjacent tissue. Average tumour purity was estimated to be approximately 70%, measured visually [434], using RandomSpot [435] and ESTIMATE [436]. Blood samples were also collected from the patients. DNA and RNA were simultaneously extracted using Qiagen AllPrep DNA/RNA FFPE kit. RNA was used for transcriptome sequencing using Illumina WG-DASL HT arrays, performed at University of Leeds; and not part of my project.

DNA from 544 tumours and 487 germline blood normal (henceforth referred to as normal) samples were sent from collaborators at University of Leeds. All samples were subjected to in-house DNA pull-down using the targeted capture custom design described in Section 2.2.2, with the Agilent SureSelectXT baits diluted 1/24. In short, DNA was sheared to obtain shorter fragments, followed by indexing and 8-14 cycles of PCR amplification. Next, DNA libraries were pooled at equimolarity and hybridised with the biotin-tagged bait library. Finally, bait-hybridised DNA was captured using streptavidin-coated magnetic beads, followed by purification and 13 cycles of PCR amplification.

Sequencing was done using the Illumina HiSeq4000 platform, using 75 base pairs (bp) Paired End (PE) sequencing, at 24 samples per lane to yield 2 Gb of sequence reads per sample. Reads were mapped to the human reference genome assembly GRCh37d5 using BWA-mem version 0.7.15 [437] and duplicates marked using biobambam bamsormadup version 2.0.72 [438]. All steps from library preparation, including sequencing and alignment were performed by the Wellcome Sanger Institute (henceforth referred to as Sanger) pipelines teams.

2.2.4 Variant calling

Single Nucleotide Variant (SNV) mutation calling was performed using Caveman v.1.11.2 [439], with Vagrent to annotate the variants. 59 tumours did not have a matched normal, for which a random normal sample, PD36169b, was used as the normal sample for the somatic variant calling of these samples. The initial post-processing filters used are specified in Table A.3, and include removal of variants with insufficient number of reads containing the variant, high proportion of low quality of reads with the variant and presence of the mutation only on one strand. The Exome Aggregation Consortium (ExAC) release 0.3 was then used to filter out known polymorphic variants at a population frequency of <0.001. Only variants falling in the targeted capture region designed for variant calling were considered. Additional filters to reduce false positive calls were applied as follows: First, variants called in positions where the coverage was <10x in either tumour or normal sample were excluded. Next, variants which both had a MAF <0.10 and a coverage <30x in the tumour sample were excluded. Coverage across the *TERT* promoter region was lower than average, as this region is particularly difficult to sequence because of high GC-rich content [440]. Therefore, for variant calling in known recurrently mutated promoter positions of *TERT*, the slightly less stringent criteria of three or more reads with the mutated base was applied instead.

Pindel v.2.2.2 [441] in conjunction with Variant Effect Predictor (VEP) [442] was used to call and annotate functional consequences of indels. The post-processing filters used for Pindel are specified in Table A.4. The indel calls were further filtered to exclude calls with low mapping quality (SUM_MQ < 150 for 1-10 bp indels and SUM_MQ < 100 for indels of at least 11 bp), coverage <10x or MAF <0.10 in combination with <30x coverage. Additionally, variants with only one supporting read were filtered out. As recurrent indels are highly unlikely providing the low frequency of such variants, a sample recurrence filter was also applied, to remove any identical indel variant being discovered in three or more samples.

2.2.5 Copy number calling

Copy number calling was done by Dr. Kerstin Haase (Section C.1), using allele-specific copy number analysis of tumours (ASCAT) [443]. Briefly, alleleCount was run on all tumour-normal pairs. Based on the 1000 genomes phase 3 SNPs, and using a coverage threshold of minimum 8 reads, the B allele frequency (BAF) and logR were calculated. The BAF is the frequency of the non-reference allele, and logR is the read depth ratio between the tumour and matched normal, subjected to normalisation and log2-transformation. Then

the samples were subjected to GC correction, followed by removal of SNP positions showing homozygous BAF data for germline samples. Finally, segmentation data was obtained using the Allele-Specific Piecewise Constant Fitting algorithm and a grid search performed to retrieve purity, ploidy and copy number data for the samples.

Samples estimated to be 100% pure were removed from the analysis, as such samples were likely cases of bad fitting of the model. The following filters were then applied to identify copy number events:

1. Samples showing whole genome duplication were defined at a ploidy cutoff of >2.7.
2. Regions showing homozygous deletion had a total copy number of 0.
3. Regions showing high level amplification were classified as a total copy number ≥ 5 for diploid samples and ≥ 9 for tetraploid samples.
4. Regions showing loss of heterozygosity had a minor copy number of 0.

Copy number events at a gene level were set using a strict filter: only in cases where the whole gene was affected by the change was the event assigned on a gene level. Loss of heterozygosity (LOH) events affecting chromosome X, were removed for all male samples.

2.3 Assessment of data quality

All quality control values were generated for reads mapping to the target region, excluding genes in the HLA typing panel (probe group 7). Samtools stats was used to generate information about PCR duplicate rates and samtools bedcov was used to generate coverage information for unique reads across all positions in the targeted capture design, see Table A.1 for further details of software parameters used.

2.3.1 PCR duplicate rates

All sequenced samples show an average PCR duplication rate of 28%, irrespective of tumour or normal origin. Several of the tumour samples had very low amounts of input DNA. Subsequently, 10 samples failed library preparation and high duplicate rates were observed for samples with very low DNA amounts (Fig. 2.1A,B). Only 1 of the normal samples had low DNA input, which yielded 70% duplicate reads. All 24 samples sequenced in library pool 43 showed a slightly higher duplicate rate than the other samples, potentially reflecting batch processing effects.

High duplicate rates are expected of samples with low input DNA, and will have an impact on sequencing coverage, with lower coverage of such samples being inevitable.

2.3.2 Sequence read coverage

The average coverages, calculated for each sequenced sample across the entire capture region excluding HLA genes, were 49x for tumours and 68x for normal samples, with tumours of low input DNA showing lower coverage. 12 tumour samples were excluded from further analysis due to an average coverage below 10x. No normal samples were excluded, as all normal samples had an average coverage of 10x or above. Samples of library pool 43 showed a lower than average coverage, probably due to the increased PCR duplicate rate. All genes in the panel had an average coverage across tumour samples above 10x, including important melanoma driver genes (Fig. 2.1E). Overall, the genes in the panel show a mean coverage across tumour samples of 53x, which should be sufficient for mutation calling.

The coverages were lower than the desired coverage, which can be explained by the following reasons: First and foremost, in the bait design, no masking criteria was used for the melanoma genes (probe group 1), which made up over 90% of the whole region size being selected in my design. UTRs commonly contain small repetitive regions such as Alu elements, which comprise over 10% of the human genome [444]. Therefore, including probes in the design which targets such segments will lead to the sequencing of very scattered reads across the genome, and loss of bait target coverage. Secondly, many of the tumour samples had very low amounts of input DNA, as can be expected from tiny primary melanomas. This led to an increased number of duplicate reads, and subsequently lower coverage. Thirdly, the very polymorphic HLA regions were not included when calculating the desired sequencing coverages. The total genomic size of these genes were 215.5kb; however, the different probes used to capture these regions amounted to 2Mb of bases. The total size of the genome targeted by the probes in my design, excluding the HLA probe group, was 6Mb. Therefore, these probes made up 25% of the total design, rather than 3%. The impact this has on the sequencing coverage would be approx. 10% loss compared to the original calculation. Despite these caveats, several early landmark genomic sequencing studies in other cancer types relied on coverages similar to my study or lower [445–447], with success.

To estimate the impact of the sequencing coverage, ABSOLUTE [448] was used to calculate the power to detect mutations at different subclonality levels with an average coverage of 49x, see Fig. 2.1F. Dr. Vivek Iyer (Section C.1) adapted the power calculation to fit my dataset, using inputs of 70% purity and assuming a ploidy of 2n, sequencing error rate of

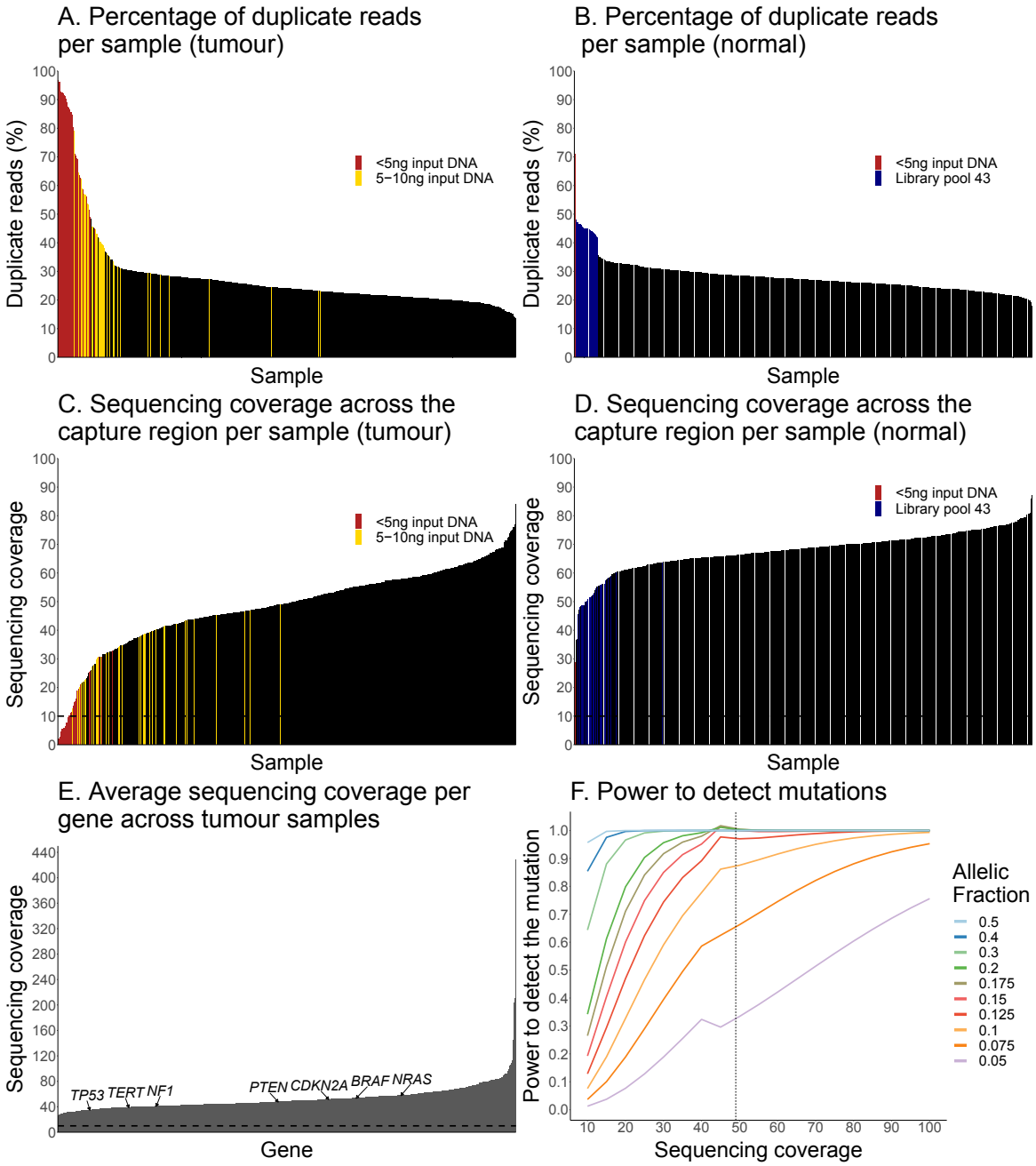


Figure 2.1: **Assessment of sequence coverage measures. A, B)** The PCR duplicate rate (%) across all sequenced samples of tumour or normal origin. Samples with low amounts of input DNA is highlighted in red or yellow, and those processed together as part of library pool 43 are marked in blue. **C, D)** The average sequencing coverage across the target region for all tumour and normal samples, respectively. The coverage cutoff of 10x is highlighted with a dashed line**. E)** The average sequencing coverage for each gene in the panel design, with 10x coverage marked with a dashed line. **F)** The power to detect mutations of varying allelic fractions at different coverages, with 49x coverage marked with a dotted line.

$1x10^{-3}$ and using a defined false positive rate of $5x10^{-7}$. With my average coverage, this results in >80% power to detect variants with an allelic fraction of 10%.

Other researchers have also investigated the impact of sequencing coverage on the proportion of mutations called using a high coverage whole genome sequencing (WGS) dataset [449]. With downsampling to simulate 50x coverage, 86% of all mutations were identified in a 100% pure tumour sample. A further 10% of mutations were lost when the tumour purity dropped to 67%. However, the biggest proportion of mutations lost as a consequence of sequencing coverage, were at lower allele frequencies. Higher coverages are necessary to pick up rarer subclonal mutations, but important mutations in driver genes are often clonal. Hence, although I will inevitably miss a small proportion of mutations, the average tumour coverage in my dataset will be sufficient for my project purpose.

2.3.3 FFPE artefact estimation

Mutations in sun-exposed melanomas often arise from UV damage, which most frequently gives rise to C>T mutations. UV damage causes cross-linking of pyrimidine bases, and when such positions undergo DNA replication, an adenine is incorporated on the opposite strand, causing a C>T base change [450].

The process of preserving tumour tissue though FFPE processing, inevitably results in minor DNA damage such as formaldehyde-induced crosslinks, deamination of cytosine molecules and the creation of abasic sites [451–453]. When subjected to PCR amplification, such damage results in erroneous nucleotides, in particular adenine, being incorporated opposite the damaged base [454, 455]. The resulting DNA would therefore have an artificial C>T mutation after sequencing, which would be difficult to distinguish from UV-induced C>T mutations.

Because of this, FFPE samples are more prone to sequencing errors than fresh tumour tissues. However, most errors will be random, and is not expected to repeatedly target the same position of all DNA molecules. Therefore, variants found in several samples are not likely caused by FFPE artefacts. Similarly, FFPE artefacts most often arise as C>T mutations at low MAFs, in particular at MAFs below 10% [456], of which I have already excluded much of due to the stringent mutation calling filters. To assess the impact of FFPE artefacts among the final mutation calls, I looked at the proportion of low allele frequency C>T mutations. Only 3% of all C>T mutations were at $\leq$10% MAF. Additionally, the percentage of mutations at $\leq$10% MAF which were C>T base changes were 51%, compared to 68% of all mutations being C>T. Therefore, I conclude that the contribution of erroneous

mutations caused by FFPE damage should be minimal in my dataset.

Distinguishing FFPE artefact variants from true mutations can be a challenge. But as discussed, using strict filters and focusing on recurrent variants will aid in removing a proportion of such false positives. That being said, any exceptional finding in this cohort should be treated with care and validated in a separate cohort, preferably from non-FFPE tissue origins.

2.3.4 Genotype concordance analysis

When dealing with a large sequencing project, there is a small possibility of sample swaps occurring, and it is general good practice to check for such errors. A mix up of tumour DNA and normal DNA was deemed highly unlikely, as these samples had been extracted and handled separately at all times by our collaborators at Leeds University. However, a mix up among the tumour DNA or among the normal DNA samples, causing an erroneously labeled tumour-normal pair, could in theory happen. Therefore, I used bcftools gtcheck (Table A.1) to compare the raw output of all unfiltered SNPs from the Caveman variant caller for all matched tumour and normal samples. Each tumour sample was compared against all normal samples, allowing me to identify sample swaps resulting in incorrect tumour-normal pairs . The error rate, calculated as the number of discordances divided by the number of shared sites between the samples compared, were used to identify the sample pair with the lowest dissimilarity. As expected, the tumour-normal pair of the same patient, showed the lowest error rate for all comparisons (see Fig. 2.2). Therefore, no sample tumour-normal pair were identified as incorrect.

A subset of the patients in this cohort previously had their normal samples analysed using a SNP array. Therefore, these data could be used as an additional genotype check. Mutation calling for the normal samples in my cohort was performed using samtools mpileup followed by bcftools call (Table A.1). I obtained the SNP array data from Leeds University (Section C.1), and filtered for the overlapping regions between the array and my target region. Unfortunately, a total of only 18 positions were shared between the 2 methods, for which SNPs were reported in more than 1 sample. In most cases, the comparison of SNP concordance was based on just one or two positions, which limits the effectiveness of this assessment.

A total of 226 samples had SNPs in any of the 18 positions. 223 samples (98.7%) showed the same SNPs analysed with both methods. Only three samples were identified to have discrepancies between the SNP array and my targeted capture method, and these

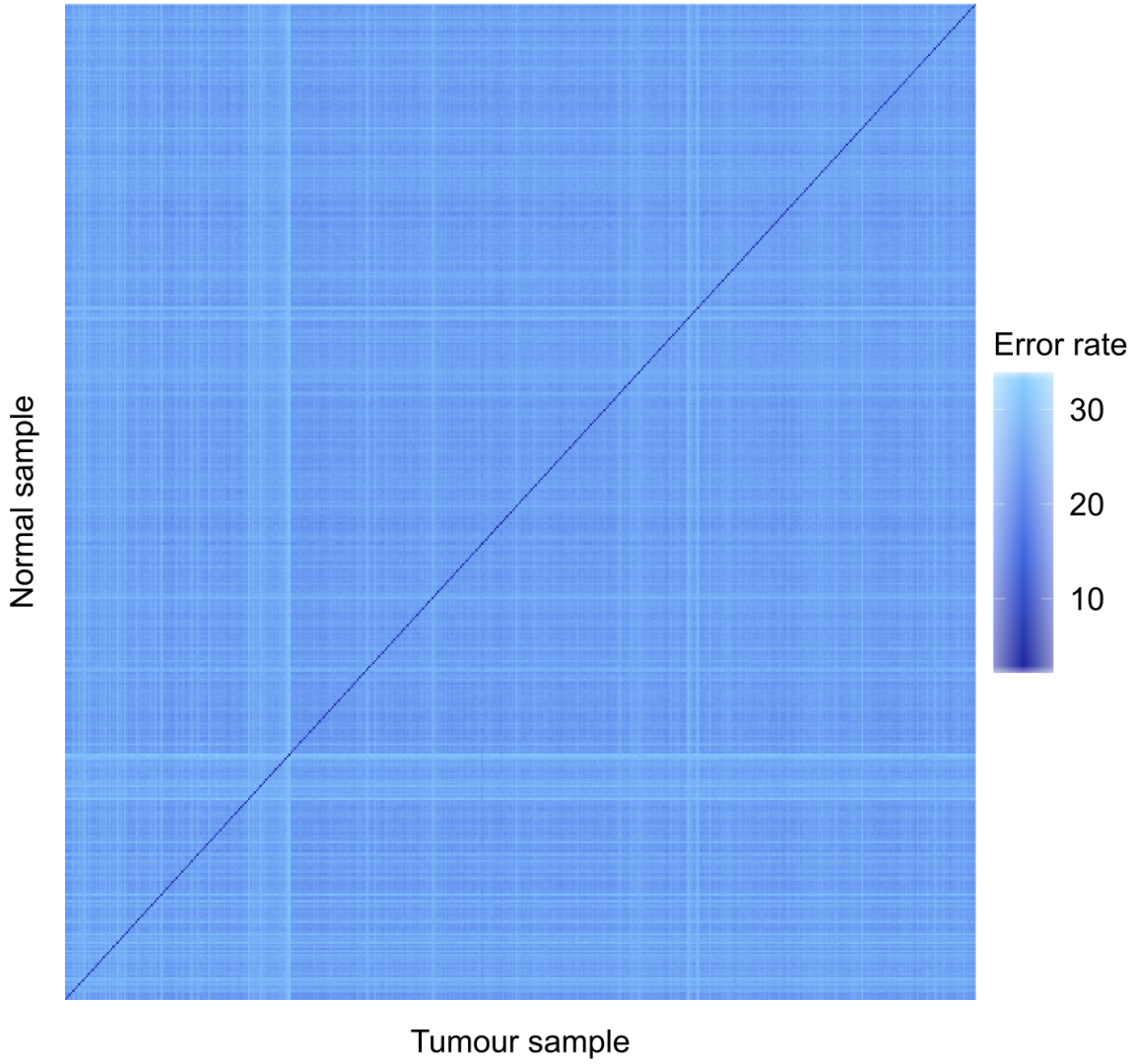


Figure 2.2: **Genotype concordance analysis to assess sample swaps affecting tumour-normal pairs.** The error rate calculated for each tumour-normal pair, with tumour samples on the x axis and normal samples on the y axis. The pair with the darkest blue colour corresponds to the tumour-normal pair with the lowest error rate.

discrepancies were investigated. For 2 samples, PD36252b and PD36311b, a variant in 1 of the 18 positions were found in the SNP array data, but not using the targeted capture method. This position had a high coverage, with no reads supporting the variant, confirmed by visual inspection using Integrative Genomics Viewer (IGV) [457]. For the third sample, PD36491b, the SNP array data called a variant in two different positions, whereas in the same sample, the targeted capture data did not identify this mutation. Absence of the mutation was again confirmed by visual inspection of the regions, which also proved to have high sequencing coverage and good quality reads. It is therefore unlikely these SNPs exist in these samples without being picked up using my method. Intriguingly, I did observe the presence of the two variants in the matched tumour sample from the same patient, rather than the normal sample. But, as mixups between tumour and normal samples were deemed impossible, and since the tumour-normal genotype check did not identify any sample swaps, I think it is highly unlikely the three discrepant samples are results of sample swaps. The exact same swap would have had to take place for both tumour and normal samples, which is extremely unlikely. More realistically, these discordances affecting 3 out of 226 samples would be due to limitations of either assay.

In summary, the genotype concordance analyses using matched tumour-normal pairs and the Leeds SNP array data showed a high concordance of 100% and 98.7% for respective analysis. This does suggest the likelihood of any sample swaps having occurred during sample processing to be extremely unlikely.

2.3.5 Variant calling concordance with patient clinical records of *BRAF* and *NRAS* mutation status

In this project, I identified patients with mutations in the clinically relevant positions *BRAF* V600E, *NRAS* Q61, and *NRAS* G12/G13. These positions are commonly analysed as part of routine clinical tests, and as such our collaborators at Leeds University provided me with a list of patients which had mutations in either of these positions, measured by pyrosequencing or from clinical records of the patients. I compared this list with the corresponding variants from my mutation calls and calculated the concordance between our methods to assess the quality of my mutation calls. 96.5% of samples (384 out of 398) showed the same genotype for *BRAF* V600E. For the *NRAS* positions, codon 61 showed a 97.4% concordance (373 out of 383 agreed) and codon 12/13 a 100% concordance (352 out of 352 agreed), see Fig. 2.3.

Of the *BRAF* V600E discordant calls, one was for a patient who had a mutation according to my analysis, but was not reported to have this mutation by Leeds University. The

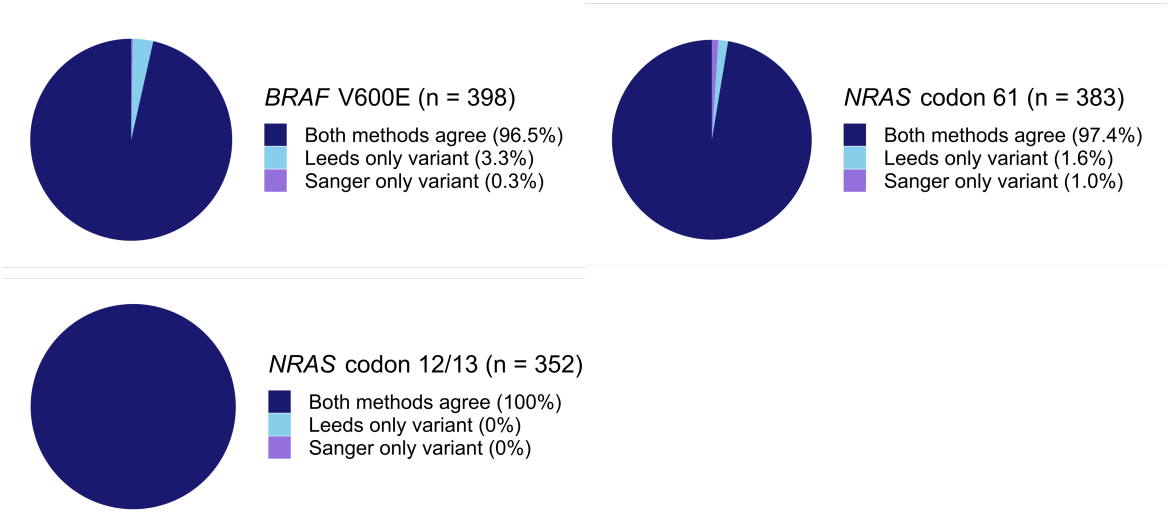


Figure 2.3: **Variant calling concordance with patient clinical records of *BRAF* and *NRAS* mutation status.** Comparison between my mutation calls (Sanger) and Leeds University records (Leeds) of *BRAF* and *NRAS* hotspot variants.

position was covered by 46 reads, with 24 of them reporting the variant, therefore I believe the mutation to be accurately called. The remaining 13 discordant calls were cases where the patient was reported to have a mutation by Leeds University, but not using my analysis. 2 of the samples had a coverage below 10 for this position, which falls below the filter threshold for my mutation calling (Section 2.2.4). One of these patients had three reads out of nine supporting the mutation, and the second patient did not have any of the nine reads supporting the mutation. In these two cases the coverage was too low in my analysis to conclude the presence of the variant, and illustrates either tumour heterogeneity or a limitation of my analysis. The remaining discordant samples had a coverage between 44-80 for this position, indicating the coverage is sufficient for mutation calling. For three of these samples, no reads supporting the variant was present, and for an additional six samples the combination of low number of alternative reads and stringent mutation calling criteria resulted in not enough evidence to assign the variant. The last two discordant samples had the mutation called, but they were filtered out due to the presence of reads supporting the variant in the germline, suggesting it might not be a somatic variant.

For *NRAS* codon 61, a total of 10 samples showed discrepant results between the methods. 4 of the 10 discordant samples had a mutation called using my analysis, whereas the mutation had not been reported by Leeds University. These 4 samples had a decent number of supporting reads, 14-24, and a coverage of 50-89 for the variant position. Therefore, I believe they are rightfully classified as having a mutation. The samples which showed a

discrepancy where the mutation was not found in my analysis were checked manually using IGV, for presence of reads containing the variant. For one sample, the variant was filtered out because the only reads containing the variant had the variant at the end of the reads, hence it was removed because of the likelihood of it being a sequencing error. The remaining five samples had no or low number of reads supporting the variant, again a consequence of stringent mutation calling criteria. Upon further investigation by the Leeds team, 5 out of the 6 samples which were classified as having a *NRAS* codon 61 mutation by them, were only classified as positive after repeating the assay, initially being classified as negative. This suggests that different assay thresholds can account for some of the discrepancies found between the data. Excluding these 5 samples from the comparison increases the concordance value to 98.7%. *NRAS* codon 12 and 13 showed 100% concordance between the two methods.

In conclusion, the few discrepant samples are likely the consequence of different assay thresholds and sensitivity. An overall high concordance between the methods, confirms that any errors occurring through sample preparation, sequencing and analysis are limited, and more importantly provides confidence for my mutation analysis suggesting the coverage and filters applied, although stringent, are appropriate.

2.3.6 Variant calling concordance with Mutect

A second algorithm was used to evaluate the false positive rate of my mutation calls. A whole range of variant callers have been developed over the years, and for the comparison I chose Mutect, because it is one of the most frequently used methods for somatic mutation calling, and has been shown to be reliable and of good performance [458–461]. Additionally, the fundamental algorithm Mutect is based on is different from Caveman. Instead of estimating genotype probabilities, Mutect works by estimating the probability of MAFs being different between the tumour and normal samples. Dr. Rashid Mamunur (Section C.1) ran the cake pipeline to yield the Mutect calls, which I used for this comparison [462].

With both methods, only nonsynonymous variants were included in the analysis. The same post-processing filters were applied to both callers: removing positions with low coverage (<10x), and low coverage in combination with low allele frequency (<30x and MAF <10%). 94% of genes which had a nonsynonymous mutation called using the Caveman algorithm, also showed the presence of a nonsynonymous mutation using Mutect (Fig. 2.4). This suggests I have a high rate of robust calls using my method.

When examining where the two methods were discordant, Mutect picked up a lot of

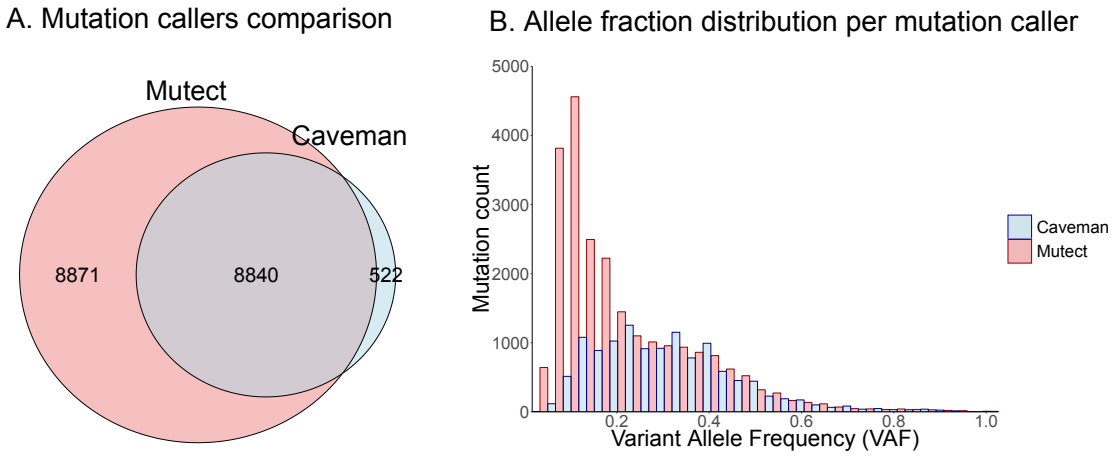


Figure 2.4: **Concordance between mutation calls using Caveman and Mutect. A)** The number of gene level nonsynonymous mutations calls with Mutect compared to Caveman, and the overlap between both methods. **B)** The allele fraction distribution of mutations called for each mutation caller.

mutations which were not found using the stricter Caveman method. However, the vast majority of the mutations called using Mutect only, were found to be variants of low allele-frequency (Fig. 2.4B). This is consistent with its algorithm design, with Mutect performing well in calling subclonal mutations. For our analyses, subclonal mutations are of less importance. While the data suggests Mutect has a higher sensitivity to pick up subclonal mutations, this comes at the expense of more 'noise'. This is especially prominent in my dataset, as the power to detect variants reduce drastically with lower allele frequency, in particular below 10% MAF (Section 2.3.2). I find it better to be strict than to increase my rate of potential false positives, therefore I conclude Caveman with strict filters applied is the right mutation calling method to use. Furthermore, as the rate of genes with nonsynonymous mutations confirmed using Mutect were 94%, the false positive calls using my method is likely to be small.

2.3.7 Samples without matched normal

59 tumours did not have matched normal samples. For these samples, PD36169b was selected randomly to run the mutation calling against. The unmatched tumour samples had a median nonsynonymous mutational load of 25.3 (range between 4 and 101), compared to the matched tumour samples which had a median nonsynonymous mutational load of 26.3 (range between 0 and 371). As no bias was found towards either low or high numbers of mutations called, this suggests the specific germline sample used as matched normal does

not cause a general over- or underestimation of mutation numbers. Additionally, low coverage of the germline sample could lead to an overestimation of somatic variants called. The selected normal sample had an average coverage of 76x, which is slightly above the average for all normal samples (68x). Therefore, an overestimation of somatic mutations as a consequence of low germline coverage is deemed unlikely.

To ensure important somatic mutations in the unmatched tumours were not missed due to specific SNPs in the selected normal sample, two approaches were taken. Firstly, germline mutation calling was performed using mpileup in conjunction with bcftools call (Table A.1). Only nine nonsynonymous germline mutations were found in the target capture region. None of these positions were reported to have somatic variants in any of the tumour samples. Secondly, because a small number of mutant reads in the germline sample can result in a potential somatic variant being discarded, the Caveman mutation calling was repeated for the 59 unmatched tumours with 2 additional, randomly selected normals. As a result, looking across all 59 tumour samples, only 1 out of the total 1235 nonsynonymous mutations called using the originally selected normal, PD36169b, were not called using any of the other two normals (see Fig. 2.5). Upon closer inspection, this is a variant in *SPTA1*, which in this tumour sample had 9 reads supporting the variant out of 70. In PD36169b all 126 reads show the reference allele, whereas PD36521b and PD36283b each have 1 out of 104 reads supporting the variant. Therefore, this variant is likely filtered out in the latter cases. Without the actual matched normal samples for these tumours, it is difficult to assess which mutations are true. Nevertheless, this comparison suggest a very low error rate, considering only 2.4% of all mutations across the 59 unmatched tumour samples were not jointly called using all 3 normal samples. Additionally, when inspecting the genes with variants only called using PD36169b, or called exclusively by the other two except PD36169b, none include well-established melanoma genes and recurrent positions. Therefore, the risk of erroneous mutation calls using PD36169b as the dedicated matched normal sample for these 59 samples can be considered fairly small.

2.3.8 ASCAT SNP distribution assessment

Analysis of copy number alterations with ASCAT relies on capturing sufficient SNPs overlapping between the sequenced genomic region and the SNP panel used by ASCAT. To investigate this, I intersected the SNP panel used by ASCAT with the targeted capture region bed file, and found a total of 170179 SNPs which could be used for copy number analysis. These SNPs were distributed across the chromosomes as shown in Table 2.3. The

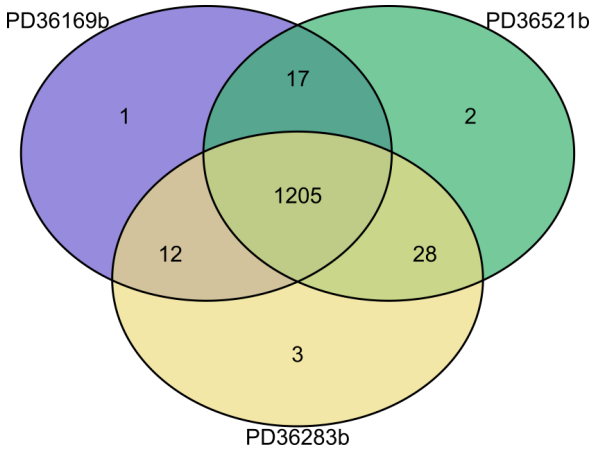


Figure 2.5: **Shared mutations in unmatched tumours called using three normal samples.** Distribution of shared and unique mutations in the 59 samples without matched normal, comparing mutation calls using 3 different normal samples.

median number of SNPs per gene in my bait panel were 215, and only 10 genes had 10 or fewer SNPs within the region of the gene. Further information can be found in Section B.2. However, not all SNPs will be informative in BAF calculations. The number of heterozygous SNPs per patient tumour-normal pair is individual, and therefore the number of SNPs used for BAF calculation will be lower than the total and different between samples. In contrast to BAF estimations, all SNPs can be used to calculate logR, and therefore, I believe the input to ASCAT is sufficient to call copy number variations in the regions containing the majority of my genes of interest. Unfortunately, focal copy number calls for genes falling in any region not included in the bait design would not be possible, as such genes would be captured as part of a larger segment with no pronounced change across specific genes as such gene-specific SNPs will be absent. However, as there are SNPs widely distributed across the genome, this still allows for whole genome low resolution copy number analysis as well as more detailed analysis of over 500 genes.

2.4 Experimental work to validate findings

2.4.1 siRNA-mediated knock-down

RVH421 is an adherent cell line, cultured in RPMI-1640 supplemented with 10% FBS and 2mM L-glutamine. Cells were seeded at 200,000 cells/well in 6-well plates and incubated at 37°C overnight. The following day cells were transfected using siRNA (Dharmacon, ON-TARGETplus SMARTpool) designed against *IRF4*, *ERH* (positive control) or a

Table 2.3: **Number of ASCAT SNPs per chromosome.** Number of SNPs per chromosome that could be used by ASCAT for copy number analysis. Further information can be found in Section B.2.

Chromosome	Number of SNPs	Chromosome	Number of SNPs
chr1	19163	chr13	2284
chr2	15478	chr14	3426
chr3	12638	chr15	4569
chr4	6194	chr16	7354
chr5	7396	chr17	9846
chr6	9941	chr18	3296
chr7	10366	chr19	6597
chr8	6593	chr20	3792
chr9	12735	chr21	3772
chr10	4796	chr22	2117
chr11	7065	chrX	1615
chr12	9143	chrY	3

nontargeting pool (negative control) according to manufacturer's instructions. Cells were retransfected after 3 or 6 days, and harvested for analysis after 10 days.

2.4.2 Flow cytometry viability assay

Cells were collected following trypsin treatment (all culture media and PBS washes were also kept), washed once with PBS and once with Annexin V binding buffer. The samples were then subjected to Annexin V-PE staining (Biolegend, 5μl/sample for 15 minutes at room temperature, followed by the addition of DAPI (Sigma, 1:5000 dilution). The samples were then centrifuged, resuspended in Annexin V binding buffer and analysed using a flow cytometer (BD Fortessa II), followed by data analysis using FlowJo v.10.

2.4.3 Confirmation of gene knock-down using Western blot

The cell culture dish was placed on ice, and the cells were washed twice with ice cold PBS, followed by the addition of ice cold RIPA lysis buffer supplemented with protease inhibitor (Sigma). Cells were scraped and transferred to an eppendorf tube, and incubated at 4°C for 30 minutes with agitation. Cells were then centrifuged at 12,000 rpm for 30 minutes, and the supernatant collected. Protein concentrations were measured using the Pierce BCA assay (Thermo Fisher) and a western blot analysis performed as described previously [463]. Samples were run under reducing conditions, stained with rabbit primary antibodies against

GAPDH (Cell Signaling, clone 14C10), IRF4 (Cell Signaling, #4964) or c-Myc (Abcam, clone Y69) followed by a Horseradish Peroxidase-conjugated goat anti-rabbit secondary antibody (Abcam, #ab6721). Protein chemiluminescence detection was performed using ImageQuant (GE LifeSciences).

2.5 Evaluation of chapter aims

- Ensuring the low amounts of DNA from primary melanomas will provide enough coverage to assess the driver gene landscape
  - Despite a loss in coverage due to limited amount of tumour DNA and limitations in the bait library design, this should not significantly reduce the possibility for driver gene discovery.
  - A high concordance between my mutation calls in known melanoma hotspots *BRAF* V600E, *NRAS* codon 61 and codon 12/13 and patient clinical records brings high confidence in my data.
- Assessing the consequences of FFPE preservation on the data quality
  - The strict filters applied limits the mutation calls at lower allele frequencies, which are more likely to arise from FFPE artefacts.
  - The final mutation calls contain low frequencies of C>T mutations at $\leq$10% MAF, therefore limiting FFPE-induced false positives.
- Ensuring the choice of method used for variant calling will provide the highest accuracy and is best suited for my dataset
  - 94% of mutations could be verified using a second algorithm, proving a high rate of robust calls.
  - Strict filters using Caveman yields lower sensitivity to detect subclonal mutations, but this selective approach will reduce false positives, and is therefore the best choice going forward.

Chapter 3

Key genetic alterations in primary melanoma

3.1 Introduction

In this chapter I will present the analyses I have performed to discover and understand some of the most important genetic alterations found in primary melanoma. Genetic alterations in cancer cells can promote tumourigenesis, but whilst melanoma is a cancer type with a high burden of neutral mutations, it is important yet challenging to uncover the critical genes and variants driving disease development. The advantages of using large patient cohorts include the possibility to utilise recurrence as a way of separating essential perturbations from irrelevant events. However, this criteria alone is insufficient, and other methods including patterns of positive selection and mutual exclusivity have also been explored in this chapter of my thesis.

Somatic mutations can have three outcomes for a cell: they can provide a survival advantage, disadvantage or exert no effect. Most mutations will not have any consequence on the selective advantage of a cell, while acquired mutations providing a survival benefit for the tumour will be selected for. An example of a positively selected mutation is the activating *BRAF* V600E mutation. This results in up to 700x increased MAPK pathway activity [94, 464], which leads to a proliferation and cell survival benefit for the cell with this mutation. These patterns of positive selection can be utilised when determining which genes are of greater importance in melanoma development. In non-driver genes, I expect to find a consistent pattern of silent, missense and loss of function mutations, with fewer nonsense mutations as these tend to be rarer events, compared to missense and silent mutations. However, positively selected driver genes will have an elevated ratio of nonsynonymous

mutations compared to silent mutations [430], and this mechanism can be utilised to define the genes driving melanoma development.

Another key method of discovering important genes contributing to melanoma development, is to investigate patterns of mutations across tumours. Mutations repeatedly affecting the same gene, or pathway, in different patients are more likely to be associated with pathogenicity. Additionally, positively selected cancer mutations do not occur randomly, as these often target pathways which provide a survival advantage for the cell. When mutations are not random, there are two main mutational patterns: mutual exclusivity and co-occurrence. Mutual exclusivity looks at the occurrence of multiple mutations with the same effect on one specific pathway. Such mutations may not be selected for as their mutual presence does not provide an additional advantage to the cell. Furthermore, the coexistence of such mutations might even be harmful for the cell, as shown in several studies [465–467] as well as being evident in patient cohorts including mine (Fig. 4.2). By studying the presence of mutually exclusive gene pairs, novel genes involved in important oncogenic pathways can be discovered. Co-occurring mutations on the other hand, suggest a co-operative or even synergistic effect of the genetic mutations involved. This mutational pattern highlights important pathways acting together to promote tumourigenesis. However, for this type of analysis, one must be cautious, especially when dealing with tumour types with a high background mutation load. Tumours with a high mutation burden have an elevated chance of any two genes being co-mutated. Additionally, due to the mutational heterogeneity across the genome, certain genes are more likely to harbour mutations due to, for example, chromatin accessibility or variations in DNA repair susceptibility [86, 468–471]. When performing analyses of mutational patterns, it is important to take into account all of the factors mentioned above. These are the main reasons behind choosing DISCOVER for my analysis, as this method employs measures to account for such variability [472].

With sun exposure being one of the main causative factors in melanoma development, it is evident that primary tumours arising at different locations across the body can have different properties. Tumours found on body sites frequently exposed to sunlight such as the head, are sometimes referred to as chronically sun-damaged (CSD) melanomas. These tumours show a particular pattern of high mutation load and a specific composition of genetic alterations, such as MAPK activation independent of *BRAF* V600E [18, 45, 54, 473]. Tumours arising on intermittently sun-exposed sites or sites commonly shielded from sunlight also show distinct features, such as a higher frequency of *BRAF* V600E mutations or a higher degree of chromosomal aberrations [54]. In this chapter, I will therefore also examine the role sun exposure has on the melanoma driver gene landscape.

3.1.1 Chapter aims

The aim of Chapter 3 is to present the key genetic alterations in primary melanoma, including:

- Outline the landscape of mutations including the top mutated genes and copy number alterations
- Identify melanoma driver genes and important genetic interactions
- Understand how sun exposure can shape the genetic composition in melanomas

3.2 Somatic mutations in the Leeds melanoma cohort

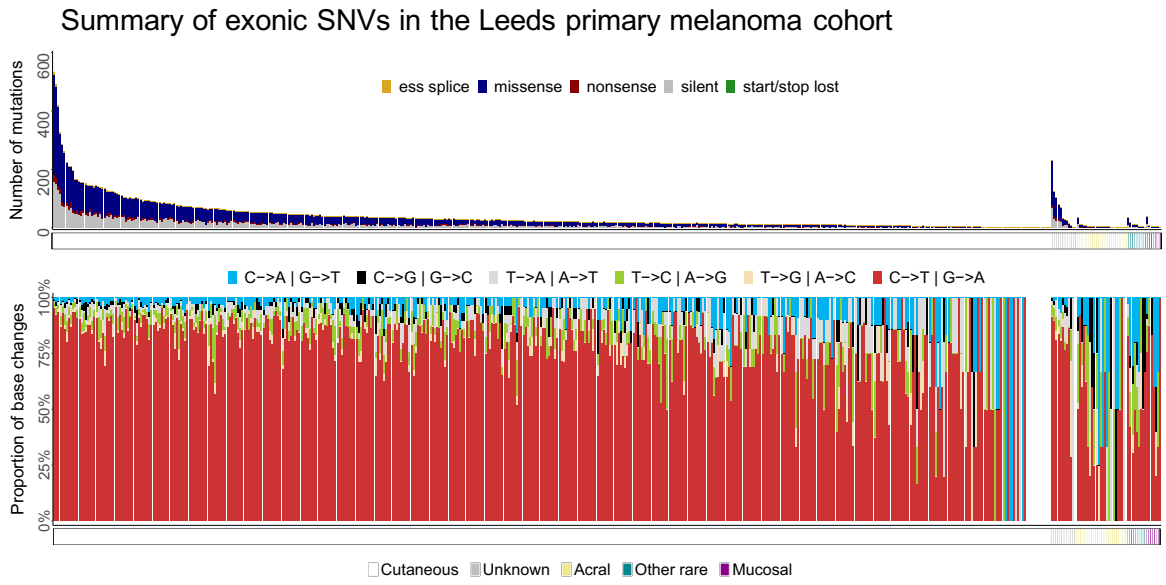


Figure 3.1: **Summary of coding SNVs in the Leeds melanoma cohort.** 524 samples were grouped based on the type of primary tumour: cutaneous, unknown, acral, other rare or mucosal, and by decreasing mutation load. Top panel: Number of exonic mutations and the distribution of variant consequences. Bottom panel: Proportion of various base changes across the samples.

Melanoma is a cancer type with a high mutational load (Fig. 3.1). As expected, as a consequence of UV damage, a high proportion of C>T base changes were found in almost all samples, with the exception of low mutational load samples and tumours classified as acral melanoma. Mutational load was calculated as the total number of nonsynonymous

mutations per sample. This value was then divided by 5.2, which is the target capture region size in MB of the melanoma genes probe group 1, to give the nonsynonymous mutation load per MB. The average nonsynonymous mutation rate was 5 mutations/MB (range of 0 to 72 mutations/MB).

I have categorised the primary tumours into cutaneous (excluding the acral), acral, mucosal, "other rare" and "unknown" subtypes. As vulval, vaginal, anal and penile primaries could not be accurately assigned as cutaneous or mucosal, they were grouped together as "other rare". A small number of tumours had been classified as nodal with an unknown primary according to the clinical records. Judging by their mutational profiles (Fig. 3.1), most contained a large proportion of C>T mutations, and were therefore assumed to be cutaneous in origin.

The top recurrently mutated positions include well-known variants in the driver genes *BRAF*, *NRAS* and *TERT*, but also hotspot mutations associated with sun exposure such as *RAC1* p.P29S (Fig. 3.2). In addition, *RQCD1* p.131L mutations were seen at a similar frequency in another cohort of primary melanomas [474]. However, lower alteration frequencies have been reported in large cohorts of metastatic melanoma (TCGA pan-cancer 2.5%) [73, 85, 161, 475], which could reflect a difference between primary and metastatic tumours. Previously undescribed hotspot variants discovered in this cohort include positions in *PCDHA2*, *TPTE* and *AHCTF1*. The S337L mutation in *PCDHA2* targets one of five Cadherin domains, while the S447L variant lies in the C2 domain of PTEN tumour suppressor protein [476] of *TPTE*. All mutation changes in noncoding regions were reported with respect to the strand location of the gene, where the position is shortened to the last three digits. As an example, the common promoter mutation in *TERT* chr5:1295228 (human genome assembly GRCh37), which is located on the reverse strand, is a C to T mutation in respect to the gene but a G to A mutation in respect to the forward strand. It is thus referred to in my thesis as a C228T mutation.

Interestingly, the C271T hotspot mutation in *AT-Hook Containing Transcription Factor 1 (AHCTF1)* affecting 3% of melanoma patients in this cohort, targeted a highly conserved region in a GABPA transcription factor binding site (Fig. 3.3). In addition to the 13 patients with this C271T mutation, 1 patient had a C271A mutation and 2 patients had the dinucleotide CC271-272TT mutation. A further six patients had mutations targeting adjacent bases. Although this region is in the promoter of *AHCTF1* and is annotated as a transcription start site by Fantom 5, the location of the GABPA transcription factor binding site suggests it will unlikely affect *AHCTF1* gene expression. *AHCTF1* is located on the reverse strand, while the transcription factor binding site is on the forward strand, making it

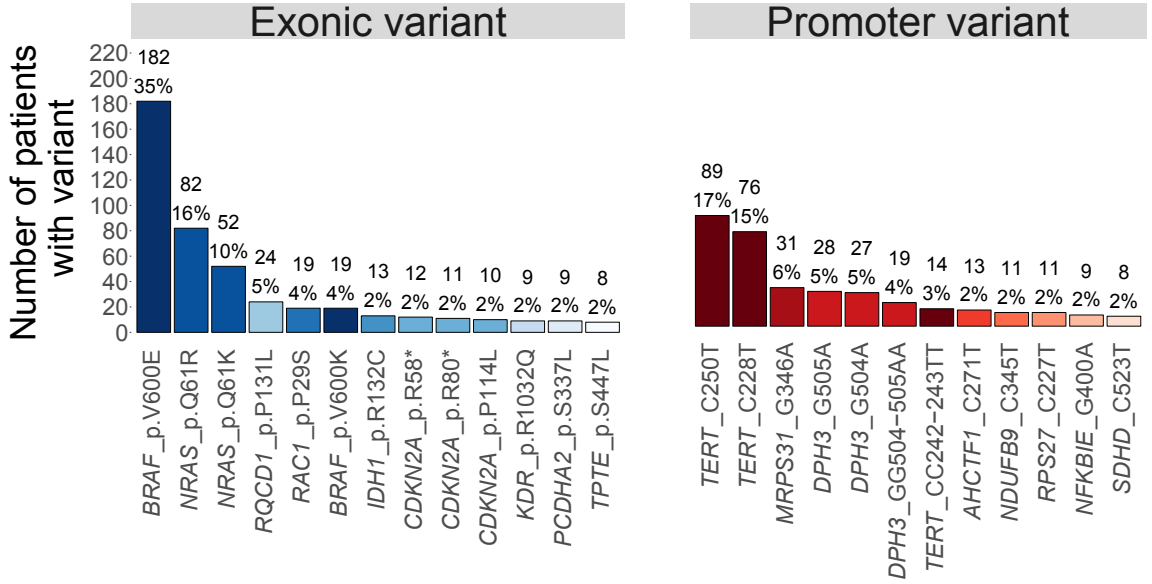


Figure 3.2: **Top recurrent mutations in the Leeds melanoma cohort.** The top recurrent exonic and promoter variants are shown, together with the alteration frequency of each variant in this cohort (n = 524). All mutation changes in noncoding regions were reported in respect to the strand location of the gene, where the position is shortened to the last three digits.

upstream and reverse in relation to the gene. However, some transcription factors have bi-directional activity, and GABPA has been reported to have this feature [477]. Additionally, disruption of ETS transcription factor binding sites such as GABPA have been reported to be enriched in tumours as a consequence of UV damage [478, 479] (discussed in Section 3.6), and as such the effect of these variants on *AHCTF1* gene expression and subsequently its role in melanoma development remains unclear and will only be elucidated following experimental validation.

5'-UTR hotspot mutations in *Ribosomal Protein S27 (RPS27)* have previously been described [475]; however, the recurrent mutation I found in position C227T affecting 11 patients in this cohort, was only recently reported [91]. In this whole genome sequencing study, patients with mutations in that position showed an almost two-fold increased expression of *RPS27*. In addition to the C227T position, an additional eight patients had variants in the promoter region of *RPS27*, across a stretch of only 25 bp. Most large-scale sequencing projects have used exome sequencing, hence would not retrieve information about noncoding regions. Looking at genetic alterations through Cbioportal [480, 481], using data generated by TCGA (TCGA SKCM, Section C.2), a large proportion of the changes in *RPS27*

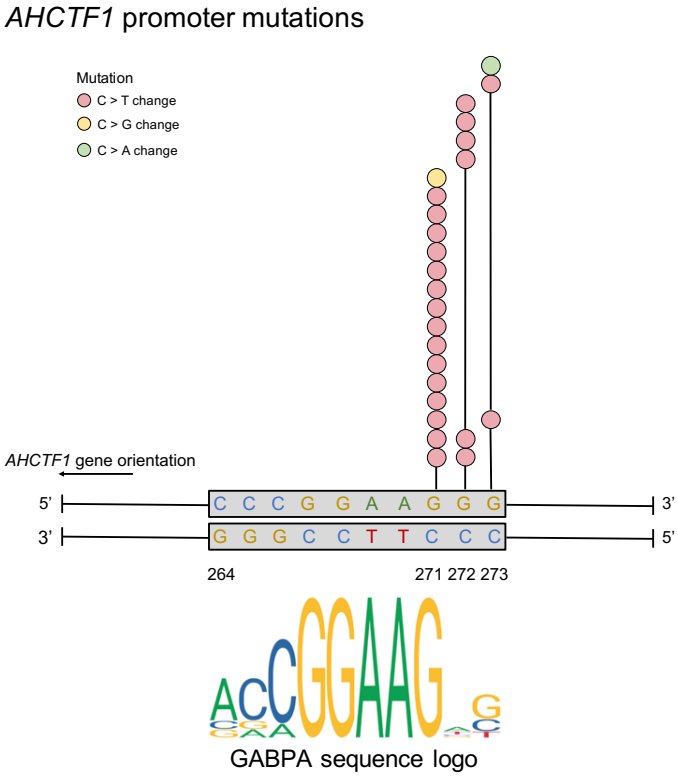


Figure 3.3: **Location of novel *AHCTF1* hotspot promoter mutations affecting the GABPA transcription factor binding site.** Recurrent mutations in the promoter region of *AHCTF1*. Each row of circles represents a sample, and the colour of the circle represents the base change. The majority of patients with *AHCTF1* mutations have the C271T variant, coinciding with a conserved GABPA binding motif base; however, other variants exists and some patients have multiple mutations across the GABPA binding site. All mutation changes in noncoding regions were reported with respect to the strand location of the gene, where the position is shortened to the last three digits.

were amplifications, which may suggest the gene acts as an oncogene.

The frequency of nonsynonymous mutations in candidate driver genes were also studied (Fig. 3.4). Hotspot mutations in *BRAF, NRAS, KRAS, HRAS, RAC1* and *IDH1* were defined as recurrent missense variants known to be associated with an increased oncogenic activity. A high proportion of loss of function variants were found in tumour suppressor genes *CDKN2A, TP53, NF1, ARID2, PTEN, FBXW7, RB1* and *RASA2*. This is in contrast to oncogenes *PPP6C, MAP2K1, KIT, CTNNB1, EZH2* and *CDK4,* which had an overwhelming majority of missense mutations. Compared to other large metastatic melanoma sequencing studies, the alteration frequencies in candidate driver genes of my primary melanoma dataset was similar [73, 85, 91]. Some reported mutation frequencies varied between datasets,

which could be reflective of cohort biases or sequencing coverage. Shain *et al.* reported alterations in cell cycle genes, the p53 pathway and multiple MAPK pathway genes to arise in later stages of melanoma [48]. Therefore, slightly lower frequencies of mutations in genes such as *ARID2* in my dataset comprising early stage melanoma tumour samples, could be reflective of melanoma disease progression.

Variants in the *TERT* promoter region were dominated by C250T, CC242-243TT and C228T changes. No other recurrent mutations were found in *TERT*. All *DPH3*, *NDUFB9* and *NFKBIE* variants found were in positions which have been reported previously. In addition to known positions, one recurrent mutation in the promoter region of *SDHD*, in chr11:111957596, was found in two samples of this cohort.

3.3 Copy number alterations in the Leeds melanoma cohort

Copy number data were generated using ASCAT [443, 482] by Dr. Kerstin Haase (Section C.1). 401 samples successfully passed all filters, including the removal of 9 samples which were excluded from the mutational analysis due to low average coverage. The samples showed a mean purity of 66%, with 274 (68%) samples found diploid, and 127 (32%) tetraploid. To estimate the amount of copy number change in each sample, I looked at the total length of amplified or deleted regions per sample (Fig. 3.6). Copy number alterations affect most samples to some extent, with many samples showing a high amount of homozygous deletions in particular.

Genetic aberrations in melanoma commonly include deletions involving chromosome 9p (*CDKN2A)*, 10q (*PTEN*), 6q and 1p, whereas amplifications often involve chromosome 1q, 6p, 7 (*BRAF*) and 8 (*MYC*) [483–486]. A genomic overview of copy number alterations across all samples in the Leeds melanoma cohort, along with a comparison with the TCGA SKCM dataset are shown in Fig. 3.6A,B. Copy number gains and losses where the proportion of all segments which differ more than 0.6 from a sample's ploidy were used to generate the figure. Similar alterations could be seen across the two datasets, suggesting similar copy number profiles between primary and metastatic melanoma. Chromosome 1 shows loss of the short arm, along with amplification of its long arm affecting genes such as *AHCTF1* and *NOTCH1*. Chromosome 6p is commonly amplified, whereas 6q were more often deleted. Other frequently amplified regions include chromosome 7 and 8, while chromosome 9 and 10 were more regularly lost.

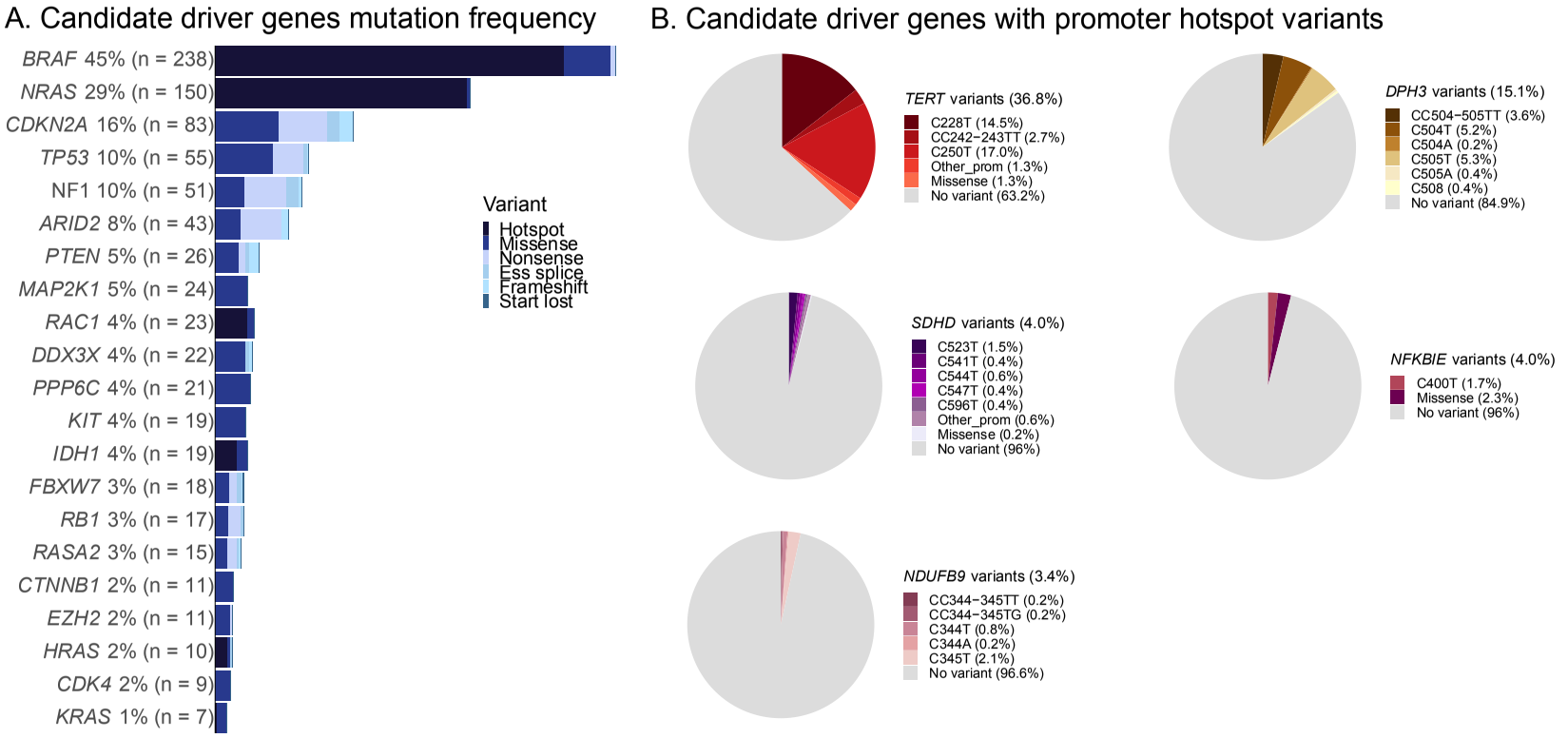


Figure 3.4: **Frequency of mutations in candidate melanoma driver genes**. **A)** Mutation frequencies in the Leeds melanoma cohort of candidate melanoma driver genes, coloured by the proportion of nonsynonymous coding variant consequences for each gene. **B)** Hotspot promoter mutations in candidate driver genes, coloured by the proportion of each recurrent variant. All mutation changes in noncoding regions were reported in respect to the strand location of the gene, where the position is shortened to the last three digits.

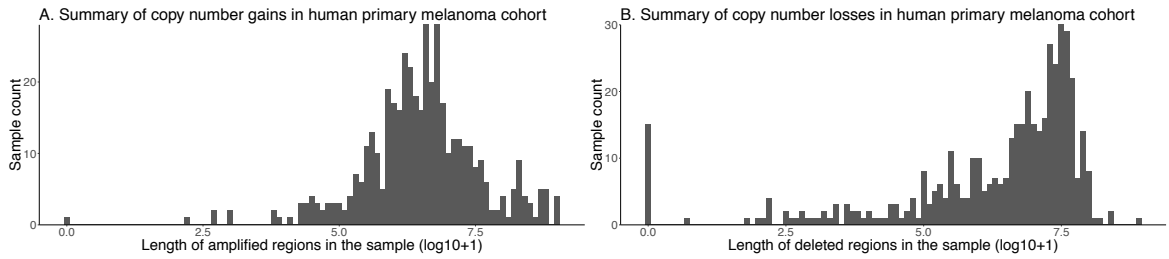


Figure 3.5: **Amount of copy number alterations in each sample.** Histogram showing the per sample total length of all **A)** high level amplifications or **B)** homozygous deletions.

3.3.1 High level amplifications

The most frequently amplified genes in my dataset were *CCND1* (n = 20, 5%), *HIST1H2BG* (n = 16, 4%), *TERT* (n = 15, 4%), *IRF4* (n = 11, 3%), *UBR5* (n = 11, 3%), *MYC* (n = 10, 2%), *NDUFB9* (n = 10, 2%) and *SNX31* (n = 10, 2%). 7 samples showed high level amplifications across the entire *BRAF* gene, with an additional 60 samples having a part of *BRAF* amplified. *UBR5, MYC*, *NDUFB9* and *SNX31* are all located on chromosome 8q, where *MYC*-amplifications have, in particular, been pointed out as the important disease-driving event. Other genes reported to be amplified in different cohorts of melanoma had a lower incidence in my dataset, with *CDK4* and *KIT* showing amplifications in only five samples each while *MITF* was amplified in just two samples. The lower frequency of copy number events in my cohort could be due to stricter filters for copy number calls, as I elected to only report high level amplifications in samples where the entire gene was affected by the event. If choosing to look at high level amplifications affecting any part of the gene*: KIT*, *MITF* and *CDK4* would be affected in 18, 17 and 5 cases, respectively.

HIST1H2BG and *IRF4* are both located on chromosome 6p, a region frequently amplified in melanoma. As this region spans many genes, the specific amplification of these two genes could therefore be a case of "guilt by association", and not driven by a potential biological advantage of amplification of any of these two genes specifically. Dr. Rashid Mamunur (Section C.1) performed an analysis of the DepMap CRISPR-Cas9 dataset [487–489], whereby he identified *IRF4,* one of the top amplified genes in my dataset, as 1 of 35 genes significantly associated with lethality in skin cancer cell lines compared to cell lines of other tissue origin. In brief, this analysis was done by processing the publicly available DepMap data, where the effect of gene knock-down in CRISPR-Cas9 dropout screens across a range of cancer cell lines were studied. Each gene and cell line were assigned a CRISPR lethality score, indicating the dependency of that cell line on the gene being studied. A higher lethality score means a larger reduction in cell viability when the specific gene

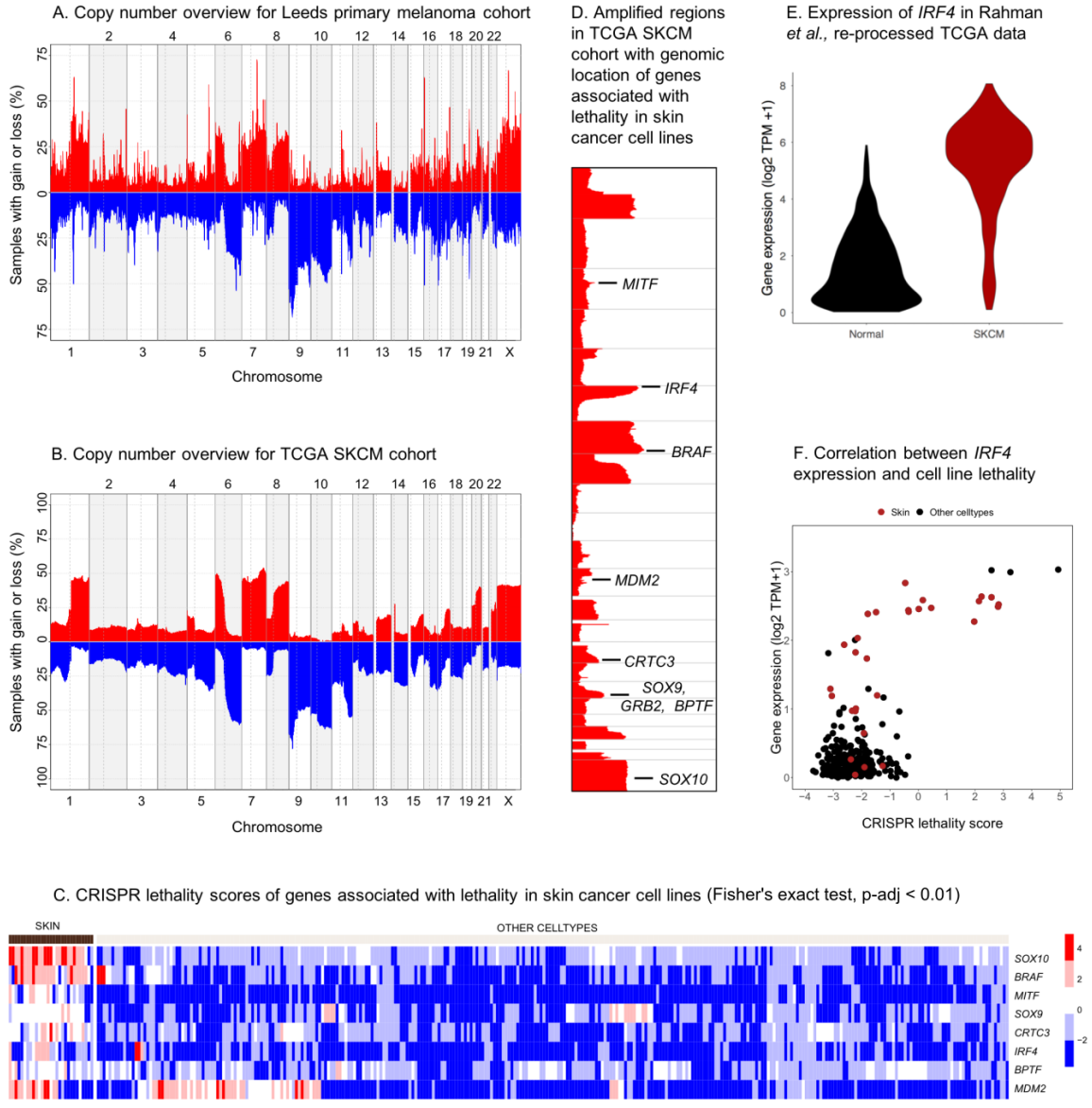


Figure 3.6: **Whole genome copy number overview highlighting interesting genes.** An overview of copy number alterations for **A)** the human primary melanomas (Leeds melanoma cohort) and **B)** the TCGA SKCM dataset. All segments with a copy number differing more than 0.6 from the sample average were used to generate the figures. Red illustrates gains and blue losses. Both figures were originally created by Kerstin Haase. **C)** CRISPR lethality scores (higher scores corresponds to a larger reduction in cell viability when the specific gene is silenced) of eight genes associated with lethality in skin cancer cell lines (Fisher's exact test, p-adj < 0.01). Red = more lethal, blue = less lethal. Other cell types include cell lines of tissue origin other than skin. **D)** Amplified regions in the TCGA SKCM cohort overlaid with the genomic location of the eight genes associated with lethality in skin cancer cell lines. **E)** Expression of *IRF4* in Rahman *et al.,* reprocessed TCGA data. **F)** Correlation between *IRF4* expression and cell line CRISPR lethality scores.

is silenced. Lethality significance scores were used to perform a Fisher's exact test comparing number of cell lines being lethal versus non-lethal in skin cancer cell lines versus cell lines of other tissue origin. All genes with a Benjamini-Hochberg multiple testing corrected p-value of 0.01 were considered significant. Eight of the significant genes were analysed as part of my sequencing panel, and their respective lethality scores across all cell lines are visualised in Fig. 3.6C.

I found *Interferon Regulatory Factor 4 (IRF4)* particularly interesting (Fisher's exact test adjusted p-value 0.000015), as this is one of the top amplified genes in my dataset. It maps to chromosome 6p, which is amongst the most frequently amplified regions in melanoma [486, 490], but for which any specific oncogene associated with this amplification has yet to be established [491, 492]. I first looked at the expression of *IRF4* in TCGA tumours [96, 493], including melanoma and normal tissue. I discovered an increased expression of *IRF4* in melanomas compared to normal tissues (Fig. 3.6.), which would support that *IRF4* amplifications could be a potential oncogenic event in melanomas. This observation could however be associated with *IRF4* having an increased gene expression in the skin compared to other organs, and indeed *IRF4* does have a high expression among primary melanocytes (Section A.6, C.2). However, The Human Protein Atlas does not report higher *IRF4* protein expression in the skin compared with other tissues (Fig. A.6) [494–496]. Finally, analysis of the DepMap data showed that melanoma cell lines with high *IRF4* expression were more often associated with lethality when that gene was lost compared to other cancer types (Fig. 3.6). Another three cell lines also showing high expression of *IRF4* with a high lethality score were of haematopoietic and lymphoid tissue origin. Immune cells also have high expression of *IRF4* (Fig. A.6), and interestingly in multiple myeloma, inhibition of *IRF4* was shown to reduce expression of the oncogene *MYC*, and cause cellular toxicity [497]. Taken together, I hypothesise that *IRF4* amplification could be an oncogenic event in melanoma, whereby the loss of *IRF4* expression through chemical perturbation might expose a vulnerability of the tumour cell.

To test my hypothesis, I knocked down *IRF4* in the human melanoma cell line RVH421. This cell line shows high expression of *IRF4,* with the gene also being amplified [488, 489]. I confirmed through 3 independent experiments, that siRNA-mediated knock-down of *IRF4* in this cell line indeed resulted in apoptosis and cell death, but through a *MYC*-independent mechanism (Fig. 3.7). Interestingly, I also observed that knock-down of *IRF4* caused these cells to increase in size, which I speculate could be linked to senescence or cell cycle regulatory effects. Collectively, the mechanism behind this *IRF4*-dependency seen in a subset of melanomas is interesting to further interrogate, and these data suggest *IRF4*

might be an interesting clinically actionable target [498].

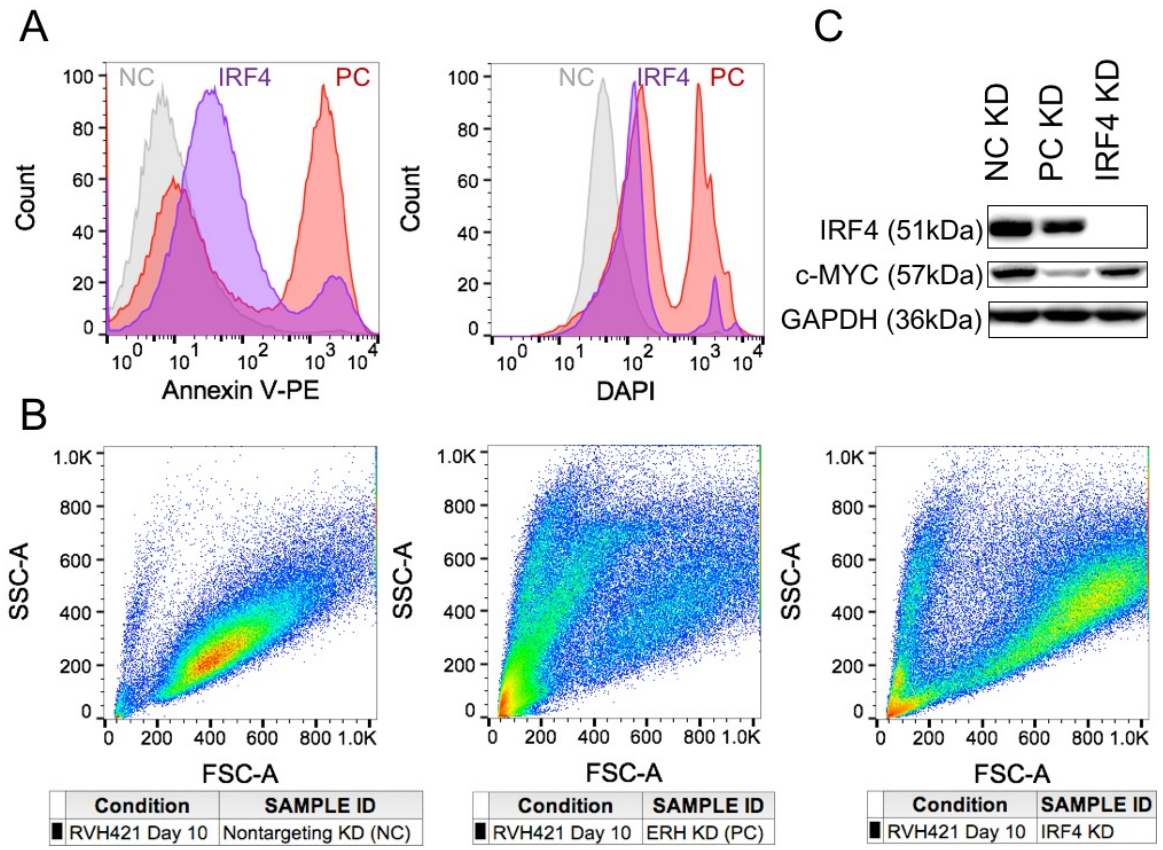


Figure 3.7: **The effect of *IRF4* knock-down in the human melanoma cell line RVH421. A)** Apoptotic or dead cells were analysed using Annexin V-PE and DAPI staining, respectively, and their fluorescence intensities measured using flow cytometry. **B)** Flow cytometry analysis of forward and side scatter shows an increase in cell size in the cells where *IRF4* is knocked down. **C)** Western blot analysis confirming the siRNA-mediated silencing of IRF4 on a protein level and sustained c-Myc expression when IRF4 is knocked down. All data is representative of 3 individual experiments. NC = Nontargeting siRNA control, PC = Positive control *(ERH* knock-down).

3.3.2 Deletions of genomic regions

In my dataset, 80 samples (20%) had a homozygous deletion in a segment overlapping *CDKN2A*, with 50 samples (13%) having the whole gene lost. An additional 171 samples showed LOH in this region, resulting in a total of 251 samples (63%) showing copy number loss targeting any part of *CDKN2A*. Other researchers have proposed that *CDKN2A* deletions occur in the transition to invasive melanoma, therefore being less prevalent in primary

melanoma [47, 136]. However, my results show for the first time in a much larger cohort of primary melanomas than have previously been reported, that the frequency of *CDKN2A* alterations is in line with reports of metastatic melanoma: 67% in the TCGA cohort [73] and 56% in a cohort of 143 primary invasive melanomas [499]. Furthermore, loss of 9p21, the region harbouring *CDKN2A,* has been shown to occur early in melanoma progression [500]; however, this study is based on just one patient. *CDKN2A* germline variants are the single most common genetic risk factor in familial melanoma [501, 502], which suggest *CDKN2A* loss could be an early event in initiation of the disease.

After 9p21 and *CDKN2A*, the most frequently homozygously deleted entire gene was *PDCD1 (*n = 6*,* 1.5%*)* followed by genes in the 10q region including *PTEN* (n = 5, 1.2%). When including hemizygous deletions, a total of 147 (37%) and 89 (22%) samples showed loss of *PTEN* and *PDCD1*, respectively. *PTEN* is a key tumour suppressor, frequently lost in various cancers and an important melanoma driver gene. In the TCGA SKCM cohort, *PTEN* is homozygously deleted in 28 out of 442 patients (6%), which is higher than in my cohort of primary melanoma (1.2%).

PDCD1 encodes PD-1, an important inhibitory receptor expressed on immune cells, which plays a key role in mounting a host anti-tumour immune response [242, 273]. As PD-1 normally is not expressed on tumour cells, the gene being deleted could reflect its loss in the immune infiltrate of the tumours. This observation is interesting, as loss of the PD-1/PD-L1 signal would activate immune cells, and could be the result of mechanisms in favour of tumour eradication by the immune cells. However, there have been reports in literature of an existing subpopulation of melanoma cells, which do express PD-1 [503], and this trait is curiously associated with enhanced tumourigenic properties. It is not possible to determine from my data whether the loss would have occurred in tumour or immune cells; however, the samples showing deletion of this gene did not have a higher proportion of immune cells than other samples, as their average tumour purity were not different (Mann-Whitney test, p-value = 0.33).

3.3.3 Loss of heterozygosity in conjunction with mutations

The two-mutation hypothesis postulated in 1971 describes how tumour suppressor genes could need biallelic inactivation to drive tumourigenesis [504]. Strong selective pressures exist to select for tumour-advantageous alterations, whereby the complete lack of a wild-type allele strengthens the effect of tumour suppressor gene inactivation. Complete loss of a gene can be achieved through multiple events, such as homozygous deletion, loss of function

mutations affecting both alleles, epigenetic silencing, or a combination of these modifications. Therefore, I investigated the concurrent presence of mutations and LOH in the same gene. I chose to focus on recurrently mutated genes with a minimum 5% mutation rate. Fig. 3.8 shows the genes with the highest fraction of samples having simultaneous mutation and loss of the second allele. In the top are several well-known melanoma tumour suppressors, including *PTEN, CDKN2A* and *TP53,* genes recognised to drive melanoma development [145, 505–508]. Next on the list is *SVEP1,* a gene involved in cell adhesion [509, 510], with very limited reports linking the gene to melanoma, other than it being frequently mutated in cutaneous melanoma [511]. The gene also shows enrichment in functional mutations by OncodriveFM analysis [511]. In my dataset, 76 samples show nonsynonymous mutations in this gene, 139 samples have hemizygous loss across this gene, but no samples show homozygous deletion. As such, it appears inactivation of this gene mainly results from the combination of mutation and hemizygous deletions. Therefore, analysis combining several different genetic alterations is necessary to pick up this signal. Various other melanoma-associated genes also show loss of heterozygosity with a mutation in the other allele, including tumour suppressors *PTPRK, NF1* and *ARID1A*. Several genes described in other sections of my thesis, such as *TLR4* and *TPTE* also show this feature. Interestingly, although the important tumour suppressor *TP53* shows frequent mutation and loss of the other allele, not a single sample has a biallelic homozygous deletion of *TP53*.

Through concurrent analysis of multiple important genetic alterations, such as mutations in combination with copy number alterations, one might get a more comprehensive view of melanomas. Most researchers study mutations and copy number alterations individually; however, I find it valuable in particular when studying tumour suppressor genes, to look at such changes all together. Genes can be silenced through a range of mutational, complex structural variant, copy number and epigenetic events, either in isolation or in combination. As I don't have information on structural variants or on an epigenetic level, I can only focus on the former two alterations. One caveat with this analysis though, is the accuracy with which it can be performed. Most mutations, in particular loss of function events, are difficult to predict based on the change observed in the gene. Additionally, allelic loss is a common genetic event in cancers [512], and therefore the specificity towards individual genes can be debated. Therefore, there will be inherent noise in the data, making follow-up studies even more important. Nevertheless, researchers have linked LOH to a selective advantage as well as showing prognostic value [513–516]. In my analyses, I've discovered a high fraction of samples which show concurrent mutation and loss of the other allele in key tumour suppressor genes such as *PTEN, CDKN2A* and *TP53*. Several other genes showing

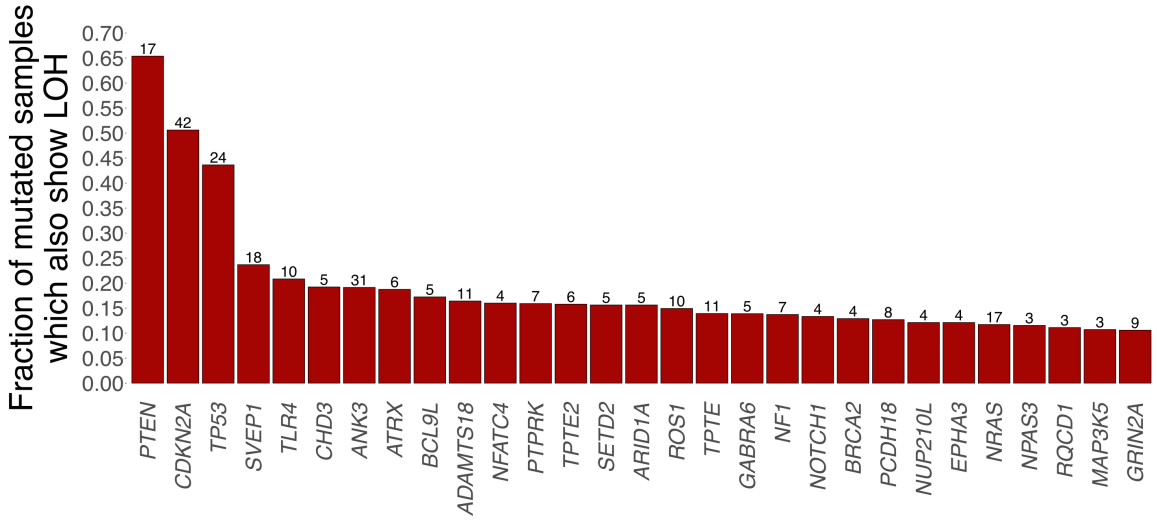


Figure 3.8: **Top genes with coincident mutation and loss of the other allele.** Genes with the highest proportion of mutations with simultaneous loss of heterozygosity. The number above each bar denotes the total number of samples showing both mutation and loss of the other allele.

this pattern were also discovered, such as *SVEP1*, *TLR4* and the novel candidate driver gene *TPTE,* discussed in Section 3.4.

3.4 Driver gene discovery

I used both SNVs and indels to feed into dNdScv (Table A.1), to identify cancer driver genes based on positive selection in my primary melanoma dataset (Fig. 3.9). Reassuringly, well-established melanoma driver genes dominate the top of the list; however, several less known and novel melanoma driver genes were also found. *FAM58A*, *RQCD1* and *MSR1* have recently been proposed to harbour properties that could make them potential driver genes [161, 428, 474]. *TPTE* on the other hand, has not been described in association with melanoma previously. Relaxing the driver gene discovery FDR threshold to 10%, an additional five driver genes reached significance. Two of these genes are known to be important for melanoma (*PPP6C* and *ARID1A)* [85, 155, 430, 517], while *ZFX, IFRD2* and *IFNL2* are additional novel candidates.

The novel gene *Transmembrane Phosphatase With Tensin Homology (TPTE)* showed a mutational pattern resembling those of known driver genes. The gene has several recurrently mutated positions, including deletions and truncating mutations (Fig. 3.10A). In

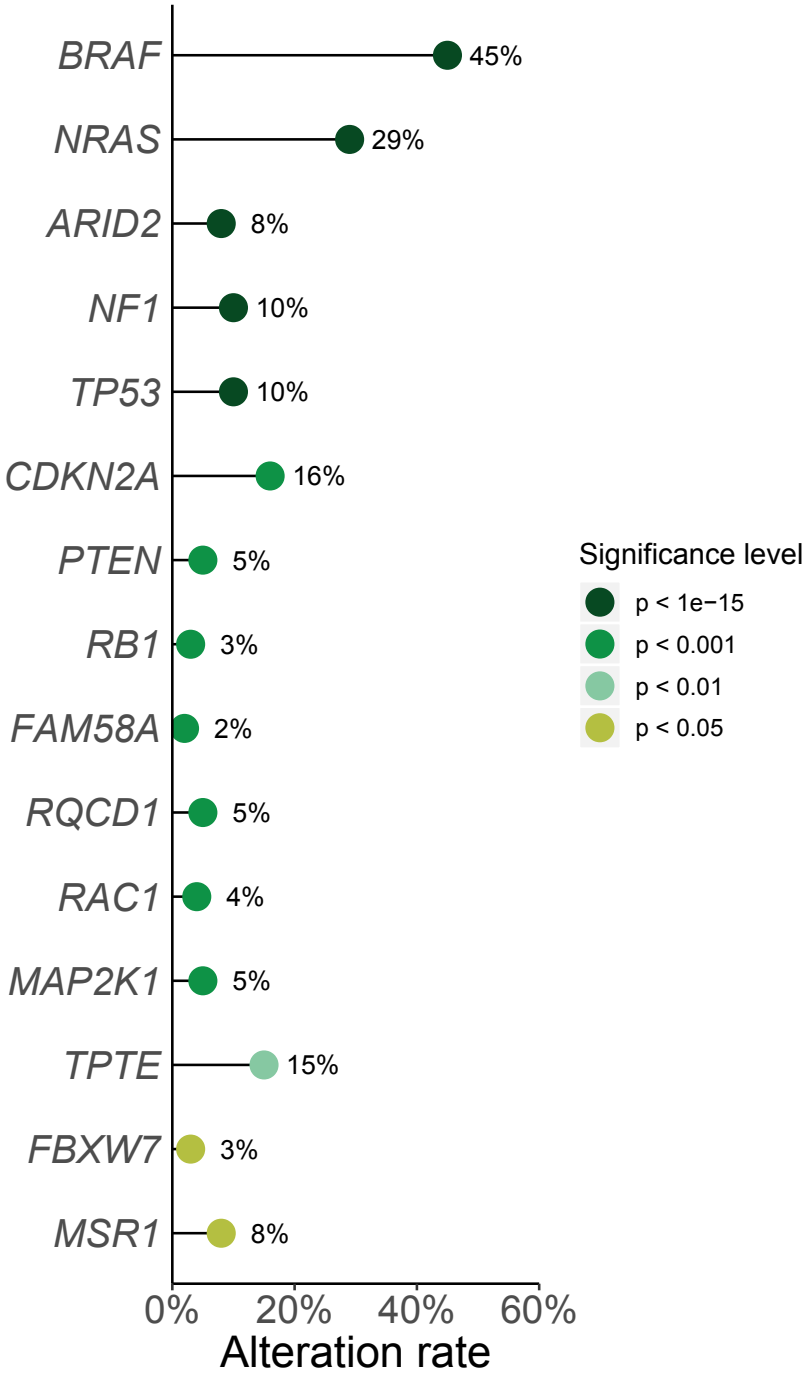


Figure 3.9: **Human primary melanoma driver genes.** The alteration rate of driver genes identified in the human primary melanomas (Leeds melanoma cohort) using dNdScv with FDR-adjusted p-values below 0.05.

particular, the S447L missense mutation, affected eight tumours making it one of the top altered positions in this cohort (Fig. 3.2). This mutation is located in the PTEN C2 domain [476, 518], a functional domain which is conserved between the paralogues *TPTE, TPTE2* and the essential tumour suppressor gene *PTEN*. Intriguingly, this functional domain is not the only shared structure between these proteins. All three proteins also have a Dual specificity phosphatase catalytic domain (DSPc) [519]. Upon closer inspection, I also discovered that mutations across *PTEN* and *TPTE2* (and also *PTEN* and *TPTE* to some extent), are largely mutually exclusive (Fig 3.10B,C), indicating similar function and therefore functional redundancy of co-occurring mutations. It is thus possible that by disrupting *TPTE* or *TPTE2* one might accomplish a similar effect to *PTEN* silencing, which has such an important role in cancer progression. In normal tissue, TPTE is mainly expressed in the testis; however, TPTE2 has more widespread distribution [520]. Research has shown that *TPTE2* is catalytically active and might be involved in negative regulation of cell growth and proliferation [521, 522], while TPTE on the other hand has not been shown to possess any phosphatase activity to date [523]. However, studies have revealed an aberrant expression of *TPTE* in tumour tissues, as well as the presence of autoantibodies in a subset of those patients, indicating a potential immune-associated role of this target gene [524, 525]. Thus, *TPTE2* might be more similar to *PTEN* in function, while the cancer-associated function of *TPTE* is less clear. As *TPTE* mutations are positively selected for by melanoma cells, and the closely related *TPTE2* possesses *PTEN*-related functions, both these genes would be interesting candidates for functional studies beyond the scope of my thesis.

3.5 Analysis of mutational patterns

I used the R package DISCOVER (Table A.1, A.5), a tool to assess mutational patterns, which deals with the mutational load confounder by assigning both an individual tumour alteration rate and gene alteration rate [472]. Eight gene pairs were found to show a mutually exclusive pattern (Table 3.1). Six involved *BRAF* (Fig. 3.11) and two included *CDKN2A* (Fig. 3.12). No co-occurring gene pairs were found.

3.5.1 *BRAF*-mutually exclusive gene pairs

Of the six genes showing mutual exclusivity with *BRAF*, the top two genes (*NRAS* and *NF1)* are well-known members of the MAPK pathway, and key driver genes in melanoma. Their mutual exclusivity with *BRAF* is established and reflects redundant activation of the MAPK

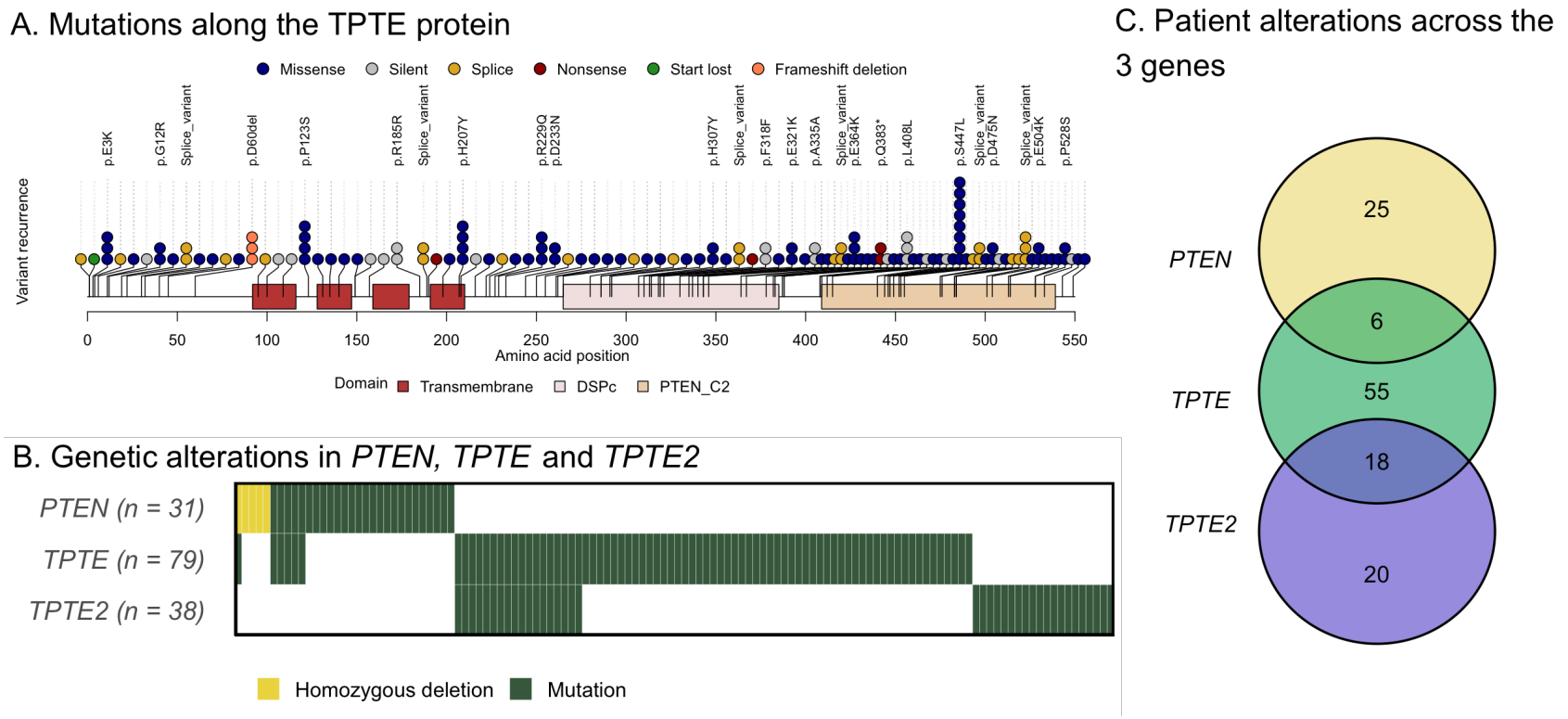


Figure 3.10: **Genetic alterations in important paralogues to *PTEN*. A)** Mutations along the TPTE protein. All mutations along TPTE are shown. For splice variants, the closest amino acid to the splice site was used to annotate its position. Only variants with a recurrence of two or more is marked with variant information. Domain information was extracted from Pfam. **B)** Distribution of homozygous deletions and mutations in the three genes: *PTEN, TPTE* and *TPTE2* in the Leeds melanoma cohort. **C)** Number of patients with shared or unique mutations in the three genes: *PTEN, TPTE* and *TPTE2* in the Leeds melanoma cohort.

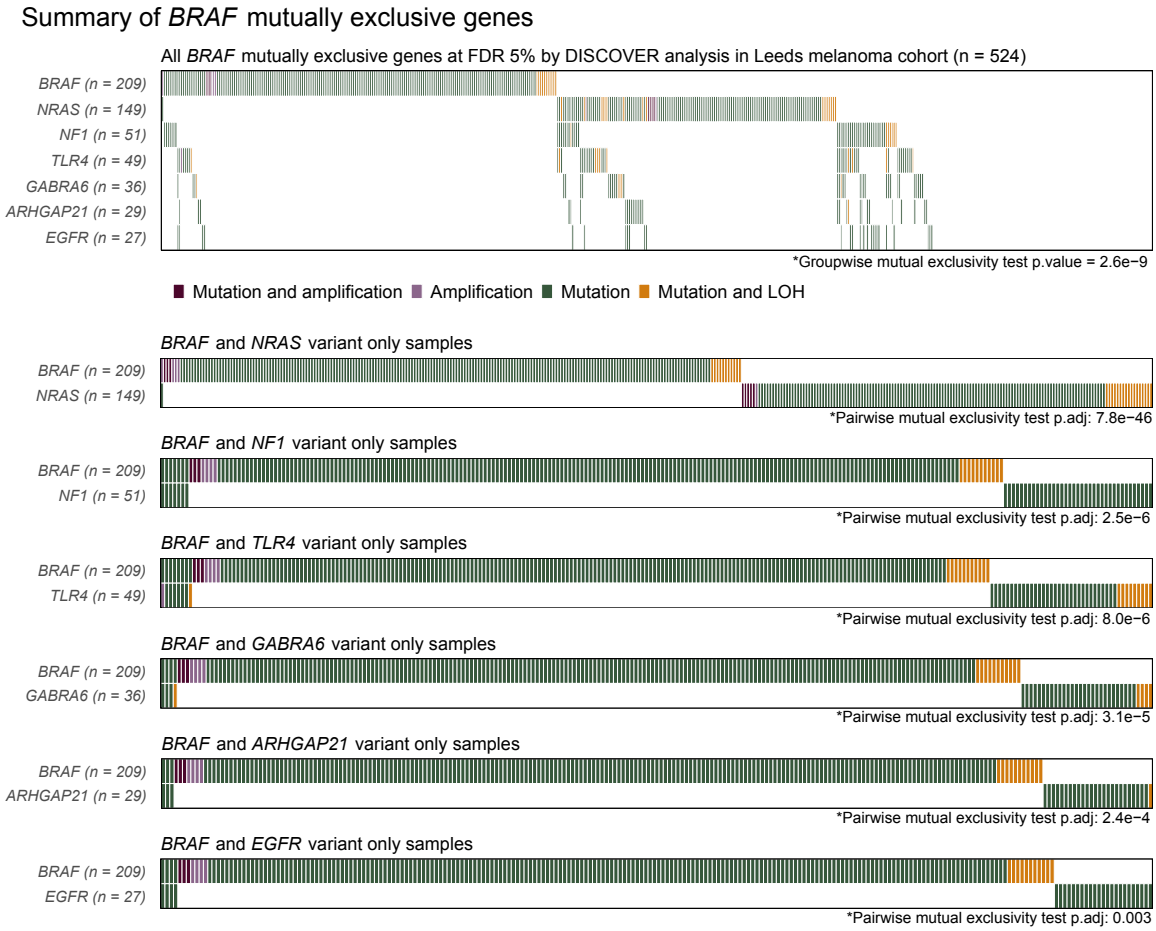


Figure 3.11: ***BRAF* mutually exclusive gene pairs found using DISCOVER.** Genes showing significant mutually exclusive patterns (FDR-adjusted p-value < 0.05) with *BRAF*. Top panel: All significant gene pairs together with their alteration patterns, with the DISCOVER groupwise mutual exclusivity test p-value shown (Table A.1). Bottom panel: Each significant gene individually with *BRAF*, and their alteration patterns, with pairwise mutual exclusivity test p-values shown. For visualisation, relevant copy number events are included in the figure, although these were not used in the analysis. Amplification refers to high level amplifications.

Table 3.1: **Mutually exclusive gene pairs found using DISCOVER.** Gene pairs showing significant mutually exclusive patterns (FDR-adjusted p-value < 0.05) found using DISCOVER analysis of the Leeds melanoma cohort (n = 524). The number of patients with mutations in each or both genes are also shown.

Gene 1	Gene 2	P-adj	Patients with mutation in gene 1 only	Patients with mutation in gene 2 only	Patients with mutation in both genes
BRAF	*NRAS*	7.8×10^{-46}	205	148	0
BRAF	*NF1*	2.5×10^{-6}	198	44	7
BRAF	*TLR4*	8.0×10^{-6}	198	41	7
BRAF	*GABRA6*	3.1×10^{-5}	201	32	4
BRAF	*ARHGAP21*	2.4×10^{-4}	202	26	3
BRAF	*EGFR*	0.003	201	23	4
CDKN2A	*ADAMTS18*	0.011	76	60	7
CDKN2A	*PRDM2*	0.041	82	25	1

pathway. Hotspot mutations in both *BRAF* and *NRAS* are rarely discovered in the same patient, confirming either event is enough to activate the MAPK pathway and promote tumourigenesis. My data agrees with this concept and provides further support to the theory of a synthetic lethal nature between the presence of *BRAF* and *NRAS* hotspot variants, as these events are almost always completely mutually exclusive [465–467]. Furtheremore, if the coexistence of both events is not harmful to the tumour, I would expect a small number of patients to have both variants by chance although it would not provide additional survival advantages to the tumour. Loss of function mutations in *NF1* also activates the MAPK signalling pathway, which explains the mutual exclusivity seen between *BRAF* and *NF1* in mine and other studies [73, 161, 526]. However, as different mutations will disrupt gene function to a varying extent, it is expected that a small number of patients will have mutations in both genes, as seen in my dataset.

Of the remaining four *BRAF*-mutually exclusive genes, *Toll Like Receptor 4 (TLR4)* and *Epidermal growth factor receptor* (*EGFR)* are both involved in *BRAF*-associated signalling pathways including the MAPK pathway [527–529], which also makes their mutually exclusive pattern with *BRAF* natural by function. However, these genes both encode receptors, which have a diverse range of ligands and can exert varying downstream effects not exclusive to MAPK signalling [530–535]. Therefore the impact of specific mutations might differ, and absolute mutual exclusivity cannot be expected of these gene pairs. To validate my findings in a separate dataset, I used the TCGA SKCM dataset through Cbioportal and discovered a trend towards mutual exclusivity between *BRAF* and *EGFR,* as well as *BRAF* and *TLR4* alterations. I chose to only look at the trend rather than the significance levels, as their mutation pattern analysis method is based on the flawed assumption of independent and identically distributed alterations across tumours [472]. Taken together, these results suggest

there might be a fraction of melanoma patients with the MAPK pathway activated through *EGFR* or *TLR4* rather than *BRAF*, and brings forth the possibility of new possible treatment options for melanoma. *EGFR* inhibitors have previously been tested without much success in clinical trials on unselected melanoma patients [536], and there are cases where resistance to *BRAF* inhibition have led to increased activity through *EGFR* [537–539]. This reactivation of the MAPK cascade through *EGFR,* is evident in colorectal cancers, but is minimal in melanomas which often express low levels of *EGFR*. Therefore, it would be interesting to test whether existing *EGFR* inhibitors such as erlotinib, gefitinib or cetuximab could be repurposed to effectively treat the subpopulation of *EGFR*-altered melanoma patients only. Similarly, for *TLR4,* several inhibitors are under development [540–543], and could be used to test treatment efficacy on the subset of patients showing alterations in *TLR4*.

The final two *BRAF*-mutually exclusive genes in my analysis were *Gamma-Aminobutyric Acid Type A Receptor Alpha6 Subuni*t (*GABRA6)* and *Rho GTPase Activating Protein 21 (ARHGAP21).* When looking in the TCGA SKCM dataset, neither of these genes were reported to have a mutually exclusive tendency with *BRAF.* Hence, these findings could either be the result of differences in patient cohorts, or it might be these findings are false positives. The presence of a putative ERK phosphorylation site in GABRA6 [544], could strengthen the claim of this gene as a player in the MAPK pathway. However, this is not likely to be its main role, as *GABRA6* has important neurological functions in the brain, and the impact this gene has on activating the MAPK pathway to subsequently contribute to tumour development is likely limited. *ARHGAP21* silencing has proven to be functionally similar to that of *BRAF* activation [545–550], and with studies suggesting shared interaction partners [551–553] this could indicate *ARHGAP21* has a novel role in MAPK signalling. But, as this gene also seems to have multiple functions, the specific activation of MAPK signalling might therefore only be a small part of its role, and subsequently its contribution to promoting tumour development might be insufficient. As many of the genes discovered through my analysis have broad functions, their MAPK-specific role could therefore be diluted by this multi-functionality, and I would not expect mutations affecting these genes to completely negate any effect a simultaneous *BRAF* activating mutation would have on tumourigenesis. To follow up and confirm the novel discoveries made here, each gene's specific contribution to MAPK signalling could be tested by assessing binding partners, by studying direct downstream effects such as phosphorylation of ERK, or their final impact on cell proliferation and survival.

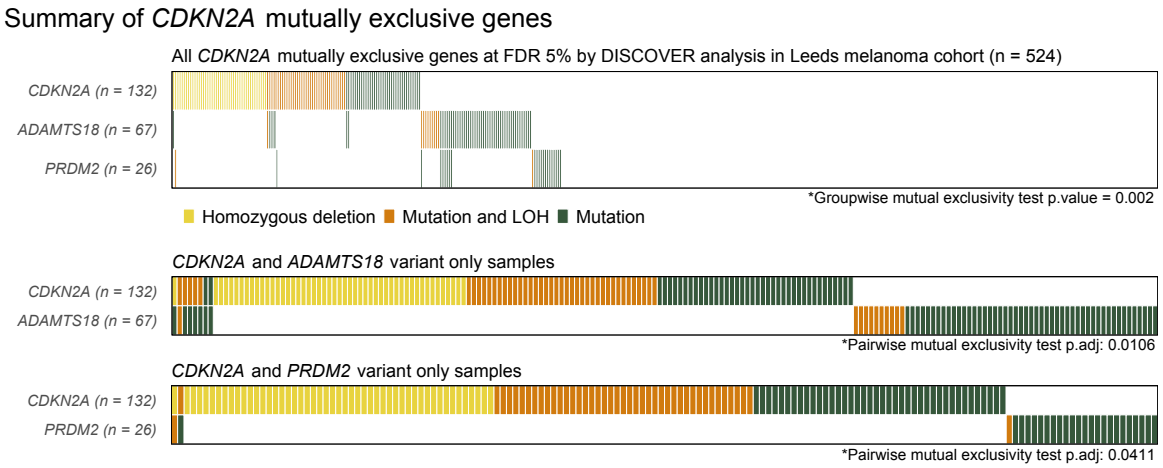


Figure 3.12: ***CDKN2A* mutually exclusive gene pairs found using DISCOVER.** Genes showing significant mutually exclusive patterns (FDR-adjusted p-value < 0.05) with *CDKN2A*. Top panel: All significant gene pairs together, with their alteration patterns, with the DISCOVER groupwise mutual exclusivity test p-value shown (Table A.1). Bottom panel: Each significant gene individually with *CDKN2A,* and their alteration patterns, with pairwise mutual exclusivity test p-values shown. For visualisation, relevant copy number events are included in the figure, although these were not used in the analysis.

3.5.2 *CDKN2A*-mutually exclusive gene pairs

My DISCOVER analysis identified two mutually exclusive gene interactions with *CDKN2A* (Fig. 3.12). Both genes are suggested tumour suppressor genes: *ADAMTS18* is a metalloproteinase shown to be highly mutated in melanoma [554] and *PRDM2* encodes a zinc-finger protein that interestingly also has Rb-binding properties [555, 556]. Using the TCGA SKCM dataset, genetic alterations between *CDKN2A* and *PRDM2* but not *CDKN2A* and *ADAMTS18* showed a trend towards mutual exclusivity.

ADAM Metallopeptidase With Thrombospondin Type 1 Motif 18 (ADAMTS18) is silenced through deletion, mutation or methylation in a range of tumours, with several members of the ADAMTS family showing tumour-suppressive capabilities [557–566]. One particular study reports evidence of *ADAMTS18* being highly altered in melanoma, with mutated cells showing increased proliferation, cell migration and metastasis [554]. Recent experimental evidence has also linked overexpression of *ADAMTS18* in cancer cell lines to cell cycle arrest, apoptosis and decreased migratory and invasive properties [567, 568]. It is therefore possible that defects in *ADAMTS18* might confer the same cell cycle mediated effects as loss of *CDKN2A*. However, *ADAMTS18* specifically has not been well-studied compared to other members of the same family, and its substrates are still unknown.

PR/SET Domain 2 (PRDM2) encodes two main proteins: RIZ1 and RIZ2, which can regulate expression of other genes through several mechanisms [569–571]. These mechanisms include histone methyl-transferase activity through its N-terminal PR-domain [572, 573], recruitment of epigenetic factors [574, 575] and by binding of DNA or proteins directly [555, 576–579]. The only region distinguishing RIZ2 from RIZ1, is the lack of the N-terminal methylation-associated PR domain, as RIZ2 is transcribed through an internal promoter at amino acid position 202 [580]. This property is sometimes referred to as "yin-yang" regulation, where the full-length protein has tumour-suppressive capabilities, while the shorter length protein instead has oncogenic properties [570, 571, 580–582]. Consequently, RIZ1 specifically is found inactivated in human malignancies [583–586]. In addition to this, *in vitro* experiments have shown a link between RIZ1 expression and cell cycle arrest, decreased proliferation and increased apoptosis [584, 585, 587, 588]. This mutually exclusive pattern found between *CDKN2A* and *PRDM2,* might therefore indicate *PRDM2* is involved in *CDKN2A*-associated cell cycle regulation. However, the mutations found in the primary melanomas are spread across the gene (Fig. 3.13), with only four patients showing mutations in or around the PR-domain. Most mutations might thus affect both RIZ1 and RIZ2 function. However, some studies have shown that the region close to the C-terminal of *PRDM2* is important for PR-binding, and thus for the PR-domain activity [589, 590]. As an example, this terminal region of *PRDM2* has been shown to be necessary for binding to the histone methyltransferase *PR-Set7*, and for conferring cell cycle regulatory and tumour-suppressive functions [589]. In my dataset, a cluster of mutations is present in this area (Fig. 3.13). Additionally, the Rb-binding capability of *PRDM2* does support a link to downstream *CDKN2A* functions [555]. Due to this shared substrate, silencing of *RB1* through both *CDKN2A* and *PRDM2* would likely be functionally redundant, and not convey an additional tumour-promoting benefit. This could therefore explain the mutually exclusive pattern of *CDKN2A* and *PRDM2*. Experimental validation to assess if silencing of *PRDM2* could substitute for *CDKN2A* loss can be tested using the following steps: First, the physical interaction between *PRDM2* and *RB1* can be verified through protein pull-down experiments or using a proximity ligation assay, for example. Next, the physiological effect of *PRDM2* loss in melanocytes can be examined through cellular senescence assays, to check whether senescence can be bypassed when *PRDM2* is silenced. This would suggest loss of *PRDM2* gives the cell cancerous properties similar to that of *CDKN2A* loss. Finally, to evaluate whether loss of both genes simultaneously is redundant for the cell, is more challenging. One possible method is to knock out both genes or either gene separately *in vitro* or *in vivo*, to then compare cell growth, death or tumour growth in mice.

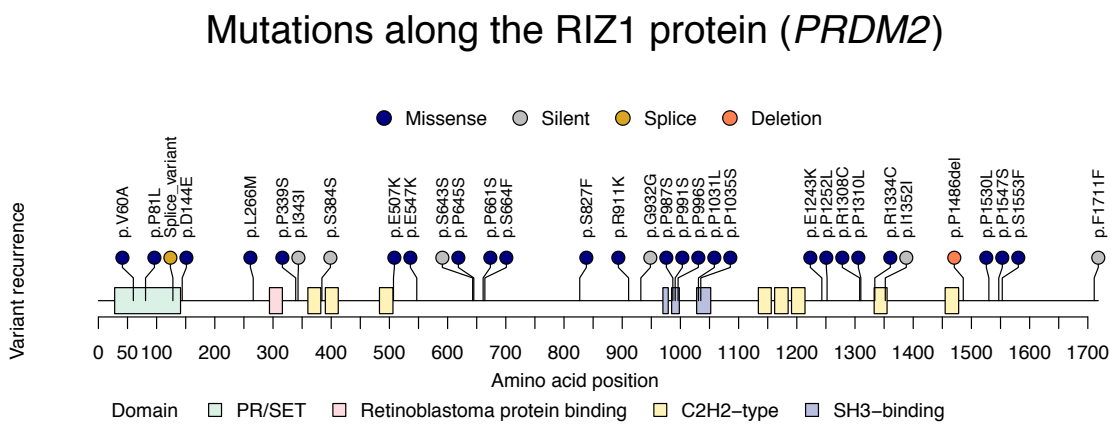


Figure 3.13: **Mutations along the RIZ1 protein (*PRDM2*).** All mutations observed in *PRDM2* in the Leeds melanoma cohort is visualised by their amino acid change. For splice variants, the closest amino acid to the splice site was used to annotate its position. Domain information was extracted from Pfam.

3.5.3 Co-occurring gene interactions

No co-occurring gene pairs were found with my analysis, which is not surprising considering the high background mutation rates in melanoma. Many former studies claiming co-occurrence of gene pairs have not taken the essential information regarding tumour-specific alteration rates into account [161, 591–595]. As an example, the previously reported co-occurrence between *NF1* and *RASA2,* looks convincing with 73% (11 out of 15) of patients with *RASA2*-variants having a simultaneous mutation in *NF1* in my dataset. However, the patients with *NF1*-mutations were both older and had a higher mutation rate than other samples, consistent with previous reports [161, 526, 596]. These patients with higher mutation rates will therefore have an increased likelihood of any two gene pairs showing co-occurrence by chance, stressing the importance of accounting for mutational load. The low patient tumour alteration rate of *RASA2* in my dataset meant this gene could not be included in my DISCOVER analysis. Therefore a final conclusion regarding the proposed co-occurring mutation pattern between *RASA2* and *NF1,* where tumour and sample alteration rates are accounted for, cannot be confirmed nor denied from my analysis.

3.6 The role of sun exposure on driver mutations

Exposure to sunlight is an important risk factor in melanoma, whereby UV-exposure driven tumours often have a higher frequency of mutations. Mutation densities are however not

equally distributed across the genome, and regional differences arise as a consequence of chromatin state [86, 478, 479, 597]. This means regions associated with a compact chromatin state or particular transcription factor binding, will have increased damage formation or be less accessible to nucleotide excision repair (NER) activity. This will subsequently cause a higher susceptibility of mutations across certain genes or regions of the genome [598, 599].

3.6.1 Pattern of UV damage in transcription factor binding sites

UV damage has been shown to selectively enrich for mutations in regions bound by the ETS family of transcription factors including ETS1, GABPA, ELF1, ELK4 and E4TF1 [478]. Well-established examples of recurrent promoter mutations in melanoma are associated with ETS transcription factor binding sites, with the mutation spanning the UV damage-associated consensus binding motif CTTCCG [478, 479]. As this pattern defines positions with elevated UV mutation vulnerability, it emphasises caution should be exerted in implementing the pathogenic role of such variants without functional studies. Several known recurrent variants overlapping ETS transcription factor binding sites have been discovered in melanoma [600–602]; however, the pathogenic effect of such variants have not been fully explored in the majority of cases. That being said, the presence of recurring mutations overlapping this pattern do not rule out the importance of the genes affected, in particular as this pattern is only seen in active transcription factor binding sites and therefore suggests an active role of the downstream gene. I interrogated the main hotspot promoter mutations discovered in my cohort for the presence of this pattern, and found it overlapped with most of the top recurrent variants (Fig. 3.14). This CTTCCG pattern was found in most frequent recurrent variants in *DPH3, SDHD, NFKBIE, NDUFB9* and *MRPS31*. The novel promoter mutations I discovered in *AHCTF1* (Section 3.2) also overlapped with such a binding site; however, the pattern was on the reverse strand. Again, this finding stresses the importance of functional validation of the effect of these promoter mutations on tumourigenesis. On the contrary, all of the variants in *RPS27* and less frequent variants in *SDHD* did not appear to be linked to this UV damage-induced mechanism. *TERT* hotspot mutations all lacked the CTTCCG pattern, which suggests these mutations are accumulated by other mechanisms than UV damage. However, in the case of C228T and C250T, these mutations create an ETS binding site instead of disrupting one.

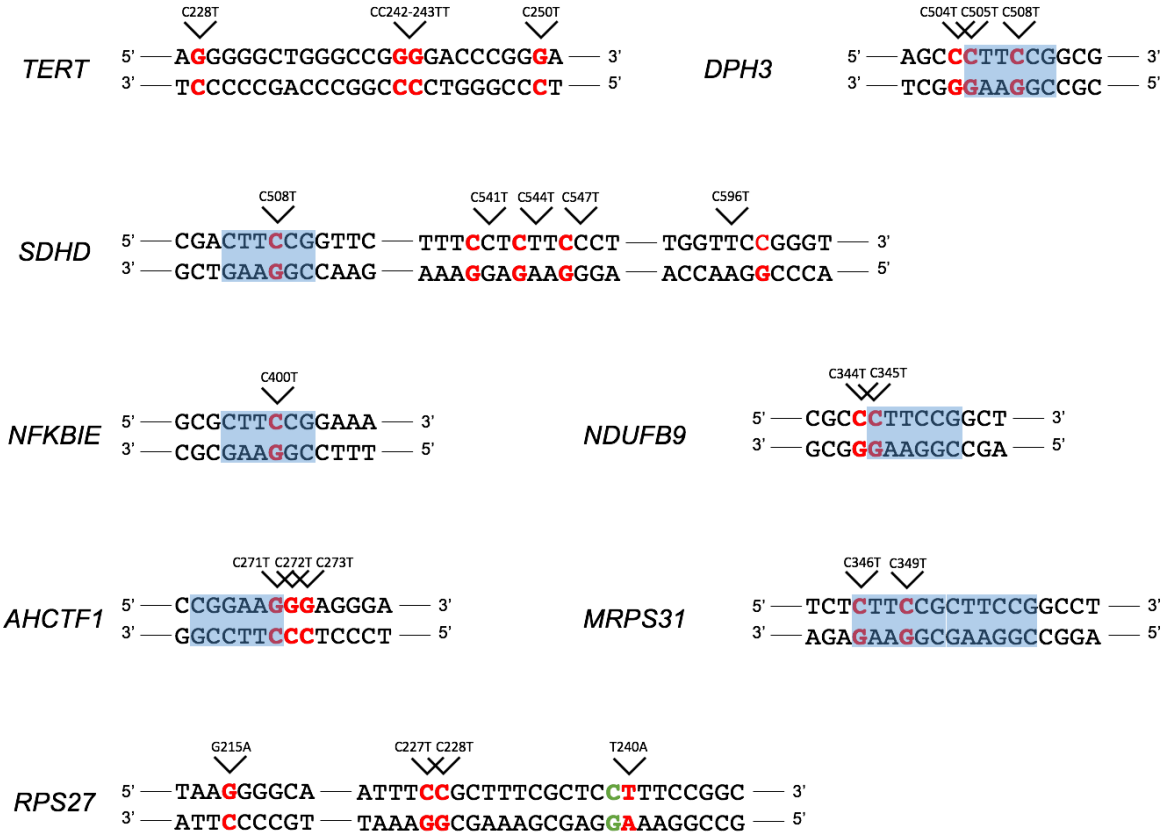


Figure 3.14: **ETS transcription factor consensus motif pattern among recurrent promoter mutations.** Recurrent variants are marked red, while the UV damage-associated CTTCCG pattern is highlighted with blue. *TERT* and *RPS27* recurrent variants were the only ones lacking this UV damage-associated pattern. The position in *RPS27* marked with green is a variant previously reported by others but filtered out in my analysis due to a population allele frequency < 0.001 in ExAC.

3.6.2 Association of frequent promoter variants with sun exposure

As I confirmed the presence of the UV damage ETS motif signature spanning many recurrent promoter variants, I wanted to check whether the top recurrent promoter mutations were enriched in primary tumours located on the head, as another indicator of sun exposure (Fig. A.7). In line with expectations, all the top recurrent promoter variants except *TERT* C250T and CC242-243TT, were highly associated with an increased mutation load (all p-values < 0.01, univariate logistic regression). The distribution of promoter mutations were also different across tumours arising on the head, limbs, trunk or other sites, with all top promoter variants shown to harbour the UV damage pattern (*MRPS31, DPH3, NFKBIE, NDUFB9, AHCTF1, SDHD)* showing a trend of enrichment in melanomas arising on the head. *TERT* hotspot mutations are likely not associated with UV exposure, as they were not

associated with a higher mutation load (except C228T). They also had a similar distribution across melanomas from the head, limbs and trunk. However, a lower frequency of *TERT* hotspot variants were observed in melanomas of other sites compared to those on the head, suggesting *TERT* activation might not be a key event in these tumours. The *RPS27* C227T variant, which did not overlap with the ETS motif, was also more common amongst primary tumours on the head. This suggests that variant could be linked to a UV-dependent mechanism not associated with the increased susceptibility of ETS transcription factor binding sites, such as mutation load.

3.6.3 Association of frequent coding variants with sun exposure

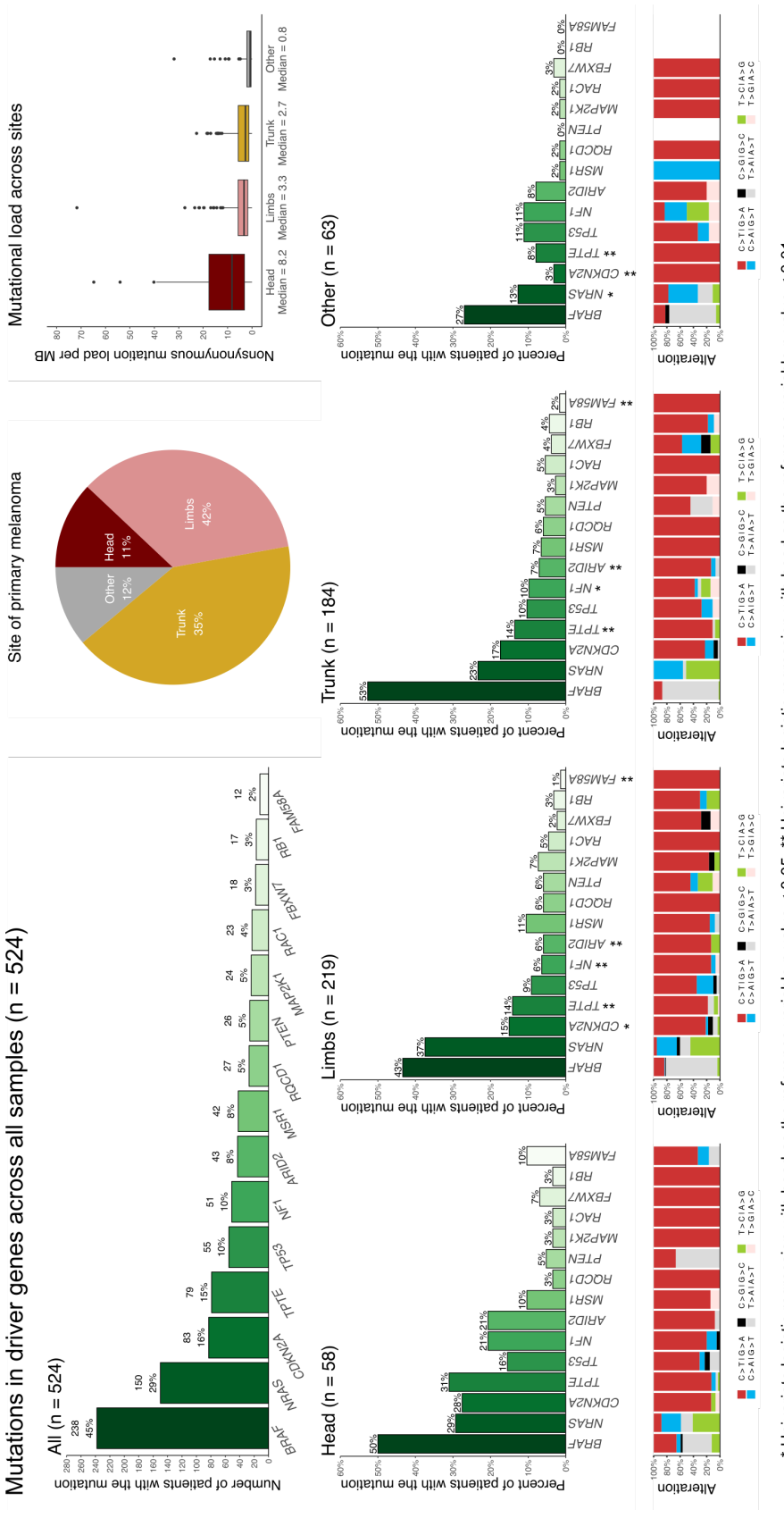
Tumours arising on body sites with varying degrees of sun exposure have been shown to feature distinct properties. Therefore, I also looked at the distribution of frequent coding changes across body sites. Of the top recurrent coding variants, no variants showed a strong preference towards the head (Fig. A.8), but a weak statistically significant trend was found for *NRAS* Q61R (head vs other site), *PCDHA2* S337L (head vs limbs) and *TPTE* S447L (head vs limbs). In addition, *RAC1* P29S, *KDR* R1032Q, *PCDHA2* S337L and *TPTE* S447L were all associated with a higher mutation burden, whereas *BRAF* V600E showed an inverse correlation with mutation burden (all p-values < 0.001). *BRAF* V600E mutations were also highly associated with site of primary on the trunk rather than the head (p-value = 0.001, OR = 3.2), in line with other reports of chronically sun-exposed versus intermittently sun-exposed sites [18, 52]. *BRAF* V600K tumours, which in addition to the c.1799T>A base change resulting in the *BRAF* V600E variant, has a concurrent c.1798G>A base change, appears to show a different mutation pattern. In concordance with other studies, *BRAF* V600K tumours had a higher mutation load than *BRAF* V600E tumours (p-value = 0.0007, OR = 1.02); however, in contrast to other studies [603, 604], *BRAF* V600K tumours were not associated with sun-exposed sites compared to *BRAF* V600E or *BRAF* wild-type tumours. In addition, it has been proposed that the higher mutation load found in *BRAF* V600K tumours compared to *BRAF* V600E tumours might result in a better response to immunotherapy in these patients [605]. However, my analysis shows that although *BRAF* V600K tumours have a higher mutation load compared to *BRAF* V600E, comparing the respective groups to the rest of the cohort, *BRAF* V600K patients did not show a significantly higher mutation burden, while *BRAF* V600E tumours correlated with a lower mutation burden (Fig. A.9). Therefore, making the comparison between the two *BRAF* variants only, would not reflect a melanoma patient cohort in its entirety and my data do not suggest that patients with *BRAF*

V600K alterations specifically would benefit from immunotherapy.

3.6.4 Association of mutations in driver genes with sun exposure

Finally, I also sought to understand the distribution of driver gene mutations across body sites commonly subjected to different sun exposure levels (Fig. 3.15). Mutations in *TPTE, NF1, ARID2* and *FAM58A* were more associated with tumours on highly sun-exposed sites (head) compared to intermittently sun-exposed areas (limbs and trunk), while *CDKN2A*-mutant tumours showed a preference towards head over limbs and other sites but not the trunk (Fig. 3.15). Tumours with mutations in the former four genes were also correlated with a higher mutation burden (all p-values < 0.001), as expected as a consequence of sun exposure. Additionally, tumours with mutations in *TP53, MSR1, FBXW7, RAC1 , RQCD1* and *RB1* were also associated with a higher mutation load. With the exception of *BRAF, NRAS,* and tumour suppressor genes *PTEN* and *TP53,* most of the other driver genes show a high proportion of the common UV radiation-induced C>T transitions. Mutations in *RQCD1* and *RAC1* were exclusively composed of C>T mutations, an effect in part mediated by each frequent hotspot variant in these genes. Interestingly, all *FBXW7* variants seen in tumours located on the head were C>T events, while those arising on the limbs and trunk showed a mix of other base changes too. The opposite effect was seen for *FAM58A,* where tumours from the limbs and trunk were all C>T mutations, while one third of tumours arising on the head had other base substitutions as well. However, these observations are all based on few events, and inferences from these analyses should therefore be treated with caution. Although not statistically significant, I observed *RQCD1* mutant tumours had a higher frequency in tumours of intermittent sun exposure compared to those on chronically sun-exposed sites, contradicting a previous study of this gene in primary melanoma and cell lines [474].

Mutations in the main melanoma genes *BRAF* and *NRAS* are generally thought of as early events in melanoma progression [47], and it is therefore not surprising to find these were not associated with sun exposure. *MAP2K1* mutations are probably also linked to UV-independent MAPK pathway activation, alongside *BRAF* and *NRAS*. *PTEN,* which was not found associated with sun exposure or particularly high proportions of C>T alterations, was interestingly completely absent in melanomas arising on sites other than the head, limbs and trunk. In Section 4.3, I also discussed the absence of *PTEN* mutations in triple wild-type, acral and mucosal melanoma. This is intriguing, as it is one of very few driver genes which appears to be constricted to a patient population where sun exposure usually plays a larger



* Univariate logistic regression with head as the reference variable p.value < 0.05, ** Univariate logistic regression with head as the reference variable p.value < 0.01

Figure 3.15: **Distribution of driver gene mutations across body sites.** Top panel shows the overall frequency of patients with nonsynonymous mutations in the driver genes, the distribution of tumours (n = 524) from the different sites of primary melanoma (head, limbs, trunk, other), and the mutation load across these four sites. The bottom panel shows the frequency of patients with nonsynonymous mutations in the driver genes, split across the four sites of primary melanoma, along with the proportion of each type of base change. Univariate logistic regression was performed to look for associations between driver mutation and the site of melanoma, with the head as the reference variable.

role (non-acral cutaneous melanoma), yet appears to arise independently of UV damage mechanisms.

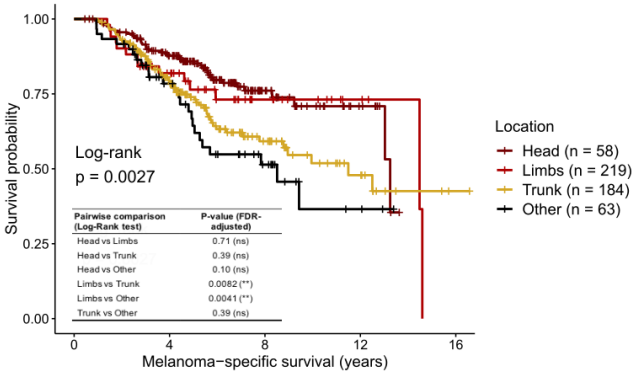
Previous smaller studies have found alterations in *KIT* to occur exclusively in melanomas of acral and mucosal origin, or skin with chronic sun-induced damage [604, 606, 607]. *KIT* mutations in my cohort were distributed amongst patients with tumours across all the four categories of body sites; however, *KIT* amplifications were found exclusively in melanomas arising on the head and other sites. This is an interesting observation, as I expected the mechanism driving rarer subtypes of melanoma to be distinct from those arising as a consequence of UV damage. Focal gains of *CCND1* and *CDK4* have also been reported to occur more frequently in tumours with chronic sun-induced damage [54, 428]; however, this result was not replicated in my larger cohort (*CCND1* head n = 2, limbs n = 6, trunk n = 7, other n = 5 and *CDK4* limbs n = 1, trunk n = 2, other n = 2).

Melanomas arising on frequently sun-exposed sites such as the head, had a significantly higher mutation load than those arising on limbs (p-value = $3x10^{-12}$, OR = 0.0008), trunk (p-value = $3x10^{-14}$, OR = 0.0004) or other sites (p-value = $6x10^{-14}$, OR = $9x10^{-5}$) in a multivariate model taking into account age (p-value = $2x10^{-6}$, OR = 1.12) which is known to correlate with mutation load [82, 608, 609]. Furthermore, I investigated the difference in survival depending on the anatomical site of primary tumour (Fig. 3.16). In a univarate model using the Kaplan-Meier method and a global log-rank statistic, I observed a similar survival curve between melanomas arising on the head and limbs; however, patients with tumours on the trunk or other sites showed significantly worse survival. When including other variables known to correlate with survival into a multivariate model, the only statistically significant difference in survival between body sites was higher survival of patients with tumours on the limbs compared to the trunk (Fig. 3.16B).

3.7 Evaluation of chapter aims

- Outline the landscape of mutations including the top mutated genes and copy number alterations
  - Frequent somatic mutations were profiled which include the following results:
    - ≻ Primary melanomas generally have a high mutation load with a high proportion of C>T mutations.
    - ≻ The most frequently mutated genes include *BRAF* (45%), *NRAS (29%)*, *CDKN2A* (16%), *TP53* (10%) and *NF1* (10%).

A. Univariate survival analysis: Site of primary melanoma



B. Cox proportional hazards regression analysis

	Multivariate analysis		
Variable	**HR**	**95% CI**	**P-value**
Site of primary melanoma			
Limbs vs Head	1.02	0.52-1.98	0.96
Limbs vs Trunk	1.64	1.05-2.57	0.03
Limbs vs Other	1.26	0.69-2.33	0.45
Age at diagnosis	1.03	1.01-1.05	0.0004
Sex, male	1.40	0.96-2.06	0.08
Ulceration, present	1.82	1.20-2.75	0.005
Breslow thickness	1.09	1.02-1.16	0.007
Stage			
II vs I	1.02	0.57-1.80	0.96
III vs I	2.28	1.27-4.09	0.006

Figure 3.16: **Survival differences between primary melanomas of different anatomical location. A)** Kaplan-Meier survival analysis stratified by the site of primary melanoma. The global log-rank statistic is shown as well as the result from pairwise comparisons. **B)** Multivariate survival analysis using Cox proportional hazards regression analysis, looking at the site of primary melanoma together with known contributors to melanoma survival as additional covariates.

  - ≻ Promoter mutations most frequently involve *TERT* (37%) and *DPH3* (15%).
  - ≻ The top recurrent exonic variants are *BRAF* V600E (35%), *NRAS* Q61R (16%) and Q61K (10%), *RQCD1* P131L (5%), *RAC1* P29S (4%) and *BRAF* V600K (4%).
  - ≻ Novel recurrent mutations identified include *PCDHA2* S337L (2%), *TPTE* S447L (2%) and *AHCTF1* C271T (2%).
- – Frequent copy number alterations were assessed as showcased by the following results:
  - ≻ Copy number alterations are found in most primary melanomas, with homozygous deletions being most common.
  - ≻ The most frequently amplified genes were *CCND1* (5%) and *TERT* (4%), followed by genes located on chromosome 6p such as *IRF4* (3%) and 8q such as *MYC* (2%).
  - ≻ *IRF4* might be an interesting clinically actionable target, showcasing a potential tumour cell vulnerability.
  - ≻ *CDKN2A* was the most repeatedly deleted gene (13% of samples showed homozygous loss across the whole gene), followed by *PDCD1* (1.5%) and *PTEN* (1.2%).

- Identify melanoma driver genes and important genetic interactions

  - 15 melanoma driver genes were identified in this primary melanoma cohort:
    - ≻ Established melanoma driver genes were confirmed as drivers: *BRAF, NRAS, ARID2, NF1, TP53, CDKN2A, PTEN, RB1, RAC1, MAP2K1* and *FBXW1*.
    - ≻ Genes implicated to play a role in melanoma development were identified as driver genes: *FAM58A, RQCD1* and *MSR1*.
    - ≻ One novel melanoma driver gene was identified for the first time in this cohort: *TPTE*, which is a paralogue to *PTEN*.
  - 8 mutually exclusive gene pairs were found, potentially reflecting novel genes involved in important oncogenic pathways:
    - ≻ 6 genes were mutually exclusive with *BRAF,* where in particular *TLR4* and *EGFR* might indicate an alternative activation of the MAPK pathway driving oncogenesis in these tumours.
    - ≻ 2 genes were mutually exclusive with *CDKN2A,* where *PRDM2* is an interesting candidate to study as both encoded proteins interact with Rb.
  - No co-occurring gene pairs were identified, which could reflect the challenge of accounting for a high background mutation rate in melanomas.

- Understand how sun exposure can shape the genetic composition in melanomas
  - Primary melanomas generally have a high mutation load with a high proportion of C>T mutations, which is particularly profound in tumours arising on sun-exposed sites.
  - UV damage targets ETS transcription factor binding sites, giving rise to recurrent mutation patterns in promoter regions, which highlights the importance to further study the oncogenic impact of the mutations involved.
  - *BRAF* V600E mutations were inversely correlated with mutation load, and more frequently arise on body sites with intermittent sun exposure.
  - 11 of the 15 driver genes showed links to sun exposure. The exceptions were *PTEN* and three MAPK pathway genes: *BRAF, NRAS* and *MAP2K1,* potentially reflecting a difference in melanoma development in these tumours compared to UV damage-associated melanoma.

Chapter 4

Dysregulated biological pathways in primary melanoma

4.1 Introduction

It is now widely established, that studying the underlying genetic events in cancers, brings valuable insights into the key mechanisms driving disease development. Additionally, it helps us understand the complex molecular heterogeneity that exist between tumours. Although previous efforts have taught us a great deal about melanoma biology, melanoma as a disease is incredibly complicated and we still have far more to learn about how tumour growth is regulated. The properties a cell need to acquire in order to develop into a tumour, include the ability to grow and divide uncontrollably, thereby evading regular growth suppression and cell death mechanisms. Melanoma cells select for mutations providing such favourable traits; therefore, by studying the genes and biological pathways altered in tumours we can understand how tumours are wired, and reverse these mechanisms. An overview of the genetic landscape of primary melanoma from my dataset (Fig. 4.1) shows tremendously diverse alteration patterns between patients, with large variabilities not only in the genes harbouring alterations, but also the total number of events impacting each patient. There are several key pathways associated with melanomagenesis, where the MAPK pathway, PI3K/AKT pathway, and *CDKN2A*-associated cell cycle control are amongst the most important and will be discussed in greater detail throughout this chapter [93, 610, 611].

Several different methods of melanoma classification exist. Melanoma in general is a heterogenous cancer type, commonly impacted by a high burden of genetic alterations, and often characterised by the activation of the MAPK pathway. In 2015, the TCGA network established four genomic classes of melanoma, based on hotspot mutations in *BRAF, RAS*

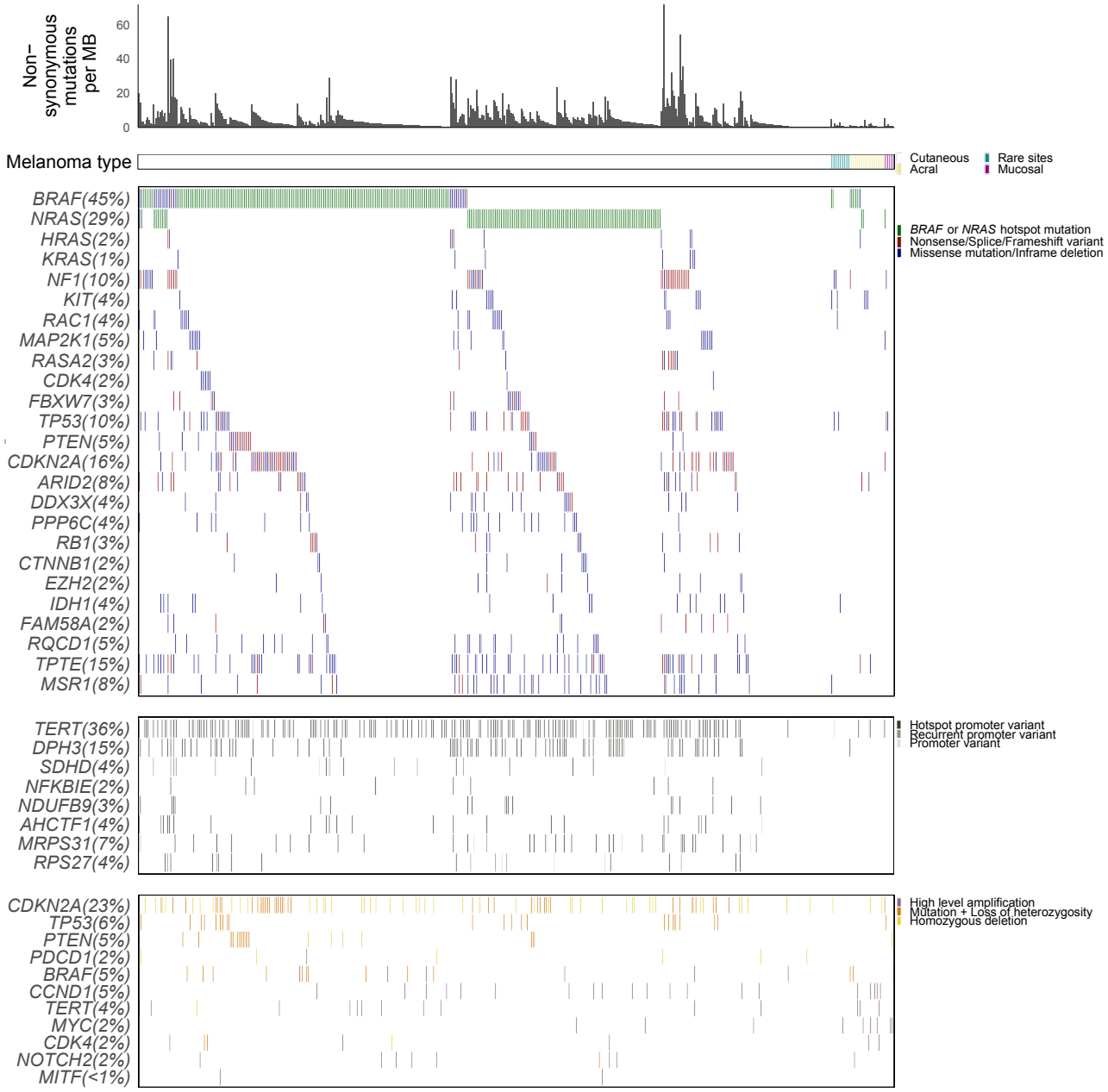


Figure 4.1: **An overview of the genetic landscape of primary melanoma.** The key alterations in each patient (Leeds melanoma cohort) are shown in this tile plot format. From top to bottom the panels show: mutation load (nonsynonymous mutation load per MB), melanoma type (cutaneous, acral, mucosal or other rare sites), coding mutations in candidate driver genes, mutations in the promoter regions of genes with recurrent variants, copy number alterations across relevant genes. The tiles are coloured by the variant consequence. Other rare sites refers to vulval, vaginal, anal and penile primaries which could not be accurately assigned as cutaneous or mucosal.

genes or *NF1*. Patients without mutations in these genes were classified as triple wild-type. Another classification approach is based on histopathological presentation [56, 57, 612]. Superficial spreading, nodular, lentigo maligna and acral lentiginous are some of the most common types of melanoma, which are often described together with other less common subtypes arising from the skin, as cutaneous melanoma. Acral melanoma differs from the former three by its presentation on non hair-bearing skin such as the palms, soles and under the nails. Occasionally, melanoma can also arise in the eye or from mucosal membranes, and is then termed uveal or mucosal melanoma, respectively. In this chapter, I will examine the composition of genetic alterations across patients belonging to the subgroups of these two melanoma classification systems. Finally, I will also explore how genetic alterations can affect the immune response in these patients.

Unless otherwise stated, all statistical analyses of gene associations comprising two categorical variables (i.e. the association of *BRAF* mutation with site of primary melanoma) were done using logistic regression. In the cases with one categorical and one continuous variable (i.e. the association of *BRAF* mutation with mutational load), linear regression was used instead. All analyses were univariate, with their reported p-value and odds-ratio (OR) without multiple testing correction, except when elsewise explicitly indicated.

4.1.1 Chapter aims

The aim of Chapter 4 is to study the heterogeneity of melanomas:

- Analyse the different genetic changes across melanoma subtypes
- Investigate the main genetic alterations and signalling pathways altered in different tumours
- Assess the impact of genetic alterations on the immune response in treatment-naive patients
- Evaluate the possible prognostic potential of genetic alterations

4.2 Analysis of mutational subtypes and pathway alterations

Mutations in key MAPK components are important events towards melanoma development. Hotspot mutations in *BRAF* and the *RAS* genes *NRAS, HRAS* and *KRAS* were completely

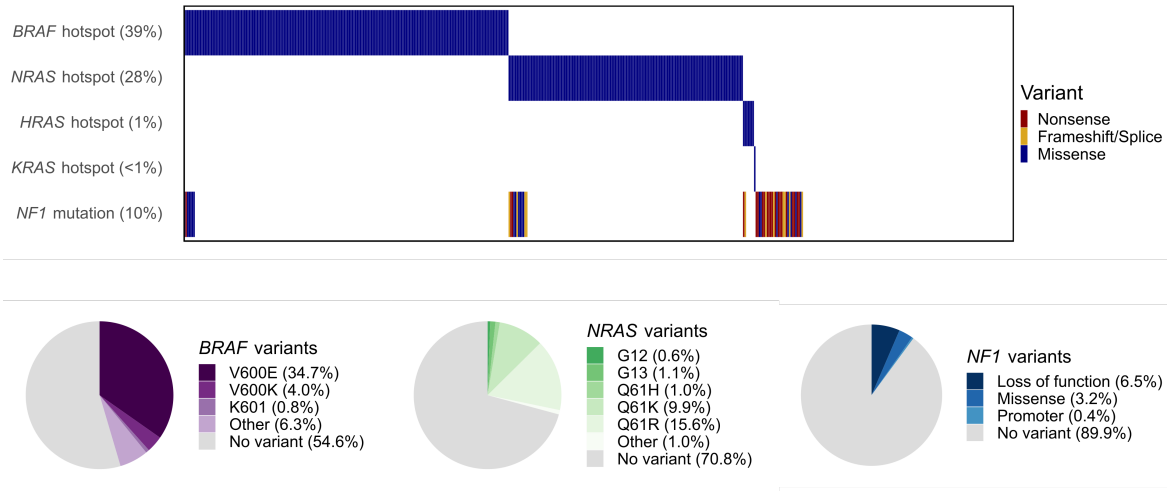


Figure 4.2: **Melanoma subtype variant distribution**. *BRAF, RAS* and *NF1* mutations in the Leeds melanoma cohort. Top panel: Distribution across patients of known hotspot mutations in *BRAF, NRAS, HRAS* and *KRAS* and any mutation in *NF1*. Bottom panel: Proportion of specific variants in *BRAF, NRAS* and nonsynonymous mutations in *NF1* in the Leeds melanoma cohort. Nonsynonymous and promoter variants were counted.

mutually exclusive (Fig. 4.2). Most non-missense mutations in *NF1* followed a mutually exclusive pattern with *BRAF* and *RAS* hotspot mutations; however, there were a few occurrences where both genes were co-mutated. This is consistent with previous studies of melanoma [73, 85, 161], and might be due to the presence of less damaging *NF1* variants (discussed in Section 3.5.1). The majority of *BRAF* variants were V600E, whilst a small proportion had V600K alterations. For *NRAS*, mutations in codon 61 were the most common, followed by a few percent of patients showing mutations in codon 12 and 13. The tumour suppressor gene *NF1* was mainly affected by loss of function mutations, represented by nonsense, frameshift variants or mutations in positions affecting splicing.

4.2.1 Sambar: Pathway-level mutational subtypes

Patient *BRAF* mutation status is used to inform targeted therapy decisions in the clinic [207, 613], but unfortunately, the current established mutational subtypes in melanoma can not distinguish between patient outcome [73]. Therefore, I set out to investigate if my mutational data can be used to outline further, more defined, mutational subtypes, that could partly explain the biological or clinicopathological differences among patients. The structure of these analyses were three-fold: First, I strived to compare the new classes with currently established mutational subtypes in melanoma, to understand the value of further

dividing patients based on mutational profiles. Second, I explored the differences between the new classes on a pathway-level, to understand how patients could be distinguished based on the composition of genetic alterations in their tumours. Lastly, I set out to test if these new classes could be used to identify particularly favourable or adverse patterns of genetic alterations in terms of patient survival.

This work was done in collaboration with Dr. Marieke Kuijjer (Section C.1), who developed the tool Sambar, which utilises mutation data projected on a pathway-level to classify patients into subtypes [614]. Sambar successfully grouped 85% of all patients into six classes based on the presence of pathway-associated mutations (Fig. A.10, Table A.1).

NF1-mutant and triple wild-type tumours comprised a large proportion of the samples which could not be subtyped. *NF1*-mutant patients are often older, harbouring a higher number of total mutations, which could overcrowd and complicate the clustering algorithm, explaining why they were particularly difficult to subtype. 17 of the 84 patients (20%) that could not be subtyped had a nonsynonymous mutation load above 20, with several patients showing very high mutation burden (Fig. 4.3). Triple wild-type patients on the other hand, reflect the other extreme - often affected by very few mutational events, which would provide too limited data for subtyping purposes. All 17 cases where no nonsynonymous mutations were found, and an additional 7 each having 5 or less nonsynonymous mutations in total (29%), were triple wild-type melanomas. To summarise, I can conclude at least 50% of the tumours unsuccessfully subtyped were likely due to them having very high or low numbers of mutations.

4.2.2 Tumour distribution into Sambar classes

Sambar clustered the patients into six groups comprising three larger (black, red, green) and three smaller (blue, cyan, magenta) classes (Fig. 4.3). To compare the new classes with the previously established, known mutation subtypes in melanoma, I looked at the distribution of *BRAF, NRAS, NF1* and WT subtypes across the Sambar classes. The black Sambar class almost exclusively contained wild-type patients, reflecting the distinct molecular composition of these patients. Previous research has shown triple wild-type melanoma patients, compared to other subtypes, have low numbers of mutations, different driver genes and a larger proportion of genetic alterations other than mutational events [73, 91, 615]. My data strengthens these findings, as the mutational composition of these patients clustered them together into one separate group.

The other five Sambar classes were more difficult to interpret, although they seem to

be driven not only by pathway data but also mutation load, with the cyan and magenta clusters in particular showing high mutation burden. Patients with hotspot mutations in *BRAF* seemed to dominate the red and blue clusters, which indicates that there is heterogeneity within the *BRAF* subtype, as a small portion of *BRAF*-mutant patients presented a different pathway mutation composition. Similarly, the magenta group contained largely the *NRAS* subtype, but *NRAS* hotspot mutant patients also comprised large fractions of the green and cyan classes. In view of these results, I conclude that the underlying genetic events in primary melanoma patients are very heterogenous, and cannot be purely explained with the currently defined mutational subtypes.

Interestingly, *NF1*-mutant samples were mainly distributed into three classes, each showing similarities with the *BRAF* (blue), *NRAS* (magenta) and WT (black) groups, respectively. This indicates *NF1*-mutant samples do not comprise a distinct molecular subtype by itself. However, as over half of the *NF1*-mutant samples in my cohort could not be subtyped via Sambar, many of which show very high mutation burden, it cannot be completely ruled out that a subset of *NF1*-mutant patients belong to a separate group with a different molecular background. In my opinion, it is also important to distinguish between some triple wild-type patients and *NF1*-mutant patients, because the genetic composition between these groups of patients is vastly different. While triple wild-type patients often show very few mutations, *NF1*-mutant patients are rather characterised by a much higher mutation load. It is possible, this latter group is driven by the high mutational burden, which distinguishes them from other patients, rather than attributing this separation specifically to harbouring a *NF1*-mutation. To conclude regarding *NF1*-mutant patients, my analysis suggest these patients can be divided into four subsets, of which the first three groups show similarities with *BRAF, NRAS* and WT subtypes, respectively, while the fourth group is driven by high mutational load.

4.2.3 Pathways operating in the Sambar classes

Next, I wanted to see which pathways were enriched in each Sambar class. For this analysis, Dr. Marieke Kuijjer made word clouds using the following steps: First, she assigned Sambar class-specific pathways by selecting pathways mutated in over 95% of all samples of one specific Sambar class, but for which less than 5% of samples in other Sambar classes showed mutations. Next, she calculated the number of times a specific word was observed per Sambar class in relation to the total number of words observed in that class. Finally, a correction was performed by normalising the word frequency to the background mutation

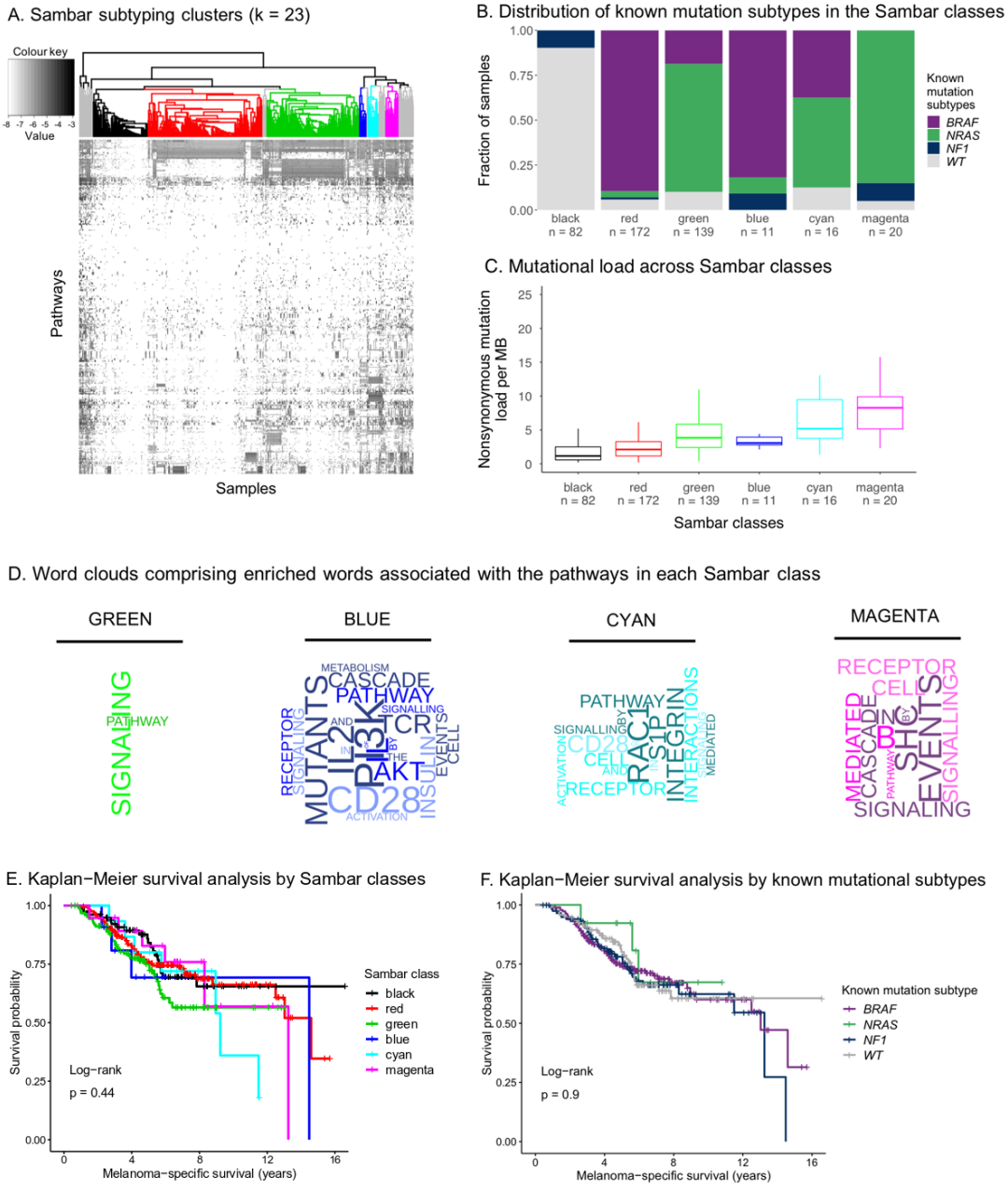


Figure 4.3: **Mutational subtypes identified using Sambar. A)** Six clusters identified by Sambar based on the composition of pathway mutations in the tumours (Figure prepared by Marieke Kuijjer). **B)** The proportion of known mutation subtypes across the six Sambar classes. **C)** The mutational load across the six Sambar classes. **D)** Word clouds comprising enriched words associated with the pathways of the green, blue, cyan and magenta Sambar classes (Figure prepared by Marieke Kuijjer). **E)** Survival curves showing melanoma-specific survival stratified by Sambar class. **F)** Survival curves showing melanoma-specific survival stratified by the current known mutation subtypes in melanoma.

pathway list. For a more detailed description, see the methods section of Marieke's paper [614]. Unfortunately, the black and red classes did not yield any words, whilst the green class did not provide any informative words (Fig. 4.3D). However, the remaining three classes did show some specific pathway-enrichment of certain words. CD28, an important co-stimulatory receptor of T cells, is shared between the blue and cyan class, although it is more pronounced in the blue class. This indicates an immune component is active in these groups. The blue class also seem to be dominated by PI3K, IL2 and to some extent AKT, the downstream effector molecule to PI3K, which suggests the PI3K pathway is altered in this group of patients, potentially through CD28-mediated activation [616, 617]. The cyan class on the other hand is enriched in RAC1, integrin and S1P, which could suggest an enhanced importance of cell adhesion, migration and trafficking [618–623]. Lastly, the top enriched words in the magenta class were SHC and B, reflecting SHC-mediated events and B cell signalling. This analysis concludes the heterogenous landscape of mutations in primary melanoma, showing a diverse range of biological pathways playing a role in different patients. The currently established mutational subtypes in melanoma do not comprehensively explain the underlying genetic events is these patients. The Sambar classes further unravels patient's biological differences, by suggesting one group of patients with enriched alterations in the PI3K pathway (blue), a second group showing links to cell migration (cyan), and a third group of patients having mutations associated with SHC and B cell signalling (magenta).

4.2.4 Prognostic value of the Sambar classes

Finally, I also wanted to check if Sambar could help identify mutational differences between patients impacting their survival. The established mutational subtypes in melanoma are not capable of predicting patient survival, in the Leeds melanoma cohort (Fig. 4.3F) or in other cohorts [73, 624, 625]. Unfortunately, the six Sambar classes also failed in providing prognostic value in a univariate model (Fig. 4.3E). I also investigated whether any of the classes were dominated by other clinicopathological features, such as site of primary melanoma, sun exposure, ulceration, Breslow thickness and relapse. However, there appears to be no bias towards any of the tested variables in any one class. This suggests the diversity in mutation composition, does not in itself explain the observed difference in clinical characteristics. However, this does not completely rule out the possibility of genetic alterations contributing to such factors. Little research to date has focused on building models based on extensive somatic mutation data [614], and such analyses are especially difficult to do for

melanoma because of the high background mutation rate. Other groups have successfully shown alterations in a small number of genes to associate with survival probability, as well as using transcriptomic data to find biological differences amongst patients of a particular cancer type [73, 626–631]. It is possible more sophisticated methods capable of separating key mutations from the excess of uninformative alterations are necessary to tease out a true signal. Furthermore, it would be valuable to also explore the possibility to improve the current clinical prognostic factors with information regarding the mutational landscape of melanoma.

4.2.5 Prognostic value of other pathway-level alterations

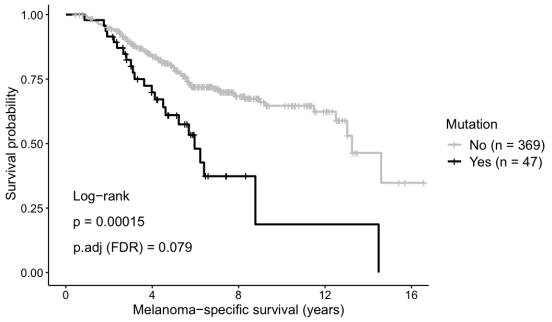
To conclude my pathway-level analysis, I wanted to check if any prognostic effect, unrelated to the Sambar classes, could have been missed. I used a simple model whereby I defined in a binary format whether a patient had alterations in any of the MSigDb canonical pathways (file "c2.cp.v5.0.edges.gmt") individually, and then I did a log-rank test to detect if alterations in any one pathway could explain a patient's survival difference. After adjusting for multiple testing (n = 1025) and removing pathways where the alteration rate were below 5%, the only pathway falling below the FDR-adjusted p-value threshold of 10% was the Reactome T cell receptor signalling pathway (Fig. 4.4A). The increased risk associated with mutations in the Reactome TCR signalling pathway held true also in a multivariate model taking into account covariates known to be associated with survival (Fig. 4.4B). It would be intriguing to test these patterns in other datasets of primary melanoma.

4.3 Genetic changes across melanoma subtypes

In the Leeds melanoma cohort, 90% of patients were diagnosed with a subtype classified as non-acral cutaneous melanoma, but there were also cases of acral (n = 24) or non-conjunctival mucosal (n = 7) melanoma. Additionally, 13 patients had vulval, vaginal, anal or penile melanoma. As their clinical records did not state whether the primary tumour arose on the skin or mucosal surfaces, I have grouped these separately as melanoma from rare sites.

In comparing the genetic alterations present across different subtypes of melanoma, it is clear a distinction can be made between cutaneous and non-cutaneous melanoma (Fig. 4.5). Firstly, the number of mutational events per patient, although varied, were a lot higher in patients with non-acral cutaneous melanoma, with patients showing as high as 72 nonsyn-

A. Univariate survival analysis:
Reactome TCR signalling pathway



B. Cox proportional hazards regression analysis

	Multivariate analysis		
Variable	**HR**	**95% CI**	**P-value**
Reactome TCR signalling, mutation	1.92	1.17-3.15	0.010
Age at diagnosis	1.03	1.01-1.05	0.006
Sex, male	1.56	1.06-2.29	0.024
Ulceration, present	1.76	1.13-2.73	0.012
Breslow thickness	1.08	1.01-1.15	0.023
Stage			
II vs I	0.93	0.51-1.71	0.815
III vs I	2.11	1.15-3.88	0.016

Figure 4.4: **Survival analysis by mutations in the Reactome TCR signalling pathway. A)** Kaplan-Meier survival analysis stratified by presence of mutations in genes of the Reactome TCR signalling pathway. The FDR-adjusted p-value was computed after correcting for multiple testing across all 1025 pathways analysed. **B)** Multivariate survival analysis using Cox proportional hazards regression analysis, looking at the presence of mutations in genes of the Reactome TCR signalling pathway together with known contributors to melanoma survival as additional covariates.

onymous mutations per MB. The mutation load in non-acral cutaneous melanoma samples (average 5.4 mutations per MB) were significantly different from that of acral (average 0.83 mutations per MB, Mann-Whitney test, p-value = 7.5 x 10^{-10}) and mucosal (average 1.5 mutations per MB, Mann-Whitney test, p-value = 0.008) melanoma. This observation is however not surprising [632], as tumours arising on sun-shielded body sites would not show the high UV mutation pattern of melanomas arising on sun-exposed locations. Secondly, mutations in candidate driver genes were more sparse in acral and mucosal melanoma compared to non-acral cutaneous melanoma. This is likely a consequence of lower mutation burden but reflects a difference in disease biology between these melanoma subtypes.

4.3.1 Mucosal melanoma

Hotspot mutations in *BRAF* are the most frequent somatic changes in non-acral cutaneous melanoma patients, whereas mucosal melanoma patients completely lacked *BRAF* mutations, and only one patient had a high level amplification of *BRAF*. No recurrent mutations were found in mucosal melanoma samples, and only one mutation each was found in *NRAS* (Q61R) and in *NF1* (M817V). In contrast, genes recurrently mutated in mucosal melanoma samples include *LZTR1* (n = 3), *ATRX* (n = 2) and *TP53* (n = 2). Single cases of coding mutations were also observed in *MAP2K1* and *CDKN2A*. Activating mutations in *GNAQ, GNA11* and *SF3B1* commonly found in uveal melanoma, with few cases reported in mu-

cosal melanomas [91, 633–636], were completely absent in my cohort of mucosal melanoma. *SF3B1* was amplified in one sample. Two samples showed oncogenic alterations in *TERT,* with one being a hotspot mutation (C250T), and the other a high level amplification. Over 50% of the mucosal melanoma samples had high level amplifications of chromosome 8q, including genes such as *MYC* (n = 3), *PREX2* (n = 4), *TRPA1* (n = 4) and *NDUFB9* (n = 3). Only one case of a homozygous deletion was found amongst the mucosal melanomas, targeting *PTEN* with one patient affected. Previous reports of mucosal melanoma have shown frequent *KIT* alterations [606, 607]; however, the small number of mucosal samples in my cohort showed neither *KIT* mutations nor amplifications. However, when examining the group of tumours from other rare sites, cases where it could not be determined whether the tumour originated on mucosal surfaces or the surrounding skin, I found another case of a *MYC* amplification. Additionally, one case harbouring a *KIT* amplification and another with both a coding *KIT* mutation and high level amplification in the same patient were also discovered. Furtheremore, I also found two cases of coding mutations in *SF3B1*. These findings suggest genetic changes in *KIT* and *SF3B1,* if present in mucosal melanoma, might be more common in tumours of genital origin.

4.3.2 Acral melanoma

Only small cohorts of acral melanomas have comprehensively studied a larger number of genes to date [91, 637–639]. Those studies report mutations affecting key members of the MAPK pathway such as *BRAF, NRAS* and *NF1* but also genes less frequently mutated in non-acral cutaneous melanoma. Recurrent alterations affecting *KIT* [606, 607, 638] and *CDK4* pathway genes [640] have been reported, as well as an abundance in copy number alterations and structural variants in general [91]. The genetic changes observed in acral melanomas of the Leeds melanoma cohort included six cases of *BRAF* V600E, two cases of *NRAS* Q61K and two cases of *TERT* C228T promoter mutations. Although the frequency of these changes were lower compared to melanomas originating from non-glabrous skin, this still suggests the genetic drivers in acral melanoma are more similar to non-acral cutaneous melanoma than mucosal melanoma. The top mutated genes in acral melanoma were *BRAF* (n = 8), *TTN* (n = 7) and *KIT* (n = 3). The most frequent copy number aberrations affecting acral melanomas were amplifications targeting chromosome 11q, harbouring *CCND1* and *GAB2* (n = 5), followed by chromosome 8q where *MYC* is located (n = 3). *TERT* was found amplified in three samples, while *KIT* and *CDK4* were amplified in two samples respectively. One of the samples with *KIT* alterations showed the presence of both a mutation and high level amplification. Two out of three patients with *MYC* amplifications showed

co-presence of *CCND1* amplifications. Homozygous deletions were found in *CDKN2A* affecting two patients with acral melanoma. Interestingly, homozygous deletions were also found in genes associated with increased risk of developing cancer: the breast cancer susceptibility gene *BRCA1* and the familial melanoma gene *BAP1* were each found deleted in one patient.

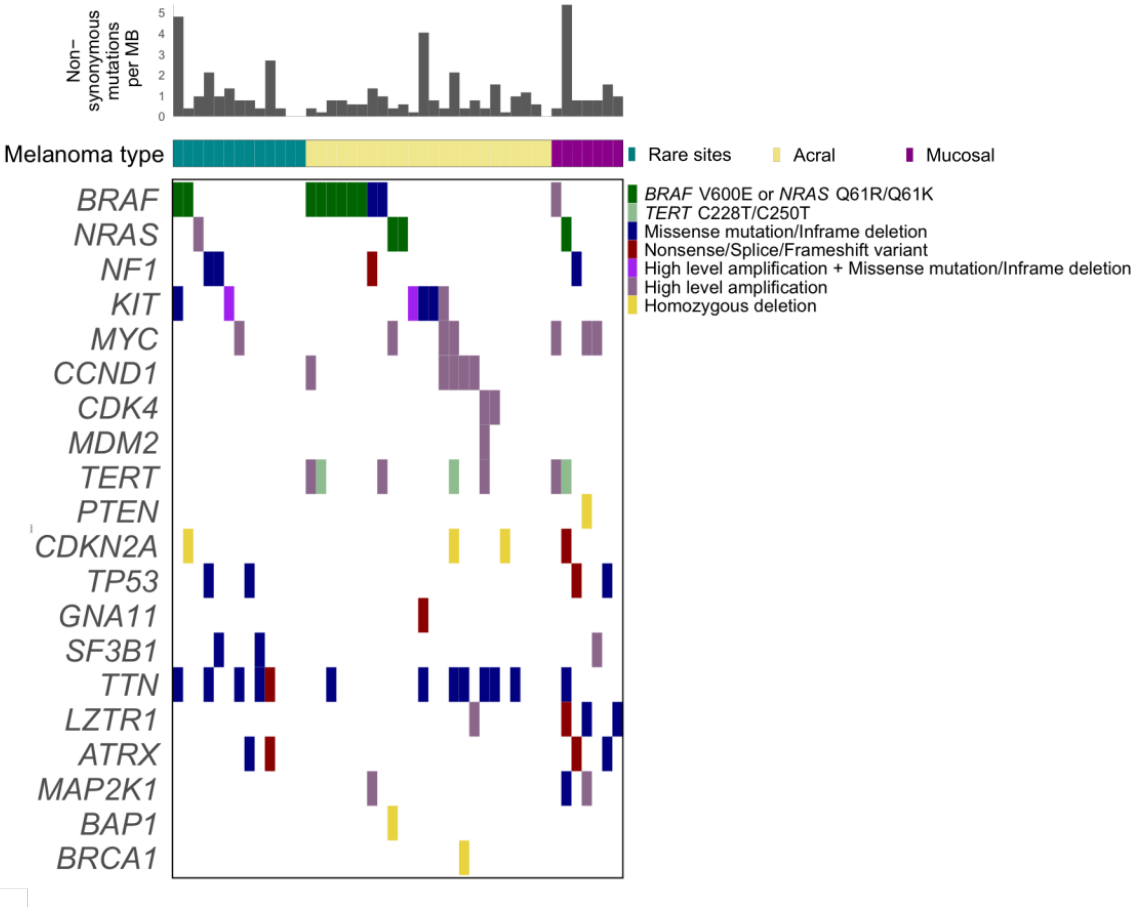


Figure 4.5: **Key genetic alterations in acral, mucosal and other rare subtypes of melanoma.** Tile plot showing an overview of the genetic alterations for each patient with acral, mucosal and other rare subtypes of melanoma. From top to bottom the panels show: mutation load (nonsynonymous mutation load per MB), melanoma type (acral, mucosal or other rare sites), and important genetic alterations in candidate driver genes coloured by their respective variant consequence.

4.3.3 Copy number profiles across major melanoma subtypes

To further study copy number alteration differences between melanoma subtypes, I looked at a stratified genomic overview of such alterations by visualising the proportion of all seg-

ments which differ more than 0.6 from a sample's ploidy (Fig. 4.6). Striking similarities were found between acral and non-acral cutaneous melanoma, where the copy number profiles looked almost identical. The acral melanomas had slightly more alterations including more frequent amplification of 1q and 8q, and deletions in 6q and 11q. Mucosal melanomas were also found to have abundant copy number changes, with 100% of samples in particular showing amplification of 6p and 8q. Similarly to acral melanomas, a high number of mucosal samples showed amplifications in 1q and deletions in 6q. Mucosal melanomas were also enriched in deletions affecting the X chromosome; however, this is likely a bias because all patients of this cohort presented with mucosal melanoma were female. The other rare sites group of melanoma samples showed a copy number profile with resemblance to both cutaneous and mucosal melanoma.

4.3.4 Copy number profiles across established mutational subtypes

To disentangle differences between patient groups of non-acral cutaneous melanomas, I also looked at the whole genome copy number profiles of each mutational subtype *BRAF*, *RAS, NF1* and WT. Most subgroups had some distinct copy number alterations specific for respective subtype. The *BRAF* subtype showed a sharp peak comprising samples which had deletions in chromosome 10, which was not seen amongst the other subtypes. The top alterations in the *NRAS* subtype were amplifications of chromosome 1q, 6p and 8q, thereby showing similarities with the acral and mucosal subtypes of melanoma. *NF1*-mutant melanomas showed a skewed copy number profile favouring amplifications across chromosome 5, as well as deletions targeting 11q and 17p. In contrast to the literature, the WT group did not seem to harbour abundant copy number alterations compared to the other subtypes [73, 85, 91]. However, the design of the copy number probes for my project allowed only for analysis of focal copy number alterations across specific genes, and therefore events outside of these regions could not be studied at high resolution. In addition, I did not investigate complex structural variants such as gene fusions, rearrangements or larger insertions. Therefore, it cannot be concluded whether such genetic aberrations not analysed as part of this project would contribute to the pathogenesis in this group of patients, as proposed by other researchers [73, 85, 91].

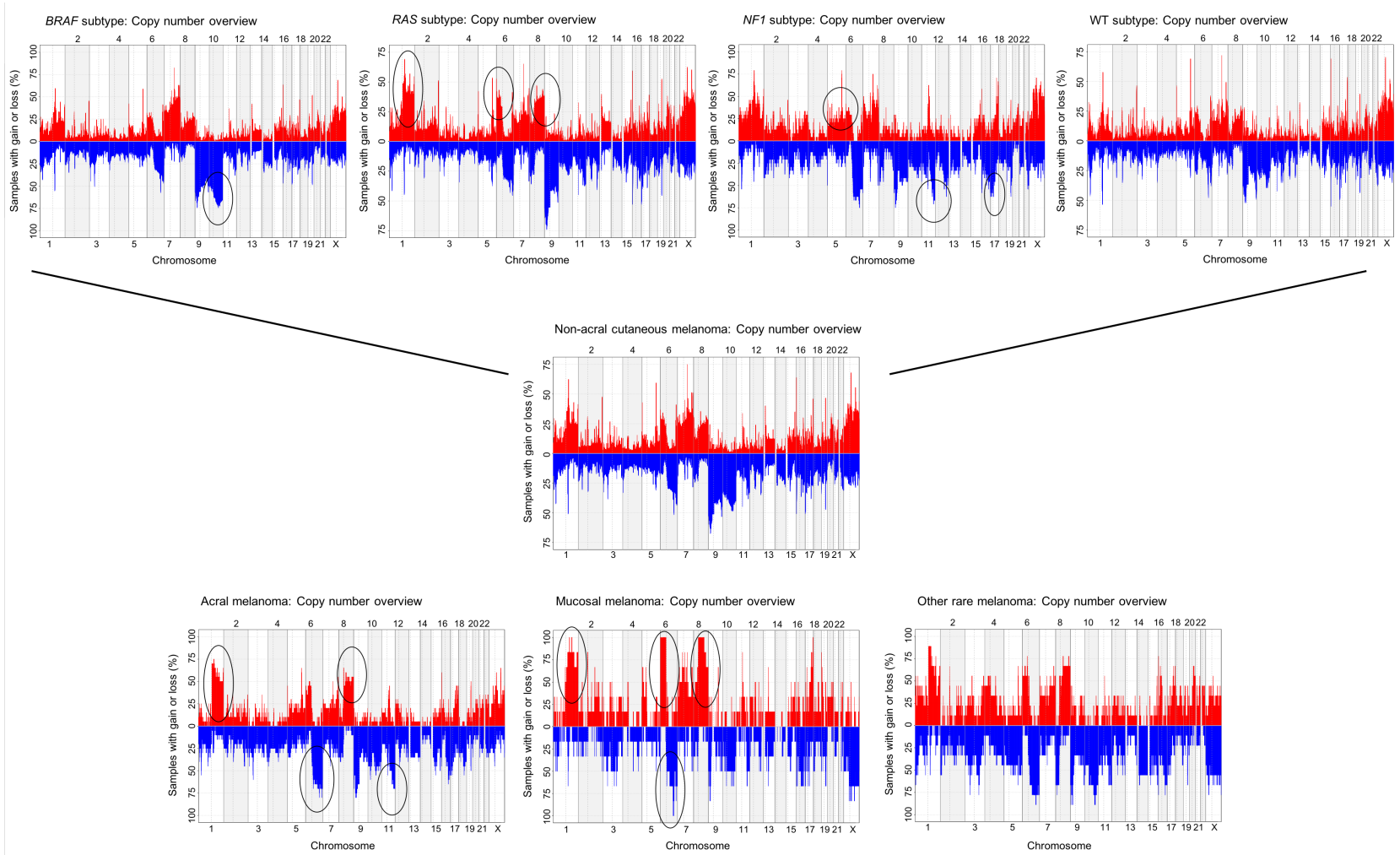


Figure 4.6: **Copy number alterations across different melanoma subtypes.** Frequency of samples showing gain (red) or loss (blue) across the whole genome. All segments with a copy number differing more than 0.6 from the sample ploidy were used to generate the figures. The top panel shows the copy number overview for *BRAF, NRAS, NF1* and *WT* subtypes of the non-acral cutaneous melanoma group. The middle panel shows the copy number profile for non-acral cutaneous melanoma. The bottom panel shows the copy number profile for acral, mucosal and melanomas from other rare sites. Interesting differences between the subtypes are marked with circles.

4.4 The MAPK pathway

Gain of function mutations activating the MAPK pathway are key events in progression towards melanoma development, leading to uncontrolled cell proliferation [102, 103, 641]. Following stimulation of receptor tyrosine kinases such as *KIT,* the cascade starts with the stepwise activation of key signalling molecules RAS and RAF*,* which sequentially phosphorylates MEK and thereafter ERK, resulting in the transcriptional activation of genes involved in promoting cell proliferation [642, 643], survival [644], and differentiation [645, 646]. Activation of the MAPK pathway is the most common alteration found in melanoma patients, with approximately 40% of patients in the Leeds melanoma cohort having obtained a recognised activating mutation in *BRAF*, and another 30% of patients affected by hotspot mutations in *NRAS*. In contrast to other cancers which are often regulated by *KRAS* mutations [647, 648]*, NRAS* is the dominant *RAS* gene altered in melanoma. This could reflect differences between genes in oncogenic propensity between the cells of origin giving rise to the tumour [649]. Hotspot mutations in *HRAS* were found in seven patients, while for *KRAS* only one case was found.

Mutations in genes encoding the downstream pathway components MEK and ERK, are less frequent in melanoma and other cancers, possibly owing to redundancy when the commonly altered *BRAF* or *NRAS* genes are already activating the cascade. In the Leeds melanoma patients, *MAP2K1* and *MAP2K2,* encoding MEK1 and MEK2, were mutated in 4.6% and 2.3% of patients, respectively. Several recurrent mutations were found across *MAP2K1*, with amino acid position 124 of *MAP2K1* altered in 10 patients, including 7 patients with the P124L variant and 3 patients with the P124S variant. Additionally, a missense mutation targeting E203K, a 6 bp in-frame deletion of L101-E102 and amino acid position 53 (single cases of F53L and F53S) were mutated in two patients respectively. In the TCGA melanoma cohorts (Pan-cancer and SKCM) [96], amino acid 124 and 203 of *MAP2K1* were the top altered positions. Additionally, the TCGA pan-cancer cohort reports recurrent in-frame deletions targeting amino acid position 102 and 103, suggesting the positions around amino acid 102 might be important for activation of the MAPK pathway through MEK1. All of the recurrent mutations in *MAP2K1* found in my primary melanoma samples, with the exception of the specific L101-E102 in-frame deletion which has not been tested to my knowledge, have been shown to cause constitutive activation of the MAPK pathway [650, 651]. Deletions targeting amino acid position 101-102 have been reported to increase ERK phosphorylation [651], therefore this in-frame deletion is likely to also induce MAPK signalling. No positions across *MAP2K2* were recurrently mutated; however, one F57C

mutation which corresponds to position F53 in *MAP2K1*, has been shown to increase phosphorylation of ERK [652]. In addition, the TCGA SKCM cohort reported two variants in the same position, albeit the amino acid change differed (F57V and F57L). Both of these alterations were deemed as likely oncogenic by OncoKB [653]. Mutations in both *MAP2K1* and *MAPK2K2* were never found in the same patient (Fig. 4.7).

Mutations in *MAPK3* (ERK1) and *MAPK1* (ERK2) are rare both in melanoma and other cancers [73, 96], therefore I chose not to include them in my bait design. Hence, I cannot study alterations affecting these genes in my cohort. Sporadic mutations in these genes affected less than 2% of melanoma patients in the TCGA SKCM cohort.

NF1, a tumour suppressor and negative regulator of *RAS*, is one of the top mutated genes in melanoma, and has gained increasing interest the past five years [654]. It has now been established as an important regulator contributing to melanoma development by activation of the MAPK pathway [73, 161]. *NF1* loss of function mutations mainly occurred in samples which did not harbour hotspot mutations in *BRAF* and *NRAS* (Fig. 4.1). A strong correlation was also observed between *NF1* mutations and higher mutation burden (univariate logistic regression p-value = 4.62 x 10^{-13}, OR = 1.19), suggesting a UV-associated disease mechanism in these patients. Patients with *NF1* mutations were also found to be older (univariate logistic regression p-value = 2.23 x 10^{-4}, OR = 1.06), which could be another factor adding to the connection with high mutation load. A weak correlation was observed between *NF1* mutation and site of primary melanoma, with tumours arising on sun-exposed areas such as the head being more common than those arising on limbs (univariate logistic regression limbs over head p-value = 0.0017, OR = 0.26) or the trunk (univariate logistic regression trunk over head p-value = 0.03, OR = 0.42). This observation led me to explore a possible correlation with sun sensitivity, and indeed *NF1* mutations were more prevalent in patients reported to be sensitive compared to non-sensitive (univariate logistic regression p-value 0.0055, OR = 2.65). Taken together into a multivariate model: mutation load, age and sun sensitivity were all correlated with *NF1* mutation (age p-value = 0.04, OR = 1.04, mutation load p-value = 1.61 x 10^{-9}, OR = 1.18, sun sensitivity p-value = 0.006, OR = 3.11). I therefore hypothesise, that *NF1* mutations arise as a consequence of UV damage in people who are sensitive to sunburn.

RAC1, another established melanoma oncogene, was mutated in 23 primary melanoma samples, of which 19 were the known gain of function alteration P29S, which has been shown to increase cell proliferation through increased ERK phosphorylation [85, 155]. Alterations in *RAC1* have been associated with UV exposure; however, in my dataset *RAC1* changes only weakly correlated with mutational load (univariate logistic regression, p =

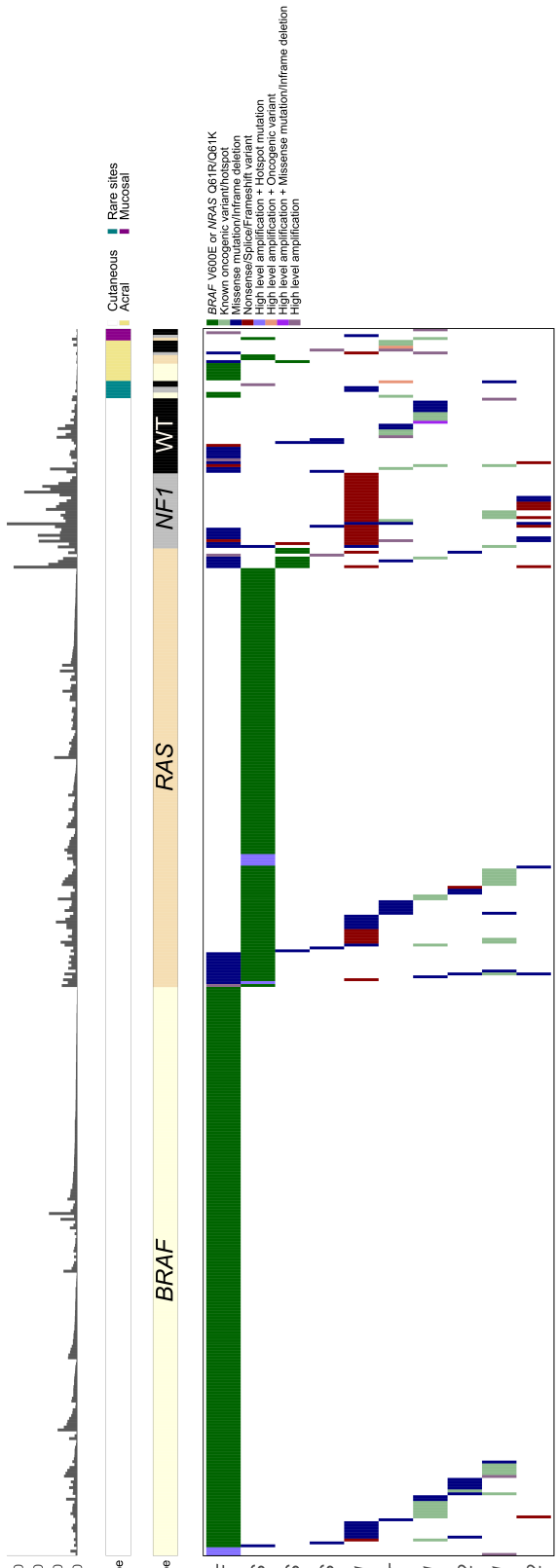


Figure 4.7: **Genetic alterations targeting selected important components of the MAPK pathway.** From top to bottom the panels show: mutation load (nonsynonymous mutation load per MB), melanoma type (cutaneous, acral, mucosal or other rare sites), mutational subtypes (*BRAF, NRAS, NF1* or *WT*), important genetic alterations in key genes of the MAPK pathway coloured by their respective variant consequence. Known oncogenic variants/hotspot denotes *MAP2K1/MAP2K2/KIT* variants with oncogenic annotations by OncoKB or *RAC1* P29S variants.

0.013, OR = 1.04). Tumours with this alteration were not associated with either gender, site of primary melanoma, sun sensitivity, suntan or sunburn susceptibility. *RAC1* was mutated in only one melanoma from rare sites, and mutations were completely absent in the triple wild-type, acral and mucosal melanoma subtypes (Fig. 4.1).

Receptor tyrosine kinases such as *KIT* are important upstream components of the main members of the MAPK signalling cascade. One known oncogenic recurrent mutation in *KIT*, L576P, was found in five patients [655, 656]. Single cases of V559A, K642E and N822Y mutations overlapped with TCGA-reported melanoma variants, with reported oncogenic or potential oncogenic capability [653]. Additionally, one in-frame deletion affecting E554, was located within a cluster of other recurrent oncogenic deletions in gastrointestinal stromal tumours [96]. Mutations targeting *KIT* have been shown by previous studies to be less prevalent in non-acral cutaneous melanoma compared to other subtypes [91], except for tumours arising on chronically sun-damaged skin [606], or in triple wild-type melanomas [73]. In my cohort, this observation was in agreement with previous studies. Out of 19 *KIT* mutations, 1 occurred in the *BRAF* subtype and 5 in the *RAS* subtype (5 with *NRAS* and 1 with *HRAS)*. Two patients showed co-mutated *KIT* and *NF1,* five triple wild-type patients had *KIT* mutations, and the remaining five patients had melanomas classified as acral or melanoma from other rare sites. Again, because of the high prevalence of *BRAF* and *NRAS* mutations activating the MAPK pathway among the non-acral cutaneous melanoma patients, additional mutations of *KIT* would be excessive and not selected for. However, this does suggest gain-of-function mutations in *KIT* might be important in wild-type and rarer subtypes of melanoma in driving tumourigenesis through activation of the MAPK pathway. It is intriguing to find *KIT* mutations in such a diverse range of tumours: from high mutation load sun-exposed tumours to rarer melanomas arising on sun-protected sites. This further exemplifies the complexity and disease heterogeneity of melanoma.

KIT might be the most important receptor tyrosine kinase in the context of melanoma; however, there are other upstream receptors of the MAPK pathway which are mutated in melanoma (Fig. A.11). Mutations in these receptor tyrosine kinases were of a mutually exclusive pattern; however, a larger fraction of co-mutated genes were observed with higher mutation burden, such as in the *NF1*-mutant patients. This suggests the presence of mutations in multiple receptor tyrosine kinases are not selected for, yet does not provide a disadvantage to the tumours.

4.5 The PI3K/AKT pathway

Activation of the PI3K/AKT pathway is another crucial step towards melanoma development [99, 118, 657, 658], conferring enhanced cell proliferation and survival [100, 118]. Although the PI3K/AKT pathway is often described as an independent route towards melanoma progression, there are several components overlapping with other important pathways, most notably the MAPK pathway, including receptor tyrosine kinase activation and regulation through RAS proteins.

With the exception of *RAS,* which modulates both the MAPK and PI3K/AKT pathway, melanomas seem to activate the PI3K/AKT pathway mainly through inactivating mutations or deletions of *PTEN* [73, 85, 91, 132, 659]. Mutations in *PIK3CA*, which encodes the p110α catalytic subunit of PI3K, or overexpression of *AKT3* have also been reported but are less frequent [660–662]. In my cohort of primary melanomas, the key components of the PI3K/AKT pathways, including *RAS*, were modulated through mutation, high level amplification or homozygous deletion in 55% of all tumours (Fig. 4.8).

PTEN was altered through mutation or copy number loss in a total of 153 samples (38% of all samples analysed for both mutation and copy number events), including 31 samples where the gene was either mutated or homozygously lost, emphasising the essential role *PTEN* has in controlling oncogenic processes. Of the class IA PI3Ks: *PIK3C*A, *PIK3CB* and *PIK3R1* were possible to study due to my sequencing design. No recurrent exonic mutations were found in *PIK3CA*, and of the eight exonic mutations found, two single cases were in previously discovered hotspot positions [663, 664], although none included the most common E545K variant [96, 665]. *PIK3CB* and *PIK3R1* were altered in nine samples each. *AKT1*, *AKT2* and *AKT3* were mutated in four, seven and five samples, respectively, with no shared variants between patients. The known oncogenic hotspot E17K was altered in one case in *AKT1*, while both *AKT2 (*3 out of 7) and *AKT3* (3 out of 5) had altered amino acid positions overlapping with reports from other cancer studies [73, 96]. Of the *AKT* genes, only *AKT3* was amplified (n = 5), and these cases were free of *AKT3* mutations. Most mutations targeting the main PI3K/AKT pathway components were mutually exclusive (Fig. 4.8), emphasising the central role of these genes in activating this pathway. Additionally, this shows that excessive stimulation of this pathway is unnecessary for the development of melanoma in these cases. Interestingly, no cases of *PTEN* mutations were found in triple wild-type melanomas, an observation that I confirmed by looking at the largest study of major melanoma subtypes to date [91]. This indicates that this group of patients might activate this pathway in a *PTEN*-independent way. As *TP53* and *PTEN* regulate each other

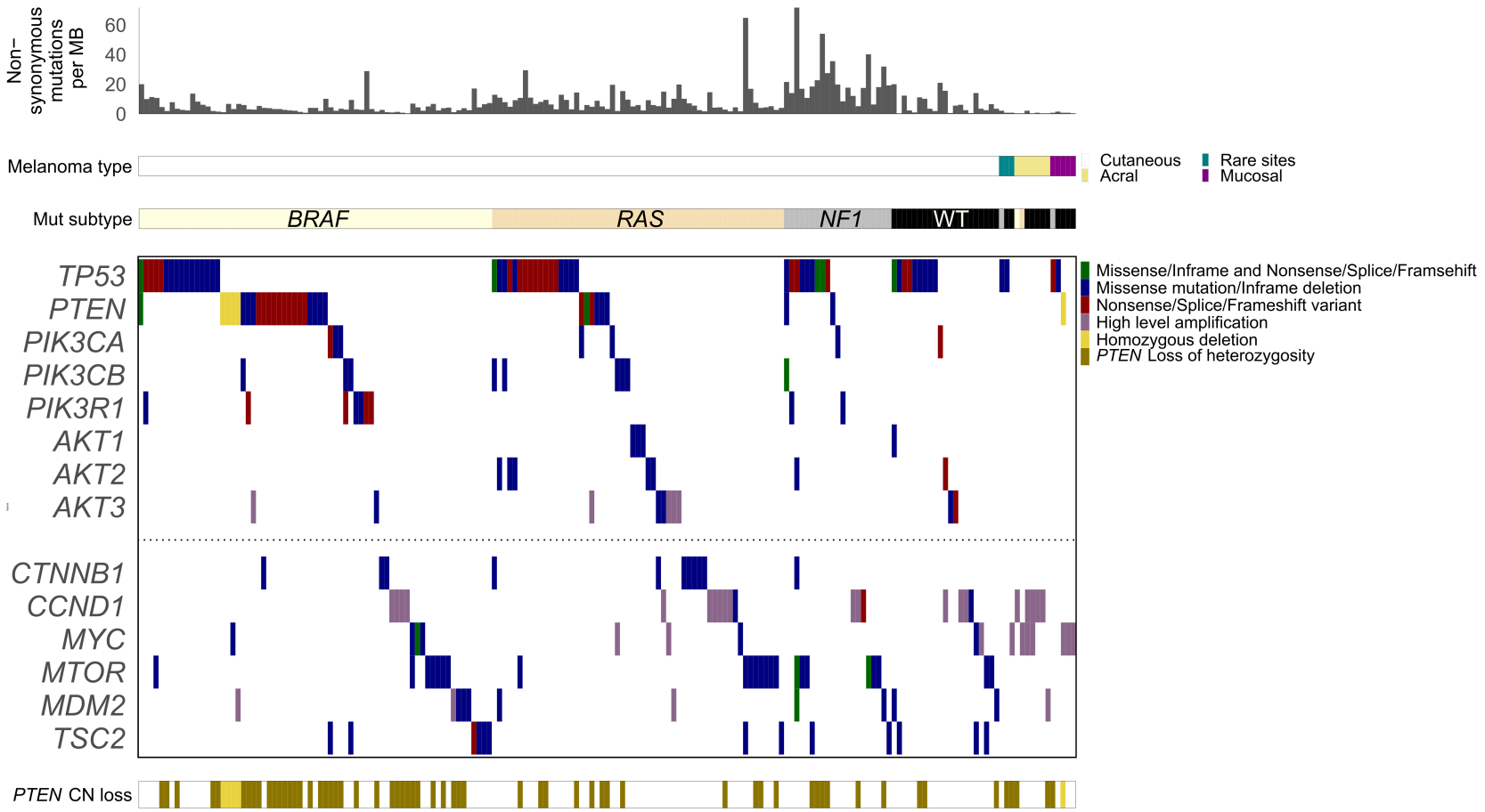


Figure 4.8: **Genetic alterations targeting selected important components of the PI3K/AKT pathway.** From top to bottom the panels show: mutation load (nonsynonymous mutation load per MB), melanoma type (cutaneous, acral, mucosal or other rare sites), mutational subtypes (*BRAF, NRAS, NF1* or *WT),* important genetic alterations in key genes of the PI3K/AKT pathway coloured by their respective variant consequence, and finally copy number loss (homozygous deletions and loss of heterozygosity) in *PTEN*.

in a dual dependent manner, it is not surprising to see alterations in these genes show a trend of mutual exclusivity. *PTEN* is also more frequently co-mutated with *BRAF* compared to *NRAS*, which has been observed previously [666, 667] and could reflect both the functional redundancy double mutation of *PTEN* and *NRAS* would confer in respect to activating the PI3K/AKT pathway, but also the pro-tumourigenic potential of dual activation of the MAPK and PI3K pathway. Earlier studies have shown that mice only develop aggressive melanoma after both *BRAF* activation and *PTEN* loss [668]. Using my cohort of primary melanomas, I could show that patients with mutations in both *PTEN* and *BRAF* had a lower survival compared to none or each mutation alone (Fig. 4.9).

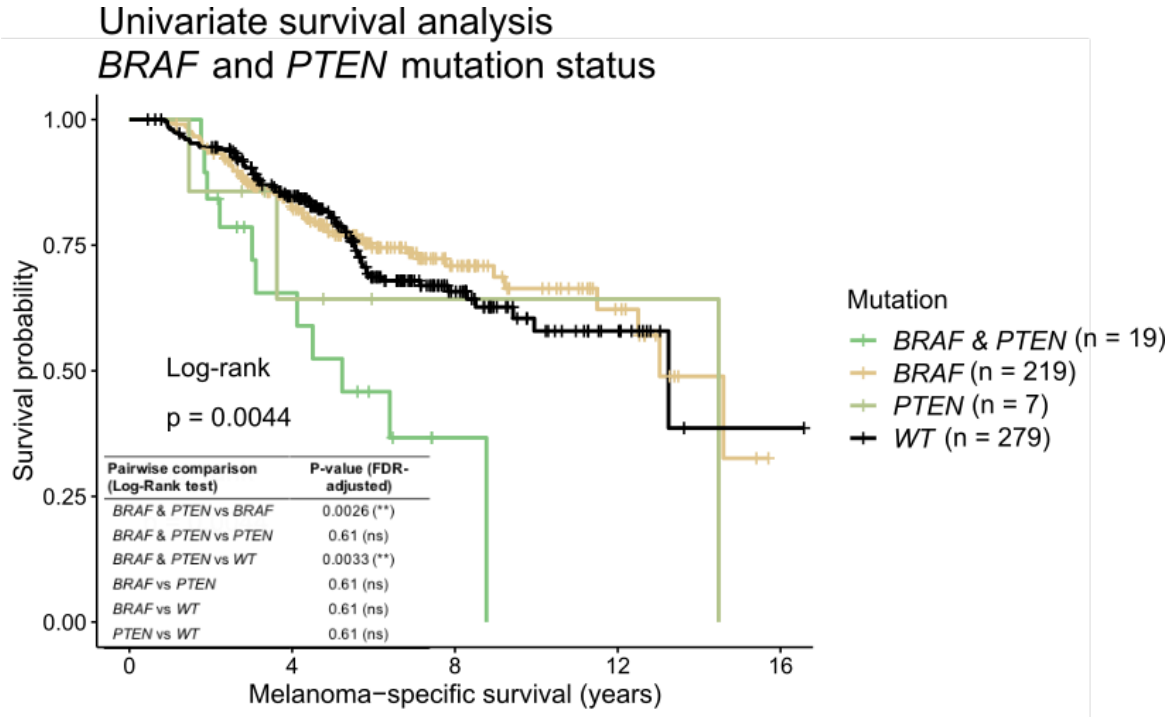


Figure 4.9: **Survival comparison between patients with mutations in *BRAF* and *PTEN*.** Kaplan-Meier survival curves showing melanoma-specific survival stratified by the presence of mutations in *BRAF* , *PTEN, BRAF* and *PTEN*, or neither gene. The global log-rank statistic is shown as well as the result from pairwise comparisons.

Looking at some effectors downstream of *AKT* (Fig. 4.8), frequent amplifications were found in *CCND1* and *MYC (*presented in Section 3.3.1), as well as four cases affecting *MDM2*. Mutations in these genes were rare, with *CCND1, MYC* and *MDM2* variants discovered in three, six and eight cases, respectively. Only one missense mutation in *MDM2* was recurrent (S304F, n = 2). *TSC2* was sporadically mutated in 13 samples, *mTOR* in 23 samples and *CTNNB1* mutations targeted 11 patients, with 2 recurrent changes (S45F n = 3

and S37F n = 2). The mutation pattern of these effectors, both in respect to each other but also to the upstream key regulators of pathway initiation, suggests activation of this pathway could be successful irrespective of the specific gene targeted. Therefore, it is also not surprising medical interventions trying to block enhanced signalling through this pathway have proven extremely challenging, in particular with regards to drug resistance [98, 669–671].

4.6 *CDKN2A*-associated regulatory pathways

CDKN2A is arguably the most important gene when discussing melanoma development, both in terms of familial melanoma [133, 134], but it is also one of the most frequently silenced genes in sporadic cases of melanoma [73, 672]. This tumour suppressor gene is also the top homozygously deleted gene in my cohort of primary melanomas. In addition to the 50 samples with complete loss of the whole gene, 42 samples showed co-mutation and loss of heterozygosity of the entire gene. A further 24 samples had a nonsynonymous mutation in the gene alone, and an additional 129 samples showed loss of heterozygosity alone across the entire *CDKN2A* region. This brings the total percentage of samples with *CDKN2A* alterations in my cohort to 61%. Several recurrent deleterious variants in this gene were found, including hotspot and known loss of function mutations P114L (n = 10), R58* (n = 12), R80* (n = 11).

CDKN2A encodes two proteins with distinct functions: p14ARF and p16INK4A. The p16INK4A protein is involved in the regulation of cell cycle progression and replicative senescence [137], through negative regulation of CDK4 and CDK6, thus preventing complex formation with Cyclin D (*CCND1*) and subsequent phosphorylation of Rb (*RB1*) [138]. Seven out of nine mutations in *CDK4* clustered to amino acid positions 22 and 24 (K22M n = 2, K22Q n = 2, R24L n = 2, R24S n = 1). Mutations at these positions have been reported to disrupt binding to the p16INK4A protein product of *CDKN2A* [673], and would therefore circumvent its inhibitory effect and enable cell cycle progression. An additional five samples had high level amplifications of the oncogene *CDK4*. *CDK6* alterations in melanoma were less common and its tumourigenic role has not been as well-studied as *CDK4* [73, 91, 674]. However, the gene was amplified in four samples and mutated in another two samples in my cohort, in a mutually exclusive manner to both *CDK4* and *RB1* alterations, which would speak in favour of its importance in tumour progression (Fig. 4.10). Amplifications of *CCND1* are often more frequent than mutations in this gene [73], and in my cohort it was the top amplified gene (n = 20) whereas only three samples were affected by mutations. 17 samples had alterations in *RB1,* with the majority being loss of function-associated

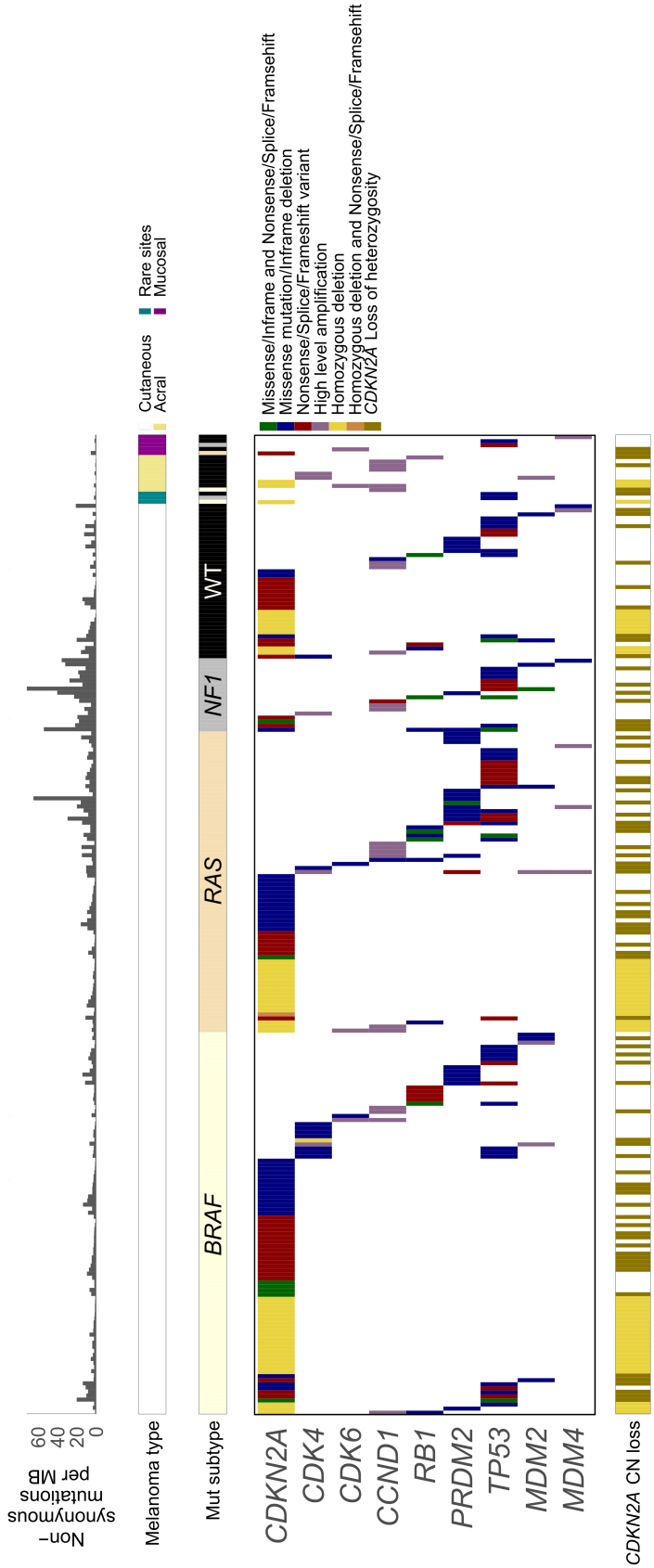


Figure 4.10: **Genetic alterations targeting selected important components of *CDKN2A*-associated regulatory pathways.** From top to bottom the panels show: mutation load (nonsynonymous mutation load per MB), melanoma type (cutaneous, acral, mucosal or other rare sites), mutational subtypes (*BRAF, NRAS, NF1* or *WT*), important genetic alterations in key genes of *CDKN2A*-associated pathways coloured by their respective variant consequence, and finally copy number loss (homozygous deletions and loss of heterozygosity) in CDKN2A.

consequences. In addition, *RB1* was never mutated or lost in acral or mucosal melanomas of my cohort. *PRDM2*, a gene discovered to be mutually exclusive to *CDKN2A (*discussed in Section 3.5), and which has a Rb-binding domain, might therefore play a role in *CDKN2A*-associated cell cycle regulation of melanoma cells. A mutual exclusivity pattern could be seen between all key members of the p16INK4A-mediated cell cycle regulatory pathway (Fig. 4.10). This pattern held true also for *PRDM2*, emphasising its potential contribution to tumour regulation through the *CDKN2A/CDK4/RB1* cell cycle regulatory pathway.

p14ARF, the alternative protein produced by *CDKN2A,* on the other hand has tumour-suppressive properties through its activation of *TP53* directly or through inhibition of *MDM2* [139–143]. These three key components of this pathway are commonly altered in various cancers, with a pattern of mutual exclusivity (Fig. 4.10). However, direct inactivation of *TP53* is less common in melanoma compared to other tumour types [73, 85, 91, 96, 675]. *MDM2* was amplified in four samples and mutated in an additional eight. *MDM4,* which also negatively regulates *TP53,* has been implicated in melanoma [676]; however, its overexpression is mainly achieved on a protein level, and would therefore not be picked up through mutational analysis. Nevertheless, two cases of missense mutations and five samples showing high level amplifications targeting *MDM4* were found amongst my primary melanoma samples, with all but one case showing mutual exclusivity with *TP53* and *MDM2*. *TP53* was mutated in 10% of my primary melanomas, with several known and recurrent loss of function consequences [653]. No samples showed biallelic copy number deletion of the entire gene; however, 98 samples showed heterozygous loss. As *CDKN2A* is frequently silenced in melanoma, this event alone might suffice to dampen the activity of *TP53* and thereby overcome tumour suppression, which might explain the low frequency of targeted loss of *TP53* itself. Moreover, when tumours progress, more alterations might be necessary to maintain tumour aggressiveness, which could explain why other researchers have found higher frequencies of *TP53* mutations in later stages of melanoma [48]. Several cases of double alteration of *CDKN2A* and *TP53* could be observed in *BRAF* mutant melanoma patients compared to only a few sporadic cases in most other subtypes of melanoma. This could be the result of specific *CDKN2A* mutations giving rise to loss of only the p16INK4A product, or the increased need to deactivate *TP53* compared to *RAS* mutant tumours, which could already modulate *TP53* though the PI3K/AKT pathway (discussed in Section 4.5). Additionally, in the absence of *CDKN2A* alterations, co-mutation of key members of both the *CDKN2A/CDK4/RB1* pathway and the *CDKN2A/MDM2/TP53* pathway could be beneficial for the melanoma, as several cases of same-patient *CDK4* and *TP53* alterations were found amongst the *BRAF* subtype melanomas. I also observed frequent co-amplification of

MDM2 with *CDK4* (3 out of 4 cases), another feature that might activate both *CDKN2A*-associated pathways while *CDKN2A* itself is still functional.

4.7 Additional driver genes and the interplay between key biological pathways

MYC and *MITF* are important genes encoding transcription factors which play diverse roles through the modulation of several biological pathways. *MYC* is tightly regulated by a multitude of transcription factors and signalling pathways including the MAPK and WNT/β-catenin pathway, and is itself a master regulator of a great number of target genes including *CDKN2A* and *TP53* [677, 678]. *MYC* is commonly overexpressed in melanoma and other tumour types [679], and it was amongst the top amplified genes in my dataset. *MYC* amplifications were more frequent in rarer types of melanoma (43% in mucosal, 13% in acral and <1% in non-acral cutaneous melanoma), consistent with other studies which have showed a higher frequency of *MYC* alterations in rarer subtypes [73, 680–682]. This suggests *MYC* might play a particularly important role in the progression of less common subtypes such as acral and mucosal melanoma.

MITF is an important melanocyte master regulator, controlling a range of cellular programs including melanocyte differentiation, proliferation, and survival [683]. Mutations in *MITF* are relatively rare in melanoma, but tumours tend to instead upregulate expression through amplification, transcriptional or epigenetic modifications [646, 684]. In addition, *MITF* expression has been used as a way to distinguish between different categories of melanoma, some which has prognostic value [73, 646, 685–687]. In my primary melanomas, I observed very few cases of *MITF* alterations (amplification n = 2, mutation n = 3), but this finding does not exclude possible other mechanisms of *MITF* overexpression in other tumours of my cohort. However, other researchers have reported a lower frequency of *MITF* alterations in primary tumours and melanocytic nevi [646] compared to metastatic lesions, and this observation is in line with *MITF* expression correlating with more advanced high-grade tumours [686, 687].

Genes involved in epigenetic regulation such as chromatin remodelling and histone modification have also been implicated in melanoma. These include *ARID2*, *EZH2*, *IDH1* and *DDX3X*. *ARID2* is mutated in 8% of my samples, and show a clear tumour suppressor mutation pattern with half of the variants being nonsense mutations. Both oncogenes *EZH2* and *IDH1* have recurrent hotspot variants (*EZH2* Y646N n = 2*, IDH1* R132C n = 13*)*. Of

the mutations observed in *DDX3X,* 83% were missense mutations.

PPP6C and *FAM58A* are both genes involved in cell cycle regulation and survival [517, 688], which show a positive selection mutation composition (discussed in Section 3.4). *PPP6C* is a negative regulator of *CCND1,* and mutations in *PPP6C* (n = 21) occurred only in samples which lacked detectable alterations of *CCND1*. Additionally, mutations in *PPP6C* were found exclusively in *BRAF, NRAS* or *NF1*-mutant samples of non-acral cutaneous origin. *FAM58A* on the other hand, was mutated in 12 patients, and was found in all mutational subtypes of non-acral cutaneous melanoma.

The emerging driver genes *RQCD1* and *MSR1* were altered in non-acral cutaneous melanoma exclusively. 89% of mutations in *RQCD1* were in the hotspot P131L. *FBXW7*, another driver gene identified in Section 3.4, is yet another broad-acting tumour suppressor gene, regulated by *TP53* [689] and which modulates several important oncogenes including *MYC, MTOR* and *MITF* [690–693]. Mutations in this gene were found exclusively in patients with non-acral cutaneous melanoma.

In addition to coding changes, melanomas harbour frequent hotspot mutations in promoter regions of some genes. *TERT,* an important driver gene both in sporadic and familial melanoma, can be activated by amplification, but is more often activated by mutations in specific regions of its promoter [148, 151]. *TERT* alterations were found in 202 samples, and comprised of 44% C250T, 38% C228T, and 7% CC242-243TT hotspot variants, and 7% high level amplifications. All hotspot mutations were completely mutually exclusive (Fig. 4.11). Three patients had co-occurring missense mutations and hotspot C228T variants, and a fourth patient harboured both a high level amplification and the C250T promoter variant. Interestingly, these four patients were all of the *NF1* mutant subtype, with a high mutation burden, which suggests these missense mutations were obtained at random and do not confer an additional advantage for the tumour cell.

Several melanoma driver genes have dual roles, and can alter more than one biological function. *NRAS* is upstream and activates both the MAPK and PI3K/AKT pathway, while *CDKN2A* modulates cell cycle progression and apoptosis by encoding two different proteins. *TP53* can be activated through a range of events, including regulation through *PTEN* and the PI3K/AKT pathway or by *CDKN2A* and *MDM2*-dependent mechanisms. Furthermore, *TP53* is also activated by cellular stress signals, including DNA-damage response genes *ATM* and *CHEK2* or hypoxia-induced mechanisms involving *ATR* and *VHL* [146]. These pathways of *TP53*-associated regulation are not considered to drive melanoma progression; however, they play a major role in some other cancers [694–696]. No coding changes or copy number alterations in *VHL* were found; however, few cases of mutations

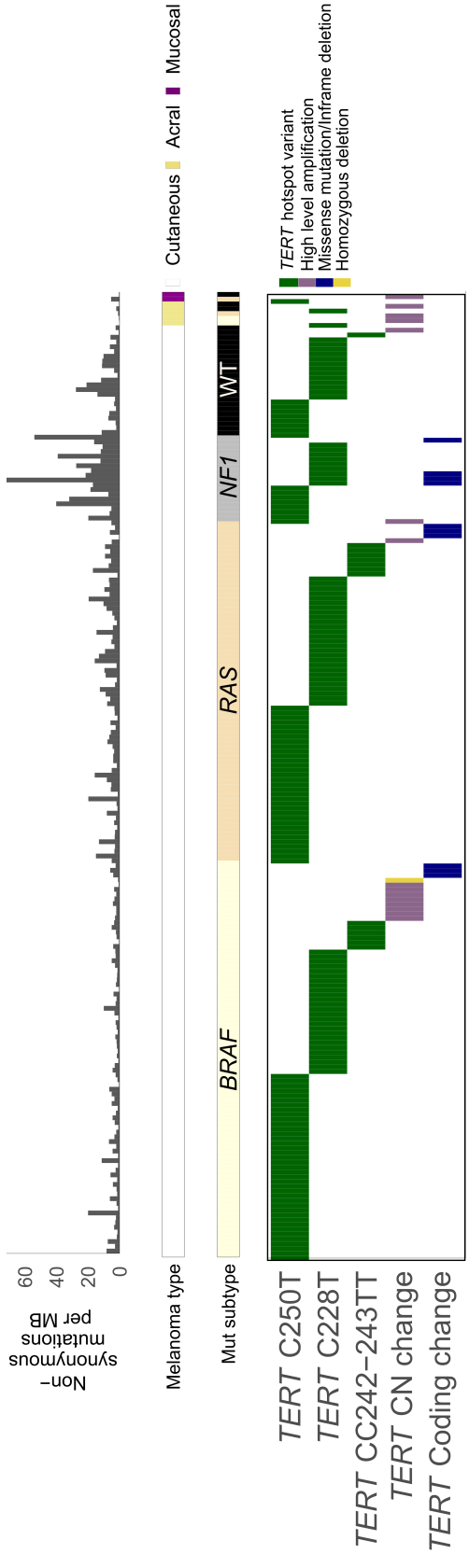


Figure 4.11: **Genetic alterations in the *TERT* gene.** From top to bottom the panels show: mutation load (nonsynonymous mutation load per MB), melanoma type (cutaneous, acral or mucosal), mutational subtypes (*BRAF, NRAS, NF1* or *WT*), and important genetic alterations in *TERT* coloured by their respective variant consequence.

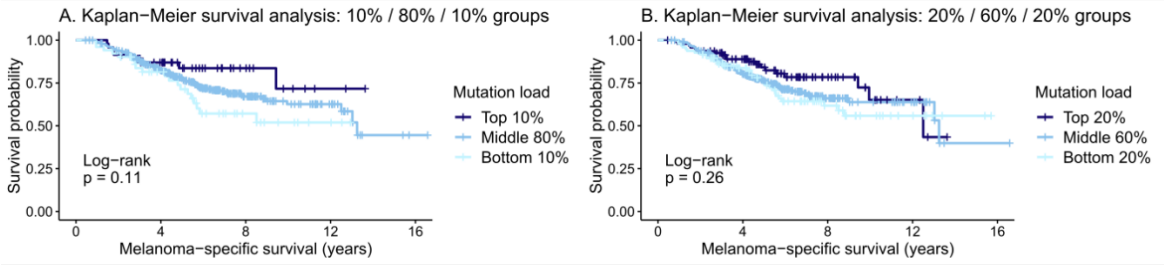
were seen in *ATM* (n = 23, plus 1 amplification), *CHEK2* (n = 12) and *ATR* (n = 24)[677] in my cohort. A proportion of these alterations were found in patients which did not show inactivation of *TP53* through the more commonly altered genes *PTEN* and *CDKN2A,* which could suggest modulation of *TP53*-activity in some patients might be dependent on alternative pathways such as through *ATM/CHEK2* or *ATR/VHL* (Fig. A.12).

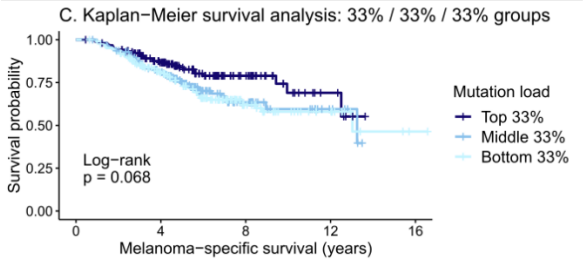
4.8 Immunological impact of genetic alterations

With the exciting development of immune checkpoint inhibitors, and their success in treating a subset of melanoma patients, comes the endeavour to further understand and utilise the immune system to combat cancer. Immunotherapies have had the highest success rates in tumour types such as melanoma, non-small cell lung cancer and bladder cancer, which share features such as a high mutation load [82]. Subsequently, tumour mutation load has been used in immunotherapy decision-making as it has been shown to have predictive value [220, 263, 264, 697–699]. Yet, the more mutations a tumour acquires should give the immune system more chances to discover and eliminate the tumour. Therefore, I aimed to investigate whether such an intrinsic capability exist, and whether particular alterations could disable or weaken the host immune response.

First, I tested whether mutation load would be prognostic in my cohort of primary melanoma patients. 507 patients were immunotherapy-naive (the 17 patients who had received immunotherapy were excluded from the analysis). Therefore, this analysis should reflect the baseline host capability to recognise tumour antigens. Although a trend could be seen towards better survival in the patient group which showed the highest mutation load (Fig. 4.12), this finding was not statistically significant. My hypothesis is therefore, that the mechanisms the tumour employs largely disables the immune system, to an extent where this suppressed state is incapable of tumour eradication irrespective of mutation burden. This state can however be reversed upon administration of checkpoint inhibitors, which would explain the benefits seen in immunotherapy-treated patients with a high mutation load.

In order for the host immune system to recognise a tumour as foreign, the acquired mutations need to give rise to mutant peptides to be presented on the tumour cell surface. As not all mutations would give rise to such neoantigens, my next step was therefore to study the relationship between neoantigen burden and patient survival. This analysis was done in two steps. First the patient HLA types were computed against the IMGT database v.3.12 using HLAssign v.3.21 [433]. Then, the pVAC-Seq pipeline (Table A.1) with NetMHCpan v.2.8 was used to predict the MHC class I-associated neoantigen load [700]. As expected,





D. Cox proportional hazards regression analysis

	Multivariate analysis		
Variable	**HR**	**95% CI**	**P-value**
Nonsynonymous mutation load per MB	0.97	0.94-1.01	0.091
Age at diagnosis	1.03	1.01-1.05	0.003
Sex, male	1.79	1.23-2.62	0.002
Ulceration, present	1.79	1.16-2.76	0.009
Breslow thickness	1.11	1.04-1.19	0.002
Stage			
II vs I	0.97	0.54-1.73	0.905
III vs I	1.98	1.06-3.70	0.033

Figure 4.12: **The effect of mutational load on patient survival in an immunotherapy-naive cohort of 507 patients. A)** Kaplan-Meier survival analysis stratified by mutation load grouped into top 10%, middle 80% or bottom 10%. **B)** Kaplan-Meier survival analysis stratified by mutation load grouped into top 20%, middle 60% or bottom 20%. **C)** Kaplan-Meier survival analysis stratified by mutation load grouped into tertiles. **D)** Multivariate survival analysis using Cox proportional hazards regression analysis, looking at mutational load as a continuous variable together with known contributors to melanoma survival as additional covariates.

there was a high degree of correlation between mutation load and predicted neoantigen load (Fig. A.13). Similarly to mutation load, neoantigen load also failed to explain a significant difference in survival between patients. Additionally, tumour aneuploidy have also been shown to correlate with immunotherapy treatment response, although this relationship is inverse [701]. I therefore sought to test this association in my treatment-naive cohort. For each sample, I calculated the percentage of the captured genomic length comprising gains and losses, as defined by segments showing a log2 fold change $\geq |0.1|$ [702]. The tumour copy number burden estimate was not prognostic in my cohort (Fig. A.14), and this measure did not correlate with mutational load (Spearman correlation coefficient rho = -0.06, p-value = 0.24), an observation in agreement with other reports [701]. To conclude, none of the factors mutation load, neoantigen load or copy number burden had a prognostic impact on survival in my cohort of immunotherapy-naive patients.

Tumour mutation load and neoantigen load are measures of alterations by which the tumour risks being detected by the immune system. Yet, tumours will only be vulnerable towards destruction by the immune system provided the host defence is functional. One such

measure, which can also be used to predict immunotherapy response, is the cytolytic score [703], defined as the geometric mean of *Granzyme A (GZMA)* and *Perforin 1 (PRF1)* gene expression. The expression of these genes are linked to the anti-cancer cytotoxic activity of Cytotoxic T Lymphocytes (CTLs, also known as CD8 T cells) and Natural Killer (NK) cells, as a higher cytotoxic activity is linked to more potent tumour eradication [704, 705]. As this measure would indicate whether the patient's immune system is active, I thought it might be prognostic either alone or in combination with mutation load even in patients not receiving immune-stimulating intervention. Therefore, I decided to study this variable in my cohort. After excluding 7 patients which had received immunotherapy, analysis was possible for 311 patients for which both mutation data and gene expression data were available. The cytolytic score did not appear to better explain a difference in survival compared to tumour mutation load alone. Interestingly, the 311 patient subset of the original 507 appeared to show a more visible separation of the survival curves whereby tumour mutation burden was significantly prognostic in the 311 patient subgroup. This is likely due to tumour-specific biases such as larger tumours yielding sufficient input material for both genomic and transcriptomic sequencing; however, no major differences were found between these two subsets in the main clinical variables (Fig. A.15). Cytolytic score, categorised into high, medium and low, did not show a significant survival difference in a univariate model; however, multivariate survival analysis with cytolytic score as a continuous variable showed a significant association with survival (Fig. 4.13). Interestingly, grouping by both mutation load and cytolytic score showed more distinct differences in the respective survival curves, in particular for the patients with either very high or very low mutation load. This effect was reduced by using a broader tertile classification of cytolytic score, which suggests this observation is limited to few extreme cases on both ends of the mutation load spectrum. Therefore, I looked into these patients and found the low mutation and cytolytic score group to be dominated by triple wild-type (78%), including 36% acral melanomas, which are known to be associated with a low mutation load and poor survival [66]. Therefore, the poor survival seen in the low mutation and cytolytic score group is likely influenced by these factors already known to be important contributors to patient outcome. Conversely, the high mutation and cytolytic score group were all non-acral cutaneous melanomas, mostly from body sites with high UV exposure, and were all found to lack copy number deletions in key tumour suppressor genes *CDKN2A, TP53* and *PTEN*. To further investigate genetic alterations associated with the cytolytic score, I performed genetic association tests for driver mutations or key copy number alterations. Interestingly, *CDKN2A* mutations were associated with a high cytolytic score (logistic regression p-value = 9.93 x 10^{-5}, OR = 1.55), whereas *CDKN2A* deletions

were linked to lower cytolytic score (logistic regression p-value = 0.011, OR = 0.74). Weak correlations were also found between lower cytolytic scores and mutations in *MSR1* (logistic regression p-value = 0.023, OR = 0.76) or *RPS27* C227T (logistic regression p-value = 0.038, OR = 0.62). Finally, activation of *TERT* through hotspot promoter mutation or amplification were highly associated with a lower cytolytic score (univariate logistic regression p-value = 4.47 x 10^{-5}, OR = 0.77).

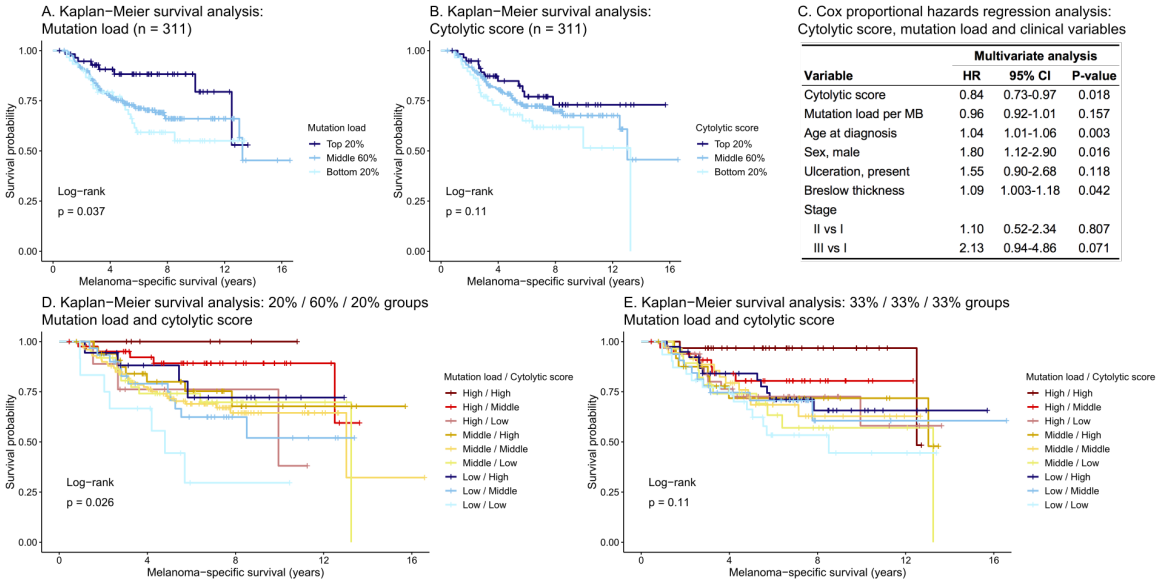


Figure 4.13: **The impact of cytolytic score on patient survival in an immunotherapy-naive cohort of 311 patients. A)** Kaplan-Meier survival analysis stratified by mutation load grouped into top 20%, middle 60% or bottom 20%. **B)** Kaplan-Meier survival analysis stratified by cytolytic score grouped into top 20%, middle 60% or bottom 20%. **C)** Multivariate survival analysis using Cox proportional hazards regression analysis, looking at mutational load (continuous variable) and cytolytic score (continuous variable) together with known contributors to melanoma survival as additional covariates. **D)** Kaplan-Meier survival analysis stratified by both mutation load and cytolytic score, each grouped into high (top 20%), middle (middle 60%) or low (bottom 20%). **E)** Kaplan-Meier survival analysis stratified by both mutation load and cytolytic score, each grouped into tertiles defining high, middle and low subgroups.

Furthermore, I wanted to also test whether mutations in immune-associated genes could explain specific patient survival differences. Very few patients had alterations affecting *B2M* (deletion n = 3, missense mutation n = 1, amplification n = 2), *CD274* (nonsense mutation n = 1) and *CTLA-4 (*no samples with alterations*).* Biallelic loss of *B2M* was associated with poor survival (Fig. A.16); however, loss of only one allele did not show any difference. *B2M* plays an essential role in antigen presentation and the anti-tumour immune response [706],

which could explain why such alterations would have a negative effect on patient survival in my cohort but also why other studies have linked *B2M*-deficiency to immunotherapy resistance [157, 707]. However, the data from my cohort on *B2M* is based on too few observations for any final conclusions to be drawn.

Tumour loss of IFN-γ pathway genes mediates resistance towards anti-CTLA-4 immunotherapy [543]. I therefore asked, whether this mechanism would impair intrinsic host immune responses already prior to immunotherapy, which could consequently serve as a marker to highlight tumours that have escaped immune regulation. Neither mutation nor copy number status of IFN-γ pathway genes made a significant difference towards patient survival in my treatment-naive cohort (Section B.3). I further investigated whether stratifying by mutation load or cytolytic score might make a difference, as I hypothesised that IFN-γ pathway defects should hinder IFN-γ-mediated upregulation of PD-L1, subsequently making tumours less protected against an immune response. A high mutation load or cytolytic score should therefore in theory be associated with an active immune environment where tumours with flawed IFN-γ pathway signalling might be more prone to elimination. Alterations of any kind did not show a prognostic value in my cohort (Section B.3). As IFN-γ also suppresses tumour growth and triggers apoptosis [708, 709], the relationship between IFN-γ defects, immune regulation and patient survival might be far more complex than could be explained by my analysis.

4.9 Evaluation of chapter aims

- Analyse the different genetic changes across melanoma subtypes

  - The simplest established molecular subtypes: *BRAF, RAS* and *NF1* account for 73% of all primary melanomas in this cohort, whilst the remaining 27% were grouped together as triple wild-type. Heterogeneity within these classes of tumours is present. Based on the mutational composition in each tumour, six biological pathway-defined Sambar subtypes could be defined. 1. The black group showed limited mutation-driven events. 2. The blue group showed an enrichment in PI3K pathway-associated mutations. 3. The cyan group was dominated by pathways associated with cell adhesion, migration and trafficking. 4. The magenta group had an overrepresentation in SHC-mediated events and B cell signalling. 5. The red and green groups did not yield any meaningful enrichments, potentially reflecting further diversity amongst these groups.

  - Cutaneous (non-acral) melanomas are distinct from non-cutaneous melanomas in the mutation load and the composition of genetic alterations. Cutaneous melanomas have the highest mutation load, with *BRAF* mutations being the most common genetic alteration. Non-conjunctival mucosal melanomas completely lacked *BRAF* mutations and showed a different driver gene composition (recurrently mutated genes include *LZTR1, TP53* and *ATRX*) compared to cutaneous melanomas. These tumours were also targeted by frequent copy number alterations including high level amplification of chromosome 8q (where *MYC* is located). Acral melanomas had the lowest mutation load, with a pattern of mutations showing shared similarities with both the cutaneous (e.g. hotspot *BRAF* and *NRAS* mutations) and mucosal (e.g. copy number alterations including *MYC* amplification) melanomas.
  - The most frequently altered genes were: *BRAF, NRAS, TERT* and *CDKN2A* in cutaneous melanoma, *BRAF, TTN*, *CCND1* and *TERT* in acral melanoma and *MYC, TTN, TERT* and *MAP2K1* in mucosal melanoma.

- Investigate the main genetic alterations and signalling pathways altered in different tumours

  - MAPK pathway activation is the most common genetic alteration in primary melanomas, with 78% of tumours harbouring alterations in the main pathway components. Hotspot mutations in *BRAF* (39%) and *NRAS* (28%) were most common, followed by *NF1* mutations (10%). Mutations in *HRAS, KRAS, MAP2K1, MAP2K2* and *KIT* accounted for the remaining 5% of MAPK-associated alterations. High level amplifications in any of these MAPK-associated genes affected in total 5% of all copy-number analysed tumours.
  - The PI3K/AKT pathway is another important process frequently targeted by genetic alterations in melanoma. Modulation of *PTEN* is the most common event, where 38% of tumours showed mutation or copy number loss. Additionally, a small number of tumours harboured mutations in *PI3K* and *AKT* genes (5.2%). As *RAS* genes also regulate the PI3K/AKT pathway, this brings the total genetic alterations affecting the PI3K/AKT pathway to 66%.
  - 61% of all primary melanomas showed alterations in *CDKN2A,* making it the single most frequently modulated gene in primary melanoma. The regulatory pathway linked to p16INK4A activity is modulated in an additional 5% of tumours through alterations in *CDK4, CDK6* and *CCND1*, while *RB1* is mutated

or show copy number loss in a another 5%. The p14ARF-associated pathway is further targeted through *TP53* mutation or copy number loss in another 10%, followed by 1.2% with alterations in *MDM2* and *MDM4*. The p16INK4A and p14ARF pathways are therefore mainly altered in 70% and 73% of tumours, respectively. Collectively, any of the *CDKN2A*-associated regulatory pathways are altered in 78% of all primary melanomas.

  - *TERT* is one of the most important melanoma genes commonly targeted through non-coding mutations. 36% of samples had hotspot promoter mutations, 3.5% high level amplifications and another 1.7% coding mutations in this gene, resulting in a total of 42% of samples with *TERT* alterations.

- Assess the impact of genetic alterations on the immune response in treatment-naive patients

  - The burden of genetic alterations including mutation load, neoantigen load and copy number burden are associated with immunotherapy response, but did not have a prognostic impact in my cohort of melanoma patients not receiving immunotherapy. I therefore hypothesise that the suppressive immune microenvironment in tumours remain dysfunctional irrespective of its genetic composition, to a certain extent. Only upon administration of immune-activating agents can a prognostic difference be seen between patients with varying degrees of genetic alterations.
  - Stratification by immune cell activity mediated by the cytolytic score, was mildly significant in a multivariate model. Interestingly, categorisation by cytolytic score showed a larger effect on survival with varying mutation load than using mutation load alone. This observation is in line with my hypothesis, where mutation load will be prognostic only in a state where the immune system is not fully inhibited.
  - *TERT*-activating genetic alterations were associated with a lower cytolytic score. Weak correlations were also found between tumours with lower cytolytic score and *CDKN2A* deletions, *MSR1* mutations or *RPS27* C227T variants. In contrast, *CDKN2A* mutations were linked to a high cytolytic score.
  - Genetic alterations targeting the IFN-γ pathway have been shown to mediate resistance to immunotherapy; however, no prognostic value was found in my treatment-naive cohort, either alone or in combination with mutation load or

cytolytic score.

- Evaluate the possible prognostic potential of genetic alterations
  - Neither the established mutational subtypes (*BRAF, RAS, NF1,* WT) nor the Sambar pathway-level subtypes could provide prognostic value.
  - Mutations targeting components of the TCR signalling pathway was associated with worse survival.
  - Patients with mutations in both *BRAF* and *PTEN* had a survival disadvantage compared to patients with none or one of these mutations alone.
  - Mutation load, neoantigen load, copy number burden or alterations of the IFN-γ pathway are not prognostic in immunotherapy-naive patients.
  - Cytolytic score showed a weak prognostic effect, with a higher cytolytic score being associated with better survival.

Part II

Regulators of PD-L1 tumour expression

The overarching goal of Part II is to understand how PD-L1 expression on tumours can be downregulated in order to reverse immune cell exhaustion and facilitate tumour cell killing. Immunotherapy targeting the PD-1/PD-L1 axis is currently at the forefront of melanoma therapies, proving the power of inhibiting this signal to trigger tumour eradication. However, antibody-based therapeutics come with ckeaveats including low bioavailability, high costs and specific resistance mechanisms involving antibody internalisation and ectodomain shedding. Additionally, at the start of my PhD, the cellular regulation of PD-L1 expression had not been adequately studied. Through understanding the pathways and mechanisms involved in limiting cell-surface PD-L1 expression, it is possible to investigate if small molecule druggable targets interfering with these processes could be identified. Coupling this research question to the powerful approach of whole genome CRISPR-Cas9 screening, I am hoping to identify key processes controlling PD-L1 regulation, with the ultimate goal to identify novel druggable targets.

Chapter 5

Design and application of a CRISPR-Cas9 screen to identify regulators of PD-L1

5.1 Introduction

Melanoma and other cancer cells can upregulate cell surface PD-L1 expression to avoid elimination by the host immune system. Such a dysfunctional state can be induced through processes including genetic alterations, oncogenic signalling or a change in the tumour microenvironment [226, 237, 248, 273]. Inhibiting the interaction between PD-1 and PD-L1, through blocking antibodies or by reducing the expression levels of these components, would therefore reverse immune cell exhaustion and facilitate tumour cell killing. Tumour PD-L1 expression can be controlled on several levels, which include modulation of gene expression, targeting protein stabilisation or degradation and by extrinsic or epigenetic regulation. Some of these processes can be explored using a phenotypic marker-based CRISPR-Cas9 screening approach whereby the effect of specific gene knock-outs on PD-L1 expression can be studied using a fluorescently labelled antibody against PD-L1. By understanding the biological processes PD-L1 undergoes in order to be presented on the cell surface, we may not only learn more about PD-L1 biology, but we could potentially identify novel druggable targets.

In this chapter, I will outline the experimental design, rationale and set up of controls that provide the foundation for a successful whole genome CRISPR-Cas9 screen. After establishing the optimal conditions to perform a screen aimed at identifying regulators of

PD-L1 expression, I continue by presenting the screening experiment, quality control and results. Finally, I will discuss the discoveries from the screen, the follow-up validation experiments and how this information could be used to understand and target the modulation of cell surface PD-L1 levels.

5.1.1 Chapter aims

The aim of Chapter 5 is to ensure an optimal experimental design and discuss the results obtained, including:

- Selection of experimental conditions best suited for my project
- Confirm adequate controls are set up and their use validated
- Verify that the screening conditions are appropriate for hit discovery
- Identify the various intracellular pathways by which PD-L1 is presented on the cell surface
- Assess the novel hits and their potential druggability
- Present top candidate hits for follow-up validation

5.2 Methods

5.2.1 Cell culture

All cell lines are adherent and express Cas9. C092, SKMEL25, HCC44 and LCLC103H were cultured in RPMI-1640 supplemented with 10% FBS and 1% Penicillin-Streptomycin. 5637 was cultured in RPMI-1640 supplemented with 10% FBS, 1% Sodium pyruvate and 1% Penicillin-Streptomycin. UKEMEL118C was cultured in DMEM supplemented with 10% FBS and 1% Penicillin-Streptomycin, and 647V in DMEM/F12 supplemented with 10% FBS and 1% Penicillin-Streptomycin. UBLC1 was cultured in 1/3 RPMI-1640 and 2/3 DMEM media, supplemented with 10% FBS and 1% Penicillin-Streptomycin. All cell lines were maintained at 37°C in 5% CO_2.

5.2.2 Generation of Cas9-expressing cell lines using lentiviral transduction

The parental cell line was transduced in suspension with Cas9 virus (prepared from the Addgene pKLV2-EF1a-Cas9Bsd-W plasmid) and 8μg/ml Polybrene (Merck). Cells were seeded in a T75 culture flask, and incubated at 37°C. The virus was removed through a media change after 20 hours. Blasticidin was added 3 days post infection and the cells grown in Blasticidin-supplemented media for an additional 10 days, until the cells stably express Cas9.

Cas9 efficiency was assessed using lentivirus carrying reporter vectors BFP and GFP (Addgene, pKLV2-U6gRNA5(Empty)-PGKBFP2AGFP-W) or BFP, GFP and a gRNA targeting GFP (Addgene, pKLV2-U6gRNA5(gGFP)-PGKBFP2AGFP-W). Cas9-expressing cells were transduced in suspension with the lentivirus carrying either reporter vector, and incubated overnight. The virus was removed the following day through a media change, and the cells grown for another two days. Cells were analysed using a flow cytometer (BD Fortessa II), followed by data analysis using FlowJo v.10. Cas9 efficiency was calculated as the ratio between the BFP/GFP positive cells using the empty vector and the BFP^+/GFP^- population using the BFP/GFP/gGFP vector. A minimum Cas9 efficiency of 90% was required for all cell lines generated.

5.2.3 Generation of PD-L1 and OR14A16 knock-out control cell lines

gRNAs against *CD274* (PD-L1) and *OR14A16* were designed by Dr. Martin Del Castillo Velasco-Herrera (Section C.1) according to the guidelines described in [710]. Complementary oligos for each gRNA sequence (PD-L1: GGCTGCACTAATTGTCTATT and OR14A16: CAAAGAATTGGCGATAGATT) were purchased (Merck, Easy Oligos) and ligated into the pKLV2-U6gRNA5(BbsI)-PGKpuroBFP-W vector following the protocol from [411]. Proper incorporation of each gRNA sequence into the plasmids were confirmed using Sanger sequencing (Eurofins).

HEK293T cells were used for lentivirus production. Low passage HEK293T cells were seeded in 10cm round dishes one day prior to transfection to allow cells to reach 80% confluency. 2 separate tubes were prepared: the first containing 40μl Lipofectamine 3000 (ThermoFisher) mixed with 1ml Opti-MEM. In the second tube 35μl P3000 and 1ml Opti-MEM were mixed with plasmids: 3.74μg pMD2.G, 5.6μg psPAX and 7.5μg of the transfer plasmid. Both tubes were vortexed, allowed to stand for 5 minutes at room temperature, followed by the addition of the first mixture to the second and a further 20 minute incuba-

tion at room temperature. Immediately prior to transfection, media was removed from the HEK293T cells and replaced with 3.5ml pre-warmed Opti-MEM, followed by addition of the transfection mixture. Cells were incubated at 37°C approximately 15 hours, after which the media was replaced with 8ml pre-warmed IMDM. 48 hours after transfection, virus was harvested and filtered through a low protein-binding filter.

The parental cell line was transduced in suspension according to the protocol in Section 5.2.2, except Puromycin was used instead of Blasticidin and selection was complete after 7 days. Validation of knock-outs were performed using the Surveyor mismatch cleavage assay according to manufacturer's instructions (IDT). Primers were designed to yield a 561 bp product cut into 200 bp and 361 bp fragments using DNA from the OR14A16 KO cell line and a 610 bp product cut into 426 bp and 184 bp fragments from the PD-L1 KO cell line.

5.2.4 Titration of lentivirus to achieve optimal multiplicity of infection (MOI)

Cells were infected in suspension with varying volumes of virus stock and 8μg/ml Polybrene (Merck). Cells were seeded at 250,000 cells/well in 6-well plates and incubated at 37°C for 24 hours. The virus were then removed through a media change and incubated for a further 48 hours. Cells were harvested, fixed with 4% paraformaldehyde and analysed using a flow cytometer (BD Fortessa II), followed by data analysis using FlowJo v.10. The percentage of cells expressing BFP was analysed, and the virus volume corresponding to a BFP percentage of 30% calculated (Fig. A.19).

5.2.5 CRISPR-Cas9 screen in C092 cells

60 million Cas9-transduced C092 cells were infected in triplicates with the genome-wide Yusa human v.1.1 lentiviral gRNA library at a MOI close to 0.3 (Fig. A.18). Successfully infected cells were selected using Puromycin, and expanded for a total of two weeks. A library representation of at least 200x were maintained at all times. 200 million cells were harvested on day 14 and day 15 respectively, due to logistic reasons. Half of the cells were kept as the unsorted control population, while the other half were stained with APC-conjugated anti-PD-L1 antibody (eBioscience, clone MIH1) and e780 fixable viability dye (eBioscience, #65-0865-14). A FACS two-way sort was performed, capturing BFP^+, live and the lowest 1% (low) or 1-6% (dim) of PD-L1-expressing clones. The sorts were performed on a MoFlo XDP (Beckman Coulter) cell sorter with help from the Sanger cytometry

core facility. Sample processing varied between control and sorted samples due to different cell numbers and library coverage. DNA was extracted according to manufacturer's instructions using the QIAamp DNA mini kit (Qiagen) for the sorted samples and the Qiagen Blood & Cell Culture DNA Maxi Kit (Qiagen) for the control samples. DNA concentrations were quantified using Qubit, after which sorted samples were split into 200ng aliquots for PCR amplification. DNA from control samples were split into 36 x 2μg reactions. Final PCR products were purified with beads (Beckman Coulter, Agencourt AMPure XP), multiplexed at equimolarity and run using a custom sequencing primer on an Illumina HiSeq 2500 machine using 19 bp Single End (SE) reads, yielding an average sequencing depth of 36.6 million reads per sample. Samples were processed from second round of PCR through to sequencing by the CGP and sequencing support facilities at Sanger. The whole screen was performed once.

5.2.6 Small-scale validation of selected genes

Complementary oligos for the best-performing gRNAs from my screen were selected and purchased (Merck, Easy Oligos), see Section B.4 for detailed information. Stable knockouts of each gene were made following the protocol outlined in Section 5.2.3. Cell surface PD-L1 expression was assessed post viral transduction, by staining with APC-conjugated anti-PD-L1 antibody (eBioscience, clone MIH1) and e780 fixable viability dye (eBioscience, #65-0865-14). Cells were analysed using a flow cytometer (BD Fortessa II), followed by data analysis using FlowJo v.10. Validation experiments were performed once in C092 and twice in LCLC103H (Once at day 9 and 14 by me and once at day 14 with help from Dr. Victoria Harle, see Section C.1).

5.2.7 Pooled validation using a custom CRISPR-Cas9 screening library

A custom CRISPR-Cas9 gRNA library (Section B.4) comprising 1000 gRNAs was designed as follows:

1. 60 genes from screen (low or dim hits, FDR < 10%) at 8 gRNAs per gene
2. 36 additional PD-L1-associated genes from literature review at 8 gRNAs per gene
3. 4 safe targeting guides per chromosome, excluding the Y chromosome
4. Control gene *OR14A16* not present in the original library, at 8 guides per gene

5. 17 cancer drivers or top mutated genes from literature review at 8 gRNAs per gene

The gRNA sequences were selected from three commonly used genome-wide libraries [399, 405, 411, 711], and where the sequence did not start with a G, this was added to the 5' end of the gRNA sequence as this has been shown to improve efficacy when the gRNA is transcribed from the human U6 promoter [712]. Forward oligos for each gRNA, with 20 nucleotides on each flanking side matching the empty vector backbone, were purchased as an oligo pool (TWIST Bioscience). The assembly and production of the gRNA library were carried out by Dr. Victoria Harle (Section C.1). In brief, the oligos were PCR amplified using forward TCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAA primers. The empty vector backbone was cut using BbsI restriction enzymes. The gRNAs were then inserted into the digested backbone by Gibson assembly (NEB, Gibson Assembly Master Mix), followed by electroporation of electrocompetent cells (Lucigen Endura), culture of bacteria and plasmid DNA extraction (Qiagen). The plasmid DNA library was sequenced on an Illumina MiSeq machine using 21 bp SE reads. Lentiviral production was performed using HEK293T cells as described in Section 5.2.3, and where Cas9-lines were not already established (UKEMEL118C, 5637, 647V and UBLC1) this was done as described in Section 5.2.2. Lentiviral production, Cas9-transductions for SKMEL25, UKEMEL118C, 5637, 647V and UBLC1 cell lines, antibiotic and library titration were done by Dr. Victoria Harle (Section C.1). The screen was performed with eight technical replicates, where 4.5 million cells of each replicate were transduced with the validation library at a MOI close to 0.3 (Section B.5). Successfully infected cells were selected using Puromycin, and expanded for a total of 9, 14 or 28 days. A minimum library representation of 3000x were maintained at all subculturing steps. Each sorting day, 3-5 million cells were harvested as control, while 7-20 million cells were stained with APC-conjugated anti-PD-L1 antibody (eBioscience, clone MIH1) and e780 fixable viability dye (eBioscience, #65-0865-14), and subjected to FACS. The number of sorted cells differed between conditions due to technical difficulties. A two-way sort was performed, where the wild-type (uninfected) cells were used to set the gates comprising the lowest 1% (low) and 1-6% (dim) PD-L1-expressing clones. These gates were then maintained for the screening replicates, collecting the cells falling within the respective gates. The validation screen was executed together with Dr. Victoria Harle. The sorts were performed on a MoFlo XDP (Beckman Coulter) cell sorter by the Sanger cytometry core facility. The final 576 samples were then subjected to DNA extraction, PCR amplification, purification and sequencing which were performed by the CGP and sequencing support facilities at Sanger. The whole validation screen was performed once.

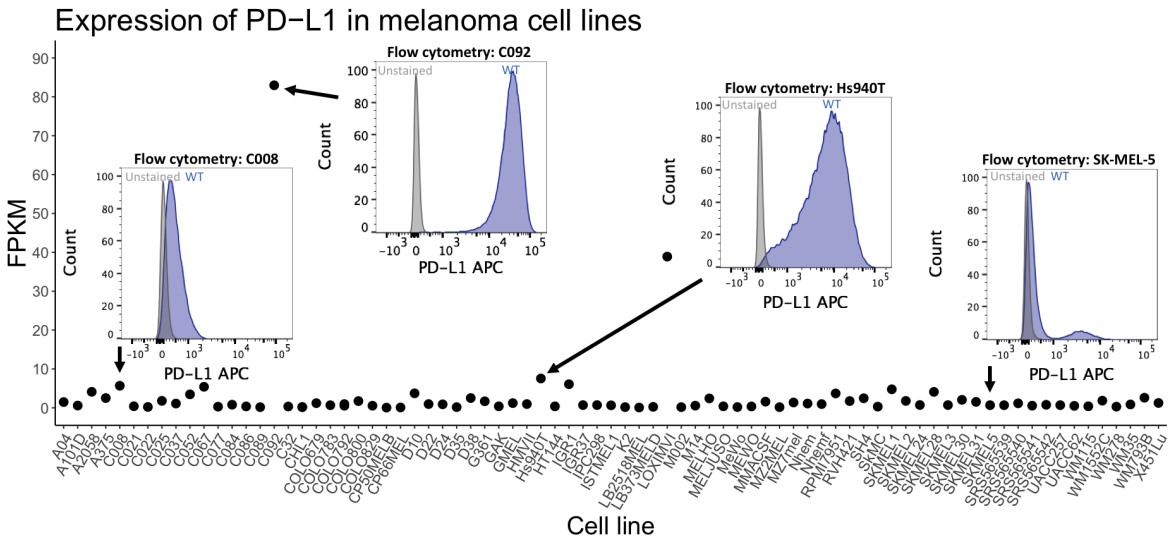


Figure 5.1: **PD-L1 expression across melanoma cell lines.** Scatter plot showing *CD274* (PD-L1) gene expression in fragments per kilobase of transcript per million mapped reads (FPKM) across melanoma cell lines, overlaid with flow cytometry plots showing PD-L1 cell surface protein expression for selected cell lines.

5.3 Experimental design

To ensure the success of the screen, a number of factors need to be optimised, and will be discussed in the following subsections.

5.3.1 Selection of cell line

The goal of this project was to identify targets involved in downregulating PD-L1 expression. It is therefore critical that the cell line used to perform the screen in has a high expression of PD-L1, ensuring the highest resolution for discovering genes and pathways involved in modulating PD-L1 cell surface expression. I evaluated a range of melanoma cell lines for PD-L1 gene expression using RNA-sequencing data generated in-house by Dr. Marco Ranzani (Section C.1). The majority of cell lines showed low PD-L1 gene expression, with a few exhibiting moderate to high expression. I selected some of the cell lines with the highest gene expression, and conducted protein level expression analysis using flow cytometry (Fig. 5.1). The ideal cell line for my project should not only show a high endogenous cell surface expression of PD-L1, but this expression also needs to be uniform across the cells as a clonal cell line would limit noise in the downstream analyses. Out of the tested cell lines, C092 was the ideal candidate which fulfilled both of these criteria.

General practicalities such as the growth properties and size of the cell line would also impact the feasibility of the screen. As such, C092 cells are reasonable cells to work with as they are medium to large, show adherent growth and have a doubling time of approximately 30 hours. In addition, it is a human melanoma cell line, derived from a lymph node metastasis of a Caucasian male diagnosed with metastatic melanoma where the primary was never identified [713]. Notably, it is characterised by a low number of mutations (13 non-synonymous mutations in total), none in any established melanoma driver genes or genes associated with regulation of PD-L1 such as IFN-γ pathway genes.

A cell line can only be selected to be used in the screen if it can be effectively transduced and functional in terms of generating gene knock-outs. Therefore, the next step in preparation for the pooled genome-wide CRISPR-Cas9 screen, was to generate a cell line expressing Cas9. This was done using lentiviral transduction of the parental C092 line. Cas9 cleaving efficiency was assessed through the transduction of the Cas9-expressing cells with lentivirus containing BFP, GFP and a gRNA targeting GFP (Fig. 5.2). Uninfected cells and cells infected with BFP and GFP but without a gRNA (empty) were used as controls. Using this system, Cas9 efficiency was estimated to be very high, with 94% of the transduced cells showing Cas9 activity. Therefore, it could be concluded that this cell line can be efficiently transduced and successfully used to generate gene knock-outs. Taken together, the C092 cell line fulfils all of the above-mentioned criteria and was deemed suitable for use in the screen.

5.3.2 Set up of appropriate controls

The set up and use of appropriate controls is paramount in maximising the chance of a successful screen. First, I wanted to make sure loss of PD-L1 would not have a big impact on the growth properties of the cells. Additionally, creating a cell line where PD-L1 is knocked out, is useful as a positive control for the screen, as it would showcase the maximum loss of PD-L1 cell surface expression that is achievable in the C092 cell line. Second, I sought to test whether PD-L1 expression would be influenced by the CRISPR-Cas9 screening process, i.e. viral infection, Cas9 cleavage and antibiotic selection. Therefore, I selected *OR14A16* as a control gene from the BAGEL list of non-essential genes [714], as I deemed it unlikely to interfere with PD-L1 function.

To mimic the screening conditions, gRNAs against *CD274* (PD-L1) or *OR14A16* were cloned into the empty lentiviral backbone (Fig. 5.3A), in the same position as the gRNAs part of the genome-wide gRNA library. This was then followed by lentiviral production and

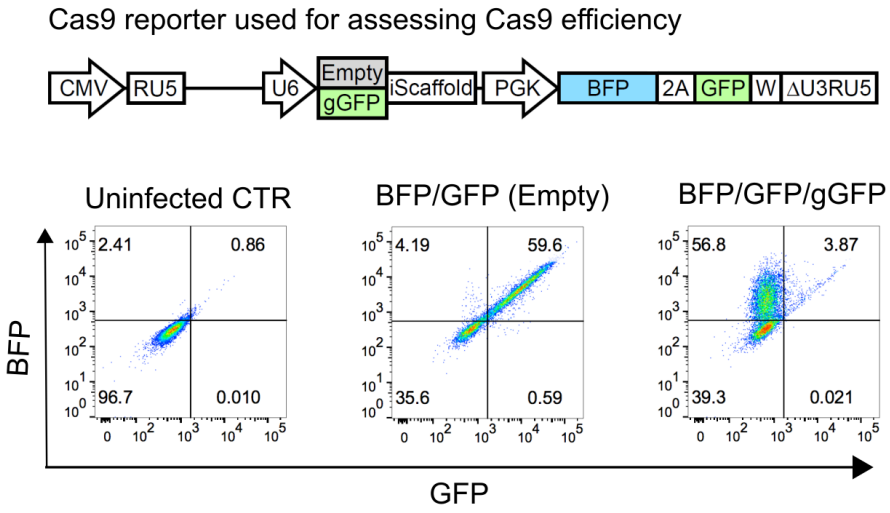


Figure 5.2: **Cas9 reporter and efficiency test.** Top panel shows the schematics of the lentiviral Cas9 reporter vector used. Bottom panel shows the results from the Cas9 efficiency test. Reporter schematics adapted from [411], with permission under the Creative Commons Attribution License (CC BY).

transduction of C092 cells to create knock-out cell lines. The successful knock-out of the respective genes were confirmed using the Surveyor assay and their impact on cell surface PD-L1 expression along with the incorporation of BFP was assessed using flow cytometry (Fig. 5.3B,C). Knock-out of PD-L1 showed a large reduction in PD-L1 cell surface expression, showcasing the high signal to noise ratio which will be advantageous in capturing PD-L1 loss using this cell line. The *OR14A16* KO did not show any difference in PD-L1 cell surface expression compared to the parental wild-type cell line, confirming the negligible impact of the aforementioned CRISPR-Cas9 screening steps on PD-L1 expression.

5.3.3 Establishing optimal screening conditions

A number of factors are important to consider when assessing the optimal variables for a successful screen. A whole range of commercial gRNA libraries exist [405, 407, 408, 411], and it is also a possibility to tailor-design and create a custom library from scratch. For my project, I did not want to limit the discovery by using specialised libraries focused on e.g. kinase, nuclear, chromatin-regulatory or cell cycle-associated genes specifically [399, 715]. Therefore, I sought a wide and unbiased approach whereby it is possible to identify targets on a genome-wide scale. I selected the Yusa human v.1.1 genome-wide gRNA library which was designed and gifted by Prof. Kosuke Yusa (Section C.1). The original v.1 of this

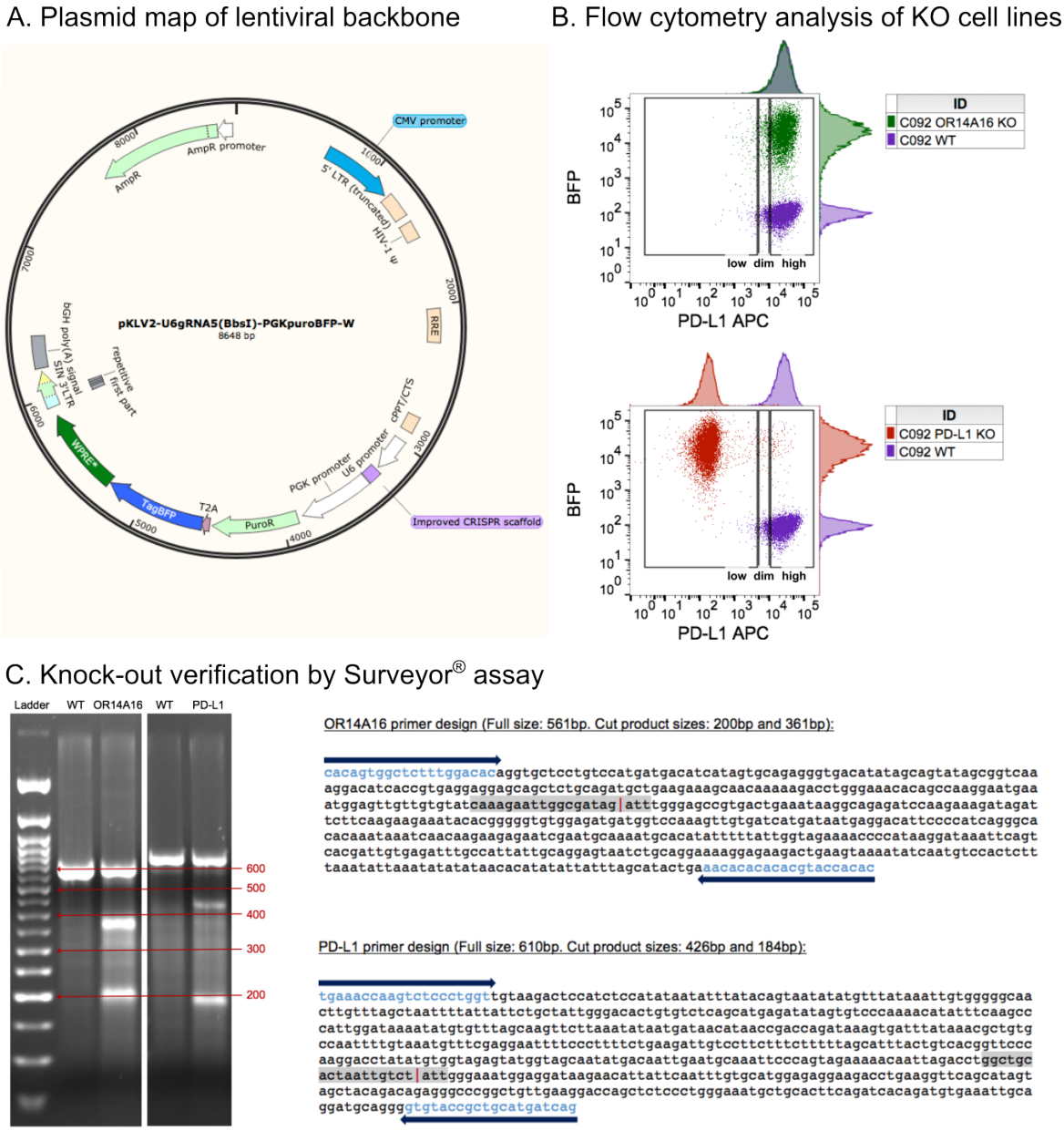


Figure 5.3: **Generation of knock-out controls for the screen. A)** Plasmid map showing the empty lentiviral backbone. To generate the controls, gRNAs were ligated upstream of the U6 promoter, into the improved CRISPR scaffold (purple). The plasmid also contains markers for puromycin resistance and a BFP tag. **B)** Expression of BFP and PD-L1 by flow cytometry shows the successful transduction of lentivirus carrying the plasmid with inserted gRNAs against *CD274* (PD-L1) or *OR14A16*. PD-L1 cell surface expression is reduced drastically in the *CD274* (PD-L1) KO cells, whereas expression is unchanged with *OR14A16* loss. **C**) Knock-out of the respective genes were confirmed using the Surveyor mismatch cleavage assay. Left figure shows the product sizes on an agarose gel stained with ethidium bromide. Right figure shows the primer sequences (blue), the sizes and the sequences of the full or cleavage products. The gRNA sequence is highlighted in grey, where the red line denotes the cleavage (mismatch) site.

library has shown high performance and success in previous screens [405, 411, 716–718]. In addition to the original library, this extended v.1.1 contains 1004 nontargeting guides and an additional 9380 gRNAs targeting 1876 selected genes at 5 guides per gene from the Lander and Sabatini gRNA library [399, 719].

After deciding which library to use, the size of the screen would be the next variable to establish. As with any screening approach, it is important to ensure a high library complexity from start to finish to avoid biasing and losing gRNA representation. The gRNA library I selected contains 101,095 gRNAs in total, and when transducing cells in a pool, a multiplicity of infection (MOI) of 0.3 would be ideal as it strengthens the probability of each cell being infected with only one gRNA [720, 721]. This means more than 300,000 cells would be required to ensure each gRNA is represented once, but it is common practice to use a library coverage in the range of 200x-500x [411, 719, 721]. However, the maximum plausible library representation becomes a trade-off between boosting screening performance and ensuring the practicalities of the experiment. As approximately 10 million C092 cells reaches confluency in a T150 culture flask, I decided to transduce and maintain a library coverage of 200x at all times. This meant an initial infection of 60 million cells, and I opted to perform the screen in triplicates. For logistical reasons, as it was not feasible to sort all replicates in one day, at the final subculturing step each replicate was split into two sets, where one set was harvested on day 14 and the other on day 15. I selected a final assay time point of two weeks, as this would allow enough time for the gene disruption to manifest and subsequently exert its effects on PD-L1 cell surface expression. Additionally, this time point is short enough to reduce the extent of faster-growing clones to overtake and skew the population.

To reduce the probability of each cell being infected with multiple gRNAs, it is essential to infect at a low MOI. A MOI of 0.3 is aimed for as it maximises the number of cells receiving just one lentiviral particle, while minimising the number of uninfected cells and cells infected by more than one genetic perturbation [720, 721]. To achieve this, each lentiviral pool was titrated and customised for each cell line and assay setup in preparation for the large screen (Fig. A.19).

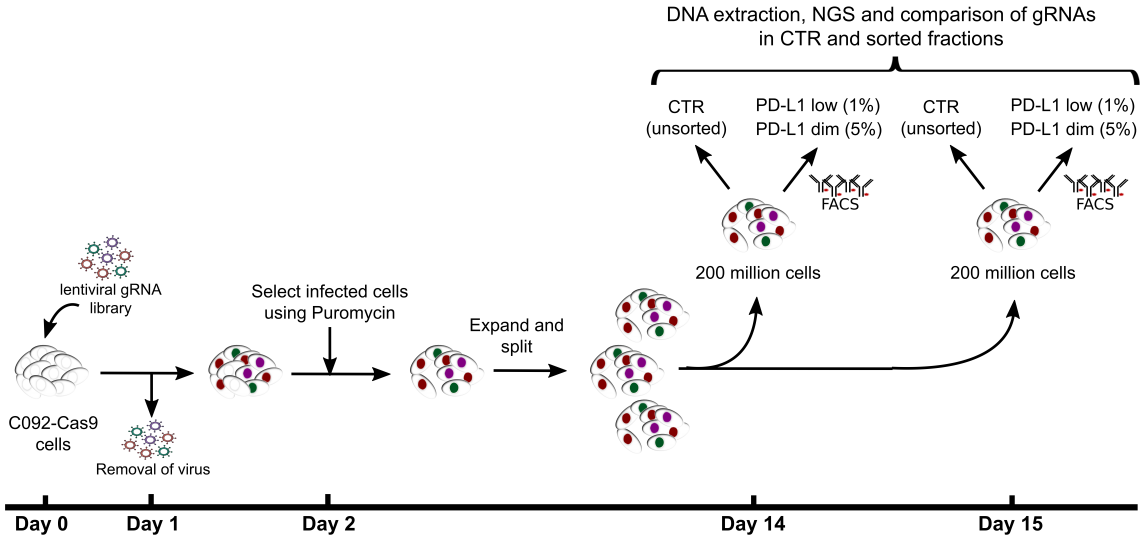


Figure 5.4: **Screening approach to identify regulators of PD-L1.** C092-Cas9 cells were infected in suspension with the Yusa human v.1.1 lentiviral gRNA library at a MOI close to 0.3. Successfully transduced cells were then selected using Puromycin, followed by continuous expansion and cell culture for a total of two weeks. Due to logistic reasons, at the final subculturing step each replicate was split into two sets, where one set was harvested on day 14 and the other on day 15. 200 million cells each were harvested on day 14 and day 15, where half of the cells were kept as the unsorted control population, while the other half were subjected to FACS. A two-way sort was performed, capturing the lowest 1% (low) and 1-6% (dim) PD-L1-expressing clones. DNA was extracted from each sample, subjected to sequencing and analysis where gRNAs enriched in the PD-L1 low and dim sorted fractions compared to the unsorted control were identified.

5.4 Pooled genome-wide CRISPR-Cas9 screen to identify regulators of tumour PD-L1 expression

After establishing the optimal parameters for a successful screen, I set out to perform a pooled genome-wide CRISPR-Cas9 screen to identify regulators of cell surface PD-L1 expression in the human melanoma cell line C092. The schematics of the screening process is depicted in Fig 5.4.

Cells were infected in triplicates, at a MOI close to 0.3 (Fig. A.18) and with a library coverage of 200x. Successfully transduced cells were selected using Puromycin, and the cells were cultured for a total of two weeks. On day 14 and 15, cells were harvested for FACS to identify cells infected with gRNAs causing a phenotype of reduced cell surface PD-L1 expression. 100 million cells were kept as the control population, while another

100 million cells were subjected to FACS, selecting the lowest 1% (low) or 1-6% (dim) of PD-L1-expressing cells (Table A.21). DNA was extracted, PCR amplified and subjected to sequencing yielding an average of 36.6 million reads per sample. Second round of PCR and sequencing were performed by the Sanger core facilities (Section C.1).

5.4.1 Quality control

Two samples (Day 14 replicate 1 dim and Day 14 replicate 2 low) failed PCR amplification and a third sample (Day 15 replicate 1 low) showed very low number of total reads. These three samples were therefore excluded from further analyses. The plasmid library had been sequenced previously, providing information about the original gRNA pool counts.

The first quality control measure I sought to assess was the potential skewing of the library during the experiment. As an example, knock-down of specific genes including critical tumour suppressors such as *CDKN2A* could provide a growth advantage [722]. These cells would therefore expand faster and become overrepresented with time compared to neutral or growth-suppressing knock-out clones. If this would have happened, the possible hits from the screen would be misleading. I first checked the distribution of gRNAs across the control samples, ensuring the profiles looked similar (Fig. A.20). Then, I looked at the number of gRNAs with zero counts in the respective samples, as this would indicate whether the library representation had been compromised during the experiment. 199 gRNAs had zero counts in the plasmid library. In addition to these, 141 gRNAs had zero counts in all controls, indicating the possible absence of these gRNAs early in the selection process. The total number of zero counts across all replicates were similar, accounting for <1% of the gRNA library (Fig. 5.5). 47% of any gRNA with zero counts were uniquely lost in only 1 sample. The low number of total gRNAs with zero counts, where the majority of gRNAs lost were specific to one sample, implies that appropriate maintenance of library complexity was achieved throughout the screen.

Next, I analysed the correlation of the gRNA counts between control replicates, as this would indicate possible limitations with a specific control sample. Pearson's correlation values between all control samples ranged between 0.86 and 0.90 (Fig. 5.6A), suggesting the screening process was robust with the library complexity maintained across all control samples throughout the screen.

As the screen was performed using a genome-wide library, it also covers genes essential for the survival of a cell. Therefore, a measure of the dropout of such essential genes over time could also be used as a quality control measure. I made a comparison between the

Zero counts gRNAs in samples

Sample	Number of zero count gRNAs (% of total)
Plasmid	199 (0.20%)
Day 14 Rep 1 CTR	764 (0.76%)
Day 14 Rep 2 CTR	685 (0.68%)
Day 14 Rep 3 CTR	732 (0.72%)
Day 15 Rep 1 CTR	770 (0.76%)
Day 15 Rep 2 CTR	795 (0.79%)
Day 15 Rep 3 CTR	841 (0.83%)

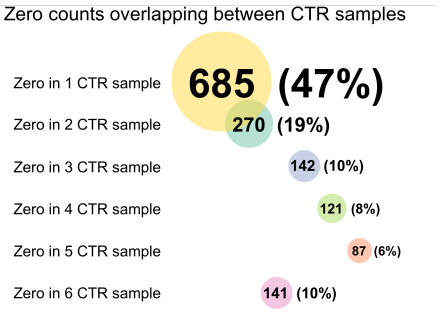


Figure 5.5: **Zero count gRNA statistics.** Left figure shows the number of gRNAs with zero counts in each control sample. Right figure shows the number of gRNAs with zero counts which are shared between control samples.

gRNAs present at the start (plasmid sample) with those found in the control samples cultured for two weeks. Genes from the Bagel core reference set of essential and non-essential human genes [714] were used to infer the effect of its knock-down. gRNAs targeting genes known to be essential were indeed found less frequently in the control samples compared to the plasmid sample. This is in comparison to the non-essential genes which did not show this dropout pattern (Fig. 5.6B).

Finally, the Bagel core reference genes were also used to assess the performance of the screen by computing the Receiver Operating Characteristic (ROC) curve, where the Area Under The Curve (AUC) is used to evaluate the sensitivity and specificity of the screen. A detailed description can be found in Section A.22. Briefly, the analysis was performed using the output from Mageck v.0.5.6 (Table A.1) comparing control samples versus the plasmid. The Bagel core reference genes were then used to estimate the false positive, true negative, sensitivity, specificity and false positive rates used for the ROC curve computation (Fig. 5.6C). The calculated AUC value of 0.91 suggests the screen performance was satisfactory, confirming its high ability to discriminate between true and false positives. Collectively, all quality control analyses performed demonstrate a robust and high-performing screen. I could therefore confidently move on to analysing the results.

5.4.2 Results

This genome-wide CRISPR-Cas9 screen allows for the discovery of new molecular pathways which downregulate PD-L1, and could therefore provide new insights into PD-L1 biology with the potential discovery of novel targets for drug development. Two weeks post infection with the pooled gRNA library, cells were harvested and FACS performed to

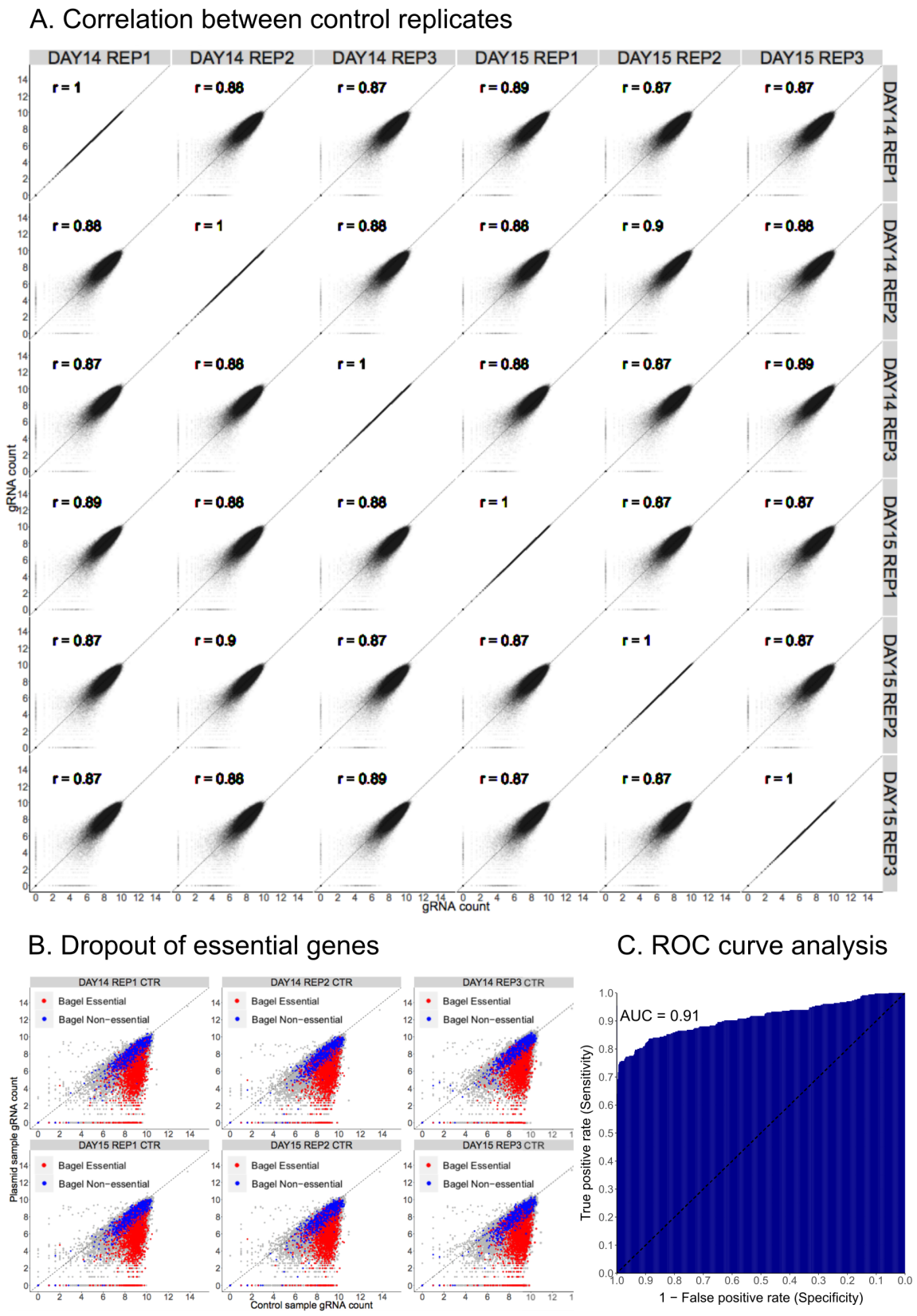


Figure 5.6: **Quality control metrics from the screen.** **A)** Correlation of gRNA counts (log2 normalised) between the control replicates, with Pearson's R values shown for each comparison. **B)** Dropout of essential genes shown by gRNA counts between plasmid sample and each control sample. Bagel core reference set of essential genes shown in red, and non-essential genes in blue. **C)** ROC curve analysis with AUC calculated.

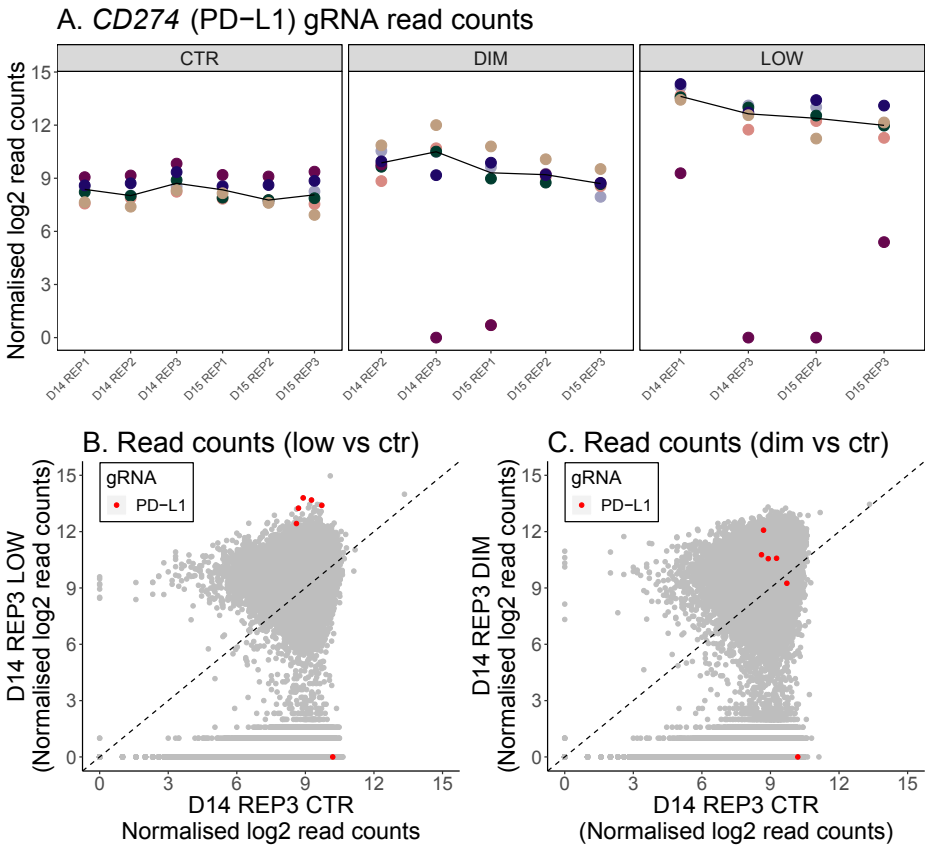


Figure 5.7: ***CD274* (PD-L1) gRNA read counts. A)** Read counts for each gRNA targeting *CD274* across all samples. Each gRNA is assigned a different colour. The black line depicts the median gRNA count. **B)** Read counts shown for one representative low vs ctr sample. All six gRNAs targeting *CD274* are highlighted in red. Dotted line shows the line of unity. **C)** Read counts shown for one representative dim vs ctr sample. All six gRNAs targeting *CD274* are highlighted in red. Dotted line shows the line of unity.

identify gRNAs enriched in cells which had reduced PD-L1 cell surface expression. The gRNAs present in the lowest 1% (low) and 1-6% (dim) were each compared with the unsorted control samples. Mageck v.0.5.6 (Table A.1) was used to score the enrichment of all genes in the library in the sorted fractions compared to the controls. Briefly, the Mageck analysis begins with a total read count normalisation. This is followed by a comparison of the gRNAs present in the control versus the sorted samples, using a negative binomial model followed by a robust rank aggregation algorithm. Genes consistently enriched in the sorted samples are given a higher rank and a permutation-based statistical test is applied.

As expected, *CD274* (encoding PD-L1) was the strongest hit in the low fraction (rank 1, FDR-adjusted p-value = 0.0002). However, although *CD274* showed enrichment in the dim fraction, it was not top ranked (rank 24, FDR-adjusted p-value = 0.023). This reflects

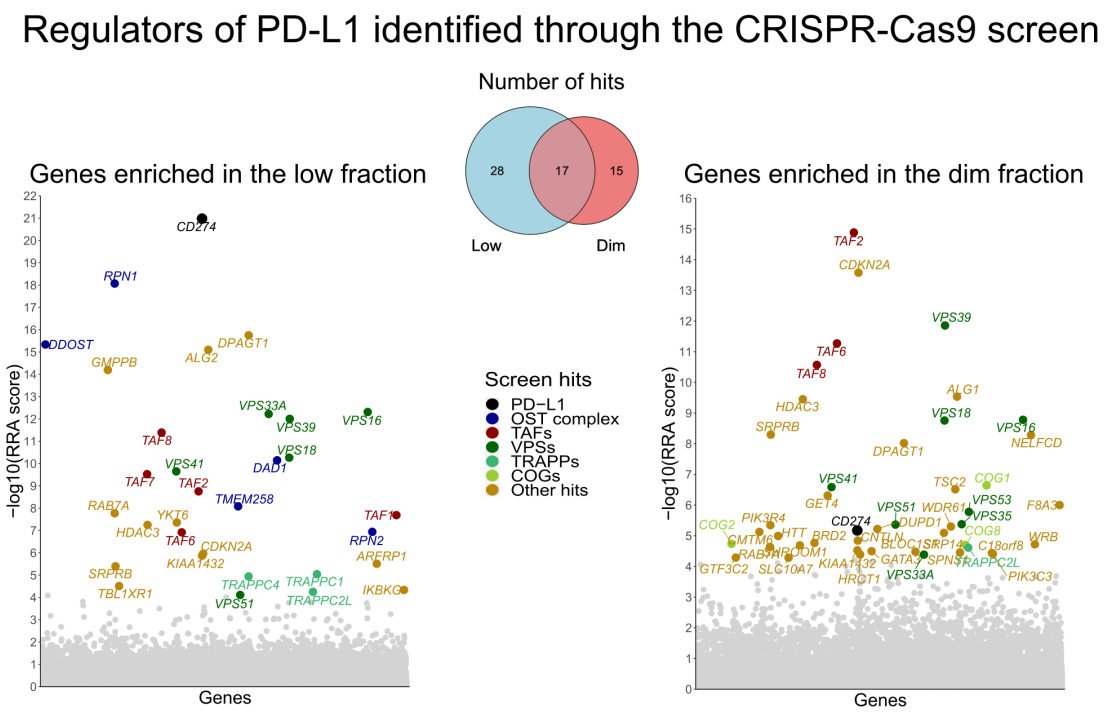


Figure 5.8: **Regulators of PD-L1 identified through the CRISPR-Cas9 screen. A)** Visualisation of the robust rank aggregation (RRA) scores (-$\log_{10}$ adjusted) for all genes comparing low with ctr samples. Genes enriched in the low fraction with FDR < 10% are marked, and coloured by function. **B)** Visualisation of the RRA scores (-$\log_{10}$ adjusted) for all genes comparing dim with ctr samples. Genes enriched in the dim fraction with FDR < 10% are marked, and coloured by function.

its drastic removal from the cell surface upon knock-out (Fig. 5.7).

A total of 60 genes were identified as regulators of PD-L1 expression (FDR-adjusted p-value <10%), with 32 genes enriched in the low fraction and 45 genes in the dim fraction (Fig. 5.8). The hits from the screen reflect a diverse range of intracellular processes involved in presentation of PD-L1 on the cell surface. Broadly, hits could be grouped into general categories covering early glycosylation, basal transcription factors and intracellular transport mediators. Furthermore, a range of novel targets were also discovered including *CMTM6*, *GTF3C2*, *SPNS1*, *F8A3*, *SHROOM1*, *HRCT1*, *DUPD1* and *HTT*. During the course of my project, *CMTM6* was identified by other researchers as an important regulator of PD-L1 through stabilisation of its cell surface expression and by protecting it from lysosomal degradation [366, 367]. Therefore, the identification of *CMTM6* in my screen provides further evidence of the screen being successful.

5.4.2.1 Basal transcription factors controlling PD-L1 expression

Five TATA box binding protein (TBP)-associated factors (TAF) genes*: TAF1, TAF2, TAF6, TAF7* and *TAF8* were discovered in my screen as positive regulators of PD-L1 surface expression. These genes encode transcription factors which are involved in many general processes to initiate gene transcription, including promoter recognition and selectivity, interaction with nucleosomes, coactivation, enzymatic or chromatin modification regulation [723–726]. TAF proteins can function as subunits in larger multi-protein complexes, where its association with Transcription factor II D (TFIID) for RNA polymerase II-mediated transcription initiation is the most well-studied. To date, thirteen human TAFs have been identified to participate in TFIID complexes [727, 728], including some which act in a tissue-specific manner [729]. Transcriptional regulation by TAFs in the context of cancer has been studied, where alterations targeting the TFIID machinery are frequent events in various cancers. In melanomas, over 30% of TCGA SKCM (Section C.2) samples had *TAF* gene mutations or copy number amplifications [73, 730]. Additionally, *TAF* genes are important for many cellular processes which are commonly hijacked to promote tumourigenesis, such as regulation of differentiation, cell cycle progression and apoptosis [731–733]. Besides the TFIID networks, TAF proteins have also been found in chromatin remodelling complexes such as SAGA in yeast and STAGA or PCAF in human [734, 735].

Interestingly, the five TAF proteins identified in my screen are all structurally connected when they participate in the TFIID complex, and together they complete the structure of the C lobe (Fig. 5.9). This suggests that the C lobe might be particularly important in regulation of PD-L1 expression. In the centre of the TFIID complex, TAF6 is found as a dimer, tethering the three lobes together [728, 736]. The C-terminal tail of TAF8 then forms a helical structure connecting TAF6 and TAF2 [736], and this structure has an important role in the incorporation of TAF2 into the complex [737]. DNA binding is likely mainly facilitated by TAF1 and TAF7 together with TAF2 [728, 738–740].

Among the *TAF* genes found in my screen, *TAF1* is the most well-studied. It has been reported to have both protein kinase activity [741, 742] and histone acetyltransferase activity [731, 743]. Additionally, it has ubiquitin-activating/conjugating functions [744–746], as well as harbouring two tandem bromodomains, indicating an additional role in chromatin remodelling [747]. *TAF1* regulates the expression of cell cycle and apoptosis-related genes [742, 748–751], as shown by late G_1 arrest, induction of apoptosis and DNA damage response in the ts13 hamster cell line which bears the G690D temperature-sensitive mutation in *TAF1* [752].

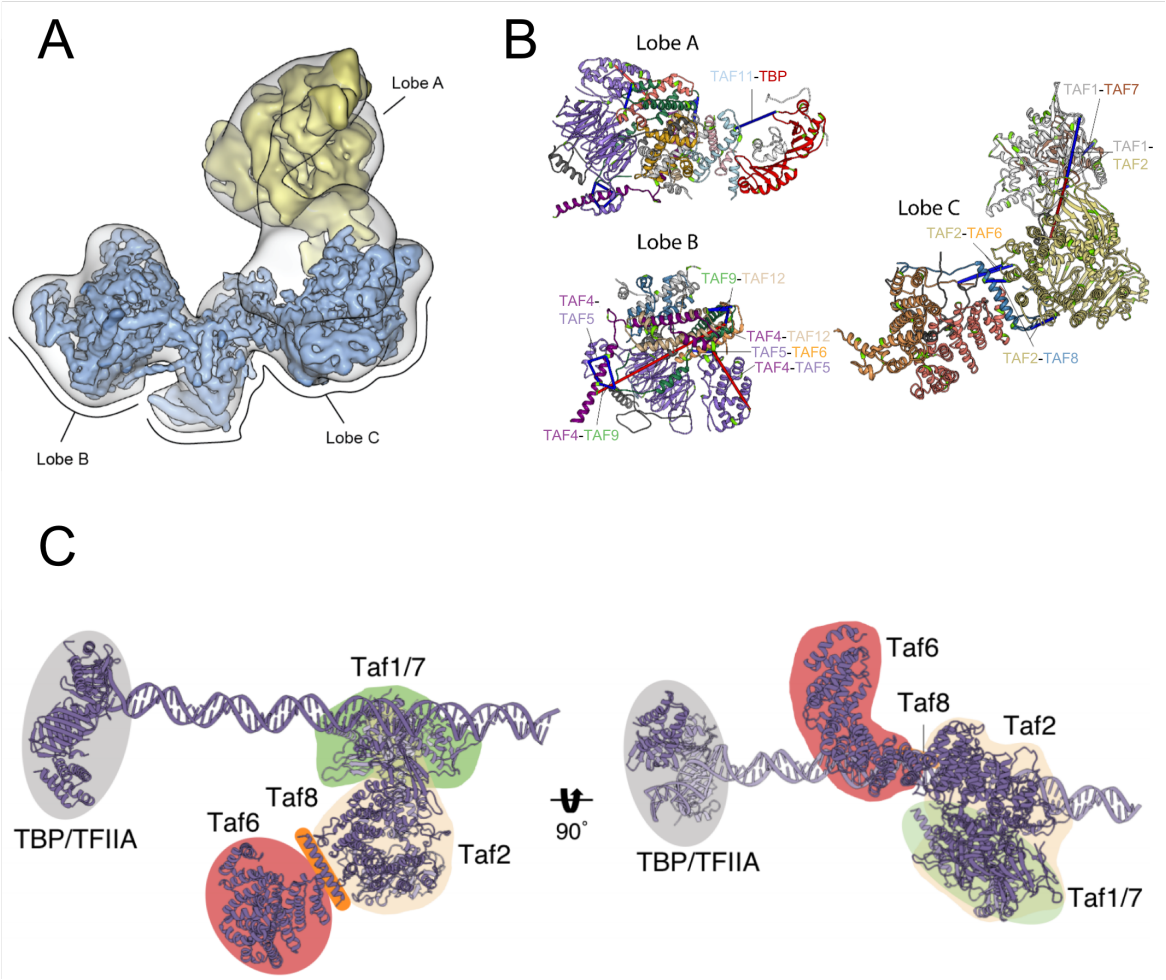


Figure 5.9: **Arrangement of TAFs in the human TFIID complex. A)** Transparent cryo-EM map showing the TFIID complex, highlighting the 3 lobes. **B)** View of the individual lobes highlighting the intermolecular crosslinks between subunits. **C)** Structural view of the TFIID subunits comprising 5 TAF, TBP/TFIIA and the DNA interaction through TAF1/7. Figure A and B printed with permission from [728] and Figure C reproduced with permission from [736] under the Creative Commons CC BY 4.0 license.

Contradictory to these data, one study has instead shown an attenuation of oxidative stress-induced apoptosis when *TAF1* was depleted in HEK293T and two human cancer cell lines [732]. This suggests *TAF1*-associated regulation might be complex and highly context-dependent. Furthermore, overexpression of *TAF1* has been reported to enhance androgen receptor activity, thereby promoting prostate cancer progression [746].

TAF2 encodes the second largest TFIID subunit, but it can also bind DNA by itself or through participation in other configurations [753]. When TAF2 assembles in a trimeric

complex with TAF1 and TBP, it shows a promoter specificity towards Inr sequences [754]. Similar to *TAF1*, *TAF2* is also required for transcription of cell cycle regulatory genes [755]. Additionally, TAF2 is overexpressed in 73% of high-grade serous ovarian cancers [730].

TAF7 has been proposed to combat premature transcription initiation by temporarily binding to TAF1, thereby inhibiting its acetyltransferase activity until the pre-initiation complex assembly is complete [756]. However, TAF7 is also capable of TAF1-independent transcriptional regulation, as exemplified by its inhibition of CIITA-mediated transcription of MHC class I and II genes [757] or regulation of c-Jun [758]. Reduced *TAF7* expression led to lower polyamine transporter activity and resistance to methylglyoxal bis(guanylhydrazone) (MGBG)-induced growth arrest in human prostate cancer cells [759].

Studies have shown that an interaction between p53 and TAF6 is important for transcriptional activation [760]. Therefore, where this interaction is disrupted *TP53* might show a hampered tumour control capability. An apoptosis-induced isoform of TAF6 has also been found, where its expression initiates transcription of a range of p53-responsive promoters including those of *GADD45* and *CDKN1A,* and its overexpression was found sufficient to trigger apoptosis [761]. In another study, the two main isoforms of TAF6 were studied in normal and cancerous breast epithelia. They found, that the growth-suppressive isoform was predominant in normal tissues, whereas cancerous samples showed an opposite, and markedly reduced ratio of this isoform [762]. The authors also found a physical interaction between GADD45 and this growth-suppressive isoform, which is postulated to be causing this growth-abrogation effect.

Taken together, my screen identified five TAF proteins which make up the C lobe of the TFIID complex (Fig. 5.9), suggesting this part of the TFIID complex could be particularly important for transcriptional regulation of *CD274*. However, it should be emphasised, that regulation of gene expression by *TAF* genes and the TFIID complex is a generic, and in many cases essential process. Nevertheless, overexpression of *TAF* genes has been associated with various cancers, with reduced expression of some genes being linked to resistance towards growth arrest and apoptosis. It therefore appears *TAF* genes might contribute to tumourigenesis; however, an immune-associated role of these genes has yet to be established. Since many *TAF* genes regulate cell cycle progression, and recent data show PD-L1 expression can be regulated by cell cycle kinases [365], it is possible that some TAF genes could have either a direct or indirect (through cell cycle regulation) effect on PD-L1 tumour expression.

5.4.2.2 Regulation of PD-L1 expression by N-linked glycosylation

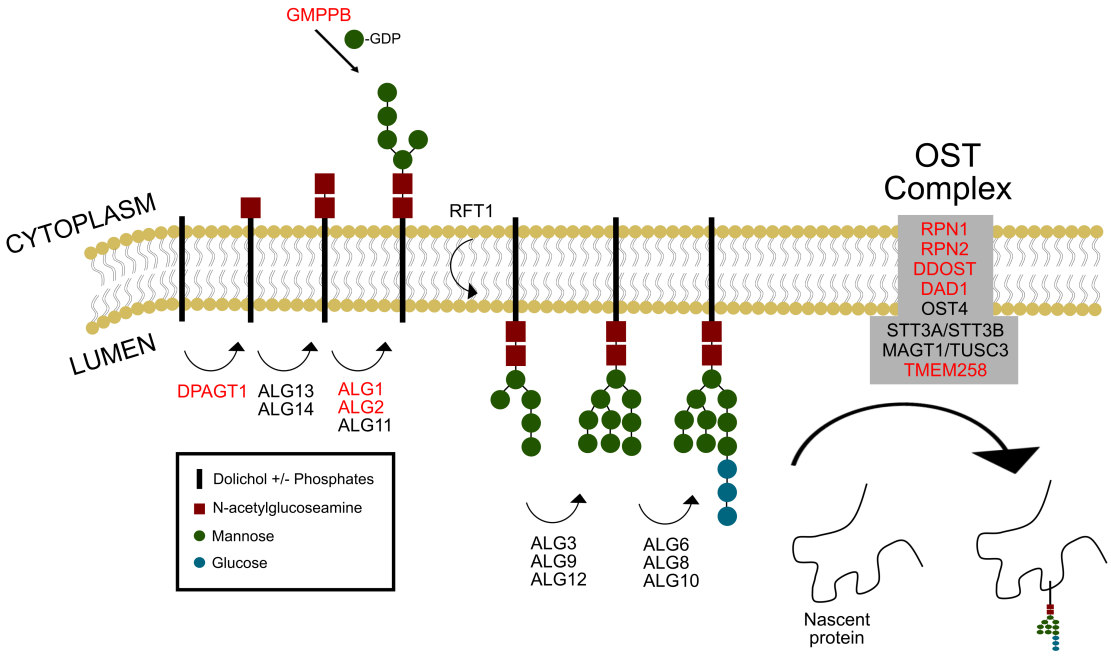


Figure 5.10: **Simple overview of the protein N-glycosylation process.** Protein N-glycosylation takes place in the ER, where a 14-sugar glycan is serially assembled, catalysed by a range of ALG enzymes. *GMPPB* is involved in the synthesis of an intermediate substrate which is added to the core N-glycan structure as part of this process. Upon completed core glycan assembly, the oligosaccharyltransferase (OST) complex mediates the transfer of this unit to newly synthesised polypeptides. Screen hits encoding the enzymes involved in this process is highlighted in red.

Several glycosylation-related genes were enriched in the PD-L1 low and dim fractions, suggesting such events are important for PD-L1 surface presentation. Of note, key members of N-linked glycosylation processes including five members of the oligosaccharyltransferase (OST) complex were found in my screen. Protein N-glycosylation is an important post-translational process taking place in the endoplasmic reticulum (ER) and Golgi apparatus, and governs protein folding and stability [763], subsequently controlling the function and activity of many proteins [764]. The first step takes place at the ER membrane (Fig. 5.10), where a 14-sugar glycan is serially assembled and mounted on a membrane-anchored dolichol phosphate [765]. These steps are catalysed by a range of enzymes encoded by *ALG* genes, which includes the screen hits *ALG1*, *ALG2* and *DPAGT1* (also known as *ALG7*). Ad-

ditionally, *GMPPB* is involved in the synthesis of an intermediate substrate which is added to the core N-glycan structure as part of this process [766]. Upon completed core glycan assembly, the OST complex mediates the transfer of this unit to newly synthesised polypeptides [763]. The human OST complex is made up of eight components [765, 767, 768], five of which (*RPN1, RPN2, DDOST, DAD1* and *TMEM258*) were hits in my screen. Taken together, these data confirm PD-L1 is a glycosylated protein, where proper cell surface presentation of PD-L1 relies on above-mentioned early N-glycosylation processes. Many proteins which are not correctly glycosylated can be targeted for degradation [769, 770]. Therefore, this process is likely generic and essential for most proteins which are glycosylated, not just PD-L1.

In the context of cancer, changes in glycosylation patterns are commonly observed [771, 772]. As an example, enhanced expression of glycosylation-regulating genes *MGAT3* and *MUC1* are linked to increased tumour aggressiveness and metastases [771, 773, 774]. In breast or colorectal cancer, high expression of either *RPN1* or *RPN2* have been associated with shorter survival and more distant metastasis in different studies [775–777]. In addition to this, *DPAGT1* can be activated through canonical Wnt signalling, and exerts direct effects on E-cadherin glycosylation, subsequently causing a reduction in cell adhesion [778]. Glycosylation is also important in the regulation of immune cell functions, playing a role in for example ligand interactions [779, 780]. A global reduction in glycosylation due to suppressed activity of early glycosylation genes might therefore be useful in mounting an anti-tumour response, both through downregulation of tumour PD-L1 expression, and by PD-L1 independent mechanisms [781, 782]. Indeed, recent studies have shown potent cancer cell eradication by hampering PD-L1 glycosylation [783], and general inhibition of glycosylation using tunicamycin does decrease PD-L1 expression [363]. However, as glycosylation is such a key process, the potential therapeutic regulation of PD-L1 expression through targeting this machinery needs to be designed with caution and warrants further investigation.

5.4.2.3 Intracellular transport for presentation of PD-L1 on the cell surface

PD-L1 goes through a range of subcellular transportation processes before being presented as a cell surface protein, and a recent study showed surface PD-L1 expression is maintained and regulated by internalisation and recycling components such as CMTM6 [367]. Several members of multi-subunit tethering complexes (MTCs) which regulate intracellular transport (Fig. 5.11) were found in my screen, including transport protein particle (TRAPP), conserved oligomeric Golgi (COG) and homotypic fusion and vacuole protein sorting (HOPS)

complex components.

The TRAPP complexes are important in controlling ER, Golgi and vesicle trafficking through regulation of membrane tethering and fusion events [784]. The screen hits *TRAPPC1* and *TRAPPC4* encode essential core TRAPP subunits, while *TRAPPC2L* is an adaptor subunit [785]. Interestingly, *TRAPPC4* is overexpressed in colorectal cancer cells, where high expression is linked to increased cell proliferation and invasiveness [786–788]. This effect is suggested to be mediated by a TRAPPC4-ERK2 interaction, which activates the MAPK pathway. Given these results, it is possible that the TRAPP complex plays a role in tumourigenesis, or that individual members could have separate independent functions. Oncogenic activation of the MAPK pathway has been shown to alter PD-L1 expression and contribute to immune escape [337, 338, 341].

COG proteins facilitate inter-Golgi and Golgi-vesicle transport by mediating the transition from tethering to vesicle fusion [789]. Cells with defective *COG1* or *COG2*, show reduced steady-state levels of some integral membrane proteins due to vesicle sorting failure and subsequent proteasomal degradation. It is therefore possible that lower PD-L1 expression by knock-down of *COG* genes is the result of a global and unspecific reduction in the level of cell surface proteins [790, 791].

The endocytic pathway also plays a role in regulating PD-L1 expression, as evident by my screen hits including members of the HOPS complex. This complex regulates tethering and fusion events involving late endosomes, mediating vesicle sorting and protein recycling [792, 793]. The core structure of the HOPS complex consists of VPS11, VPS16, VPS18 and VPS33A. gRNAs against *VPS11* were absent from the screen library, therefore this gene could not be studied in my screen. In addition to the core subunits, the HOPS complex also includes VPS39 and VPS41. All of these HOPS complex components, including its interaction partner Rab7, were found in my screen, suggesting the endolysosomal pathway is involved in managing cell surface PD-L1 expression.

My screen also identified other genes involved in vesicle trafficking and membrane recycling of proteins, such as *PIK3C3* (also known as *VPS34*), *VPS51*, *VPS53*, *VPS35* and *YKT6* [794–797]. Collectively, these results increase our knowledge of the biological mechanisms of cell surface presentation of PD-L1, emphasising how it is processed through various vesicles from the ER, to the Golgi, endosomes, lysosomes and finally the plasma membrane (Fig. 5.11). In the context of melanoma, alterations in endolysosomal genes such as *RAB7A* are enriched, and facilitate tumour progression through promoting proliferation and metastatic properties [798]. Other studies also link known oncogenic events such as *BRAF* V600E mutations or the activation of the PI3K pathway to deregulated vesicle trafficking

[799–801], further emphasising how these pathways can be altered to regulate tumourigenesis, whereby PD-L1 might also play a role.

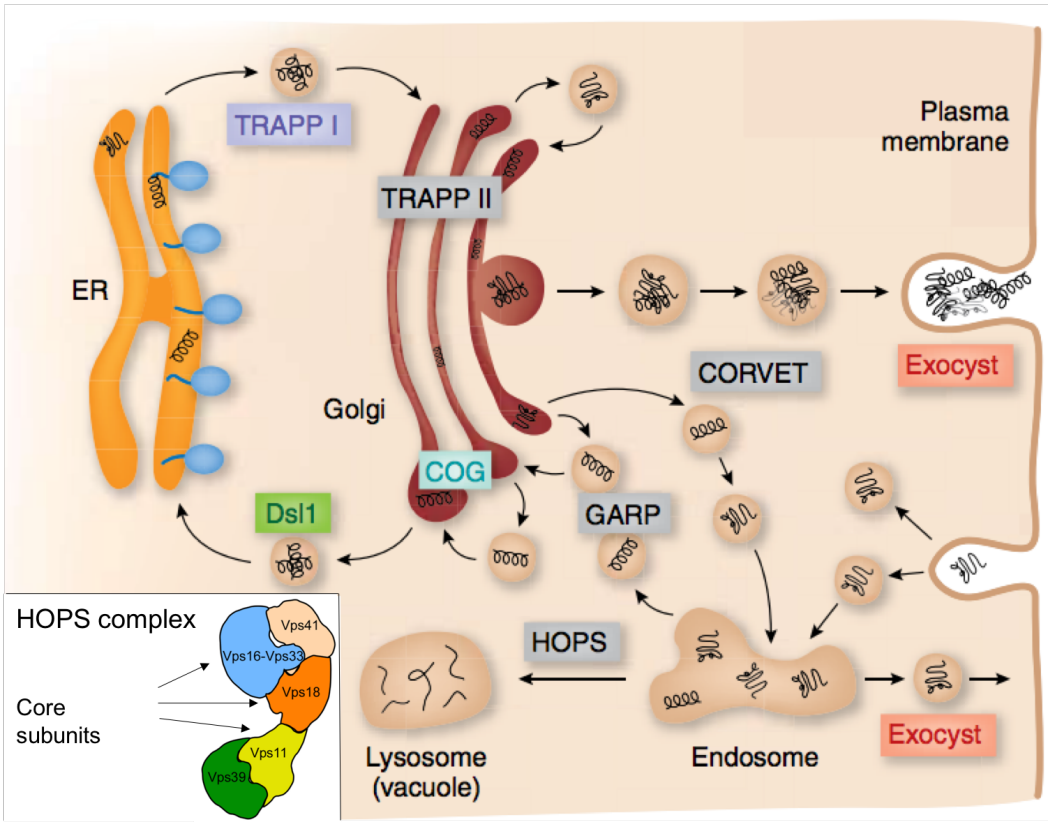


Figure 5.11: **Intracellular vesicle trafficking schematics.** Key intracellular trafficking pathways and the complexes involved. The subunits of the HOPS complex is shown in the bottom left figure. Modified with permission from [802], and [803] under the Creative Commons Attribution License CC BY 3.0.

5.4.2.4 Other processes involved in PD-L1 regulation

The results from my screen show that epigenetic modifications also impact PD-L1 expression, as exemplified with hits such as *BRD2*, *HDAC3* and *TBL1XR1*. Recent studies have shown that PD-L1 expression levels can be modulated using both histone deacetylase (HDAC) inhibitors and bromodomain and extraterminal (BET) inhibitors [804–807]. The cellular response to these epigenetic regulators is both broad and complex, involving a range of biological functions such as inflammation, immunity, apoptosis and cell cycle regulation in addition to changes in the expression levels of various genes [808–812]. This could therefore partly explain why the direction of change in PD-L1 expression appear to depend on

inhibitor specificity and cell type studied [317, 319, 805, 806, 813]. My screen supports the studies showing PD-L1 expression in melanoma cells can be diminished by the knock-down of *HDAC3* or *BRD2*. Additionally, another screen hit: the *TBL1XR1*-encoded protein associating with HDAC3 in the nuclear receptor corepressor complex is overexpressed in a range of cancers and linked to increased tumour aggressiveness [814–816] making it an interesting therapeutic target.

A recent study showed that PD-L1 levels vary during cell cycle progression, with protein expression levels peaking in the M and early G_1 phases [365]. Using inhibitors against CDK4 and CDK6, they demonstrated that PD-L1 expression levels could be increased. My screen identified that knock-down of *CDKN2A,* which encodes the Cyclin D-CDK4/6 inhibitor p16, reduces PD-L1 expression, which is in line with the data from the aforementioned study. As *CDKN2A* is such an important tumour suppressor in melanoma, it would not be feasible to exploit the PD-L1-regulatory effects of inhibiting its function therapeutically.

My screen also identified a range of novel genes, including but not limited to *BLOC1S1*, *CNTLN*, *DUPD1*, *F8A3*, *GET4*, *GTF3C2*, *HRCT1*, *HTT*, *SHROOM1*, *SLC10A7*, *SPNS1* and *WRB*. Some of these genes could be involved in PD-L1 regulation through mechanisms mentioned in this chapter, or through other functions yet to be explored.

5.5 Small-scale validation of hits from the screen

As an initial validation, I chose to investigate genes from the top hits in the PD-L1 low fraction, whilst also ensuring my choices would cover diverse biological functions. Using this selection method, the top three N-linked glycosylation genes (*RPN1*, *DPAGT1* and *DDOST*), two intracellular transport genes (*VPS16* and *VPS33A*) and two basal transcription factor genes (*TAF8* and *TAF2*) were selected. In addition to these, I also picked *YKT6*, as it is linked to the main intracellular transport complexes identified in my screen (Fig. A.23): the HOPS, COG and TRAPP complexes. Finally, as the studies showing CMTM6-mediated regulation of PD-L1 [366, 367] were published at the same time as my screen was analysed, I also chose to include *CMTM6* in my validation experiments.

The individual knock-out of each of the 9 genes in the C092 cell line resulted in a decrease in cell surface PD-L1 expression 14 days post viral transduction for all genes studied (Fig. 5.12). The extent of PD-L1 surface expression loss varied between the samples, with a heterogenous effect on individual cells. As I did not establish a clonal population for this experiment, some fluctuations in the efficiency of knock-down will inevitably be recapitulated in the results. Nevertheless, as 9/9 selected genes for validation did cause a degree of

PD-L1 expression loss, this implies a high robustness in my original screen.

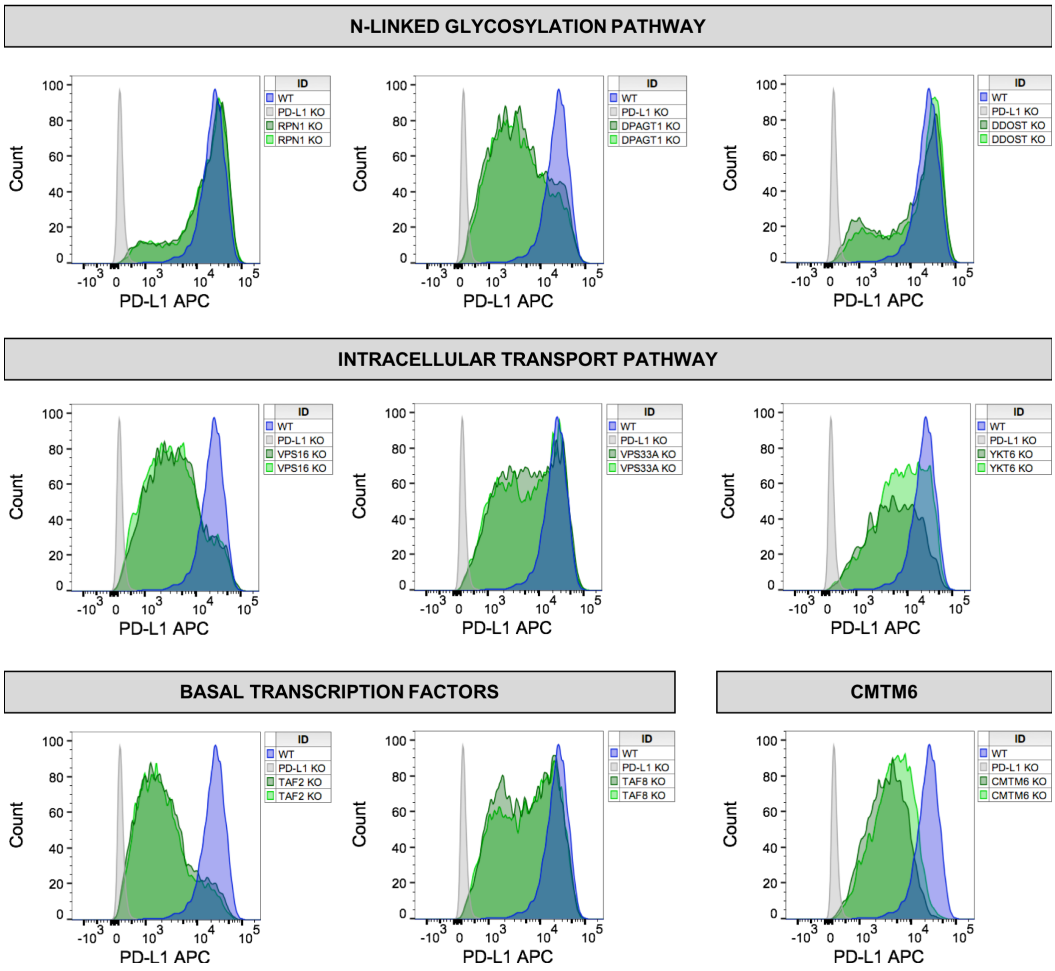


Figure 5.12: **Validation of nine genes in the C092 melanoma cell line, 14 days post viral transduction.** Histogram of PD-L1 expression measured by flow cytometry, showing the PD-L1 expression in the various conditions. Experiment was performed with technical duplicates.

Next, I wanted to assess the reproducibility of the hits, and chose to validate the same nine genes in an independent cell line. Based on endogenous PD-L1 cell surface expression, I picked the high PD-L1-expressing non-small cell lung cancer (NSCLC) cell line LCLC103H for this experiment. Although the extent varied between the cases, all nine genes when individually knocked-out by the CRISPR-Cas9 method showed some reduction in PD-L1 expression (Fig. 5.13), again confirming a high reproducibility of my original screen. Notably, in both LCLC103H and C092 cell lines, I observed a growth retardation and lower viability when *DPAGT1* or *YKT6* were knocked-out. These cells also acquired

a more spindle-shaped morphology, which is interesting as it is also a feature associated with cells undergoing epithelial-mesenchymal transition (EMT) [817–819]. In contrast, the *CMTM6*, *TAF2* and *TAF8* knock-out clones proliferated slightly faster than the other knock-outs.

Surprisingly, in the LCLC103H cell line, the amount of PD-L1 knock-down varied with time. Most genes caused a larger reduction in PD-L1 protein expression 9 days after viral transduction compared to 14 days (Fig. 5.13). This effect could be mediated by a range of different mechanisms. First, these two cell lines are of different cancer types, with varying doubling rates, genetic background and other cell line-specific properties . Second, and more importantly, a single clone was not selectively expanded, thus the knock-out experiments were performed on a non-clonal population of cells. Hence, the efficiency of the CRISPR-Cas9 machinery could vary between cells in this population, causing a variation in the degree of gene loss, leading to differences in cell proliferation and viability amongst individual clones. This could bias the population as less damaging gene disruptions are favourable in terms of survival, and these cells could outgrow clones targeted by more severe alterations. This could therefore contribute to the time-dependent discrepancy in the PD-L1 gene expression observed in my validation experiments. Third, PD-L1 expression is influenced by cell cycle regulatory factors as recently described [365]. I also observed a minor difference in PD-L1 expression based on cell confluency, where a higher confluency generally yielded slightly lower PD-L1 surface expression. It is possible that there is a joint effect where knock-down of some genes could enhance the cells' susceptibility towards cell cycle mediated regulation of PD-L1 expression. Additionally, in cases where protein expression is not completely abrogated by the CRISPR-Cas9 treatment, these cells could also show an increased sensitivity towards cell cycle mediated effects, causing the observed fluctuations in PD-L1 expression with time. Finally, it is also possible in cases where important or generic pathways are targeted, that cell-intrinsic compensatory mechanisms are in place to combat any such effect.

In hindsight, these validation experiments should all have been repeated in a FACS-selected and established monoclonal cell line with a proven deleterious homozygous knock-out in order to comprehensively assess their effects on PD-L1 expression. Nonetheless, the small-scale validation of the nine selected genes confirmed all tested genes could modulate PD-L1 expression to some extent, in the original melanoma cell line as well as a second cell line of lung cancer origin.

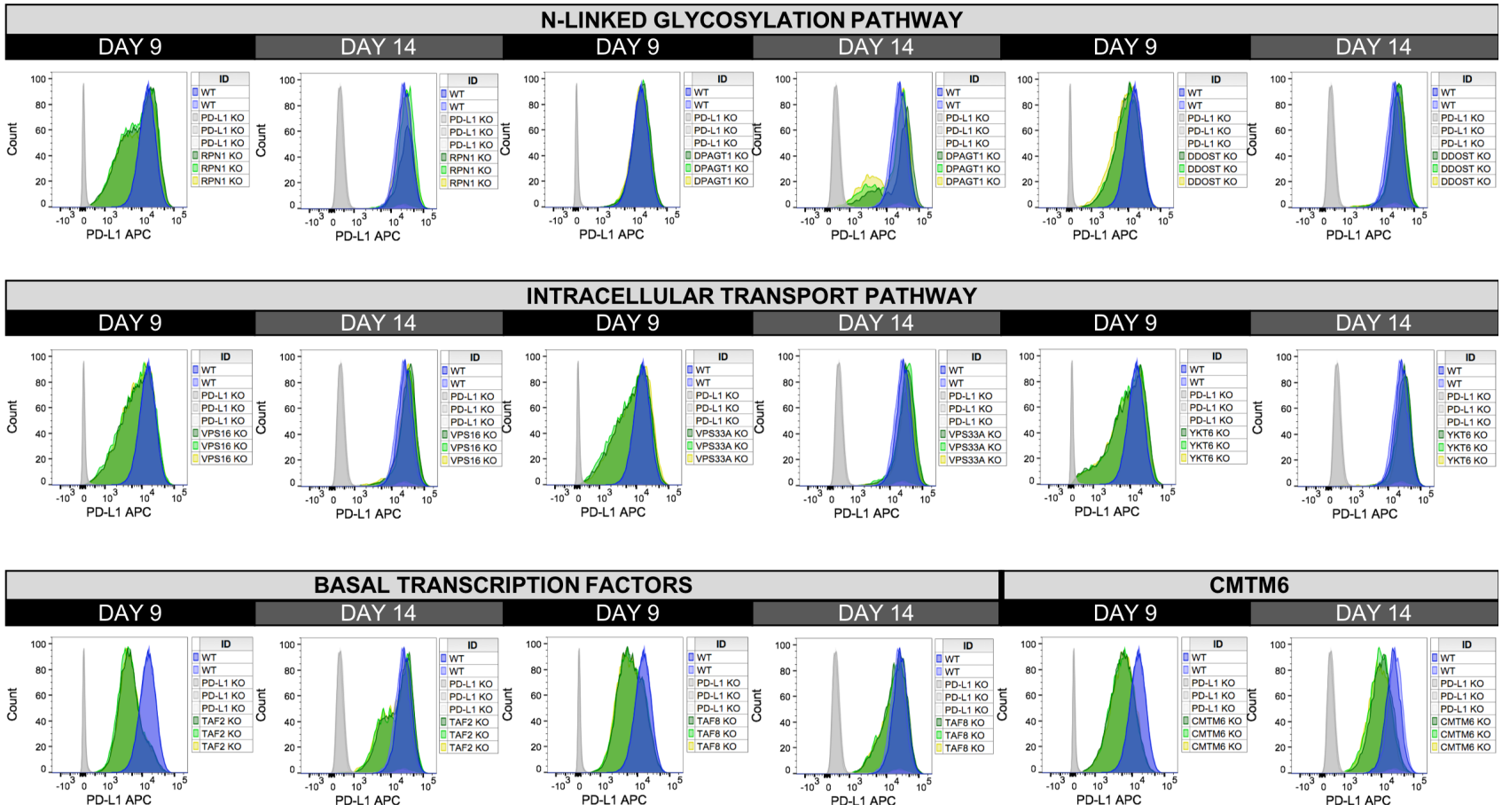


Figure 5.13: **Validation of nine genes in the LCLC103H lung cancer cell line.** Histogram of PD-L1 expression measured by flow cytometry, showing the PD-L1 expression in the various conditions. Experiment was performed with technical triplicates, and performed once for the 9-day timepoint and twice for the 14-day timepoint (post viral transduction). One representative example of the 14-day timepoint is shown.

5.6 Pooled validation using a custom CRISPR-Cas9 screening library

In conjunction with performing the small-scale validation, where the results showcased a high performance of my original screen as all nine genes validated in the two tested cell lines, I decided to perform a second more extensive validation. I designed a new validation screen library comprising 1000 gRNAs (Section B.4) to screen across 9 cell lines of 3 cancer types. Unfortunately, one NSCLC cell line lost Cas9 expression and had to be excluded from the screen, thus a total of eight cell lines were used in the validation experiments.

For each condition eight replicates were adopted, to ensure a high robustness of the experiment could be maintained even if some samples did not pass all sample processing and quality control steps. Three time points were selected (day 9, 14 and 28), as I experienced from the small-scale validation experiments that cell line-specific properties and technical issues such as confluency could impact the validation rate.

When selecting the cell lines to perform the pooled validation screen in, the following criteria were taken into consideration. First and foremost, the cell line needs to have an intermediate to high expression of PD-L1. This was challenging, as only a limited number of cell lines I had access to showed an ideal PD-L1 expression profile. Therefore I had to include a few cell lines with intermediate or broader expression. Second, I chose to focus on therapeutically relevant tumour types i.e. cases where PD-1/PD-L1 therapies are currently approved. I therefore decided to focus on three cancer types: melanoma, non-small cell lung cancer (NSCLC) and bladder cancer. Third, cell lines with adherent growth properties, reasonable cell size and doubling times were chosen. Finally, I selected cell lines that were used by other groups at Sanger, or could be easily accessible through collaborations. With these selection criteria in combination, I chose the three melanoma cell lines C092, SKMEL25 and UKEMEL118C, two NSCLC lines LCLC103H and HCC44 (a third cell line had to be excluded due to loss of Cas9 expression), and the three bladder cancer cell lines 5637, 647V and UBLC1 for my experiments. UKEMEL118C is a patient-derived melanoma cell line, gifted by Prof. Annette Paschen at Universitätsklinikum Essen (Section C.1). The PD-L1 protein expression of all cell lines used in the pooled validation screen, as measured by flow cytometry is shown in Fig. 5.14.

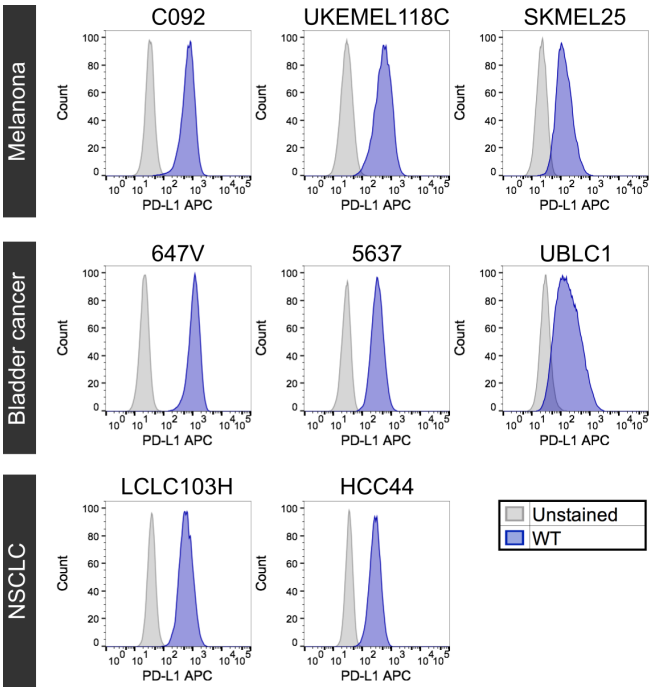


Figure 5.14: **PD-L1 expression in the cell lines used for validation.** Protein expression of PD-L1 measured by flow cytometry for the eight cell lines selected for the pooled validation screen (blue = wild-type (WT), grey = unstained). NSCLC = Non-small cell lung cancer.

5.6.1 Quality control

Unfortunately, when producing the gRNA library the Gibson reaction can introduce errors, especially at the junctions [820, 821], leading to an imprecise gRNA sequence being inserted into the plasmid. In my case, after sequencing the plasmid library, it was discovered that the first or last base of the gRNA sequence were mutated in 27% of the library. After this finding, a modified sequencing primer was designed and used to be able to identify exact gRNA sequences as well as erroneous gRNA sequences in all samples of the pooled validation screen. On average, 71% of all reads matched perfectly to the gRNA library, with an additional 12% of reads showing a G to A substitution in the first base, and thus would still work in transcription initiation [822]. This means the total number of useful gRNA reads amounted to 83% (Section B.6).

The screen was performed with eight technical replicates, obtaining a median MOI across all eight cell lines of 0.28 (Table A.24), thus ensuring a library representation of 1000x at infection. 55 out of 576 (10%) samples failed in the PCR amplification steps due to technical issues, and were not submitted for sequencing. Unfortunately, these included

all SKMEL25 day 14 dim samples, making analysis of this condition impossible. Additionally, five of the UBLC1 day nine low samples as well as six of the UBLC1 day nine control samples also failed, making it difficult to do a robust analysis for this cell line and time point. Luckily, in all other cases, at least four replicates could be maintained for each condition. A median sequencing coverage of 1200x was achieved (Fig. A.25); however, two samples had to be excluded due to low library coverage (SKMEL25 day 14 rep 6 low and 5637 day 9 rep 1 control). Correlation between gRNA counts in the control samples were high for most comparisons (Fig. A.26); however, slightly lower values can be expected due to increased noise with a small-scale library. In the cases where the sample correlation were lower, visual examination of plotted gRNA counts were performed to assess the outliers and no samples were excluded.

In summary, 519 out of 576 (90%) samples passed the quality control assessment and were used in downstream analyses.

5.6.2 Results

The goal with the pooled validation screen was to simultaneously verify the findings from the original screen, whilst exploring the robustness and biology of the hits by screening a custom 1000 gRNA library across eight cell lines of three tumour types, with analysis at three different time points. Mageck v.0.5.6 (Table A.1) was used to score the enrichment of all genes in the library in the sorted fractions compared to the controls. Each condition (cell line, time point and sort fraction) was analysed individually, followed by comparison of hits between groups.

In total, 54 genes were hits (FDR < 10%) in any of the conditions, with 35 genes identified in the low fraction and 53 genes in the dim fraction. As expected, *CD274* (PD-L1) was the top hit in most of the conditions. The recently discovered PD-L1-regulator *CMTM6* [366, 367] was also identified in the majority of the comparisons. Generally, most of the hits were identified in the dim fraction, as opposed to the low fraction, which could reflect the nature of PD-L1 control. When knocking down *CD274*, which encodes PD-L1, a complete abrogation in PD-L1 expression is observed, resulting in this gene being predominant in the low fraction but less detected in the dim fraction. The knock-down of other genes does not appear to have a similarly drastic effect on PD-L1 surface expression, explaining why there is an overall abundance of hits in the dim fraction.

I chose to mainly look at the profile of how hits validate across the different cell lines, irrespective of timepoint and sort fraction. This analysis provides information regarding the

Table 5.1: **Number of hits identified per cell line.** Statistics showing the total, dim or low number of hits (FDR <10%) found in each cell line, and its cancer type.

Cell line	Cancer type	Total number of hits	Number of hits in low fraction	Number of hits in dim fraction
C092	Melanoma	45	29	38
SKMEL25	Melanoma	12	5	10
UKEMEL118C	Melanoma	12	2	12
LCLC103H	NSCLC	9	9	1
HCC44	NSCLC	4	4	1
5637	Bladder cancer	15	5	14
647V	Bladder cancer	26	10	23
UBLC1	Bladder cancer	8	2	8

robustness of these candidates, and is independent of the degree of modulation and potential fluctuations with time. The number of genes identified across cell lines ranged between 4 and 45, with a median of 12 genes (Table 5.1). Some conditions could show lower validation rates due to technical issues, such as suboptimal cell confluences and limited number of sorted cells. As expected, the high PD-L1-expressing C092 melanoma cell line which the original screen was performed in, showed the highest number of validated genes. 45 out of the original 60 genes (75%) could be confirmed across any of the three timepoints assessed. The bladder cancer cell line 647V, which also had a high baseline PD-L1 expression, showed the second highest number of validated genes (26 genes). This does emphasise the importance of selecting cell lines with an optimal PD-L1 protein expression profile in order to get the highest possible window to identify regulators in a screen. This could also in part explain why another recently published PD-L1 screen only identified two genes (*CD274* and *CMTM6*) as PD-L1 regulators in their baseline experiment [367], as the PD-L1 expression in that cell line (BxPC-3) was only intermediate to low.

CD274 and *CMTM6* were the only two genes which validated across all eight cell lines (Fig. 5.15). These findings therefore strengthen the evidence of *CMTM6* being a general modulator of PD-L1 cell surface expression, holding true across multiple tissue types. Two genes validated across five cell lines, which included *VPS16,* an important member of the HOPS complex, and *SPNS1*, encoding a sphingolipid transporter which interestingly might play a role in lysosomal transport and autophagy [823–825]. Looking in a tissue-specific context, other interesting observations were made (Table A.27). The glycosylation-associated gene *DAD1,* validated across all melanoma cell lines but not the other tissue types, whilst *TAF2, TAF8* and *VPS39* validated in two out of three melanoma and bladder cell lines, respectively, but not in any of the lung cancer cell lines. These findings could

possibly reflect a tissue-dependent context of PD-L1 regulation; however, a negative result most likely reflects the lack of resolution to identify genes in some cell lines.

Looking at the hits in each cell line, some additional observations were made. In the melanoma cell line C092, the three main general processes identified in the original screen (early glycosylation, basal transcription factor regulation and intracellular transport mediation) could be confirmed as pathways regulating PD-L1 cell surface presentation (Fig. 5.16). Some other genes were also validated, including *SPNS1*, which was a hit across multiple cell lines of different tissue types. In the other two melanoma cell lines, glycosylation-associated genes and *TAF* genes validated in the SKMEL25 cell line, and a large number of *VPS* genes could be identified in the UKEMEL118C cell line. Collectively, these results reflect the importance of these basal processes in the regulation of PD-L1 in melanoma cells, but also suggest slight cell line specific variations exist in the dependency of respective pathways.

Out of the three bladder cancer cell lines, 5637 and 647V yielded many significant hits, possibly owing to their high endogenous expression of PD-L1, facilitating the discovery of PD-L1 regulators. UBLC1 had the poorest baseline PD-L1 expression out of all cell lines in the validation screens, and it is therefore not surprising to find a very sparse number of hits in this cell line. One common theme in the 5637 cell line appears to be the intracellular transport pathway including movement through lysosomes. Additionally, genes associated with transcriptional control such as *ARID1A, HDAC3* and *YAP1* were also identified in the dim fraction, across multiple time points. However, neither the N-linked glycosylation genes nor the basal transcription factor *TAF* genes were identified in this cell line. The 647V cell line had the second highest number of hits, after the original screen cell line C092. Genes identified in this cell line cover glycosylation, *TAF* genes and transcriptional regulators, *VPS* genes and intracellular transport mediators, as well as a few other hits such as *TSC2, CDKN2A, KIAA1432* and *WRB*. Most of the genes were identified in multiple conditions per cell line, providing added support for the robustness of these findings. The identification of tumour suppressors *CDKN2A* and *PTEN* is curious, and could arguably reflect an obtained survival advantage of the cells which lost expression of these genes. However, if this would be the only effect, these clones should be enriched in the control fraction but not specifically found in the sorted fractions. Therefore, these findings suggest they mediate some level of PD-L1 regulation; however, therapeutic silencing of these genes in a clinical setting would most likely not lead to a favourable outcome.

The NSCLC cell lines had fewer hits than the other tissue types, which was unexpected in terms of their satisfactory PD-L1 protein expression profiles. In the LCLC103H cell

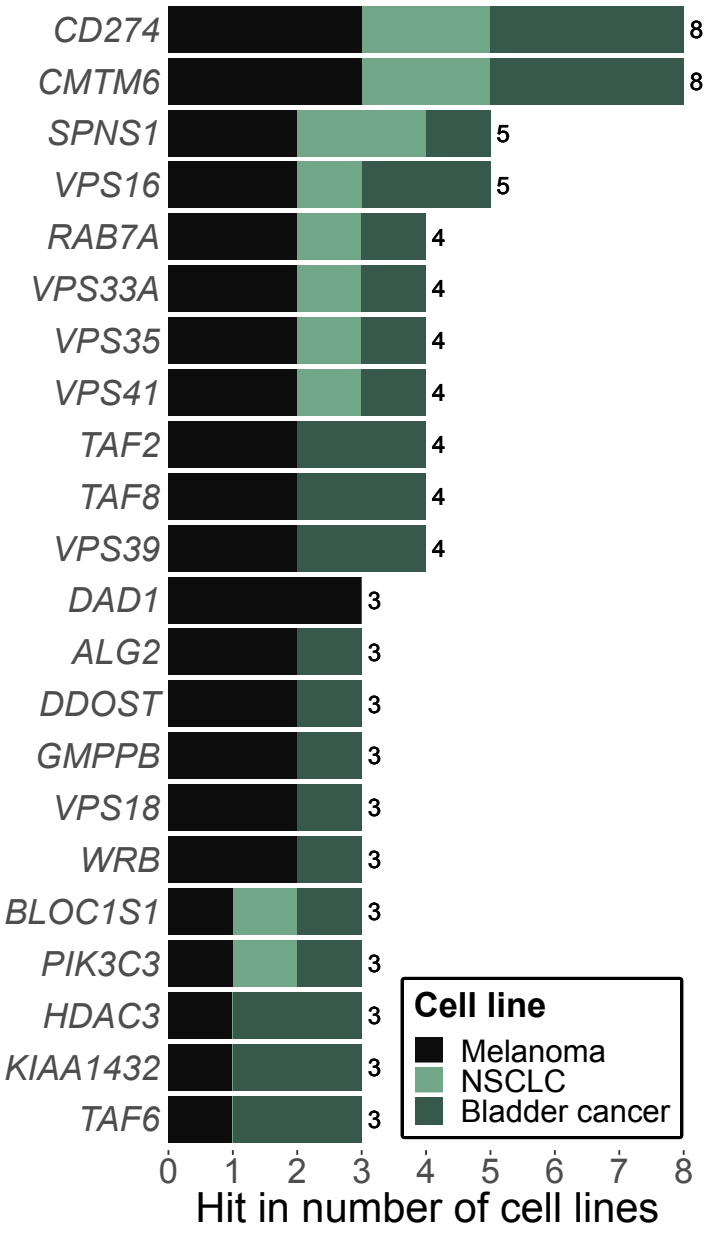
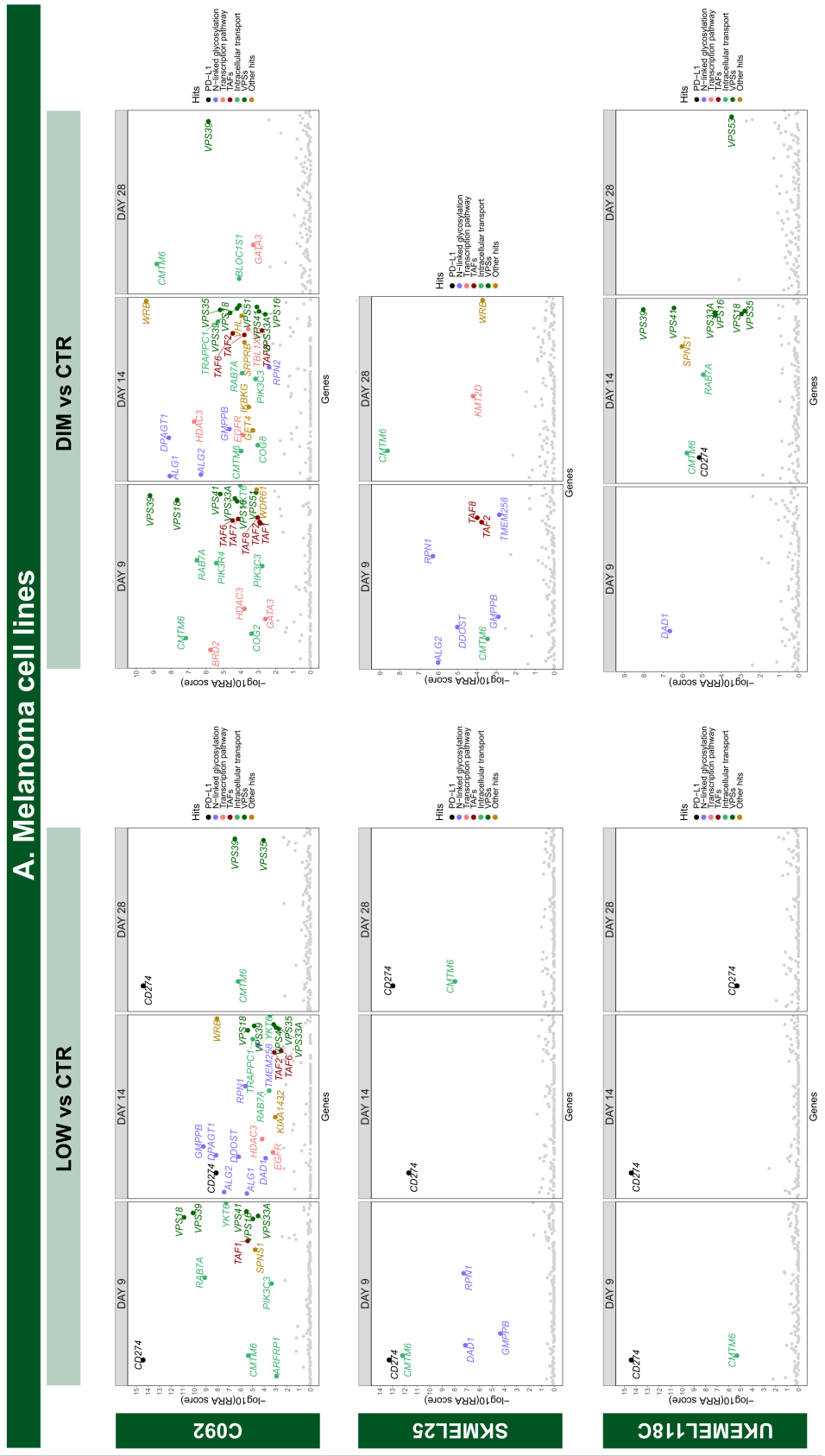
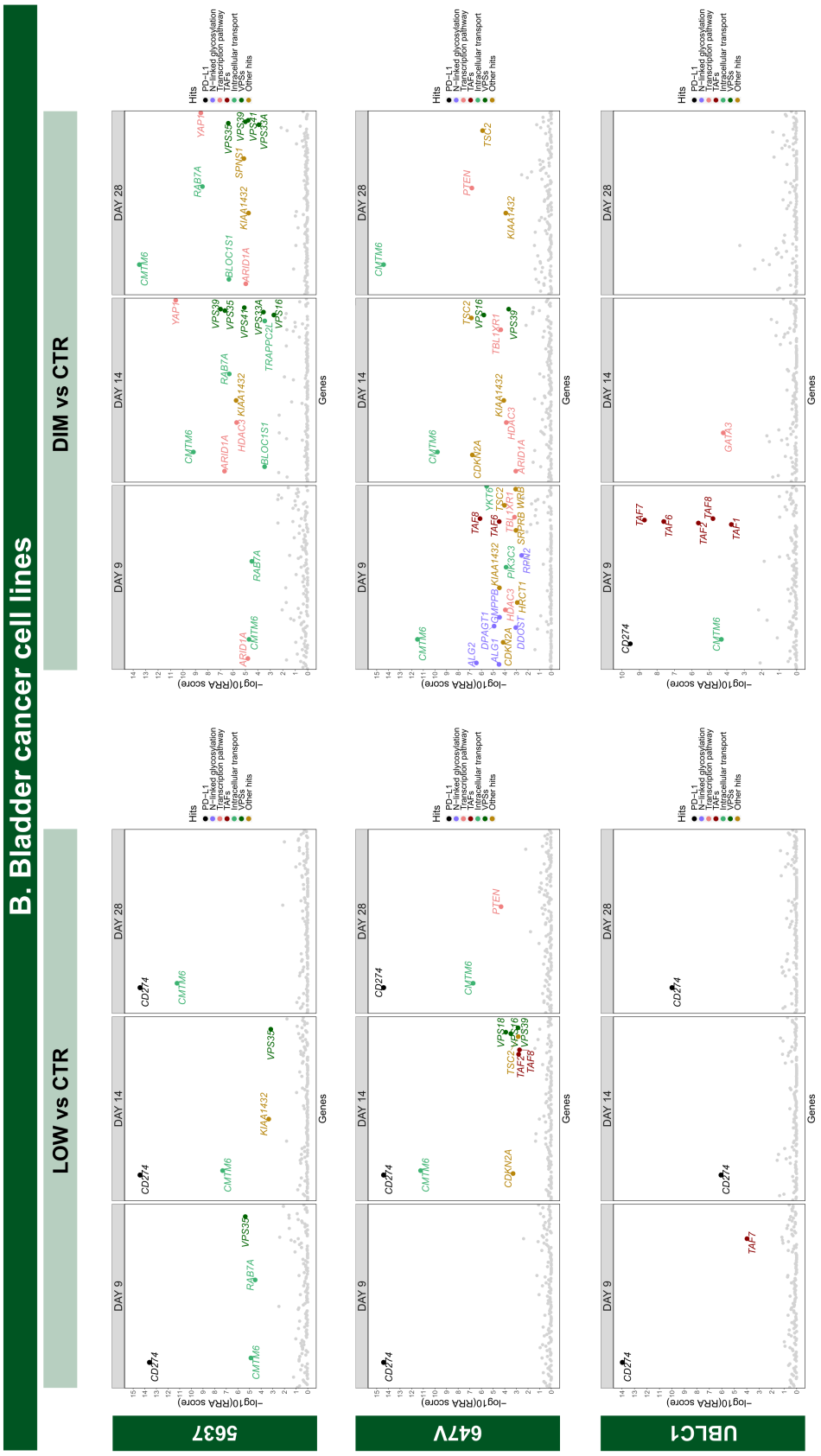


Figure 5.15: **Gene validation results across cell lines.** The number of cell lines where each gene was successfully validated in is shown for genes validating in at least three cell lines. A more detailed and extended list can be found in Table A.27.


A. Melanoma cell lines
LOW vs CTR
DIM vs CTR
C092
SKMEL25
UKEMEL18C
DAY 9
DAY 14
DAY 28
Genes
−log10(RRA score)
Hits
PD−L1
N−linked glycosylation
Transcription pathway
TAFs
Intracellular transport
VPSs
Other hits


B. Bladder cancer cell lines
LOW vs CTR
DIM vs CTR
5637
647V
UBLC1
DAY 9
DAY 14
DAY 28
−log10(RRA score)
Genes
Hits
PD−L1
N−linked glycosylation
Transcription pathway
TAFs
Intracellular transport
VPSs
Other hits

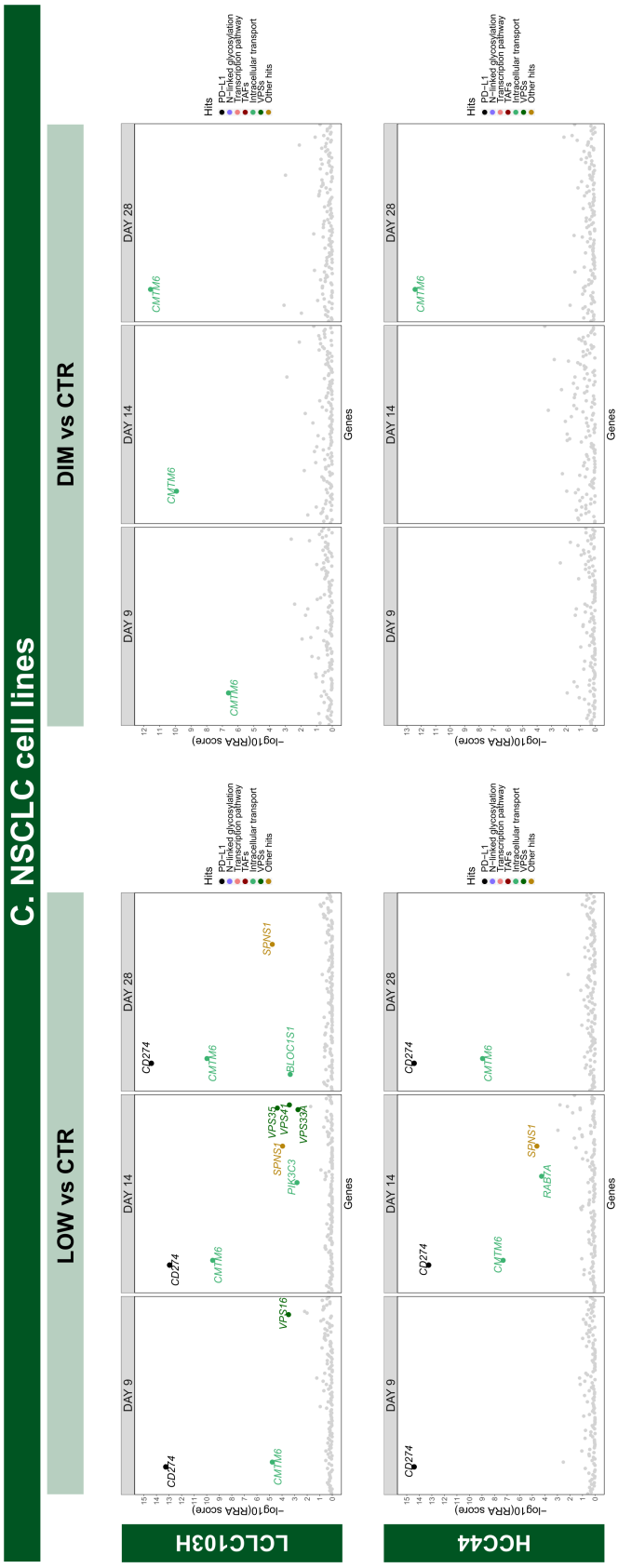


Figure 5.16: **Regulators of PD-L1 cell surface expression identified in the pooled validation screen. A)** Hits identified in the melanoma cell lines C092, SKMEL25 and UKEMEL118C (FDR < 10%). **B)** Hits in identified in the bladder cancer cell lines 5637, 647V and UBLC1 (FDR < 10%). **C)** Hits identified in the NSCLC cell lines LCLC103H and HCC44 (FDR < 10%). Genes classified as transcription factor pathway genes were obtained from the Reactome generic transcription pathway gene set, the intracellular transport genes from the GO biological process term intracellular transport (GO:0046907) and the N-linked glycosylation genes from the GO biological process term protein N-linked glycosylation (GO:0006487) but with *GMPPB* added.

line, genes associated with intracellular transport dominated. HCC44 only had four hits in total, where apart from *CD274* and *CMTM6*, *RAB7A* and *SPNS1* were also identified. These genes are all presumed to play a role in lysosomal transport, providing further support of this process in promoting PD-L1 surface presentation. The lack of validation of glycosylation and basal transcription factor genes in these two cell lines could be reflective of differences in cell-specific properties; however, as the total findings were very limited in these cell lines, false negatives cannot be ruled out.

The eight cell lines show different patterns in the type of genes which control PD-L1 expression, but also in the dynamics of their PD-L1 regulation. Genes associated with intracellular transport including *VPS* genes were robust hits in several cell lines irrespective of the time point studied. However, N-linked glycosylation genes were only hits at day 14 in the C092 cell line, whilst in the SKMEL25 and 647V cell lines they validated at day 9 only. This discrepancy could potentially reflect cell-intrinsic differences, including growth properties such as doubling rate.

The pooled validation screen succeeded in confirming individual genes as well as general pathways involved in the regulation of PD-L1. In general, three main biological processes were found to control PD-L1 expression in the various cell lines. In some melanoma and bladder cancer cell lines, *TAF* genes are involved in the control of PD-L1. Furtheremore, my results also suggest that in several cell lines, PD-L1 is targeted by N-linked glycosylation, which could possibly affect the folding and stability of the protein, explaining its lack of cell surface expression when this machinery is inoperative. PD-L1 then gets processed through the intracellular transport apparatus, including transitioning through lysosomes. Targeting these three general pathways could be a method to reduce PD-L1 expression to facilitate a host anti-tumour response, which could be tested using inhibitors such as tunicamycin (glycosylation) or chloroquine (lysosomal transport). However, the clinical utility of targeting these broad pathways warrants caution, as these systems are likely to regulate the expression of a plethora of other cell surface proteins as well as PD-L1. To conclude, my PD-L1 screen and follow-up validation have provided insights into how PD-L1 expression is intrinsically regulated, processed and presented on the surface of a tumour cell.

5.6.3 *SPNS1* as a novel regulator of PD-L1

Some novel hits which were not linked to the above described generic processes were also identified (Fig. 5.17). *Sphingolipid Transporter 1 (SPNS1)* was the most interesting hit, being a recurrent finding across five out of eight cell lines. In the melanoma cell line C092

and both the NSCLC cell lines, knock-out of *SPNS1* caused sufficient removal of cell surface PD-L1 for these cells to be identified in the low fraction. Additionally, *SPNS1* was also identified in the dim fraction of the melanoma cell line UKEMEL118C and the bladder cancer cell line 5637. It is not possible to conclude whether the absence of hit discovery in other cell lines only reflects the lack of detection capability in some settings. Several conditions had *CD274* or *CD274* and *CMTM6* as the only significant genes identified as hits, which could be due to a low number of replicate samples such as for the UBLC1 cell line at day 9 or 14, and the 647V cell line at day 9. In the original C092 screen, *SPNS1* was identified as a hit in the dim fraction (FDR-adjusted p-value = 0.064), which was the rationale for including this gene in the gRNA library used for the pooled validation screen.

The *SPNS1* validation profile suggest it is a generic regulator of PD-L1 across multiple cell lines and tissue types. The *SPNS1* gene encodes the Sphingolipid Transporter 1 (Putative), and is proposed to function in lysosomal transport and autophagy [823–826]. Genetic defects of *SPNS1* homologues in mice, zebrafish and *Drosophila* have been reported to present phenotypes associated with ageing, viability and nervous system defects [823, 824, 827–829]. Some of these mutants show altered endosome-to-lysosome trafficking, senescence and programmed cell death phenotypes or aberrant autolysosomal formation. I therefore hypothesise, that *SPNS1*-mediated control of lysosomal turnover and autophagic processes might contribute to its regulation of cell surface PD-L1 levels. Defective *SPNS1* might cause PD-L1 to accumulate in intracellular vesicles, thus hampering cell surface presentation and turnover of PD-L1. Overexpression of *Spin* in *Drosophila* triggers autophagy [826], while hypomorphic loss leads to accumulation of enlarged autolysosomes and impaired autophagic lysosome reformation [825, 830]. The capability of autolysosomes to degrade its contents were only impaired in *Spin* mutants following prolonged starvation, whilst in nutrient-rich conditions, the effect of *Spin* knock-down on autophagy were ameliorated [825]. It is therefore intriguing to compare how these studies performed under conditions following starvation, resembles that of a hypoxic and nutrient-deprived environment during tumour progression [831, 832]. In addition to reduced cell surface PD-L1 expression, knock-down of *SPNS1* might therefore confer a milder phenotype in normal cells compared to nutrient-deprived tumour cells.

SPNS1 is expressed in various tissues of both normal (Fig. A.28) and cancerous origin (Fig. A.29). The most common genetic events targeting *SPNS1* in melanoma, NSCLC and bladder cancer patients were mutations and amplifications (Fig. 5.18A). Additionally, in the TCGA SKCM and TCGA BLCA cohorts, high expression of *SPNS1* is associated with worse survival (Fig. 5.18B). In the the Leeds melanoma cohort, high expression of

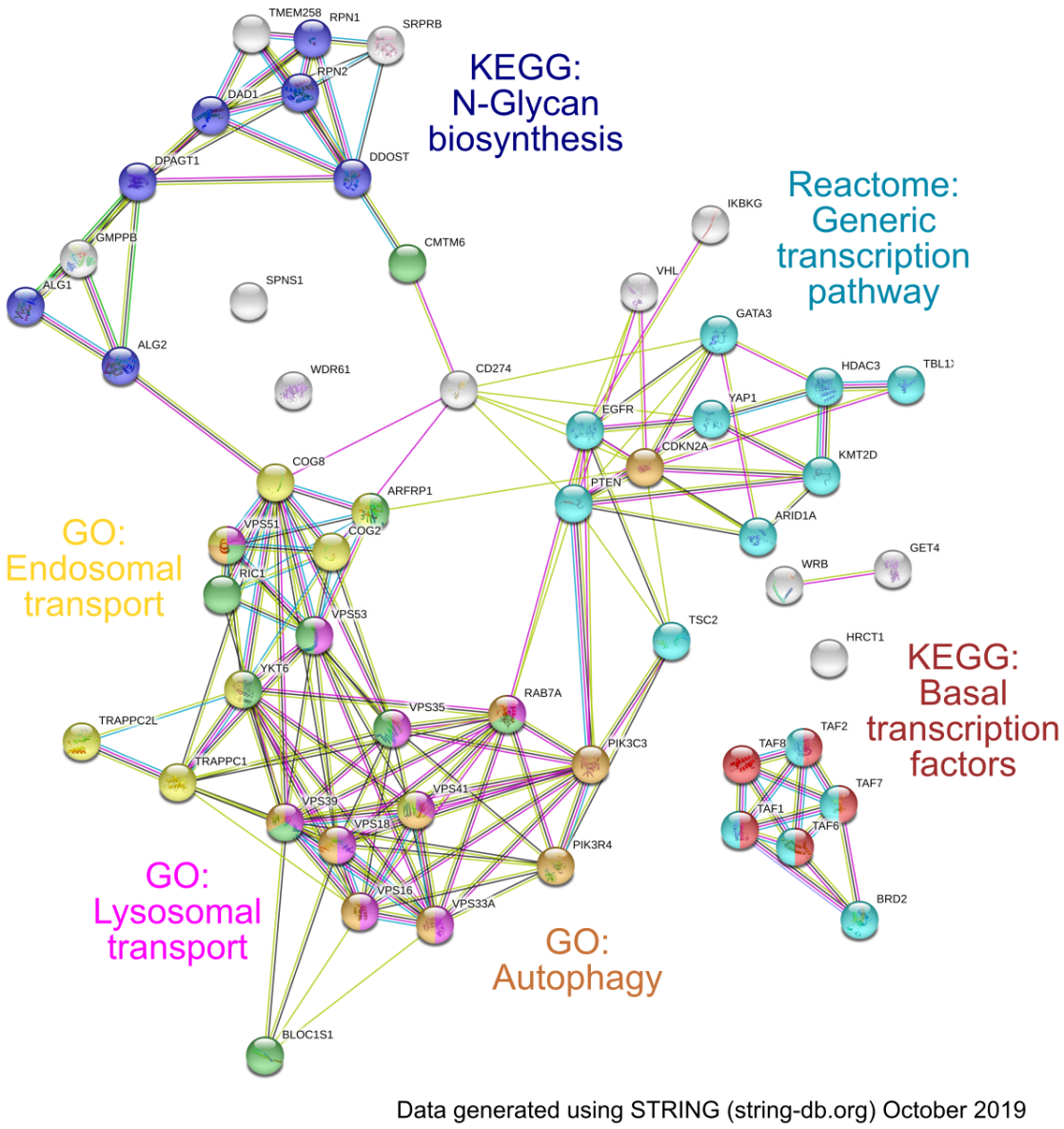


Figure 5.17: **Output from search tool for the retrieval of interacting genes/proteins (STRING) analysis of pooled validation screen hits.** All 54 genes identified as hits (FDR < 10%) in any of the conditions studied, with selected general pathways or processes highlighted. *SPNS1* was selected for follow-up partly because of it not being linked to any other hits in the screens in this analysis.

SPNS1 was also found to correlate with poor outcome, both when studying overall survival (p.value = 0.031) and melanoma-specific survival (p.value = 0.031); however, the biggest difference was found when comparing relapse-free survival (Fig. 5.18B). Collectively, these data suggest *SPNS1* function might be associated with more aggressive tumours.

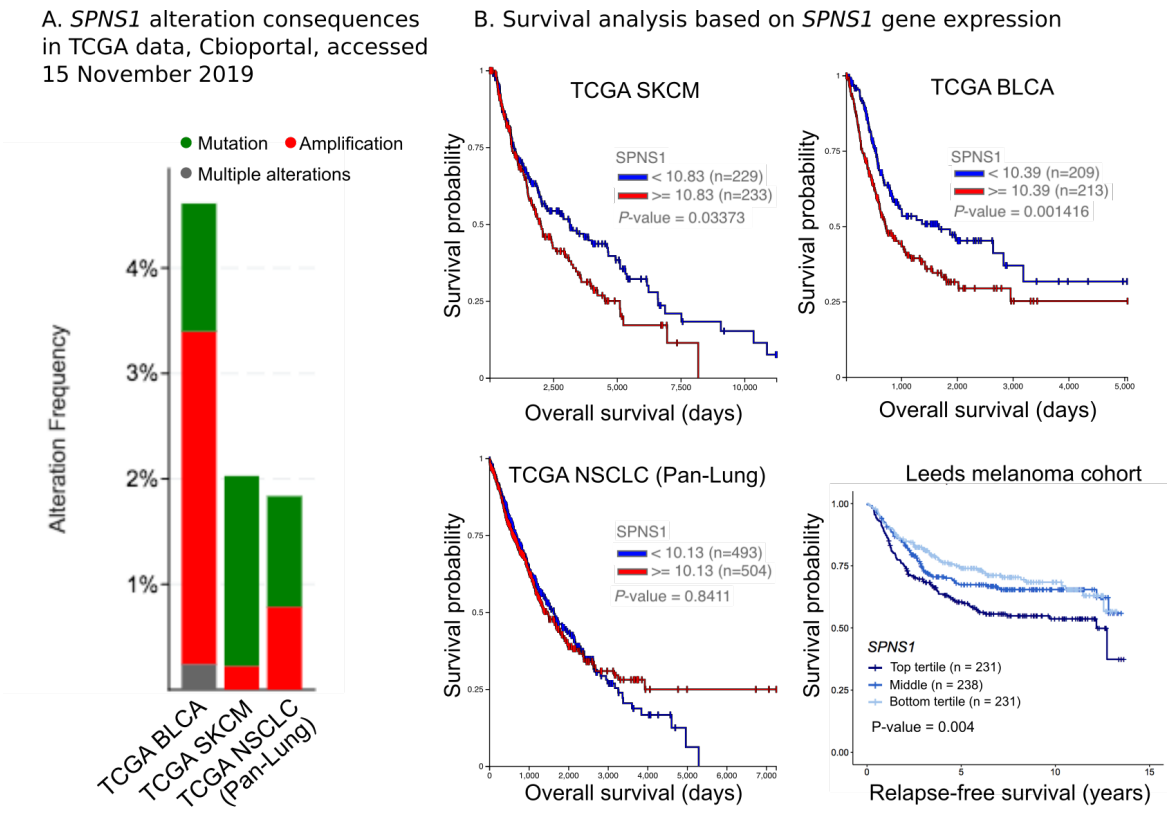


Figure 5.18: ***SPNS1* alterations in human cancers and its association with patient survival. A)** The most common genetic alterations in *SPNS1* in human bladder cancers, melanoma and NSCLC were amplifications or mutations. **B)** Univariate survival analysis show high *SPNS1* gene expression is associated with worse outcome in melanoma (TCGA SKCM and Leeds melanoma cohort [833]) and bladder cancer (TCGA BLCA). No survival difference was found in NSCLC (TCGA Pan-Lung). TCGA figures were generated using the UCSC Xena platform (www.xenabrowser.net) [834].

Before proceeding to unravel the biology behind *SPNS1*-mediated PD-L1 regulation, this finding needs to be further validated using individual and clonal knock-out experiments. In the first instance, PD-L1 expression following *SPNS1* loss should be confirmed in a range of cell lines, including those part of the validation screen panel. When this effect has been assessed, the mechanism behind its regulation can be further studied. Any direct interac-

tions or co-localisation of this sphingolipid transporter with PD-L1, as well as the effect of where PD-L1 accumulates when *SPNS1* is absent, can be tested using e.g. specific intracellular localisation markers such as RAB7 (late endosomes), LAMP1 (lysosomes) and LC3 (autophagosomes).

Furthermore, a range of compounds which block various vesicle functions could be used [835], including chloroquine, which inhibit the fusion of lysosomes with autophagosomes, and wortmannin, a PI3K inhibitor, which inhibits autophagy by blocking the formation of autophagosomes. As autophagy is an important cellular process, it would be important to also profile the specificity of *SPNS1* regulation to understand if it modulates protein expression on a global level. This would be an important factor in determining the druggability of *SPNS1* and the clinical utility of targeting PD-L1 through this interaction. With that being said, even if *SPNS1* control is not PD-L1 specific, this does not necessarily rule out the possibility of targeting its function. Interesting data have emerged which show that targeting autophagy on a broad level can mediate anti-cancer effects and it was recently proposed as a new strategy of cancer therapy [836]. Ultimately, the effect of *SPNS1* depletion on the ability to sensitise tumour cells towards immune evasion would be essential to study in co-culture or *in vivo* settings.

5.7 Evaluation of chapter aims

- Selection of experimental conditions best suited for my project
  - C092 was established as the most preferable cell line to conduct the screen in due to its high and uniform endogenous PD-L1 expression, optimal growth characteristics and verified high Cas9 efficiency.
  - To conduct an unbiased screen, a genome-wide gRNA library (Yusa human v.1.1) with conditions ensuring 200x library representation, and a final timepoint of two weeks were selected.
- Confirm adequate controls are set up and their use validated
  - A *CD274* knock-out C092 cell line was generated as a positive control, and the knock-out validated using the Surveyor mismatch cleavage assay. This cell line showed a drastic reduction in PD-L1 cell surface expression, showcasing a high resolution for the discovery of PD-L1-regulating genes and pathways.

  - An *OR14A16* knock-out C092 cell line was generated as a negative control, and the knock-out validated using the Surveyor mismatch cleavage assay. This cell line did not show any difference in PD-L1 cell surface expression compared to the parental wild-type cell line, confirming the negligible impact of the CRISPR-Cas9 screening process on PD-L1 expression.

- Verify that the screening conditions are appropriate for hit discovery
  - Low number of zero gRNA counts, where the majority of gRNAs lost were sample-specific confirmed that appropriate maintenance of library complexity was achieved.
  - High correlation between control replicates, verified dropout of essential genes and a high ROC curve AUC value demonstrates a robust and high-performing screen.
  - *CD274* (encoding PD-L1) was the top hit in both the original screen and pooled validation screens, further confirming the success of the screens.
  - Small-scale validation using individual gRNA knock-down confirmed PD-L1 regulation of all nine tested genes in two cell lines, showcasing a high reproducibility of the original screen.
  - The pooled validation screens using the 1000 gRNA library in eight cell lines of three tissue types showed acceptable control sample correlations, albeit lower than the original screen. The smaller screens yielded higher noise, which was reflected in the lower number of hits identified in many of the conditions assessed. Absence of hit validation in some conditions could therefore reflect an insufficient detection ability.

- Identify the various intracellular pathways by which PD-L1 is presented on the cell surface
  - Three main categories of genes of general pathways involved in PD-L1 presentation on the cell surface of cancer cells were identified:
    - ≻ Basal transcription factor genes: Five *TAF* genes encoding proteins which make up the C lobe of the TFIID complex were identified in the original screen, suggesting this part of the TFIID complex could be particularly important for transcriptional regulation of PD-L1. Some of these hits also validated in additional melanoma and bladder cancer cell lines.

  - ≻ N-linked glycosylation genes: My screen confirmed the need of PD-L1 to be properly glycosylated for cell surface display, where early glycosylation processes mediated by *ALG* genes and the OST complex are particularly important. Some of these hits also validated in additional melanoma and bladder cancer cell lines.
  - ≻ Intracellular transport genes: Hits from my screen show how PD-L1 transitions from the ER to the cell surface through the Golgi, endosomes, lysosomes and autophagosomes, where it is regulated by complexes including TRAPP, COG and HOPS. Genes associated with intracellular transport including *VPS* genes were found in all eight validation screen cell lines, proving the importance of these generic processes for proper PD-L1 expression.

- Assess the novel hits and their potential druggability
  - Epigenetic modulation of PD-L1 by *BRD2*, *HDAC3* and *TBL1XR1* were also discovered in my screen. *BRD2* only validated in the original C092 cell line, whilst *HDAC3* and *TBL1XR1* was confirmed in two and one bladder cancer cell lines, respectively, in addition to C092. HDAC and BET inhibitors might therefore be beneficial in conferring a reduced PD-L1 tumour expression to trigger an immune response.
  - *CDKN2A* was a hit in the original screen, and validated in the 647V cell line. The role of cell cycle on PD-L1 expression is supported in literature, but the benefit of reduced PD-L1 expression would most likely not overcome the detriment of tumour suppressor silencing in a clinical setting.
  - Several other novel genes were found, but from the validation screens *SPNS1* was deemed most interesting to follow up.

- Present top candidate hits for follow-up validation
  - *SPNS1* was a novel hit which validated across five cell lines, and is postulated to be involved in autophagy. It is commonly amplified or deleted in cancers, where a high expression is linked to a worse prognosis. This candidate was selected for follow-up experiments which are currently ongoing.

Part III

Conclusion

Chapter 6

Discussion

The overarching aim of my dissertation was to improve our understanding of melanoma as a disease: from the genetic processes that govern tumour development, to the impairment of host protection and how we can reverse such a dysfunctional state.

In the first part, I presented some of the most critical genetic aberrations found in primary melanomas. I designed my own targeted capture bait library to find alterations in genes driving melanoma disease progression, and to understand their contribution to the disease. Large cohorts of primary melanomas have not been comprehensively studied by genomic profiling before, therefore my study can help us understand how cancers evolve and are regulated in their earliest stages. Additionally, my study has the power to validate or contradict previous findings, as well as present new insights and knowledge. The first results chapter, Chapter 2, presented the methods, showed the reliability of the data generated for this project and how I dealt with caveats and limitations. In the following two chapters, Chapter 3 and 4, I investigated key genetic alterations in human primary melanoma to an extent previous research has not been able to, whilst using the unique characteristics of the Leeds melanoma cohort.

In Chapter 5, the second part of my thesis, I identified regulators of tumour PD-L1 surface expression using a CRISPR-Cas9 screening approach. First, I established the optimal screening conditions and controls, followed by the demonstration of a successful genome-wide CRISPR-Cas9 screen. I discussed the results, which covered general pathways controlling PD-L1 expression, novel regulators and new plausible therapeutic angles. To validate my results, I simultaneously performed a smaller pooled validation screen across eight cell lines of three cancer types, and could confirm many of the hits from the original screen, as well as identify one top candidate *(SPNS1)* for follow-up experiments.

In the following sections, I will highlight and discuss some of the key findings presented

throughout my thesis, and how this has improved our knowledge about the genetic landscape and immune regulation of melanoma.

6.1 Novel genetic alterations in primary melanomas

In my dissertation, I showed that most primary melanomas have a high mutation burden, replete with UV damage-associated C to T mutations. Mutations were frequently found in driver genes such as *BRAF, NRAS, NF1, TERT* and *CDKN2A*, but I also discovered hotspot variants in novel genes including *PCDHA2*, *TPTE* and *AHCTF1.* Based on positive selection, 15 melanoma driver genes were found comprising mainly well-established melanoma driver genes. However, my analysis also validated less acclaimed prospective driver genes *FAM58A, RQCD1* and *MSR1*, as well as one novel candidate gene *TPTE*. In support of my findings, a recent sophisticated method of driver gene discovery also identified *TPTE* as a driver using pan-cancer or melanoma-specific TCGA data [837]. Interestingly, my analysis also discovered a mutual exclusive mutation pattern which proposes that both *PTEN* paralogues *TPTE* and *TPTE2* might be involved in the same tumourigenic pathways as *PTEN,* showcasing a possible functional redundancy. However, whilst all of the mutation calls in this gene were high quality, there appears to be multiple copies of *TPTE* pseudogenes across the genome, a feature known to potentially confound variant detection. In spite of this, my data suggests a possible role for *TPTE* as a novel driver of melanoma development, an observation that needs to be explored functionally and lies beyond the scope of my thesis.

A more comprehensive analysis of mutation patterns identified mutual exclusivity between two key melanoma driver genes, *BRAF* and *CDKN2A,* respectively, and several other known or novel genes. The pattern between *BRAF* and *EGFR* for example*,* could indicate alternative activation of the MAPK pathway driving oncogenesis in some patients, a feature that could be utilised for personalised treatment using already approved drugs. Another interesting pattern was observed between *CDKN2A* and *PRDM2*, where the Rb-binding capability of *PRDM2* provides further support of a reciprocal function between the two genes. Additionally, experimental evidence showed that *PRDM2* might be involved in cell cycle regulation [584, 585]. It would be very interesting to further study this hypothesis using knock-out models and senescence bypass assays.

Unsurprisingly, co-occurring gene pairs were not observed in my dataset. The high mutation rate in melanomas increases the likelihood of any two genes being co-mutated by chance. Although I found no such interactions in my dataset, I believe co-occurring mutations in theory should exist, and is important to study. However, these associations

might be far more complex to unravel. As an example, they might be visible only on a pathway rather than gene level. Therefore, the discovery of such interactions through this type of analysis will prove extremely challenging. Nevertheless, studies in other cancers have shown that patients showing alterations in multiple driver genes tend to have inferior survival [631]. But again, similar studies in melanoma will be more difficult, as any effect could be clouded by the high background mutation rates.

Copy number profiles in primary melanomas mimicked those of metastatic melanomas, with chromosome 1q, 6p, 7 and 8 being frequently amplified whilst 6q, 9 and 10 were more often deleted. *IRF4,* located on chromosome 6p, was of particular interest. In addition to being one of the most frequently amplified genes in my dataset, *IRF4* associated with lethality in melanoma cell lines by analysing the DepMap CRISPR-Cas9 screen dataset. I therefore hypothesised that *IRF4* amplification could be an oncogenic event in melanoma, whereby the loss of *IRF4* expression might cause tumour cell vulnerability. This hypothesis was in part confirmed through experimental validation, as siRNA-mediated knock-down of this gene in an $IRF4^{high}$ cell line resulted in apoptosis and cell death. My data therefore suggest that a subset of melanomas might be dependent on *IRF4* expression, which suggests *IRF4* could be an interesting therapeutic target.

Any type of analysis involving genetic alterations comes with several challenges. Studying less known cancer genes is difficult because we cannot fully ascertain the functional consequence of each mutation. Generally, loss of function mutations such as nonsense, splice site and frameshift mutations target tumour suppressor genes, while recurrent missense mutations tend to cluster in oncogenes and activate them. However, this is not always true, as for example hotspot missense mutations causing disruption of gene function in the tumour suppressor gene *CDKN2A* are commonly found. Moreover, mutations of the same type, but in different positions or with different amino acid substitutions could affect the function of the gene in different ways and to varying degrees. As the impact of each mutation is difficult to predict, analyses involving driver gene discovery or mutational pattern assessment will be noisy and difficult to interpret. Comparably, large copy number events spanning multiple genes makes it difficult to assess the pathogenicity involving any of those genes in isolation, which further complicates the accuracy by which key drivers behind such events can be assigned. Additionally, there might be a dosage effect as heterozygous and less damaging alterations might not completely abrogate the function of a gene or pathway. In those cases, the occurrence of a mutation in a second component of the same pathway might be favourable for the tumour. All of these factors play a role when we try to identify the most important events driving tumourigenesis, and there is still much to learn in this

field.

Sun exposure is an important contributor to melanoma development. Another consequence of UV-driven high mutation burden in tumours, is the preference towards specific alterations in active ETS transcription factor binding sites. I showed that with the exception of *TERT* and *RPS27,* all other top recurrent promoter variants in melanoma, including both known and novel positions, overlapped with this UV damage-associated pattern. This emphasises how we cannot assign such events as pathogenic drivers purely based on recurrence. These findings require additional functional proof of the link between such modifications and melanoma development.

In summary, my work has comprehensively outlined the landscape of genetic alterations in primary melanoma. The pathways that govern primary tumours largely reflect that of metastatic melanomas, and my analyses have validated the importance of many known and prospective driver genes. In addition to this, I have also presented novel genetic alterations and how they might play a role in disease progression. Given these results, *IRF4* and *EGFR* are interesting candidates to further explore in the context of personalised melanoma therapy. The role of the *PTEN* paralogues *TPTE* and *TPTE2*, as well as the potentially functionally equivalent genetic alterations seen in *CDKN2A* and *PRDM2* are other compelling findings which warrant further investigation.

6.2 Melanoma heterogeneity

Heterogeneity across melanoma has resulted in the adoption of several classifications systems, including histopathological, mutational and UV-associated categories. Several of these existing classifications share commonalities, and it would be beneficial to stratify patients based on a combination of such features. Therefore, I have comprehensively studied the genomic composition of different primary melanomas, highlighting both similarities and clear differences between tumours.

The genomic landscape across melanomas originating from different body sites are distinct, where specific features can be found for each group. Patients with acral and mucosal melanomas had both a lower frequency of mutations and a difference in driver genes, compared to non-acral cutaneous melanomas. *BRAF* hotspot mutations were not found in a single mucosal melanoma case; instead, these tumours seemed to be more influenced by amplifications in chromosome 8q (*MYC*) and 6p. *BRAF* and *NRAS* hotspot mutations were present in acral melanomas, but acral tumours also harboured frequent amplifications targeting *MYC* and cell cycle regulatory genes such as *CCND1* and *CDK4*. Several candid-

ate melanoma driver genes, including *PTEN*, *RAC1, RB1, DDX3X* and *PPP6C* were not found mutated in any cases of acral or mucosal melanoma, in my cohort nor in the recently published large cohort of melanomas [91]. Another interesting observation was that driver genes in the rarer melanoma subtypes tended to be affected by copy number changes rather than mutations. This feature was consistent across the whole genome, with acral and mucosal melanomas generally presenting more frequent copy number alterations. These distinct differences in part reflect the varying level of sun exposure commonly received at the respective body sites, but also a fundamental difference in melanocyte biology and the surrounding tissue where the tumour originated.

The current simplified definition of mutational subtypes in melanoma involves four groups which are distinguished by alterations in *BRAF*, *RAS, NF1* or the lack of mutations in the former genes, termed triple wild-type. Although *BRAF* status is used for companion diagnostics (*BRAF* inhibitors), the subtypes themselves do not provide prognostic value. These driver genes all activate the MAPK pathway, but there are differences between patients in respective group. Patients with *NF1* mutations were generally older and had accrued the largest number of mutational events. Conversely, patients of the *BRAF* subtype were typically younger, presenting a primary tumour with lower mutation load, often originating from a body site with intermittent exposure to sun. Interestingly, patients with *NRAS* hotspot mutations were found to share similarities in their copy number profiles with the acral and mucosal subtypes, suggesting some overlap between these categorisation systems exist. Through my Sambar pathway-mutation analysis, I showed that the mutational events in patients can be further explained beyond the existing four mutational categories. The triple wild-type group show their own distinct pattern of mutations, yet the other subtypes were spread across several classes, suggesting further genetic differences exist between these patients. Unfortunately, my extended Sambar-derived mutational classes could not better explain the genetic variation contributing to a difference in survival probability amongst patients. However, my analysis did show that further diversity in genetic composition beyond the existing classification exist, and therefore patients might benefit from being assessed using a modified stratification approach.

The site of primary melanoma can be used as a guideline to evaluate the level of sun exposure received. This measure is based on the assumption that different body sites normally would have varying degrees of sun protection, with the head having the highest degree of exposure, followed by limbs, trunk and other sites which includes acral, mucosal and genital tumours. Several studies have proposed a biological difference between tumours occurring on chronically sun-exposed, intermittent or completely sun-shielded locations. Of the can-

didate melanoma driver genes identified in my study of 524 primary melanomas, 11 out of 15 genes were associated with either increased mutation load and/or increased alteration frequencies in tumours arising on sun-exposed sites compared to more sun-shielded areas. These genes might therefore be particularly important in the disease aetiology of sun-associated melanoma. The mechanism driving the development of sun-shielded melanomas is not well understood, but is evidently different. Genes such as *BRAF, NRAS* and *MAP2K1* did not appear to show a preference towards sun-related metrics, indicating a difference in disease aetiology and progression in these tumours. It is possible that this reflects the origin and trajectory of melanoma development, where genes such as *BRAF* and *NRAS* are early founding events in some tumours, whilst other driver genes linked to UV damage might promote melanomagenesis along a divergent path. Important melanoma genes such as *PTEN* and *KIT* show other intriguing mutation patterns. *PTEN*-mutant tumours presented alteration patterns distinct from UV damage, yet never targeted melanomas of the triple wild-type, acral or mucosal category. Amplifications of *KIT* on the other hand, occurred selectively in mucosal, acral or cutaneous melanomas with signs of chronic sun damage, tumours otherwise known to be genetically distinct. Additionally, the same gene could be targeted by different genetic events in different types of melanomas, highlighting the important role such genes play in melanomagenesis. The contribution of sun exposure to the pathogenesis of melanoma is multifaceted, suggesting although sun exposure can in part explain the differences between some melanomas, it cannot be used exclusively for stratification. Melanoma is a highly heterogeneous cancer, where likely the combination of many variables including UV exposure, tissue of origin, mutation composition and the biological pathways altered are needed to comprehensively explain the underlying characteristics of each patient.

6.3 Immune evasion

With the advent of checkpoint inhibitors, a large focus has been placed on the power of harnessing the immune system to battle cancer. Mutation load, neoantigen load, copy number load and cytolytic score have all successfully been used to predict response to immunotherapy. However, I was curious to see if these effects would manifest also in a treatment-naive cohort. Using several stratification methods, higher mutation load or cytolytic score showed a trend towards favourable prognosis; however, cytolytic score alone or in combination with mutation load were the only factors mildly significant in multivariate analyses. This led me to hypothesise that most tumours keep the immune system in check through suppressive

mechanisms, to an extent where not even a highly antigenic environment is enough to trigger a response. This would explain why a strong trend towards favourable survival could not be seen in treatment-naive high mutation load tumours. Luckily, this suppressed state can be reversed with immunotherapy administration, whereby high mutation load tumours in particular become visible and targeted through an awakened host anti-tumour response.

Checkpoint inhibitors, with the PD-1/PD-L1 blocking antibodies at the forefront, have revolutionised melanoma care. However, these therapies do come with limitations such as severe drug-related toxicities, high cost, low bioavailability and resistance mechanisms. Targeting PD-L1, which is often upregulated in tumour cells, rather than PD-1, could provide a more focused approach, and potentially reduce adverse events. Additionally, activation of the immune system by minimising PD-L1 expression levels can be achieved using small-molecule inhibitors interfering with the biological pathways involved in PD-L1 regulation, rather than relying on antibodies to block its physical interaction with PD-1. By improving our understanding of the processes PD-L1 undergoes to be displayed on the surface of a tumour cell, we can envision probable tumour counteractive mechanisms as well as strive to identify new drug targets. Using the powerful CRISPR-Cas9 screening technology, I identified several pathways encompassing the "life cycle of PD-L1". With these new data, I learned that TAF proteins, part of the TFIID machinery, are important for the transcription of *CD274*. PD-L1 is then further processed through a range of post-translational modifications, including N-linked glycosylation, which is mediated by an array of ALG proteins and the OST complex. PD-L1 subsequently continues its journey towards the cell surface, where its expression is controlled by equilibrium mechanisms regulating protein turnover and recycling. Expression levels of PD-L1 on the surface of tumour cells are therefore carefully monitored through intracellular transport processes, where defects in such mechanisms alters the dose of PD-L1 on the cell surface. My data demonstrates how PD-L1 is regulated by several MTCs including TRAPP, COG and HOPS complexes whilst transitioning through vesicles between the ER, Golgi, endosomes, lysosomes and the plasma membrane. Autophagosomal regulation through *SPNS1* was also a major finding, providing a novel therapeutic approach by simultaneously targeting PD-L1 expression as well as autophagic processes.

Several of the pathways I discovered to modulate PD-L1 expression, although broad and general, could still be explored in a clinical setting. Recent studies have found that N-linked glycosylation of PD-L1 is necessary to avoid subsequent internalisation and degradation by the proteasomal machinery, and treatment with glycosylation-altering drugs have been shown to diminish PD-1/PD-L1 binding, stimulate cytotoxic T cell activity and reduce tu-

mour size in mouse models [783]. Furthermore, some genes associated with the regulation of the above-mentioned PD-L1 controlling pathways are overexpressed in melanoma and other tumours, and can promote tumour progression [746, 798]. Targeting these genes and pathways could therefore in addition to controlling tumour growth, activate host anti-cancer immunity through downregulation of PD-L1.

Melanomas typically harbour an abundance of genetic alterations, where certain oncogenic events can have a secondary tumourigenic effect in promoting immune evasion, including through the regulation of PD-L1 expression. In addition, drug treatments have also been found to alter PD-L1 expression levels and skew the immune architecture. We should therefore remember to view the tumour and immune system as a dynamic network which co-evolves with time, and this makes it particularly challenging to identify strong predictive biomarkers. Nevertheless, it would be valuable to study genetic aberrations also in the context of immune regulation, in order for us to learn more about how these events shape the tumour microenvironment. As melanomas show different mechanisms of disease development, drug susceptibility and immunogenicity, the ambition is to be able to predict outcome and mechanisms of drug resistance from the genetic footprint of any individual tumour.

6.4 Concluding remarks

It is important to emphasise that many genes do not regulate any one pathway in isolation. In reality, there is an intricate network of intertwined interactions operating, and it is likely that this interplay facilitates tumour progression rather than any one gene or pathway alone. Oncogenic processes are often initiated through the stimulation of receptor tyrosine kinases, which connects to several important cascades including both the MAPK and PI3K/AKT pathways. Additionally, recognised essential melanoma driver genes such as *NRAS* and *CDKN2A* regulate multiple pathways, and could shape the mutational composition in a different way compared to patients which have alterations in other driver genes. As an example, the acquisition of *BRAF* mutations is an early event in melanoma progression; however, this activation is not enough for melanoma development, and a second alteration targeting e.g. *PTEN* could be required for oncogenesis. As activating mutations in *BRAF* and *NRAS* normally would not be present simultaneously, the sequential mutations a tumour would acquire could differ depending on whether a tumour first acquired a *BRAF* or *NRAS* mutation. As an example, *NRAS*-mutant melanomas were more genetically similar in their copy number profiles to rarer subtypes of melanomas, compared to other tumours of the cu-

taneous category. Similarly, tumours with a high mutation burden show a particular driver gene composition. The formation of sporadic tumours rather than naevi-progressed have been reported to be more frequent on skin highly subjected to UV radiation [27, 49, 63], adding further support to the existence of fundamentally different melanoma disease groups which requires further investigation.

Applying the same rationale on a general level, past events would therefore shape the final genetic heterogeneity seen in melanomas. However, it appears the simplest model of this genetic diversity, studying the *BRAF, NRAS, NF1* and WT mutational subtypes in melanoma, does not completely stratify patients based on their mutational composition, nor do these groups provide prognostic value. Studying mutations on a pathway-level is more informative than looking at single genes alone; however, the cross-talk and multifaceted role of important genes adds complexity to these types of analyses. Furthermore, on a mutation level, not only will inevitable sporadic passenger events cause difficulty in accurate interpretation of the data, but incomplete alterations of gene functions by some mutations would also add to this uncertainty. It is therefore intriguing to find across many of the main pathways altered in primary melanoma, a clear pattern where mutations in any one component obviates the need for additional alterations to take place. Only in few cases, which is more pronounced with higher mutational load, do the same patients show co-mutations of key members of the same pathway. This suggests, this model of studying mutual exclusivity to discover novel pathway components which harbour pro-tumourigenic alterations is plausible and useful despite the aforementioned limitations.

Unfortunately, my study failed to present strong evidence of particular genetic events associated with a favourable patient outcome. The important task of identifying genetic contributors to patient survival in melanoma is a challenge, as not only is it a heterogenous disease, but driver events are easily overshadowed by the vast amount of unspecific mutations evident of tumours prone to have a high alteration rate. This increase in background noise, as well as the lack in directionality (activating versus damaging) and rank (how activating or damaging) in the model, could be important factors contributing to the negligence of anticipating patient survival.

Melanoma remains one of the most curable malignancies when diagnosed at early stages. However, we are currently unable to identify the prognostic factors associated with relapse, and few studies on early stage primary melanomas have been performed. With my thesis, I have presented the most comprehensive evaluation into the somatic alteration landscape of primary melanomas to date, bringing valuable insight into the architecture of such tumours. Additionally, I introduced novel genes and key processes controlling PD-L1 gene

transcription, processing and presentation on the cell surface. These findings expand our understanding of cellular regulation of PD-L1 expression, potentially identifying new directions for drug development. I hope this new knowledge will further our comprehension of melanoma progression and diversity, as well as bringing forth novel insights and possible treatment modalities.

References

[1] Cancer incidence for common cancers. Cancer Research UK. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/incidence/common-cancers-compared.

[2] Howlader, N. *et al.* SEER Cancer Statistics Review 1975-2016. National Cancer Institute. URL: https://seer.cancer.gov/csr/1975_2016.

[3] Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians* **68**, 394–424 (2018).

[4] Melanoma of the Skin - Cancer Stat Facts. URL: https://seer.cancer.gov/statfacts/html/melan.html.

[5] Gershenwald, J. E. *et al.* Melanoma Staging: Evidence-Based Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA: A Cancer Journal for Clinicians* **67**, 472–492 (2017).

[6] Gandini, S., Autier, P. & Boniol, M. Reviews on sun exposure and artificial light and melanoma. *Progress in Biophysics and Molecular Biology* **107**, 362–366 (2011).

[7] Gandini, S. *et al.* Meta-analysis of risk factors for cutaneous melanoma: II. Sun exposure. *European Journal of Cancer* **41**, 45–60 (2005).

[8] Veierød, M. B., Adami, H.-O., Lund, E., Armstrong, B. K. & Weiderpass, E. Sun and solarium exposure and melanoma risk: effects of age, pigmentary characteristics, and nevi. *Cancer Epidemiology, Biomarkers & Prevention* **19**, 111–120 (2010).

[9] Wu, S., Han, J., Laden, F. & Qureshi, A. A. Long-term Ultraviolet Flux, Other Potential Risk Factors, and Skin Cancer Risk: A Cohort Study. *Cancer Epidemiology, Biomarkers & Prevention* (2014).

[10] Armstrong, B. K. & Kricker, A. The epidemiology of UV induced skin cancer. *Journal of Photochemistry and Photobiology B: Biology* **63**, 8–18 (2001).

[11] Olsen, C. M., Carroll, H. J. & Whiteman, D. C. Estimating the attributable fraction for melanoma: a meta-analysis of pigmentary characteristics and freckling. *International Journal of Cancer* **127**, 2430–2445 (2010).

[12] Tucker, M. A. *et al.* Clinically recognized dysplastic nevi. A central risk factor for cutaneous melanoma. *JAMA* **277**, 1439–1444 (1997).

[13] van der Leest, R. J. T., Flohil, S. C., Arends, L. R., de Vries, E. & Nijsten, T. Risk of subsequent cutaneous malignancy in patients with prior melanoma: a systematic review and meta-analysis. *Journal of the European Academy of Dermatology and Venereology* **29**, 1053–1062 (2015).

[14] Goldstein, A. M. & Tucker, M. A. Genetic Epidemiology of Cutaneous Melanoma: A Global Perspective. *Archives of Dermatology* **137**, 1493–1496 (2001).

[15] Kamb, A. *et al.* A cell cycle regulator potentially involved in genesis of many tumor types. *Science* **264**, 436–440 (1994).

[16] Soengas, M. S. & Lowe, S. W. Apoptosis and melanoma chemoresistance. *Oncogene* **22**, 3138–3151 (2003).

[17] Gilchrest, B. A., Eller, M. S., Geller, A. C. & Yaar, M. The Pathogenesis of Melanoma Induced by Ultraviolet Radiation. *New England Journal of Medicine* **340**, 1341–1348 (1999).

[18] Shain, A. H. & Bastian, B. C. From melanocytes to melanomas. *Nature Reviews Cancer* **16**, 345–358 (2016).

[19] Lin, J. Y. & Fisher, D. E. Melanocyte biology and skin pigmentation. *Nature* **445**, 843–850 (2007).

[20] D'Mello, S. A. N., Finlay, G. J., Baguley, B. C. & Askarian-Amiri, M. E. Signaling Pathways in Melanogenesis. *International Journal of Molecular Sciences* **17** (2016).

[21] Yamaguchi, Y., Brenner, M. & Hearing, V. J. The Regulation of Skin Pigmentation. *Journal of Biological Chemistry* **282**, 27557–27561 (2007).

[22] Mort, R. L., Jackson, I. J. & Patton, E. E. The melanocyte lineage in development and disease. *Development* **142**, 620–632 (2015).

[23] Gupta, P. B. *et al.* The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. *Nature Genetics* **37**, 1047–1054 (2005).

[24] Peinado, H., Olmeda, D. & Cano, A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nature Reviews Cancer* **7**, 415–428 (2007).

[25] Eggermont, A. M., Spatz, A. & Robert, C. Cutaneous melanoma. *The Lancet* **383**, 816–827 (2014).

[26] Elder, D. E. Precursors to melanoma and their mimics: nevi of special sites. *Modern Pathology* **19 Suppl 2**, S4–20 (2006).

[27] Shitara, D. *et al.* Nevus-associated melanomas: clinicopathologic features. *American Journal of Clinical Pathology* **142**, 485–491 (2014).

[28] Saida, T. Histogenesis of cutaneous malignant melanoma: The vast majority do not develop from melanocytic nevus but arise de novo as melanoma in situ. *The Journal of Dermatology* **46**, 80–94 (2019).

[29] Clark, W. H. *et al.* A study of tumor progression: The precursor lesions of superficial spreading and nodular melanoma. *Human Pathology* **15**, 1147–1165 (1984).

[30] Pollock, P. M. *et al.* High frequency of *BRAF* mutations in nevi. *Nature Genetics* **33**, 19–20 (2003).

[31] Poynter, J. *et al.* BRAF and NRAS mutations in melanoma and melanocytic nevi. *Melanoma Research* **16**, 267–273 (2006).

[32] Michaloglou, C. *et al.* BRAF E600 -associated senescence-like cell cycle arrest of human naevi. *Nature* **436**, 720 (2005).

[33] Peeper, D. S. Oncogene-induced senescence and melanoma: where do we stand? *Pigment Cell & Melanoma Research* **24**, 1107–1111 (2011).

[34] Gray-Schopfer, V. C. *et al.* Cellular senescence in naevi and immortalisation in melanoma: a role for p16? *British Journal of Cancer* **95**, 496 (2006).

[35] Chiba, K. *et al.* Mutations in the promoter of the telomerase gene TERT contribute to tumorigenesis by a two-step mechanism. *Science* **357**, 1416–1420 (2017).

[36] Miller, A. J. & Mihm, M. C. Melanoma. *New England Journal of Medicine* **355**, 51–65 (2006).

[37] Pampena, R. *et al.* A meta-analysis of nevus-associated melanoma: Prevalence and practical implications. *Journal of the American Academy of Dermatology* **77**, 938–945.e4 (2017).

[38] Hanahan, D. & Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011).

[39] Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).

[40] Ackerman, A. B. & Mihara, I. Dysplasia, dysplastic melanocytes, dysplastic nevi, the dysplastic nevus syndrome, and the relation between dysplastic nevi and malignant melanomas. *Human Pathology* **16**, 87–91 (1985).

[41] Ichii-Nakato, N. *et al.* High Frequency of BRAFV600E Mutation in Acquired Nevi and Small Congenital Nevi, but Low Frequency of Mutation in Medium-Sized Congenital Nevi. *The Journal of Investigative Dermatology* **126**, 2111–2118 (2006).

[42] Yazdi, A. S. *et al.* Mutations of the BRAF gene in benign and malignant melanocytic lesions. *The Journal of Investigative Dermatology* **121**, 1160–1162 (2003).

[43] Lin, J. *et al.* Polyclonality of BRAF Mutations in Acquired Melanocytic Nevi. *Journal of the National Cancer Institute* **101**, 1423–1427 (2009).

[44] Takata, M., Murata, H. & Saida, T. Molecular pathogenesis of malignant melanoma: a different perspective from the studies of melanocytic nevus and acral melanoma. *Pigment Cell & Melanoma Research* **23**, 64–71 (2010).

[45] Bastian, B. C. The molecular pathology of melanoma: An integrated taxonomy of melanocytic neoplasia. *Annual Review of Pathology* **9**, 239–271 (2014).

[46] Bennett, D. C. How to make a melanoma: what do we know of the primary clonal events? *Pigment Cell & Melanoma Research* **21**, 27–38 (2008).

[47] Shain, A. H. *et al.* The Genetic Evolution of Melanoma from Precursor Lesions. *New England Journal of Medicine* **373**, 1926–1936 (2015).

[48] Shain, A. H. *et al.* Genomic and Transcriptomic Analysis Reveals Incremental Disruption of Key Signaling Pathways during Melanoma Evolution. *Cancer Cell* **34**, 45–55.e4 (2018).

[49] Bevona, C., Goggins, W., Quinn, T., Fullerton, J. & Tsao, H. Cutaneous melanomas associated with nevi. *Archives of Dermatology* **139**, 1620–1624 (2003).

[50] Liu, W. *et al.* Distinct clinical and pathological features are associated with the BRAF(T1799A(V600E)) mutation in primary melanoma. *The Journal of Investigative Dermatology* **127**, 900–905 (2007).

[51] Long, G. V. *et al.* Prognostic and clinicopathologic associations of oncogenic BRAF in metastatic melanoma. *Journal of Clinical Oncology* **29**, 1239–1246 (2011).

[52] Maldonado, J. L. *et al.* Determinants of BRAF mutations in primary melanomas. *Journal of the National Cancer Institute* **95**, 1878–1890 (2003).

[53] Shitara, D. *et al.* Nevus-associated melanomas: clinicopathologic features. *American Journal of Clinical Pathology* **142**, 485–491 (2014).

[54] Curtin, J. A. *et al.* Distinct sets of genetic alterations in melanoma. *New England Journal of Medicine* **353**, 2135–2147 (2005).

[55] Elder DE, Greene MH, Bondi EE & Clark WH. *Acquired melanocytic nevus and melanoma: the dysplastic nevus syndrome. In: Ackerman AB, ed. Pathology of Malignant Melanoma.* (New York, 1981).

[56] Clark, W. H., From, L., Bernardino, E. A. & Mihm, M. C. The histogenesis and biologic behavior of primary human malignant melanomas of the skin. *Cancer Research* **29**, 705–727 (1969).

[57] Ward, W. H., Lambreton, F., Goel, N., Yu, J. Q. & Farma, J. M. Clinical Presentation and Staging of Melanoma. In Ward, W. H. & Farma, J. M. (eds.) *Cutaneous Melanoma: Etiology and Therapy* (Codon Publications, Brisbane (AU), 2017).

[58] McGovern, V. J. *et al.* The classification of malignant melanoma and its histologic reporting. *Cancer* **32**, 1446–1457 (1973).

[59] Gunderson, L. L. & Tepper, J. E. *Clinical Radiation Oncology* (Elsevier, Philadelphia, PA, 2015), 4 edition edn.

[60] Markovic, S. N. *et al.* Malignant Melanoma in the 21st Century, Part 1: Epidemiology, Risk Factors, Screening, Prevention, and Diagnosis. *Mayo Clinic Proceedings* **82**, 364–380 (2007).

[61] Wang, Y., Zhao, Y. & Ma, S. Racial differences in six major subtypes of melanoma: descriptive epidemiology. *BMC Cancer* **16** (2016).

[62] Elder, D. E., Massi, D., Scolyer, R. & Willemze, R. (eds.) *WHO Classification of Skin Tumours* (World Health Organization, Lyon, 2018), 4 edn.

[63] Whiteman, D. C. *et al.* Melanocytic nevi, solar keratoses, and divergent pathways to cutaneous melanoma. *Journal of the National Cancer Institute* **95**, 806–812 (2003).

[64] Cohen, L. M. Lentigo maligna and lentigo maligna melanoma. *Journal of the American Academy of Dermatology* **33**, 923–936 (1995).

[65] Feibleman, G. E., Stoll, H. & Maize, J. C. Melanomas of the palm, sole, and nailbed: A clinicopathologic study. *Cancer* **46**, 2492–2504 (1980).

[66] Bradford, P. T., Goldstein, A. M., McMaster, M. L. & Tucker, M. A. Acral Lentiginous Melanoma: Incidence and Survival Patterns in the United States, 1986-2005. *Archives of Dermatology* **145**, 427–434 (2009).

[67] NHS. Skin cancer (melanoma). URL: https://www.nhs.uk/conditions/melanoma-skin-cancer.

[68] Bristow, I. R. & Acland, K. Acral lentiginous melanoma of the foot and ankle: A case series and review of the literature. *Journal of Foot and Ankle Research* **1**, 11 (2008).

[69] DeMatos, P., Tyler, D. S. & Seigler, H. F. Malignant melanoma of the mucous membranes: a review of 119 cases. *Annals of Surgical Oncology* **5**, 733–742 (1998).

[70] Saida, T. *et al.* Histopathological characteristics of malignant melanoma affecting mucous membranes: a unifying concept of histogenesis. *Pathology* **36**, 404–413 (2004).

[71] Krantz, B. A., Dave, N., Komatsubara, K. M., Marr, B. P. & Carvajal, R. D. Uveal melanoma: epidemiology, etiology, and treatment of primary disease. *Clinical Ophthalmology* **11**, 279–289 (2017).

[72] Kaliki, S. & Shields, C. L. Uveal melanoma: relatively rare but deadly cancer. *Eye* **31**, 241–257 (2017).

[73] The Cancer Genome Atlas Network. Genomic Classification of Cutaneous Melanoma. *Cell* **161**, 1681–1696 (2015).

[74] Breslow, A. Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma. *Annals of Surgery* **172**, 902–908 (1970).

[75] Balch, C. M. *et al.* The prognostic significance of ulceration of cutaneous melanoma. *Cancer* **45**, 3012–3017 (1980).

[76] Balch, C. M. *et al.* Final Version of 2009 AJCC Melanoma Staging and Classification. *Journal of Clinical Oncology* **27**, 6199–6206 (2009).

[77] Gershenwald, J. E. & Scolyer, R. A. Melanoma Staging: American Joint Committee on Cancer (AJCC) 8th Edition and Beyond. *Annals of Surgical Oncology* **25**, 2105–2110 (2018).

[78] Gimotty, P. A. *et al.* Identification of high-risk patients among those diagnosed with thin cutaneous melanomas. *Journal of Clinical Oncology* **25**, 1129–1134 (2007).

[79] Thompson, J. F. *et al.* Prognostic Significance of Mitotic Rate in Localized Primary Cutaneous Melanoma: An Analysis of Patients in the Multi-Institutional American Joint Committee on Cancer Melanoma Staging Database. *Journal of Clinical Oncology* **29**, 2199–2205 (2011).

[80] Azzola, M. F. *et al.* Tumor mitotic rate is a more powerful prognostic indicator than ulceration in patients with primary cutaneous melanoma: an analysis of 3661 patients from a single center. *Cancer* **97**, 1488–1498 (2003).

[81] Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).

[82] Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).

[83] Jhappan, C., Noonan, F. P. & Merlino, G. Ultraviolet radiation and cutaneous malignant melanoma. *Oncogene* **22**, 3099 (2003).

[84] Ikehata, H. & Ono, T. The mechanisms of UV mutagenesis. *Journal of Radiation Research* **52**, 115–125 (2011).

[85] Hodis, E. *et al.* A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012).

[86] Lawrence, M. S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).

[87] Jackson, S. P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009).

[88] Hainaut, P. & Hollstein, M. p53 and human cancer: the first ten thousand mutations. *Advances in Cancer Research* **77**, 81–137 (2000).

[89] Negrini, S., Gorgoulis, V. G. & Halazonetis, T. D. Genomic instability — an evolving hallmark of cancer. *Nature Reviews Molecular Cell Biology* **11**, 220–228 (2010).

[90] Kabbarah, O. & Chin, L. Revealing the genomic heterogeneity of melanoma. *Cancer Cell* **8**, 439–441 (2005).

[91] Hayward, N. K. *et al.* Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175–180 (2017).

[92] Hocker, T. L., Singh, M. K. & Tsao, H. Melanoma Genetics and Therapeutic Approaches in the 21st Century: Moving from the Benchside to the Bedside. *The Journal of Investigative Dermatology* **128**, 2575–2595 (2008).

[93] Shtivelman, E. *et al.* Pathways and therapeutic targets in melanoma. *Oncotarget* **5**, 1701–1752 (2014).

[94] Davies, H. *et al.* Mutations of the *BRAF* gene in human cancer. *Nature* **417**, 949–954 (2002).

[95] Desideri, E., Cavallo, A. L. & Baccarini, M. Alike but Different: RAF Paralogs and Their Signaling Outputs. *Cell* **161**, 967–970 (2015).

[96] Network, T. C. G. A. *et al.* The Cancer Genome Atlas Pan-Cancer Analysis Project. *Nature Genetics* **45**, 1113–1120 (2013).

[97] Prior, I. A., Lewis, P. D. & Mattos, C. A comprehensive survey of Ras mutations in cancer. *Cancer Research* **72**, 2457–2467 (2012).

[98] Fruman, D. A. *et al.* The PI3k Pathway in Human Disease. *Cell* **170**, 605–635 (2017).

[99] Luo, J., Manning, B. D. & Cantley, L. C. Targeting the PI3k-Akt pathway in human cancer: rationale and promise. *Cancer Cell* **4**, 257–262 (2003).

[100] Siroy, A. E., Davies, M. A. & Lazar, A. J. The PI3k-AKT Pathway in Melanoma. In Torres-Cabala, C. A. & Curry, J. L. (eds.) *Genetics of Melanoma*, 165–180 (Springer, New York, NY, 2016).

[101] Pearson, G. *et al.* Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions. *Endocrine Reviews* **22**, 153–183 (2001).

[102] Zhang, W. & Liu, H. T. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research* **12**, 9–18 (2002).

[103] Amaral, T. *et al.* The mitogen-activated protein kinase pathway in melanoma part I - Activation and primary resistance mechanisms to BRAF inhibition. *European Journal of Cancer* **73**, 85–92 (2017).

[104] Dhillon, A. S., Hagan, S., Rath, O. & Kolch, W. MAP kinase signalling pathways in cancer. *Oncogene* **26**, 3279–3290 (2007).

[105] Wellbrock, C., Karasarides, M. & Marais, R. The RAF proteins take centre stage. *Nature Reviews Molecular Cell Biology* **5**, 875 (2004).

[106] Ramos, J. W. The regulation of extracellular signal-regulated kinase (ERK) in mammalian cells. *The International Journal of Biochemistry & Cell Biology* **40**, 2707–2719 (2008).

[107] Wellbrock, C. & Arozarena, I. The Complexity of the ERK/MAP-Kinase Pathway and the Treatment of Melanoma Skin Cancer. *Frontiers in Cell and Developmental Biology* **4** (2016).

[108] Vanhaesebroeck, B., Guillermet-Guibert, J., Graupera, M. & Bilanges, B. The emerging mechanisms of isoform-specific PI3k signalling. *Nature Reviews Molecular Cell Biology* **11**, 329–341 (2010).

[109] Alessi, D. R. *et al.* Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase Balpha. *Current Biology* **7**, 261–269 (1997).

[110] Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098–1101 (2005).

[111] Potter, C. J., Pedraza, L. G. & Xu, T. Akt regulates growth by directly phosphorylating Tsc2. *Nature Cell Biology* **4**, 658–665 (2002).

[112] Manning, B. D., Tee, A. R., Logsdon, M. N., Blenis, J. & Cantley, L. C. Identification of the tuberous sclerosis complex-2 tumor suppressor gene product tuberin as a target of the phosphoinositide 3-kinase/akt pathway. *Molecular Cell* **10**, 151–162 (2002).

[113] Inoki, K., Li, Y., Zhu, T., Wu, J. & Guan, K.-L. TSC2 is phosphorylated and inhibited by Akt and suppresses mTOR signalling. *Nature Cell Biology* **4**, 648–657 (2002).

[114] Datta, S. R. *et al.* Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. *Cell* **91**, 231–241 (1997).

[115] Brunet, A. *et al.* Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* **96**, 857–868 (1999).

[116] Kops, G. J. *et al.* Direct control of the Forkhead transcription factor AFX by protein kinase B. *Nature* **398**, 630–634 (1999).

[117] Altomare, D. A. & Testa, J. R. Perturbations of the AKT signaling pathway in human cancer. *Oncogene* **24**, 7455 (2005).

[118] Vivanco, I. & Sawyers, C. L. The phosphatidylinositol 3-Kinase–AKT pathway in human cancer. *Nature Reviews Cancer* **2**, 489 (2002).

[119] Zhou, B. P. *et al.* HER-2/neu induces p53 ubiquitination via Akt-mediated MDM2 phosphorylation. *Nature Cell Biology* **3**, 973–982 (2001).

[120] Mayo, L. D. & Donner, D. B. A phosphatidylinositol 3-kinase/Akt pathway promotes translocation of Mdm2 from the cytoplasm to the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11598–11603 (2001).

[121] Cross, D. A. E., Alessi, D. R., Cohen, P., Andjelkovich, M. & Hemmings, B. A. Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B. *Nature* **378**, 785 (1995).

[122] Wu, D. & Pan, W. GSK3: a multifaceted kinase in Wnt signaling. *Trends in Biochemical Sciences* **35**, 161–168 (2010).

[123] van Noort, M., Meeldijk, J., van der Zee, R., Destree, O. & Clevers, H. Wnt signaling controls the phosphorylation status of beta-catenin. *Journal of Biological Chemistry* **277**, 17901–17905 (2002).

[124] Gregory, M. A., Qi, Y. & Hann, S. R. Phosphorylation by glycogen synthase kinase-3 controls c-myc proteolysis and subnuclear localization. *Journal of Biological Chemistry* **278**, 51606–51612 (2003).

[125] Diehl, J. A., Cheng, M., Roussel, M. F. & Sherr, C. J. Glycogen synthase kinase-3beta regulates cyclin D1 proteolysis and subcellular localization. *Genes & Development* **12**, 3499–3511 (1998).

[126] Chandarlapaty, S. *et al.* AKT inhibition relieves feedback suppression of receptor tyrosine kinase expression and activity. *Cancer Cell* **19**, 58–71 (2011).

[127] Chakrabarty, A., Sánchez, V., Kuba, M. G., Rinehart, C. & Arteaga, C. L. Feedback upregulation of HER3 (ErbB3) expression and activity attenuates antitumor effect of PI3k inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 2718–2723 (2012).

[128] Steck, P. A. *et al.* Identification of a candidate tumour suppressor gene, MMAC1, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nature Genetics* **15**, 356–362 (1997).

[129] Furnari, F. B., Lin, H., Huang, H. S. & Cavenee, W. K. Growth suppression of glioma cells by PTEN requires a functional phosphatase catalytic domain. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 12479–12484 (1997).

[130] Stambolic, V. *et al.* Negative regulation of PKB/Akt-dependent cell survival by the tumor suppressor PTEN. *Cell* **95**, 29–39 (1998).

[131] Maehama, T. & Dixon, J. E. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *Journal of Biological Chemistry* **273**, 13375–13378 (1998).

[132] Tsao, H., Zhang, X., Benoit, E. & Haluska, F. G. Identification of PTEN/MMAC1 alterations in uncultured melanomas and melanoma cell lines. *Oncogene* **16**, 3397–3402 (1998).

[133] Kamb, A. *et al.* Analysis of the p16 gene (CDKN2) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nature Genetics* **8**, 22–26 (1994).

[134] Hussussian, C. J. *et al.* Germline p16 mutations in familial melanoma. *Nature Genetics* **8**, 15–21 (1994).

[135] Zuo, L. *et al.* Germline mutations in the p16ink4a binding domain of CDK4 in familial melanoma. *Nature Genetics* **12**, 97–99 (1996).

[136] Reed, J. A. *et al.* Loss of Expression of the p16/Cyclin-dependent Kinase Inhibitor 2 Tumor Suppressor Gene in Melanocytic Lesions Correlates with Invasive Stage of Tumor Progression. *Cancer Research* **55**, 2713–2718 (1995).

[137] Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16ink4a. *Cell* **88**, 593–602 (1997).

[138] Sheppard, K. E. & McArthur, G. A. The cell-cycle regulator CDK4: an emerging therapeutic target in melanoma. *Clinical Cancer Research* **19**, 5320–5328 (2013).

[139] Tao, W. & Levine, A. J. P19(ARF) stabilizes p53 by blocking nucleo-cytoplasmic shuttling of Mdm2. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 6937–6941 (1999).

[140] Zhang, Y., Xiong, Y. & Yarbrough, W. G. ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell* **92**, 725–734 (1998).

[141] Kamijo, T. *et al.* Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 8292–8297 (1998).

[142] Pomerantz, J. *et al.* The Ink4a tumor suppressor gene product, p19arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell* **92**, 713–723 (1998).

[143] Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science* **253**, 49–53 (1991).

[144] Greenblatt, M. S., Bennett, W. P., Hollstein, M. & Harris, C. C. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Research* **54**, 4855–4878 (1994).

[145] Efeyan, A. & Serrano, M. p53: guardian of the genome and policeman of the oncogenes. *Cell Cycle* **6**, 1006–1010 (2007).

[146] Horn, H. F. & Vousden, K. H. Coping with stress: multiple ways to activate p53. *Oncogene* **26**, 1306–1316 (2007).

[147] Law, M. H. *et al.* Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nature Genetics* **47**, 987–995 (2015).

[148] Horn, S. *et al.* TERT promoter mutations in familial and sporadic melanoma. *Science* **339**, 959–961 (2013).

[149] Nagore, E. *et al.* TERT promoter mutations in melanoma survival. *International Journal of Cancer* **139**, 75–84 (2016).

[150] Heidenreich, B. *et al.* Telomerase reverse transcriptase promoter mutations in primary cutaneous melanoma. *Nature Communications* **5**, 3401 (2014).

[151] Huang, F. W. *et al.* Highly recurrent TERT promoter mutations in human melanoma. *Science* **339**, 957–959 (2013).

[152] Rimm, D. L., Caca, K., Hu, G., Harrison, F. B. & Fearon, E. R. Frequent Nuclear/Cytoplasmic Localization of β-Catenin without Exon 3 Mutations in Malignant Melanoma. *The American Journal of Pathology* **154**, 325–329 (1999).

[153] Rubinfeld, B. *et al.* Stabilization of beta-catenin by genetic defects in melanoma cell lines. *Science* **275**, 1790–1792 (1997).

[154] Damsky, W. E. *et al.* β-catenin signaling controls metastasis in Braf-activated Pten-deficient melanomas. *Cancer Cell* **20**, 741–754 (2011).

[155] Krauthammer, M. *et al.* Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nature Genetics* **44**, 1006–1014 (2012).

[156] Peng, W. *et al.* Loss of PTEN Promotes Resistance to T Cell–Mediated Immunotherapy. *Cancer Discovery* **6**, 202–216 (2016).

[157] Zaretsky, J. M. *et al.* Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *New England Journal of Medicine* **375**, 819–829 (2016).

[158] The Cancer Genome Atlas. URL: https://www.cancer.gov/tcga.

[159] Cancer Genome Project. URL: https://www.sanger.ac.uk/science/groups/cancer-genome-project.

[160] The International Cancer Genome Consortium. International network of cancer genome projects. *Nature* **464**, 993–998 (2010).

[161] Krauthammer, M. *et al.* Exome sequencing identifies recurrent mutations in NF1 and RASopathy genes in sun-exposed melanomas. *Nature Genetics* **47**, 996–1002 (2015).

[162] Friedman, R. J., Rigel, D. S. & Kopf, A. W. Early detection of malignant melanoma: The role of physician examination and self-examination of the skin. *CA: A Cancer Journal for Clinicians* **35**, 130–151 (1985).

[163] Abbasi, N. R. *et al.* Early Diagnosis of Cutaneous Melanoma: Revisiting the ABCD Criteria. *JAMA* **292**, 2771–2776 (2004).

[164] Rigel, D. S. & Friedman, R. J. The rationale of the ABCDs of early melanoma. *Journal of the American Academy of Dermatology* **29**, 1060–1061 (1993).

[165] Grob, J. J. & Bonerandi, J. J. The 'ugly duckling' sign: identification of the common characteristics of nevi in an individual as a basis for melanoma screening. *Archives of Dermatology* **134**, 103–104 (1998).

[166] Swetter, S. M. *et al.* Guidelines of care for the management of primary cutaneous melanoma. *Journal of the American Academy of Dermatology* **80**, 208–250 (2019).

[167] Brunssen, A., Waldmann, A., Eisemann, N. & Katalinic, A. Impact of skin cancer screening and secondary prevention campaigns on skin cancer incidence and mortality: A systematic review. *Journal of the American Academy of Dermatology* **76**, 129–139.e10 (2017).

[168] Boniol, M., Autier, P. & Gandini, S. Melanoma mortality following skin cancer screening in Germany. *BMJ Open* **5**, e008158 (2015).

[169] Katalinic, A. *et al.* Does skin cancer screening save lives?: an observational study comparing trends in melanoma mortality in regions with and without screening. *Cancer* **118**, 5395–5402 (2012).

[170] Koh, H. K. *et al.* Evaluation of the American Academy of Dermatology's National Skin Cancer Early Detection and Screening Program. *Journal of the American Academy of Dermatology* **34**, 971–978 (1996).

[171] Skin Cancer Foundation. May is Skin Cancer Awareness Month. URL: https://www.skincancer.org/get-involved/skin-cancer-awareness-month.

[172] War on Melanoma™. URL: https://www.ohsu.edu/xd/health/services/dermatology/war-on-melanoma.

[173] Aitken, J. F., Elwood, M., Baade, P. D., Youl, P. & English, D. Clinical whole-body skin examination reduces the incidence of thick melanomas. *International Journal of Cancer* **126**, 450–458 (2010).

[174] Schneider, J. S., Moore, D. H. & Mendelsohn, M. L. Screening program reduced melanoma mortality at the Lawrence Livermore National Laboratory, 1984 to 1996. *Journal of the American Academy of Dermatology* **58**, 741–749 (2008).

[175] SPOTme® Skin Cancer Screenings | American Academy of Dermatology. URL: https://www.aad.org/public/spot-skin-cancer/programs/screenings.

[176] Wong, S. L. *et al.* Sentinel Lymph Node Biopsy and Management of Regional Lymph Nodes in Melanoma: American Society of Clinical Oncology and Society of Surgical Oncology Clinical Practice Guideline Update. *Journal of Clinical Oncology* **36**, 399–413 (2017).

[177] Morton, D. L. *et al.* Final Trial Report of Sentinel-Node Biopsy versus Nodal Observation in Melanoma. *New England Journal of Medicine* **370**, 599–609 (2014).

[178] Balch, C. M. *et al.* Sentinel node biopsy and standard of care for melanoma. *Journal of the American Academy of Dermatology* **60**, 872–875 (2009).

[179] Faries, M. B. *et al.* Completion Dissection or Observation for Sentinel-Node Metastasis in Melanoma. *New England Journal of Medicine* **376**, 2211–2222 (2017).

[180] Ross, M. I. & Gershenwald, J. E. Evidence-based treatment of early-stage melanoma. *Journal of Surgical Oncology* **104**, 341–353 (2011).

[181] van Zeijl, M. C. T., van den Eertwegh, A. J., Haanen, J. B. & Wouters, M. W. J. M. (Neo)adjuvant systemic therapy for melanoma. *European Journal of Surgical Oncology* **43**, 534–543 (2017).

[182] Garbe, C. *et al.* Diagnosis and treatment of melanoma. European consensus-based interdisciplinary guideline – Update 2016. *European Journal of Cancer* **63**, 201–217 (2016).

[183] Thirlwell, C. & Nathan, P. Melanoma–Part 2: management. *BMJ* **337**, a2488 (2008).

[184] Wheatley, K. *et al.* Does adjuvant interferon-α for high-risk melanoma provide a worthwhile benefit? A meta-analysis of the randomised trials. *Cancer Treatment Reviews* **29**, 241–252 (2003).

[185] Coit, D. G. *et al.* Cutaneous Melanoma, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network* **17**, 367–402 (2019).

[186] Batus, M. *et al.* Optimal Management of Metastatic Melanoma: Current Strategies and Future Directions. *American Journal of Clinical Dermatology* **14**, 179–194 (2013).

[187] Domingues, B., Lopes, J. M., Soares, P. & Pópulo, H. Melanoma treatment in review. *ImmunoTargets and Therapy* **7**, 35–49 (2018).

[188] Gogas, H. J., Kirkwood, J. M. & Sondak, V. K. Chemotherapy for metastatic melanoma: time for a change? *Cancer* **109**, 455–464 (2007).

[189] BHATIA, S., TYKODI, S. S. & THOMPSON, J. A. Treatment of Metastatic Melanoma: An Overview. *Oncology* **23**, 488–496 (2009).

[190] Serrone, L., Zeuli, M., Sega, F. M. & Cognetti, F. Dacarbazine-based chemotherapy for metastatic melanoma: thirty-year experience overview. *Journal of Experimental & Clinical Cancer Research* **19**, 21–34 (2000).

[191] Hill, G. J., Krementz, E. T. & Hill, H. Z. Dimethyl triazeno imidazole carboxamide and combination therapy for melanoma. IV. Late results after complete response to chemotherapy (Central Oncology Group protocols 7130, 7131, and 7131a). *Cancer* **53**, 1299–1305 (1984).

[192] Lui, P. *et al.* Treatments for metastatic melanoma: synthesis of evidence from randomized trials. *Cancer Treatment Reviews* **33**, 665–680 (2007).

[193] Cocconi, G. *et al.* Treatment of Metastatic Malignant Melanoma with Dacarbazine plus Tamoxifen. *New England Journal of Medicine* **327**, 516–523 (1992).

[194] Middleton, M. R. *et al.* A randomized phase III study comparing dacarbazine, BCNU, cisplatin and tamoxifen with dacarbazine and interferon in advanced melanoma. *British Journal of Cancer* **82**, 1158 (2000).

[195] Eigentler, T. K., Caroli, U. M., Radny, P. & Garbe, C. Palliative therapy of disseminated malignant melanoma: a systematic review of 41 randomised clinical trials. *The Lancet* **4**, 748–759 (2003).

[196] Jungnelius, U. *et al.* Dacarbazine-vindesine versus dacarbazine-vindesine-cisplatin in disseminated malignant melanoma. A randomised phase III trial. *European Journal of Cancer* **34**, 1368–1374 (1998).

[197] Cocconi, G. *et al.* Treatment of metastatic malignant melanoma with dacarbazine plus tamoxifen, or vindesine plus tamoxifen: a prospective randomized study. *Melanoma Research* **13**, 73–79 (2003).

[198] Wittes, R. E., Wittes, J. T. & Golbey, R. B. Combination chemotherapy in metastatic malignant melanoma: a randomized study of three DTIC-containing combination. *Cancer* **41**, 415–421 (1978).

[199] Chapman, P. B. *et al.* Phase III multicenter randomized trial of the Dartmouth regimen versus dacarbazine in patients with metastatic melanoma. *Journal of Clinical Oncology* **17**, 2745–2751 (1999).

[200] Kirkwood, J. M. *et al.* High-dose interferon alfa-2b significantly prolongs relapse-free and overall survival compared with the GM2-KLH/QS-21 vaccine in patients with resected stage IIB-III melanoma: results of intergroup trial E1694/S9512/C509801. *Journal of Clinical Oncology* **19**, 2370–2380 (2001).

[201] Telang, S. *et al.* Phase II trial of the regulatory T cell-depleting agent, denileukin diftitox, in patients with unresectable stage IV melanoma. *BMC Cancer* **11**, 515 (2011).

[202] Prince, H. M. *et al.* Phase III placebo-controlled trial of denileukin diftitox for patients with cutaneous T-cell lymphoma. *Journal of Clinical Oncology* **28**, 1870–1877 (2010).

[203] Ives, N. J. *et al.* Adjuvant interferon-α for the treatment of high-risk melanoma: An individual patient data meta-analysis. *European Journal of Cancer* **82**, 171–183 (2017).

[204] Flaherty, K. T. *et al.* Inhibition of Mutated, Activated BRAF in Metastatic Melanoma. *New England Journal of Medicine* **363**, 809–819 (2010).

[205] Ribas, A. *et al.* BRIM-2: An open-label, multicenter phase II study of vemurafenib in previously treated patients with BRAF V600E mutation-positive metastatic melanoma. *Journal of Clinical Oncology* **29**, 8509–8509 (2011).

[206] Chapman, P. B. *et al.* Improved survival with vemurafenib in melanoma with BRAF V600e mutation. *New England Journal of Medicine* **364**, 2507–2516 (2011).

[207] McArthur, G. A. *et al.* Safety and efficacy of vemurafenib in BRAF(V600E) and BRAF(V600K) mutation-positive melanoma (BRIM-3): extended follow-up of a phase 3, randomised, open-label study. *The Lancet Oncology* **15**, 323–332 (2014).

[208] National Cancer Institute. Drugs Approved for Melanoma. URL: https://www.cancer.gov/about-cancer/treatment/drugs/melanoma.

[209] Ballantyne, A. D. & Garnock-Jones, K. P. Dabrafenib: first global approval. *Drugs* **73**, 1367–1376 (2013).

[210] Wright, C. J. M. & McCormack, P. L. Trametinib: first global approval. *Drugs* **73**, 1245–1254 (2013).

[211] Woodman, S. E., Lazar, A. J., Aldape, K. D. & Davies, M. A. New Strategies in Melanoma: Molecular Testing in Advanced Disease. *Clinical Cancer Research* **18**, 1195–1200 (2012).

[212] Long, G. V. *et al.* Increased MAPK reactivation in early resistance to dabrafenib/trametinib combination therapy of BRAF-mutant metastatic melanoma. *Nature Communications* **5**, 5694 (2014).

[213] Shi, H. *et al.* Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy. *Cancer Discovery* **4**, 80–93 (2014).

[214] Schadendorf, D. *et al.* Three-year pooled analysis of factors associated with clinical outcomes across dabrafenib and trametinib combination therapy phase 3 randomised trials. *European Journal of Cancer* **82**, 45–55 (2017).

[215] Long, G. V. *et al.* Combined BRAF and MEK Inhibition versus BRAF Inhibition Alone in Melanoma. *New England Journal of Medicine* **371**, 1877–1888 (2014).

[216] Larkin, J. *et al.* Combined vemurafenib and cobimetinib in BRAF-mutated melanoma. *New England Journal of Medicine* **371**, 1867–1876 (2014).

[217] Wong, D. J. *et al.* Antitumor activity of the ERK inhibitor SCH722984 against BRAF mutant, NRAS mutant and wild-type melanoma. *Molecular Cancer* **13**, 194 (2014).

[218] FDA Approves YERVOY™ (ipilimumab) for the Treatment of Patients with Newly Diagnosed or Previously-Treated Unresectable or Metastatic Melanoma, the Deadliest Form of Skin Cancer | BMS Newsroom. URL: https://news.bms.com/press-release/rd-news/fda-approves-yervoy-ipilimumab-treatment-patients-newly-diagnosed-or-previousl.

[219] Matsushita, H. *et al.* Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* **482**, 400–404 (2012).

[220] McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016).

[221] Burnet, F. M. The Concept of Immunological Surveillance. *Immunological Aspects of Neoplasia* **13**, 1–27 (1970).

[222] Burnet, M. Cancer—A Biological Approach. *British Medical Journal* **1**, 841–847 (1957).

[223] Burnet, F. M. Immunological Surveillance in Neoplasia. *Immunological Reviews* **7**, 3–25 (1971).

[224] Dunn, G. P., Bruce, A. T., Ikeda, H., Old, L. J. & Schreiber, R. D. Cancer immunoediting: from immunosurveillance to tumor escape. *Nature Immunology* **3**, 991 (2002).

[225] Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews Cancer* **12**, 252–264 (2012).

[226] Zou, W. & Chen, L. Inhibitory B7-family molecules in the tumour microenvironment. *Nature Reviews Immunology* **8**, 467–477 (2008).

[227] Ribas, A. Tumor Immunotherapy Directed at PD-1. *New England Journal of Medicine* **366**, 2517–2519 (2012).

[228] Murphy, K. *Janeway's Immunobiology* (Garland Science, 2011), 8th edn.

[229] Stamper, C. C. *et al.* Crystal structure of the B7-1/CTLA-4 complex that inhibits human immune responses. *Nature* **410**, 608 (2001).

[230] Waterhouse, P. *et al.* Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4. *Science* **270**, 985–988 (1995).

[231] Jiang, Y., Li, Y. & Zhu, B. T-cell exhaustion in the tumor microenvironment. *Cell Death & Disease* **6**, e1792 (2015).

[232] Blackburn, S. D. *et al.* Coregulation of $CD8^+$ T cell exhaustion by multiple inhibitory receptors during chronic viral infection. *Nature Immunology* **10**, 29–37 (2009).

[233] Takahashi, T. *et al.* Immunologic Self-Tolerance Maintained by Cd25+Cd4+Regulatory T Cells Constitutively Expressing Cytotoxic T Lymphocyte–Associated Antigen 4. *Journal of Experimental Medicine* **192**, 303–310 (2000).

[234] Read, S., Malmström, V. & Powrie, F. Cytotoxic T Lymphocyte–Associated Antigen 4 Plays an Essential Role in the Function of Cd25+Cd4+ Regulatory Cells

That Control Intestinal Inflammation. *Journal of Experimental Medicine* **192**, 295–302 (2000).

[235] Simpson, T. R. *et al.* Fc-dependent depletion of tumor-infiltrating regulatory T cells co-defines the efficacy of anti–CTLA-4 therapy against melanoma. *Journal of Experimental Medicine* **210**, 1695–1710 (2013).

[236] Qureshi, O. S. *et al.* Trans-Endocytosis of CD80 and CD86: A Molecular Basis for the Cell-Extrinsic Function of CTLA-4. *Science* **332**, 600–603 (2011).

[237] Wherry, E. J. T cell exhaustion. *Nature Immunology* **12**, 492–499 (2011).

[238] Sharpe, A. H. & Freeman, G. J. The B7–CD28 superfamily. *Nature Reviews Immunology* **2**, 116 (2002).

[239] Agata, Y. *et al.* Expression of the PD-1 antigen on the surface of stimulated mouse T and B lymphocytes. *International Immunology* **8**, 765–772 (1996).

[240] Chen, N. *et al.* Upregulation of PD-L1 by EGFR Activation Mediates the Immune Escape in EGFR-Driven NSCLC: Implication for Optional Immune Targeted Therapy for NSCLC Patients with EGFR Mutation. *Journal of Thoracic Oncology* **10**, 910–923 (2015).

[241] Wang, X., Teng, F., Kong, L. & Yu, J. PD-L1 expression in human cancers and its association with clinical outcomes. *OncoTargets and Therapy* **9**, 5023–5039 (2016).

[242] Dong, H. *et al.* Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion. *Nature Medicine* **8**, 793–800 (2002).

[243] Robert, C. *et al.* Nivolumab in Previously Untreated Melanoma without BRAF Mutation. *New England Journal of Medicine* **372**, 320–330 (2015).

[244] Bristol-Myers Squibb Receives Accelerated Approval of Opdivo (nivolumab) from the U.S. Food and Drug Administration | BMS Newsroom. URL: https://news.bms.com/press-release/bristol-myers-squibb-receives-accelerated-approval-opdivo-nivolumab-us-food-and-drug-a.

[245] Robert, C. *et al.* Pembrolizumab versus Ipilimumab in Advanced Melanoma. *New England Journal of Medicine* **372**, 2521–2532 (2015).

[246] Merck Receives Accelerated Approval of KEYTRUDA® (pembrolizumab), the First FDA-Approved Anti-PD-1 Therapy | Merck Newsroom Home. URL: https://www.mrknewsroom.com/news-release/prescription-medicine-news/merck-receives-accelerated-approval-keytruda-pembrolizumab-f.

[247] Dummer, R., Hauschild, A., Lindenblatt, N., Pentheroudakis, G. & Keilholz, U. Cutaneous melanoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology* **26**, 8 (2018).

[248] Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. Primary, Adaptive and Acquired Resistance to Cancer Immunotherapy. *Cell* **168**, 707–723 (2017).

[249] Atezolizumab (TECENTRIQ). URL: https://www.fda.gov/drugs/resources-information-approved-drugs/atezolizumab-tecentriq.

[250] FDA approves durvalumab after chemoradiation for unresectable stage III NSCLC. URL: https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-durvalumab-after-chemoradiation-unresectable-stage-iii-nsclc.

[251] Avelumab (BAVENCIO). URL: https://www.fda.gov/drugs/resources-information-approved-drugs/avelumab-bavencio
(2019).

[252] Hellmann, M. D. *et al.* Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. *New England Journal of Medicine* **378**, 2093–2104 (2018).

[253] Wolchok, J. D. *et al.* Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *New England Journal of Medicine* **377**, 1345–1356 (2017).

[254] Hodi, F. S. *et al.* Nivolumab plus ipilimumab or nivolumab alone versus ipilimumab alone in advanced melanoma (CheckMate 067): 4-year outcomes of a multicentre, randomised, phase 3 trial. *The Lancet Oncology* **19**, 1480–1492 (2018).

[255] Huang, A. C. *et al.* A single dose of neoadjuvant PD-1 blockade predicts clinical outcomes in resectable melanoma. *Nature Medicine* **25**, 454–461 (2019).

[256] Topalian, S. L. *et al.* Survival, durable tumor remission, and long-term safety in patients with advanced melanoma receiving nivolumab. *Journal of Clinical Oncology* **32**, 1020–1030 (2014).

[257] Trunzer, K. *et al.* Pharmacodynamic effects and mechanisms of resistance to vemurafenib in patients with metastatic melanoma. *Journal of Clinical Oncology* **31**, 1767–1774 (2013).

[258] Solit, D. B. & Rosen, N. Resistance to BRAF Inhibition in Melanomas. *New England Journal of Medicine* **364**, 772–774 (2011).

[259] Hauschild, A. *et al.* Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. *The Lancet* **380**, 358–365 (2012).

[260] Sosman, J. A. *et al.* Survival in BRAF V600–Mutant Advanced Melanoma Treated with Vemurafenib. *New England Journal of Medicine* **366**, 707–714 (2012).

[261] Postow, M. A., Sidlow, R. & Hellmann, M. D. Immune-Related Adverse Events Associated with Immune Checkpoint Blockade. *New England Journal of Medicine* **378**, 158–168 (2018).

[262] FDA unveils a streamlined path for the authorization of tumor profiling tests alongside its latest product action. URL: https://www.fda.gov/news-events/press-announcements/fda-unveils-streamlined-path-authorization-tumor-profiling-tests-alongside-its-latest-product-action.

[263] Samstein, R. M. *et al.* Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nature Genetics* 1 (2019).

[264] Rizvi, N. A. *et al.* Mutational landscape determines sensitivity to PD-1 blockade in non–small cell lung cancer. *Science* **348**, 124–128 (2015).

[265] Snyder, A. *et al.* Genetic basis for clinical response to CTLA-4 blockade in melanoma. *New England Journal of Medicine* **371**, 2189–2199 (2014).

[266] Khunger, M. *et al.* Programmed Cell Death 1 (PD-1) Ligand (PD-L1) Expression in Solid Tumors As a Predictive Biomarker of Benefit From PD-1/PD-L1 Axis Inhibitors: A Systematic Review and Meta-Analysis. *JCO Precision Oncology* 1–15 (2017).

[267] Weber, J. S. *et al.* Nivolumab versus chemotherapy in patients with advanced melanoma who progressed after anti-CTLA-4 treatment (CheckMate 037): a

randomised, controlled, open-label, phase 3 trial. *The Lancet Oncology* **16**, 375–384 (2015).

[268] Riaz, N. *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171**, 934–949.e15 (2017).

[269] Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).

[270] Ji, R.-R. *et al.* An immune-active tumor microenvironment favors clinical response to ipilimumab. *Cancer Immunology, Immunotherapy* **61**, 1019–1031 (2012).

[271] The Nobel Prize in Physiology or Medicine 2018. URL: https://www.nobelprize.org/prizes/medicine/2018/press-release.

[272] Dong, H., Zhu, G., Tamada, K. & Chen, L. B7-H1, a third member of the B7 family, co-stimulates T-cell proliferation and interleukin-10 secretion. *Nature Medicine* **5**, 1365 (1999).

[273] Freeman, G. J. *et al.* Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *Journal of Experimental Medicine* **192**, 1027–1034 (2000).

[274] Garcia-Diaz, A. *et al.* Interferon Receptor Signaling Pathways Regulating PD-L1 and PD-L2 Expression. *Cell Reports* **19**, 1189–1201 (2017).

[275] Carter, L. *et al.* PD-1:PD-L inhibitory pathway affects both CD4(+) and CD8(+) T cells and is overcome by IL-2. *European Journal of Immunology* **32**, 634–643 (2002).

[276] Habicht, A. *et al.* A Link between PDL1 and T Regulatory Cells in Fetomaternal Tolerance. *The Journal of Immunology* **179**, 5211–5219 (2007).

[277] Guleria, I. *et al.* A critical role for the programmed death ligand 1 in fetomaternal tolerance. *Journal of Experimental Medicine* **202**, 231–237 (2005).

[278] Hori, J. *et al.* B7-H1-Induced Apoptosis as a Mechanism of Immune Privilege of Corneal Allografts. *The Journal of Immunology* **177**, 5928–5935 (2006).

[279] Iwai, Y. *et al.* Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12293–12297 (2002).

[280] Barber, D. L. *et al.* Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature* **439**, 682–687 (2006).

[281] Freeman, G. J., Wherry, E. J., Ahmed, R. & Sharpe, A. H. Reinvigorating exhausted HIV-specific T cells via PD-1-PD-1 ligand blockade. *Journal of Experimental Medicine* **203**, 2223–2227 (2006).

[282] Tang, J. *et al.* Trial watch: The clinical trial landscape for PD1/PDL1 immune checkpoint inhibitors. *Nature Reviews Drug Discovery* **17**, 854–855 (2018).

[283] Gong, J., Chehrazi-Raffle, A., Reddi, S. & Salgia, R. Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations. *Journal for Immunotherapy of Cancer* **6** (2018).

[284] Butte, M. J., Keir, M. E., Phamduy, T. B., Sharpe, A. H. & Freeman, G. J. Programmed Death-1 Ligand 1 Interacts Specifically with the B7-1 Costimulatory Molecule to Inhibit T Cell Responses. *Immunity* **27**, 111–122 (2007).

[285] Butte, M. J., Peña-Cruz, V., Kim, M.-J., Freeman, G. J. & Sharpe, A. H. Interaction of human PD-L1 and B7-1. *Molecular Immunology* **45**, 3567–3572 (2008).

[286] Park, J.-J. *et al.* B7-H1/CD80 interaction is required for the induction and maintenance of peripheral T-cell tolerance. *Blood* **116**, 1291–1298 (2010).

[287] Loke, P. & Allison, J. P. PD-L1 and PD-L2 are differentially regulated by Th1 and Th2 cells. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5336–5341 (2003).

[288] Yearley, J. H. *et al.* PD-L2 Expression in Human Tumors: Relevance to Anti-PD-1 Therapy in Cancer. *Clinical Cancer Research* **23**, 3158–3167 (2017).

[289] Taube, J. M. *et al.* Association of PD-1, PD-1 Ligands, and Other Features of the Tumor Immune Microenvironment with Response to Anti–PD-1 Therapy. *Clinical Cancer Research* **20**, 5064–5074 (2014).

[290] Tseng, S.-Y. *et al.* B7-Dc, a New Dendritic Cell Molecule with Potent Costimulatory Properties for T Cells. *Journal of Experimental Medicine* **193**, 839–846 (2001).

[291] Shin, T. *et al.* In vivo costimulatory role of B7-DC in tuning T helper cell 1 and cytotoxic T lymphocyte responses. *Journal of Experimental Medicine* **201**, 1531–1541 (2005).

[292] Shin, T. *et al.* Cooperative B7-1/2 (CD80/CD86) and B7-DC Costimulation of CD4+ T Cells Independent of the PD-1 Receptor. *Journal of Experimental Medicine* **198**, 31–38 (2003).

[293] Zhang, Y. *et al.* Regulation of T cell activation and tolerance by PDL2. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11695–11700 (2006).

[294] Latchman, Y. *et al.* PD-L2 is a second ligand for PD-1 and inhibits T cell activation. *Nature Immunology* **2**, 261 (2001).

[295] Dighe, A. S., Richards, E., Old, L. J. & Schreiber, R. D. Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFNγ receptors. *Immunity* **1**, 447–456 (1994).

[296] Beatty, G. & Paterson, Y. IFN-gamma-dependent inhibition of tumor angiogenesis by tumor-infiltrating CD4+ T cells requires tumor responsiveness to IFN-gamma. *The Journal of Immunology* **166**, 2276–2282 (2001).

[297] Detjen, K. M., Farwig, K., Welzel, M., Wiedenmann, B. & Rosewicz, S. Interferon gamma inhibits growth of human pancreatic carcinoma cells via caspase-1 dependent induction of apoptosis. *Gut* **49**, 251–262 (2001).

[298] Schoenborn, J. R. & Wilson, C. B. Regulation of interferon-gamma during innate and adaptive immune responses. *Advances in Immunology* **96**, 41–101 (2007).

[299] Chin, Y. E. *et al.* Cell Growth Arrest and Induction of Cyclin-Dependent Kinase Inhibitor p21waf1/CIP1 Mediated by STAT1. *Science* **272**, 719–722 (1996).

[300] Taube, J. M. *et al.* Colocalization of Inflammatory Response with B7-H1 Expression in Human Melanocytic Lesions Supports an Adaptive Resistance Mechanism of Immune Escape. *Science Translational Medicine* **4**, 127ra37 (2012).

[301] Spranger, S. *et al.* Up-Regulation of PD-L1, IDO, and Tregs in the Melanoma Tumor Microenvironment Is Driven by CD8+ T Cells. *Science Translational Medicine* **5**, 200ra116 (2013).

[302] Sun, J.-M. *et al.* PD-L1 expression and survival in patients with non-small cell lung cancer (NSCLC) in Korea. *Journal of Clinical Oncology* **32**, 8066–8066 (2014).

[303] Song, G., Yu, J. & Xue, S. The prognostic significance of PD-L1 expression in patients with glioma: A meta-analysis. *Scientific Reports* **7**, 4231 (2017).

[304] Gadiot, J. *et al.* Overall survival and PD-L1 expression in metastasized malignant melanoma. *Cancer* **117**, 2192–2201 (2011).

[305] Long, G. V. *et al.* Tumor PD-L1 expression, immune cell correlates and PD-1+ lymphocytes in sentinel lymph node melanoma metastases. *Modern Pathology* **28**, 1535 (2015).

[306] Steiniche, T. *et al.* PD-L1 expression and survival among melanoma patients treated with standard immunotherapy or chemotherapy. *Journal of the European Academy of Dermatology and Venereology* **31**, e319–e321 (2017).

[307] Hino, R. *et al.* Tumor cell expression of programmed cell death-1 ligand 1 is a prognostic factor for malignant melanoma. *Cancer* **116**, 1757–1766 (2010).

[308] Massi, D. *et al.* PD-L1 marks a subset of melanomas with a shorter overall survival and distinct genetic and morphological characteristics. *Annals of Oncology* **25**, 2433–2442 (2014).

[309] Wang, X. *et al.* Inflammatory cytokines IL-17 and TNF-α up-regulate PD-L1 expression in human prostate and colon cancer cells. *Immunology Letters* **184**, 7–14 (2017).

[310] Liang, Y. *et al.* Targeting IFNα to tumor by anti-PD-L1 creates feedforward antitumor responses to overcome checkpoint blockade resistance. *Nature Communications* **9**, 4586 (2018).

[311] Beswick, E. J. *et al.* TLR4 Activation Enhances the PD-L1-Mediated Tolerogenic Capacity of Colonic CD90+ Stromal Cells,. *The Journal of Immunology* **193**, 2218–2229 (2014).

[312] Liu, J. *et al.* Plasma cells from multiple myeloma patients express B7-H1 (PD-L1) and increase expression after stimulation with IFN-γ and TLR ligands via a MyD88-, TRAF6-, and MEK-dependent pathway. *Blood* **110**, 296–304 (2007).

[313] Wrangle, J. *et al.* Alterations of immune response of non-small cell lung cancer with Azacytidine. *Oncotarget* **4**, 2067–2079 (2013).

[314] Goltz, D., Gevensleben, H., Dietrich, J. & Dietrich, D. PD-L1 (CD274) promoter methylation predicts survival in colorectal cancer patients. *OncoImmunology* **6**, e1257454 (2017).

[315] Noman, M. Z. *et al.* PD-L1 is a novel direct target of HIF-1α, and its blockade under hypoxia enhanced MDSC-mediated T cell activation. *Journal of Experimental Medicine* **211**, 781–790 (2014).

[316] Hogg, S. J. *et al.* BET-Bromodomain Inhibitors Engage the Host Immune System and Regulate Expression of the Immune Checkpoint Ligand PD-L1. *Cell Reports* **18**, 2162–2174 (2017).

[317] Woods, D. M. *et al.* HDAC Inhibition Upregulates PD-1 Ligands in Melanoma and Augments Immunotherapy with PD-1 Blockade. *Cancer Immunology Research* **3**, 1375–1385 (2015).

[318] M, L. *et al.* Essential role of HDAC6 in the regulation of PD-L1 in melanoma. *Molecular Oncology* **10**, 735–750 (2016).

[319] Booth, L. *et al.* HDAC inhibitors enhance the immunotherapy response of melanoma cells. *Oncotarget* **8**, 83155–83170 (2017).

[320] Chen, L. *et al.* Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression. *Nature Communications* **5**, 5241 (2014).

[321] Wang, X. *et al.* Tumor suppressor miR-34a targets PD-L1 and functions as a potential immunotherapeutic target in acute myeloid leukemia. *Cellular Signalling* **27**, 443–452 (2015).

[322] Gong, A.-Y. *et al.* MicroRNA-513 regulates B7-H1 translation and is involved in IFN-gamma-induced B7-H1 expression in cholangiocytes. *The Journal of Immunology* **182**, 1325–1333 (2009).

[323] Sun, C., Mezzadra, R. & Schumacher, T. N. Regulation and Function of the PD-L1 Checkpoint. *Immunity* **48**, 434–452 (2018).

[324] Wang, W. *et al.* A miR-570 binding site polymorphism in the B7-H1 gene is associated with the risk of gastric adenocarcinoma. *Human Genetics* **132**, 641–648 (2013).

[325] Budczies, J. *et al.* Pan-cancer analysis of copy number changes in programmed death-ligand 1 (PD-L1, CD274) – associations with gene expression, mutational load, and survival. *Genes, Chromosomes and Cancer* **55**, 626–639 (2016).

[326] George, J. *et al.* Genomic Amplification of CD274 (PD-L1) in Small-Cell Lung Cancer. *Clinical Cancer Research* **23**, 1220–1226 (2017).

[327] Goodman, A. M. *et al.* Prevalence of PDL1 Amplification and Preliminary Response to Immune Checkpoint Blockade in Solid Tumors. *JAMA Oncology* **4**, 1237–1244 (2018).

[328] Green, M. R. *et al.* Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood* **116**, 3268–3277 (2010).

[329] Ansell, S. M. *et al.* PD-1 Blockade with Nivolumab in Relapsed or Refractory Hodgkin's Lymphoma. *New England Journal of Medicine* **372**, 311–319 (2015).

[330] Kataoka, K. *et al.* Aberrant PD-L1 expression through 3'-UTR disruption in multiple cancers. *Nature* **534**, 402–406 (2016).

[331] Twa, D. D. W. *et al.* Genomic rearrangements involving programmed death ligands are recurrent in primary mediastinal large B-cell lymphoma. *Blood* **123**, 2062–2065 (2014).

[332] Steidl, C. *et al.* MHC class II transactivator *CIITA* is a recurrent gene fusion partner in lymphoid cancers. *Nature* **471**, 377–381 (2011).

[333] Manguso, R. T. *et al.* In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature* **547**, 413–418 (2017).

[334] Spranger, S., Bao, R. & Gajewski, T. F. Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231–235 (2015).

[335] Parsa, A. T. *et al.* Loss of tumor suppressor PTEN function increases B7-H1 expression and immunoresistance in glioma. *Nature Medicine* **13**, 84–88 (2007).

[336] Marzec, M. *et al.* Oncogenic kinase NPM/ALK induces through STAT3 expression of immunosuppressive protein CD274 (PD-L1, B7-H1). *Proceedings of the National Academy of Sciences of the United States of America* **105**, 20852–20857 (2008).

[337] Hu-Lieskovan, S. *et al.* Improved antitumor activity of immunotherapy with BRAF and MEK inhibitors in BRAFV600e melanoma. *Science Translational Medicine* **7**, 279ra41 (2015).

[338] Jiang, X., Zhou, J., Giobbie-Hurder, A., Wargo, J. & Hodi, F. S. The activation of MAPK in melanoma cells resistant to BRAF inhibition promotes PD-L1 expression that is reversible by MEK and PI3k inhibition. *Clinical Cancer Research* **19**, 598–609 (2013).

[339] Liu, L. *et al.* The BRAF and MEK Inhibitors Dabrafenib and Trametinib: Effects on Immune Function and in Combination with Immunomodulatory Antibodies Targeting PD-1, PD-L1, and CTLA-4. *Clinical Cancer Research* **21**, 1639–1651 (2015).

[340] Qian, Y. *et al.* TLR4 Signaling Induces B7-H1 Expression Through MAPK Pathways in Bladder Cancer Cells. *Cancer Investigation* **26**, 816–821 (2008).

[341] Yamamoto, R. *et al.* B7-H1 expression is regulated by MEK/ERK signaling pathway in anaplastic large cell lymphoma and Hodgkin lymphoma. *Cancer Science* **100**, 2093–2100 (2009).

[342] Frederick, D. T. *et al.* BRAF Inhibition Is Associated with Enhanced Melanoma Antigen Expression and a More Favorable Tumor Microenvironment in Patients with Metastatic Melanoma. *Clinical Cancer Research* **19**, 1225–1231 (2013).

[343] Wilmott, J. S. *et al.* Selective BRAF Inhibitors Induce Marked T-cell Infiltration into Human Metastatic Melanoma. *Clinical Cancer Research* **18**, 1386–1394 (2012).

[344] Boni, A. *et al.* Selective BRAFV600e Inhibition Enhances T-Cell Recognition of Melanoma without Affecting Lymphocyte Function. *Cancer Research* **70**, 5213–5219 (2010).

[345] Donia, M. *et al.* BRAF inhibition improves tumor recognition by the immune system. *OncoImmunology* **1**, 1476–1483 (2012).

[346] Liu, C. *et al.* BRAF Inhibition Increases Tumor Infiltration by T cells and Enhances the Antitumor Activity of Adoptive Immunotherapy in Mice. *Clinical Cancer Research* **19**, 393–403 (2013).

[347] Cooper, Z. A. *et al.* Response to BRAF Inhibition in Melanoma Is Enhanced When Combined with Immune Checkpoint Blockade. *Cancer Immunology Research* **2**, 643–654 (2014).

[348] Deken, M. A. *et al.* Targeting the MAPK and PI3k pathways in combination with PD1 blockade in melanoma. *OncoImmunology* **5**, e1238557 (2016).

[349] Ascierto, P. A. *et al.* KEYNOTE-022 Part 3: Phase II randomized study of 1l dabrafenib (D) and trametinib (T) plus pembrolizumab (Pembro) or placebo (PBO) for BRAF-mutant advanced melanoma. *Annals of Oncology* **29** (2018).

[350] A Study of the Safety and Efficacy of Pembrolizumab (MK-3475) in Combination With Trametinib and Dabrafenib in Participants With Advanced Melanoma (MK-3475-022/KEYNOTE-022). URL: https://clinicaltrials.gov/ct2/show/NCT02130466.

[351] Coelho, M. A. *et al.* Oncogenic RAS Signaling Promotes Tumor Immunoresistance by Stabilizing PD-L1 mRNA. *Immunity* **47**, 1083–1099.e6 (2017).

[352] Song, M. *et al.* PTEN Loss Increases PD-L1 Protein Expression and Affects the Correlation between PD-L1 Expression and Clinical Parameters in Colorectal Cancer. *PLOS ONE* **8**, e65821 (2013).

[353] Lastwika, K. J. *et al.* Control of PD-L1 Expression by Oncogenic Activation of the AKT–mTOR Pathway in Non–Small Cell Lung Cancer. *Cancer Research* **76**, 227–238 (2016).

[354] Mittendorf, E. A. *et al.* PD-L1 Expression in Triple-Negative Breast Cancer. *Cancer Immunology Research* **2**, 361–370 (2014).

[355] Ota, K. *et al.* Induction of PD-L1 Expression by the EML4–ALK Oncoprotein and Downstream Signaling Pathways in Non–Small Cell Lung Cancer. *Clinical Cancer Research* **21**, 4014–4021 (2015).

[356] Akbay, E. A. *et al.* Activation of the PD-1 pathway contributes to immune escape in EGFR-driven lung tumors. *Cancer Discovery* **3**, 1355–1363 (2013).

[357] Dorand, R. D. *et al.* Cdk5 disruption attenuates tumor PD-L1 expression and promotes antitumor immunity. *Science* **353**, 399–403 (2016).

[358] Kim, M. H. *et al.* YAP-Induced PD-L1 Expression Drives Immune Evasion in BRAFi-Resistant Melanoma. *Cancer Immunology Research* **6**, 255–266 (2018).

[359] Rensburg, H. J. J. v. *et al.* The Hippo Pathway Component TAZ Promotes Immune Evasion in Human Cancer through PD-L1. *Cancer Research* **78**, 1457–1470 (2018).

[360] Casey, S. C. *et al.* MYC regulates the antitumor immune response through CD47 and PD-L1. *Science* **352**, 227–231 (2016).

[361] Gowrishankar, K. *et al.* Inducible but Not Constitutive Expression of PD-L1 in Human Melanoma Cells Is Dependent on Activation of NF-κB. *PLOS ONE* **10**, e0123410 (2015).

[362] Kortlever, R. M. *et al.* Myc Cooperates with Ras by Programming Inflammation and Immune Suppression. *Cell* **171**, 1301–1315.e14 (2017).

[363] Li, C.-W. *et al.* Glycosylation and stabilization of programmed death ligand-1 suppresses T-cell activity. *Nature Communications* **7**, 12632 (2016).

[364] Lim, S.-O. *et al.* Deubiquitination and Stabilization of PD-L1 by CSN5. *Cancer Cell* **30**, 925–939 (2016).

[365] Zhang, J. *et al.* Cyclin D–CDK4 kinase destabilizes PD-L1 via cullin 3–SPOP to control cancer immune surveillance. *Nature* **553**, 91–95 (2018).

[366] Mezzadra, R. *et al.* Identification of CMTM6 and CMTM4 as PD-L1 protein regulators. *Nature* **549**, 106–110 (2017).

[367] Burr, M. L. *et al.* CMTM6 maintains the expression of PD-L1 and regulates anti-tumour immunity. *Nature* **549**, 101 (2017).

[368] Wang, H. *et al.* HIP1r targets PD-L1 to lysosomal degradation to alter T cell–mediated cytotoxicity. *Nature Chemical Biology* **15**, 42 (2019).

[369] Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution* **60**, 174–182 (2005).

[370] Jansen, R., Embden, J. D. A. v., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* **43**, 1565–1575 (2002).

[371] Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338 (2012).

[372] Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biology Direct* **1**, 7 (2006).

[373] Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* **315**, 1709–1712 (2007).

[374] Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).

[375] Hsu, P. D., Lander, E. S. & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**, 1262–1278 (2014).

[376] Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).

[377] Makarova, K. S. *et al.* Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology* **9**, 467–477 (2011).

[378] Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).

[379] Brouns, S. J. J. *et al.* Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* **321**, 960–964 (2008).

[380] Deltcheva, E. *et al.* CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).

[381] Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).

[382] Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2579–E2586 (2012).

[383] Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**, 2281–2308 (2013).

[384] Adli, M. The CRISPR tool kit for genome editing and beyond. *Nature Communications* **9**, 1911 (2018).

[385] Cromwell, C. R. *et al.* Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity. *Nature Communications* **9**, 1448 (2018).

[386] Kleinstiver, B. P. *et al.* High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).

[387] Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).

[388] Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature Protocols* **12**, 828–863 (2017).

[389] Klann, T. S. *et al.* CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. *Nature Biotechnology* **35**, 561–568 (2017).

[390] Dow, L. E. *et al.* Inducible *in vivo* genome editing with CRISPR-Cas9. *Nature Biotechnology* **33**, 390–394 (2015).

[391] Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173–1183 (2013).

[392] Jaitin, D. A. *et al.* Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell* **167**, 1883–1896.e15 (2016).

[393] Jinek, M. *et al.* RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013).

[394] Kleinstiver, B. P. *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* **523**, 481–485 (2015).

[395] Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology* **34**, 184–191 (2016).

[396] Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nature Biotechnology* **32**, 1262–1267 (2014).

[397] Mali, P. *et al.* RNA-Guided Human Genome Engineering via Cas9. *Science* **339**, 823–826 (2013).

[398] Shalem, O., Sanjana, N. E. & Zhang, F. High-throughput functional genomics using CRISPR–Cas9. *Nature Reviews Genetics* **16**, 299–311 (2015).

[399] Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science* **343**, 80–84 (2014).

[400] Shalem, O. *et al.* Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* **343**, 84–87 (2014).

[401] Jackson, A. L. *et al.* Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* **12**, 1179–1187 (2006).

[402] Heigwer, F. *et al.* CRISPR library designer (CLD): software for multispecies design of single guide RNA libraries. *Genome Biology* **17**, 55 (2016).

[403] Park, J. & Bae, S. Cpf1-Database: web-based genome-wide guide RNA library design for gene knockout screens using CRISPR-Cpf1. *Bioinformatics* **34**, 1077–1079 (2018).

[404] Addgene: CRISPR Pooled gRNA Libraries. URL: https://www.addgene.org/crispr/libraries.

[405] Hart, T. *et al.* Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. *G3: Genes, Genomes, Genetics* **7**, 2719–2727 (2017).

[406] Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, M. D. C. & Yusa, K. Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nature Biotechnology* **32**, 267–273 (2014).

[407] Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nature Methods* **11**, 783–784 (2014).

[408] Sanson, K. R. *et al.* Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications* **9**, 5416 (2018).

[409] Kurata, M., Yamamoto, K., Moriarity, B. S., Kitagawa, M. & Largaespada, D. A. CRISPR/Cas9 library screening for drug target discovery. *Journal of Human Genetics* **63**, 179 (2018).

[410] Miles, L. A., Garippa, R. J. & Poirier, J. T. Design, execution, and analysis of pooled in vitro CRISPR/Cas9 screens. *The FEBS Journal* **283**, 3170–3180 (2016).

[411] Tzelepis, K. *et al.* A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia. *Cell Reports* **17**, 1193–1205 (2016).

[412] Hart, T. *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515–1526 (2015).

[413] Wang, T. *et al.* Identification and characterization of essential genes in the human genome. *Science* **350**, 1096–1101 (2015).

[414] Park, R. J. *et al.* A genome-wide CRISPR screen identifies a restricted set of HIV host dependency factors. *Nature Genetics* **49**, 193–203 (2017).

[415] Parnas, O. *et al.* A Genome-wide CRISPR Screen in Primary Immune Cells to Dissect Regulatory Networks. *Cell* **162**, 675–686 (2015).

[416] DeJesus, R. *et al.* Functional CRISPR screening identifies the ufmylation pathway as a regulator of SQSTM1/p62. *eLife* **5** (2016).

[417] Ren, Q. *et al.* A Dual-Reporter System for Real-Time Monitoring and High-throughput CRISPR/Cas9 Library Screening of the Hepatitis C Virus. *Scientific Reports* **5**, 8865 (2015).

[418] Hart, T. & Moffat, J. BAGEL: a computational framework for identifying essential genes from pooled library screens. *BMC Bioinformatics* **17**, 164 (2016).

[419] Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biology* **15**, 554 (2014).

[420] Pusapati, G. V. *et al.* CRISPR Screens Uncover Genes that Regulate Target Cell Sensitivity to the Morphogen Sonic Hedgehog. *Developmental Cell* **44**, 113–129.e8 (2018).

[421] Lebensohn, A. M. *et al.* Comparative genetic screens in human cells reveal new regulatory mechanisms in WNT signaling. *eLife* **5** (2016).

[422] Mezzadra, R. *et al.* SLFN11 can sensitize tumor cells towards IFN-γ-mediated T cell killing. *PLOS ONE* **14**, e0212053 (2019).

[423] Sheffer, M. *et al.* Identification and Validation of Molecular Markers of Tumor Cell Resistance to Natural Killer Cells through CRISPR-Based Screens and Large-Scale Phenotypic Screens of Pooled Tumor Cell Lines. *Blood* **130**, 623–623 (2017).

[424] Han, P. *et al.* Genome-Wide CRISPR Screening Identifies JAK1 Deficiency as a Mechanism of T-Cell Resistance. *Frontiers in Immunology* **10** (2019).

[425] Pan, D. *et al.* A major chromatin regulator determines resistance of tumor cells to T cell–mediated killing. *Science* **359**, 770–775 (2018).

[426] Kearney, C. J. *et al.* Tumor immune evasion arises through loss of TNF sensitivity. *Science Immunology* **3**, eaar3451 (2018).

[427] Newton-Bishop, J. A. *et al.* Serum 25-Hydroxyvitamin D3 Levels Are Associated With Breslow Thickness at Presentation and Survival From Melanoma. *Journal of Clinical Oncology* **27**, 5439–5444 (2009).

[428] Zhang, T., Dutton-Regester, K., Brown, K. M. & Hayward, N. K. The genomic landscape of cutaneous melanoma. *Pigment Cell & Melanoma Research* **29**, 266–283 (2016).

[429] Martincorena, I. *et al.* High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015).

[430] Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029–1041.e21 (2017).

[431] Gao, J. *et al.* Loss of IFN-γ Pathway Genes in Tumor Cells as a Mechanism of Resistance to Anti-CTLA-4 Therapy. *Cell* **167**, 397–404.e9 (2016).

[432] Iorio, F. *et al.* A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **166**, 740–754 (2016).

[433] Wittig, M. *et al.* Development of a high-resolution NGS-based HLA-typing and analysis pipeline. *Nucleic Acids Research* **43**, e70–e70 (2015).

[434] Conway, C. *et al.* Gene expression profiling of paraffin-embedded primary melanoma using the DASL assay identifies increased osteopontin expression as predictive of reduced relapse-free survival. *Clinical Cancer Research* **15**, 6939–6946 (2009).

[435] Wright, A. I., Grabsch, H. I. & Treanor, D. E. RandomSpot: A web-based tool for systematic random sampling of virtual slides. *Journal of Pathology Informatics* **6**, 8 (2015).

[436] Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications* **4**, 2612 (2013).

[437] Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

[438] Tischler, G. & Leonard, S. biobambam: tools for read pair collation based algorithms on BAM files. *Source Code for Biology and Medicine* **9**, 13 (2014).

[439] Jones, D. *et al.* cgpCaVEManWrapper: Simple Execution of CaVEMan in Order to Detect Somatic Single Nucleotide Variants in NGS Data. *Current Protocols in Bioinformatics* **56**, 15.10.1–15.10.18 (2016).

[440] Chiba, K. *et al.* Cancer-associated TERT promoter mutations abrogate telomerase silencing. *eLife* **4**, e07918 (2015).

[441] Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).

[442] McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 122 (2016).

[443] Raine, K. M. *et al.* ascatNgs: Identifying Somatically Acquired Copy-Number Alterations from Whole-Genome Sequencing Data. *Current Protocols in Bioinformatics* **56**, 15.9.1–15.9.17 (2016).

[444] Consortium, I. H. G. S. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

[445] The Cancer Genome Atlas Research Network. Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67–73 (2013).

[446] The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).

[447] The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).

[448] Carter, S. L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nature Biotechnology* **30**, 413–421 (2012).

[449] Alioto, T. S. *et al.* A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing. *Nature Communications* **6**, 10001 (2015).

[450] Cleaver, J. E. & Crowley, E. UV damage, DNA repair and skin carcinogenesis. *Frontiers in Bioscience* **7**, d1024–1043 (2002).

[451] Chalkley, R. & Hunter, C. Histone-histone propinquity by aldehyde fixation of chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **72**, 1304–1308 (1975).

[452] Kavli, B., Otterlei, M., Slupphaug, G. & Krokan, H. E. Uracil in DNA—General mutagen, but normal intermediate in acquired immunity. *DNA Repair* **6**, 505–516 (2007).

[453] Lindahl, T. & Nyberg, B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**, 3610–3618 (1972).

[454] Williams, C. *et al.* A High Frequency of Sequence Alterations Is Due to Formalin Fixation of Archival Specimens. *The American Journal of Pathology* **155**, 1467–1471 (1999).

[455] Do, H. & Dobrovic, A. Sequence Artifacts in DNA from Formalin-Fixed Tissues: Causes and Strategies for Minimization. *Clinical Chemistry* **61**, 64–71 (2015).

[456] Wang, M. *et al.* Somatic Mutation Screening Using Archival Formalin-Fixed, Paraffin-Embedded Tissues by Fluidigm Multiplex PCR and Illumina Sequencing. *The Journal of Molecular Diagnostics* **17**, 521–532 (2015).

[457] Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**, 178–192 (2013).

[458] Krøigård, A. B., Thomassen, M., Lænkholm, A.-V., Kruse, T. A. & Larsen, M. J. Evaluation of Nine Somatic Variant Callers for Detection of Somatic Mutations in Exome and Targeted Deep Sequencing Data. *PLOS ONE* **11** (2016).

[459] Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnology* **31**, 213–219 (2013).

[460] Wang, Q. *et al.* Detecting somatic point mutations in cancer genome sequencing data: a comparison of mutation callers. *Genome Medicine* **5**, 91 (2013).

[461] Xu, H., DiCarlo, J., Satya, R. V., Peng, Q. & Wang, Y. Comparison of somatic mutation calling methods in amplicon and whole exome sequence data. *BMC Genomics* **15**, 244 (2014).

[462] Rashid, M., Robles-Espinoza, C. D., Rust, A. G. & Adams, D. J. Cake: a bioinformatics pipeline for the integrated analysis of somatic variants in cancer genomes. *Bioinformatics* **29**, 2208–2210 (2013).

[463] Mahmood, T. & Yang, P.-C. Western blot: technique, theory, and trouble shooting. *North American Journal of Medical Sciences* **4**, 429–434 (2012).

[464] Wan, P. T. C. *et al.* Mechanism of Activation of the RAF-ERK Signaling Pathway by Oncogenic Mutations of B-RAF. *Cell* **116**, 855–867 (2004).

[465] Yeang, C.-H., McCormick, F. & Levine, A. Combinatorial patterns of somatic gene mutations in cancer. *The FASEB Journal* **22**, 2605–2622 (2008).

[466] Cisowski, J., Sayin, V. I., Liu, M., Karlsson, C. & Bergo, M. O. Oncogene-induced senescence underlies the mutual exclusive nature of oncogenic KRAS and BRAF. *Oncogene* **35**, 1328–1333 (2016).

[467] Cisowski, J. & Bergo, M. O. What makes oncogenes mutually exclusive? *Small GTPases* **8**, 187–192 (2016).

[468] Hodgkinson, A. & Eyre-Walker, A. Variation in the mutation rate across mammalian genomes. *Nature Reviews Genetics* **12**, 756–766 (2011).

[469] Frigola, J. *et al.* Reduced mutation rate in exons due to differential mismatch repair. *Nature Genetics* **49**, 1684–1692 (2017).

[470] Polak, P. *et al.* Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **518**, 360–364 (2015).

[471] Schuster-Böckler, B. & Lehner, B. Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* **488**, 504–507 (2012).

[472] Canisius, S., Martens, J. W. M. & Wessels, L. F. A. A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. *Genome Biology* **17**, 261 (2016).

[473] Broekaert, S. M. *et al.* Genetic and morphologic features for melanoma classification. *Pigment Cell & Melanoma Research* **23**, 763–770 (2010).

[474] Wong, S. Q. *et al.* Whole exome sequencing identifies a recurrent RQCD1 P131l mutation in cutaneous melanoma. *Oncotarget* **6**, 1115–1127 (2014).

[475] Dutton-Regester, K. *et al.* A highly recurrent RPS27 5'UTR mutation in melanoma. *Oncotarget* **5**, 2912–2917 (2014).

[476] Pfam: Family: PTEN_c2 (PF10409). URL: http://pfam.xfam.org/family/PF10409.

[477] Collins, P. J., Kobayashi, Y., Nguyen, L., Trinklein, N. D. & Myers, R. M. The ets-related transcription factor GABP directs bidirectional transcription. *PLOS Genetics* **3**, e208 (2007).

[478] Fredriksson, N. J. *et al.* Recurrent promoter mutations in melanoma are defined by an extended context-specific mutational signature. *PLOS Genetics* **13**, e1006773 (2017).

[479] Mao, P. *et al.* ETS transcription factors induce a unique UV damage signature that drives recurrent mutagenesis in melanoma. *Nature Communications* **9** (2018).

[480] Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling* **6**, pl1 (2013).

[481] Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* **2**, 401–404 (2012).

[482] Loo, P. V. *et al.* Allele-specific copy number analysis of tumors. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 16910–16915 (2010).

[483] Limon, J., Dal Cin, P., Sait, S. N. J., Karakousis, C. & Sandberg, A. A. Chromosome changes in metastatic human melanoma. *Cancer Genetics and Cytogenetics* **30**, 201–211 (1988).

[484] Mertens, F., Johansson, B., Höglund, M. & Mitelman, F. Chromosomal Imbalance Maps of Malignant Solid Tumors: A Cytogenetic Survey of 3185 Neoplasms. *Cancer Research* **57**, 2765–2780 (1997).

[485] Walker, G. J. *et al.* Deletion mapping suggests that the 1p22 melanoma susceptibility gene is a tumor suppressor localized to a 9-mb interval. *Genes, Chromosomes and Cancer* **41**, 56–64 (2004).

[486] Bastian, B. C., LeBoit, P. E., Hamm, H., Bröcker, E. B. & Pinkel, D. Chromosomal gains and losses in primary cutaneous melanomas detected by comparative genomic hybridization. *Cancer Research* **58**, 2170–2175 (1998).

[487] Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

[488] Meyers, R. M. *et al.* Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nature Genetics* **49**, 1779–1784 (2017).

[489] Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564–576.e16 (2017).

[490] Santos, G. C., Zielenska, M., Prasad, M. & Squire, J. A. Chromosome 6p amplification and cancer progression. *Journal of Clinical Pathology* **60**, 1–7 (2007).

[491] Kim, M. *et al.* Comparative Oncogenomics Identifies NEDD9 as a Melanoma Metastasis Gene. *Cell* **125**, 1269–1281 (2006).

[492] Khodadoust, M. S. *et al.* Melanoma Proliferation and Chemoresistance Controlled by the DEK Oncogene. *Cancer Research* **69**, 6405–6413 (2009).

[493] Rahman, M. *et al.* Alternative preprocessing of RNA-Sequencing data in The Cancer Genome Atlas leads to improved analysis results. *Bioinformatics* **31**, 3666–3672 (2015).

[494] The Human Protein Atlas: Tissue expression of IRF4. URL: https://www.proteinatlas.org/ENSG00000137265-IRF4/tissue.

[495] Uhlén, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

[496] Carithers, L. J. *et al.* A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation and Biobanking* **13**, 311–319 (2015).

[497] Shaffer, A. L. *et al.* IRF4 addiction in multiple myeloma. *Nature* **454**, 226–231 (2008).

[498] DGIdb. URL: http://www.dgidb.org/.

[499] Young, R. J. *et al.* Loss of CDKN2a expression is a frequent event in primary invasive melanoma and correlates with sensitivity to the CDK4/6 inhibitor PD0332991 in melanoma cell lines. *Pigment Cell & Melanoma Research* **27**, 590–600 (2014).

[500] Cowan, J. M., Halaban, R. & Francke, U. Cytogenetic analysis of melanocytes from premalignant nevi and melanomas. *Journal of the National Cancer Institute* **80**, 1159–1164 (1988).

[501] Goldstein, A. M. *et al.* Features associated with germline CDKN2a mutations: a GenoMEL study of melanoma-prone families from three continents. *Journal of Medical Genetics* **44**, 99–106 (2007).

[502] Soura, E., Eliades, P., Shannon, K., Stratigos, A. & Tsao, H. Hereditary Melanoma: Update on Syndromes and Management - Genetics of familial atypical multiple mole melanoma syndrome. *Journal of the American Academy of Dermatology* **74**, 395–407 (2016).

[503] Kleffel, S. *et al.* Melanoma cell-intrinsic PD-1 receptor functions promote tumor growth. *Cell* **162**, 1242–1256 (2015).

[504] Knudson, A. G. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **68**, 820–823 (1971).

[505] Koh, J., Enders, G. H., Dynlacht, B. D. & Harlow, E. Tumour-derived p16 alleles encoding proteins defective in cell-cycle inhibition. *Nature* **375**, 506–510 (1995).

[506] Lukas, J. *et al.* Retinoblastoma-protein-dependent cell-cycle inhibition by the tumour suppressor p16. *Nature* **375**, 503–506 (1995).

[507] Stahl, J. M. *et al.* Loss of PTEN promotes tumor development in malignant melanoma. *Cancer Research* **63**, 2881–2890 (2003).

[508] Di Cristofano, A., Pesce, B., Cordon-Cardo, C. & Pandolfi, P. P. Pten is essential for embryonic development and tumour suppression. *Nature Genetics* **19**, 348–355 (1998).

[509] Samuelov, L. *et al.* SVEP1 plays a crucial role in epidermal differentiation. *Experimental Dermatology* **26**, 423–430 (2017).

[510] Shur, I., Socher, R., Hameiri, M., Fried, A. & Benayahu, D. Molecular and cellular characterization of SEL-OB/SVEP1 in osteogenic cells in vivo and in vitro. *Journal of Cellular Physiology* **206**, 420–427 (2006).

[511] IntOGen - GABRA6 gene cancer mutations. URL: https://www.intogen.org/search?gene=gabra6.

[512] Ryland, G. L. *et al.* Loss of heterozygosity: what is it good for? *BMC Medical Genomics* **8**, 45 (2015).

[513] Healy, E., Rehman, I., Angus, B. & Rees, J. L. Loss of heterozygosity in sporadic primary cutaneous melanoma. *Genes, Chromosomes and Cancer* **12**, 152–156 (1995).

[514] Healy, E. *et al.* Prognostic significance of allelic losses in primary melanoma. *Oncogene* **16**, 2213–2218 (1998).

[515] Herbst, R. A. *et al.* A defined region of loss of heterozygosity at 11q23 in cutaneous malignant melanoma. *Cancer Research* **55**, 2494–2496 (1995).

[516] Robertson, G., Coleman, A. & Lugo, T. G. A malignant melanoma tumor suppressor on human chromosome 11. *Cancer Research* **56**, 4487–4492 (1996).

[517] Hammond, D. *et al.* Melanoma-associated mutations in protein phosphatase 6 cause chromosome instability and DNA damage owing to dysregulated Aurora-A. *Journal of Cell Science* **126**, 3429–3440 (2013).

[518] El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Research* **47**, D427–D432 (2019).

[519] Pfam: Family: DSPc (PF00782). URL: http://pfam.xfam.org/family/DSPc.

[520] Walker, S. M., Downes, C. P. & Leslie, N. R. TPIP: a novel phosphoinositide 3-phosphatase. *Biochemical Journal* **360**, 277–283 (2001).

[521] Mishra, R. R., Chaudhary, J. K., Bajaj, G. D. & Rath, P. C. A Novel Human TPIP Splice-Variant (TPIP-C2) mRNA, Expressed in Human and Mouse Tissues, Strongly Inhibits Cell Growth in HeLa Cells. *PLOS ONE* **6**, e28433 (2011).

[522] Mishra, R. R., Chaudhary, J. K. & Rath, P. C. Cell cycle arrest and apoptosis by expression of a novel TPIP (TPIP-C2) cDNA encoding a C2-domain in HEK-293 cells. *Molecular Biology Reports* **39**, 7389–7402 (2012).

[523] Rodríguez-Escudero, I. *et al.* A comprehensive functional analysis of PTEN mutations: implications in tumor- and autism-related syndromes. *Human Molecular Genetics* **20**, 4132–4142 (2011).

[524] Kuemmel, A. *et al.* Humoral immune responses of lung cancer patients against the Transmembrane Phosphatase with TEnsin homology (TPTE). *Lung Cancer* **90**, 334–341 (2015).

[525] Dong, X.-Y. *et al.* Identification of two novel CT antigens and their capacity to elicit antibody response in hepatocellular carcinoma patients. *British Journal of Cancer* **89**, 291–297 (2003).

[526] Halaban, R. & Krauthammer, M. RASopathy Gene Mutations in Melanoma. *The Journal of Investigative Dermatology* **136**, 1755–1759 (2016).

[527] Peroval, M. Y., Boyd, A. C., Young, J. R. & Smith, A. L. A Critical Role for MAPK Signalling Pathways in the Transcriptional Regulation of Toll Like Receptors. *PLOS ONE* **8**, e51243 (2013).

[528] Roberts, P. J. & Der, C. J. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene* **26**, 3291–3310 (2007).

[529] Li, Y. *et al.* MAPK signaling downstream to TLR4 contributes to paclitaxel-induced peripheral neuropathy. *Brain, Behavior, and Immunity* **49**, 255–266 (2015).

[530] Benvenuti, S. *et al.* Oncogenic Activation of the RAS/RAF Signaling Pathway Impairs the Response of Metastatic Colorectal Cancers to Anti–Epidermal Growth Factor Receptor Antibody Therapies. *Cancer Research* **67**, 2643–2648 (2007).

[531] Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nature Reviews Molecular Cell Biology* **2**, 127–137 (2001).

[532] Shaw, R. J. & Cantley, L. C. Ras, PI(3)K and mTOR signalling controls tumour cell growth. *Nature* **441**, 424–430 (2006).

[533] McKay, M. M. & Morrison, D. K. Integrating signals from RTKs to ERK/MAPK. *Oncogene* **26**, 3113–3121 (2007).

[534] Montor, W. R., Salas, A. R. O. S. E. & Melo, F. H. M. d. Receptor tyrosine kinases and downstream pathways as druggable targets for cancer treatment: the current arsenal of inhibitors. *Molecular Cancer* **17**, 55 (2018).

[535] Korneev, K. V. *et al.* TLR-signaling and proinflammatory cytokines as drivers of tumorigenesis. *Cytokine* **89**, 127–135 (2017).

[536] Patel, S. P. *et al.* A Phase II Study of Gefitinib in Patients with Metastatic Melanoma. *Melanoma Research* **21**, 357–363 (2011).

[537] Stites, E. C. The Response of Cancers to BRAF Inhibition Underscores the Importance of Cancer Systems Biology. *Science Signaling* **5**, pe46–pe46 (2012).

[538] Prahallad, A. *et al.* Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* **483**, 100–103 (2012).

[539] Corcoran, R. B. *et al.* EGFR-Mediated Reactivation of MAPK Signaling Contributes to Insensitivity of BRAF-Mutant Colorectal Cancers to RAF Inhibition with Vemurafenib. *Cancer Discovery* **2**, 227–235 (2012).

[540] Matsunaga, N., Tsuchimori, N., Matsumoto, T. & Ii, M. TAK-242 (Resatorvid), a Small-Molecule Inhibitor of Toll-Like Receptor (TLR) 4 Signaling, Binds

Selectively to TLR4 and Interferes with Interactions between TLR4 and Its Adaptor Molecules. *Molecular Pharmacology* **79**, 34–41 (2011).

[541] Monnet, E. *et al.* Evidence of NI-0101 pharmacological activity, an anti-TLR4 antibody, in a randomized phase I dose escalation study in healthy volunteers receiving LPS. *Clinical Pharmacology and Therapeutics* **101**, 200–208 (2017).

[542] Bhattacharyya, S. *et al.* Pharmacological Inhibition of Toll-Like Receptor-4 Signaling by TAK242 Prevents and Induces Regression of Experimental Organ Fibrosis. *Frontiers in Immunology* **9** (2018).

[543] Gao, W., Xiong, Y., Li, Q. & Yang, H. Inhibition of Toll-Like Receptor Signaling as a Promising Therapy for Inflammatory Diseases: A Journey from Molecular to Nano Therapeutics. *Frontiers in Physiology* **8** (2017).

[544] Bell-Horner, C. L., Dohi, A., Nguyen, Q., Dillon, G. H. & Singh, M. ERK/MAPK pathway regulates GABAA receptors. *Journal of Neurobiology* **66**, 1467–1474 (2006).

[545] Aznar, S. & Lacal, J. C. Rho signals to cell growth and apoptosis. *Cancer Letters* **165**, 1–10 (2001).

[546] Symons, M. Rho family GTPases: the cytoskeleton and beyond. *Trends in Biochemical Sciences* **21**, 178–181 (1996).

[547] Jaffe, A. B. & Hall, A. RHO GTPASES: Biochemistry and Biology. *Annual Review of Cell and Developmental Biology* **21**, 247–269 (2005).

[548] Rosa, L. R. O., Soares, G. M., Silveira, L. R., Boschero, A. C. & Barbosa-Sampaio, H. C. L. ARHGAP21 as a master regulator of multiple cellular processes. *Journal of Cellular Physiology* **233**, 8477–8481 (2018).

[549] Barcellos, K. S. A. *et al.* ARHGAP21 Protein, a New Partner of α-Tubulin Involved in Cell-Cell Adhesion Formation and Essential for Epithelial-Mesenchymal Transition. *Journal of Biological Chemistry* **288**, 2179–2189 (2013).

[550] Bigarella, C. L., Borges, L., Costa, F. F. & Saad, S. T. O. ARHGAP21 modulates FAK activity and impairs glioblastoma cell migration. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1793**, 806–816 (2009).

[551] Lu, H. *et al.* Oncogenic BRAF-mediated Melanoma Cell Invasion. *Cell Reports* **15**, 2012–2024 (2016).

[552] Klein, R. M., Spofford, L. S., Abel, E. V., Ortiz, A. & Aplin, A. E. B-RAF Regulation of Rnd3 Participates in Actin Cytoskeletal and Focal Adhesion Organization. *Molecular Biology of the Cell* **19**, 498–508 (2008).

[553] Vial, E., Sahai, E. & Marshall, C. J. ERK-MAPK signaling coordinately regulates activity of Rac1 and RhoA for tumor cell motility. *Cancer Cell* **4**, 67–79 (2003).

[554] Wei, X. *et al.* Mutational and Functional Analysis Reveals ADAMTS18 Metalloproteinase as a Novel Oncogene in Melanoma. *Molecular Cancer Research* **8**, 1513–1525 (2010).

[555] Buyse, I. M., Shao, G. & Huang, S. The retinoblastoma protein binds to RIZ, a zinc-finger protein that shares an epitope with the adenovirus E1a protein. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 4467–4471 (1995).

[556] Chadwick, R. B. *et al.* Candidate tumor suppressor RIZ is frequently involved in colorectal carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 2662–2667 (2000).

[557] Viloria, C. G. *et al.* Genetic inactivation of ADAMTS15 metalloprotease in human colorectal cancer. *Cancer Research* **69**, 4926–4934 (2009).

[558] Rocks, N. *et al.* Emerging roles of ADAM and ADAMTS metalloproteinases in cancer. *Biochimie* **90**, 369–379 (2008).

[559] Porter, S., Clark, I. M., Kevorkian, L. & Edwards, D. R. The ADAMTS metalloproteinases. *Biochemical Journal* **386**, 15–27 (2005).

[560] Porter, S. *et al.* Dysregulated Expression of Adamalysin-Thrombospondin Genes in Human Breast Carcinoma. *Clinical Cancer Research* **10**, 2429–2440 (2004).

[561] Lu, T. *et al.* Adamts18 deficiency promotes colon carcinogenesis by enhancing β-catenin and p38mapk/ERK1/2 signaling in the mouse model of AOM/DSS-induced colitis-associated colorectal cancer. *Oncotarget* **8**, 18979–18990 (2017).

[562] Jin, H. *et al.* Epigenetic identification of ADAMTS18 as a novel 16q23.1 tumor suppressor frequently silenced in esophageal, nasopharyngeal and multiple other carcinomas. *Oncogene* **26**, 7490–7498 (2007).

[563] Xu, B. *et al.* Hypermethylation of the 16q23.1 Tumor Suppressor Gene ADAMTS18 in Clear Cell Renal Cell Carcinoma. *International Journal of Molecular Sciences* **16**, 1051–1065 (2015).

[564] Li, Z. *et al.* High-resolution melting analysis of ADAMTS18 methylation levels in gastric, colorectal and pancreatic cancers. *Medical Oncology* **27**, 998–1004 (2010).

[565] Alkebsi, L. *et al.* Chromosome 16q genes CDH1, CDH13 and ADAMTS18 are correlated and frequently methylated in human lymphoma. *Oncology Letters* **12**, 3523–3530 (2016).

[566] Zhang, L., Liu, Y. & Zheng, P. Downregulation of ADAMTS18 May Serve as a Poor Prognostic Biomarker for Cervical Cancer Patients. *Applied Immunohistochemistry & Molecular Morphology* **26**, 670–675 (2018).

[567] Zhang, Y. *et al.* Inactivation of ADAMTS18 by aberrant promoter hypermethylation contribute to lung cancer progression. *Journal of Cellular Physiology* **234**, 6965–6975 (2019).

[568] Xu, H. *et al.* Epigenetic silencing of ADAMTS18 promotes cell migration and invasion of breast cancer through AKT and NF-κB signaling. *Cancer Medicine* **6**, 1399–1408 (2017).

[569] Hohenauer, T. & Moore, A. W. The Prdm family: expanding roles in stem cells and development. *Development* **139**, 2267–2282 (2012).

[570] Sorrentino, A., Rienzo, M., Ciccodicola, A., Casamassimi, A. & Abbondanza, C. Human PRDM2: Structure, function and pathophysiology. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1861**, 657–671 (2018).

[571] Mzoughi, S., Tan, Y. X., Low, D. & Guccione, E. The role of PRDMs in cancer: one family, two sides. *Current Opinion in Genetics & Development* **36**, 83–91 (2016).

[572] Derunes, C. *et al.* Characterization of the PR domain of RIZ1 histone methyltransferase. *Biochemical and Biophysical Research Communications* **333**, 925–934 (2005).

[573] Kim, K.-C., Geng, L. & Huang, S. Inactivation of a histone methyltransferase by mutations in human cancers. *Cancer Research* **63**, 7619–7623 (2003).

[574] Carling, T. *et al.* A Histone Methyltransferase Is Required for Maximal Response to Female Sex Hormones. *Molecular and Cellular Biology* **24**, 7032–7042 (2004).

[575] Brayer, K. J. & Segal, D. J. Keep your fingers off my DNA: protein-protein interactions mediated by C2h2 zinc finger domains. *Cell Biochemistry and Biophysics* **50**, 111–131 (2008).

[576] Huang, S., Shao, G. & Liu, L. The PR Domain of the Rb-binding Zinc Finger Protein RIZ1 Is a Protein Binding Interface and Is Related to the SET Domain Functioning in Chromatin-mediated Gene Expression. *Journal of Biological Chemistry* **273**, 15933–15939 (1998).

[577] Medici, N. *et al.* Identification of a DNA binding protein cooperating with estrogen receptor as RIZ (retinoblastoma interacting zinc finger protein). *Biochemical and Biophysical Research Communications* **264**, 983–989 (1999).

[578] Xie, M., Shao, G., Buyse, I. M. & Huang, S. Transcriptional Repression Mediated by the PR Domain Zinc Finger Gene RIZ. *Journal of Biological Chemistry* **272**, 26360–26366 (1997).

[579] Muraosa, Y., Takahashi, K., Yoshizawa, M. & Shibahara, S. cDNA Cloning of a Novel Protein Containing two zinc-finger Domains that may Function as a Transcription Factor for the Human Heme-oxygenase-1 Gene. *European Journal of Biochemistry* **235**, 471–479 (1996).

[580] Liu, L., Shao, G., Steele-Perkins, G. & Huang, S. The Retinoblastoma Interacting Zinc Finger Gene RIZ Produces a PR Domain-lacking Product through an Internal Promoter. *Journal of Biological Chemistry* **272**, 2984–2991 (1997).

[581] Huang, S. The retinoblastoma protein-interacting zinc finger gene RIZ in 1p36-linked cancers. *Frontiers in Bioscience* **4**, D528–532 (1999).

[582] Jiang, G. L. & Huang, S. The yin-yang of PR-domain family genes in tumorigenesis. *Histology and Histopathology* **15**, 109–117 (2000).

[583] Jiang, G.-l., Liu, L., Buyse, I. M., Simon, D. & Huang, S. Decreased RIZ1 expression but not RIZ2 in hepatoma and suppression of hepatoma tumorigenicity by RIZ1. *International Journal of Cancer* **83**, 541–546 (1999).

[584] Chadwick, R. B. *et al.* Candidate tumor suppressor RIZ is frequently involved in colorectal carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 2662–2667 (2000).

[585] He, L. *et al.* RIZ1, but not the Alternative RIZ2 Product of the Same Gene, Is Underexpressed in Breast Cancer, and Forced RIZ1 Expression Causes G2-M Cell Cycle Arrest and/or Apoptosis. *Cancer Research* **58**, 4238–4244 (1998).

[586] Du, Y. *et al.* Hypermethylation in Human Cancers of the RIZ1 Tumor Suppressor Gene, a Member of a Histone/Protein Methyltransferase Superfamily. *Cancer Research* **61**, 8094–8099 (2001).

[587] Lakshmikuttyamma, A. *et al.* RIZ1 is potential CML tumor suppressor that is down-regulated during disease progression. *Journal of Hematology & Oncology* **2**, 28 (2009).

[588] Cui, Y., Ding, M., Dong, S., Wang, Y. & Zhang, P. The unusual yin-yang fashion of RIZ1/RIZ2 contributes to the progression of esophageal squamous cell carcinoma. *Open Life Sciences* **11**, 136–141 (2016).

[589] Congdon, L. M., Sims, J. K., Tuzon, C. T. & Rice, J. C. The PR-Set7 binding domain of Riz1 is required for the H4k20me1-H3k9me1 trans-tail 'histone code' and Riz1 tumor suppressor function. *Nucleic Acids Research* **42**, 3580–3589 (2014).

[590] Wu, S. & Rice, J. C. A new regulator of the cell cycle. *Cell Cycle* **10**, 68–72 (2011).

[591] Linos, K. & Tafe, L. J. Isocitrate dehydrogenase 1 mutations in melanoma frequently co-occur with NRAS mutations. *Histopathology* **73**, 963–968 (2018).

[592] Garman, B. *et al.* Genetic and genomic characterization of 462 melanoma patient-derived xenografts, tumor biopsies and cell lines. *Cell Reports* **21**, 1936–1952 (2017).

[593] Remy, E. *et al.* A Modeling Approach to Explain Mutually Exclusive and Co-Occurring Genetic Alterations in Bladder Tumorigenesis. *Cancer Research* **75**, 4042–4052 (2015).

[594] Houillier, C. *et al.* Prognostic impact of molecular markers in a series of 220 primary glioblastomas. *Cancer* **106**, 2218–2223 (2006).

[595] Arafeh, R. *et al.* Recurrent inactivating *RASA2* mutations in melanoma. *Nature Genetics* **47**, 1408–1410 (2015).

[596] Kiuru, M. & Busam, K. J. The NF1 gene in tumor syndromes and melanoma. *Laboratory Investigation* **97**, 146–157 (2017).

[597] Adar, S., Hu, J., Lieb, J. D. & Sancar, A. Genome-wide kinetics of DNA excision repair in relation to chromatin state and mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America* **113**, E2124–E2133 (2016).

[598] Sabarinathan, R., Mularoni, L., Deu-Pons, J., Gonzalez-Perez, A. & López-Bigas, N. Nucleotide excision repair is impaired by binding of transcription factors to DNA. *Nature* **532**, 264–267 (2016).

[599] Perera, D. *et al.* Differential DNA repair underlies mutation hotspots at active promoters in cancer genomes. *Nature* **532**, 259–263 (2016).

[600] Denisova, E. *et al.* Frequent DPH3 promoter mutations in skin cancers. *Oncotarget* **6**, 35922–35930 (2015).

[601] Zhang, T. *et al.* SDHD Promoter Mutations Ablate GABP Transcription Factor Binding in Melanoma. *Cancer Research* **77**, 1649–1661 (2017).

[602] Poulos, R. C. *et al.* Systematic Screening of Promoter Regions Pinpoints Functional Cis-Regulatory Mutations in a Cutaneous Melanoma Genome. *Molecular Cancer Research* **13**, 1218–1226 (2015).

[603] Menzies, A. M. *et al.* Distinguishing Clinicopathologic Features of Patients with V600E and V600K BRAF-Mutant Metastatic Melanoma. *Clinical Cancer Research* **18**, 3242–3249 (2012).

[604] Mar, V. J. *et al.* BRAF/NRAS wild-type melanomas have a high mutation load correlating with histologic and molecular signatures of UV damage. *Clinical Cancer Research* **19**, 4589–4598 (2013).

[605] Silva, I. P. d. *et al.* Distinct Molecular Profiles and Immunotherapy Treatment Outcomes of V600E and V600K BRAF-Mutant Melanoma. *Clinical Cancer Research* **25**, 1272–1279 (2019).

[606] Curtin, J. A., Busam, K., Pinkel, D. & Bastian, B. C. Somatic activation of KIT in distinct subtypes of melanoma. *Journal of Clinical Oncology* **24**, 4340–4346 (2006).

[607] Beadling, C. *et al.* KIT Gene Mutations and Copy Number in Melanoma Subtypes. *Clinical Cancer Research* **14**, 6821–6828 (2008).

[608] Chalmers, Z. R. *et al.* Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Medicine* **9**, 34 (2017).

[609] Milholland, B., Auton, A., Suh, Y. & Vijg, J. Age-related somatic mutations in the cancer genome. *Oncotarget* **6**, 24627–24635 (2015).

[610] Wangari-Talbot, J. & Chen, S. Genetics of melanoma. *Frontiers in Genetics* **3** (2013).

[611] Schadendorf, D. *et al.* Melanoma. *The Lancet* **392**, 971–984 (2018).

[612] Kuk, D. *et al.* Prognosis of Mucosal, Uveal, Acral, Nonacral Cutaneous, and Unknown Primary Melanoma From the Time of First Metastasis. *The Oncologist* **21**, 848–854 (2016).

[613] DailyMed - ZELBORAF. URL: https://dailymed.nlm.nih.gov/dailymed/drugInfo.cfm?setid=38eea320-7e0c-485a-bc30-98c3c45e2763.

[614] Kuijjer, M. L., Paulson, J. N., Salzman, P., Ding, W. & Quackenbush, J. Cancer subtype identification using somatic mutation data. *British Journal of Cancer* **118**, 1492–1501 (2018).

[615] Rajkumar, S. & Watson, I. R. Molecular characterisation of cutaneous melanoma: creating a framework for targeted and immune therapies. *British Journal of Cancer* **115**, 145–155 (2016).

[616] Okkenhaug, K., Bilancio, A., Emery, J. L. & Vanhaesebroeck, B. Phosphoinositide 3-kinase in T cell activation and survival. *Biochemical Society Transactions* **32**, 332–335 (2004).

[617] Garcon, F. *et al.* CD28 provides T-cell costimulation and enhances PI3k activity at the immune synapse independently of its capacity to interact with the p85/p110 heterodimer. *Blood* **111**, 1464–1471 (2008).

[618] Spiegel, S. & Milstien, S. The outs and the ins of sphingosine-1-phosphate in immunity. *Nature Reviews Immunology* **11**, 403–415 (2011).

[619] Pyne, N. J. & Pyne, S. Sphingosine 1-phosphate and cancer. *Nature Reviews Cancer* **10**, 489–503 (2010).

[620] Messias, C. V. *et al.* Sphingosine-1-Phosphate Induces Dose-Dependent Chemotaxis or Fugetaxis of T-ALL Blasts through S1p1 Activation. *PLOS ONE* **11** (2016).

[621] Hamidi, H. & Ivaska, J. Every step of the way: integrins in cancer progression and metastasis. *Nature Reviews Cancer* **18**, 533–548 (2018).

[622] Keely, P. J., Westwick, J. K., Whitehead, I. P., Der, C. J. & Parise, L. V. Cdc42 and Rac1 induce integrin-mediated cell motility and invasiveness through PI(3)K. *Nature* **390**, 632–636 (1997).

[623] Katoh, H., Hiramoto, K. & Negishi, M. Activation of Rac1 by RhoG regulates cell migration. *Journal of Cell Science* **119**, 56–65 (2006).

[624] Carlino, M. S. *et al.* Correlation of BRAF and NRAS mutation status with outcome, site of distant metastasis and response to chemotherapy in metastatic melanoma. *British Journal of Cancer* **111**, 292–299 (2014).

[625] Heppt, M. V. *et al.* Prognostic significance of BRAF and NRAS mutations in melanoma: a German study from routine care. *BMC Cancer* **17**, 536 (2017).

[626] Khan, J. *et al.* Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nature Medicine* **7**, 673–679 (2001).

[627] Reis-Filho, J. S. & Pusztai, L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. *The Lancet* **378**, 1812–1823 (2011).

[628] Golub, T. R. *et al.* Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. *Science* **286**, 531–537 (1999).

[629] Bucheit, A. D. *et al.* Complete loss of PTEN protein expression correlates with shorter time to brain metastasis and survival in stage IIIB/C melanoma patients with BRAFV600 mutations. *Clinical Cancer Research* **20**, 5527–5536 (2014).

[630] Jao, K. *et al.* Prognostic effect of single versus multiple somatic mutations in non-small cell lung cancer (NSCLC). *Journal of Clinical Oncology* **33**, 7521–7521 (2015).

[631] Papaemmanuil, E. *et al.* Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**, 3616–3627 (2013).

[632] Furney, S. J. *et al.* Genome sequencing of mucosal melanomas reveals that they are driven by distinct mechanisms from cutaneous melanoma. *The Journal of Pathology* **230**, 261–269 (2013).

[633] Cosgarea, I. *et al.* Targeted next generation sequencing of mucosal melanomas identifies frequent NF1 and RAS mutations. *Oncotarget* **8**, 40683–40692 (2017).

[634] Sheng, X. *et al.* GNAQ and GNA11 mutations occur in 9.5% of mucosal melanoma and are associated with poor prognosis. *European Journal of Cancer* **65**, 156–163 (2016).

[635] Hintzsche, J. *et al.* Whole-exome sequencing identifies recurrent SF3b1 R625 mutation and comutation of NF1 and KIT in mucosal melanoma. *Melanoma Research* **27**, 189–199 (2017).

[636] Wong, K. *et al.* Cross-species genomic landscape comparison of human mucosal melanoma with canine oral and equine melanoma. *Nature Communications* **10**, 353 (2019).

[637] Moon, K. R. *et al.* Genetic Alterations in Primary Acral Melanoma and Acral Melanocytic Nevus in Korea: Common Mutated Genes Show Distinct Cytomorphological Features. *The Journal of Investigative Dermatology* **138**, 933–945 (2018).

[638] Zebary, A. *et al.* KIT, NRAS, BRAF and PTEN mutations in a sample of Swedish patients with acral lentiginous melanoma. *Journal of Dermatological Science* **72**, 284–289 (2013).

[639] Mazurenko, N. N. *et al.* The spectrum of oncogene mutations differs among melanoma subtypes. *Molecular Biology* **49**, 917–923 (2015).

[640] Kong, Y. *et al.* Frequent Genetic Aberrations in the CDK4 Pathway in Acral Melanoma Indicate the Potential for CDK4/6 Inhibitors in Targeted Therapy. *Clinical Cancer Research* **23**, 6946–6957 (2017).

[641] Santarpia, L., Lippman, S. L. & El-Naggar, A. K. Targeting the Mitogen-Activated Protein Kinase RAS-RAF Signaling Pathway in Cancer Therapy. *Expert Opinion on Therapeutic Targets* **16**, 103–119 (2012).

[642] Marais, R., Wynne, J. & Treisman, R. The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell* **73**, 381–393 (1993).

[643] Zhao, J., Yuan, X., FrÃ¶din, M. & Grummt, I. ERK-Dependent Phosphorylation of the Transcription Initiation Factor TIF-IA Is Required for RNA Polymerase I Transcription and Cell Growth. *Molecular Cell* **11**, 405–413 (2003).

[644] Bonni, A. *et al.* Cell Survival Promoted by the Ras-MAPK Signaling Pathway by Transcription-Dependent and -Independent Mechanisms. *Science* **286**, 1358–1362 (1999).

[645] Hemesath, T. J., Price, E. R., Takemoto, C., Badalian, T. & Fisher, D. E. MAP kinase links the transcription factor Microphthalmia to c-Kit signalling in melanocytes. *Nature* **391**, 298–301 (1998).

[646] Garraway, L. A. *et al.* Integrative genomic analyses identify *MITF* as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117–122 (2005).

[647] Haigis, K. M. KRAS Alleles: The Devil Is in the Detail. *Trends in Cancer* **3**, 686–697 (2017).

[648] Sanchez-Vega, F. *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321–337.e10 (2018).

[649] Whitwam, T. *et al.* Differential oncogenic potential of activated RAS isoforms in melanocytes. *Oncogene* **26**, 4563–4570 (2007).

[650] Nikolaev, S. I. *et al.* Exome sequencing identifies recurrent somatic *MAP2K1* and *MAP2K2* mutations in melanoma. *Nature Genetics* **44**, 133–139 (2012).

[651] Gao, Y. *et al.* Allele-Specific Mechanisms of Activation of MEK1 Mutants Determine Their Properties. *Cancer Discovery* **8**, 648–661 (2018).

[652] Rodriguez-Viciana, P. *et al.* Germline Mutations in Genes Within the MAPK Pathway Cause Cardio-facio-cutaneous Syndrome. *Science* **311**, 1287–1290 (2006).

[653] Chakravarty, D. *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precision Oncology* **1**, 1–16 (2017).

[654] Cirenajwis, H. *et al.* NF1-mutated melanoma tumors harbor distinct clinical and biological characteristics. *Molecular Oncology* **11**, 438–451 (2017).

[655] Antonescu, C. R. *et al.* L576p KIT mutation in anal melanomas correlates with KIT protein expression and is sensitive to specific kinase inhibition. *International Journal of Cancer* **121**, 257–264 (2007).

[656] Conca, E. *et al.* Activate and resist: L576p-KIT in GIST. *Molecular Cancer Therapeutics* **8**, 2491–2495 (2009).

[657] Altomare, D. A. & Testa, J. R. Perturbations of the AKT signaling pathway in human cancer. *Oncogene* **24**, 7455 (2005).

[658] Manning, B. D. & Toker, A. AKT/PKB Signaling: Navigating the Network. *Cell* **169**, 381–405 (2017).

[659] Aguissa-Toure, A.-H. & Li, G. Genetic alterations of PTEN in human melanoma. *Cellular and Molecular Life Sciences* **69**, 1475–1491 (2012).

[660] Omholt, K., Krockel, D. & Hansson, J. PIK3ca mutations in cutaneous melanoma. *Cancer Research* **66**, 273–273 (2006).

[661] Curtin, J. A., Stark, M. S., Pinkel, D., Hayward, N. K. & Bastian, B. C. PI3-Kinase Subunits Are Infrequent Somatic Targets in Melanoma. *The Journal of Investigative Dermatology* **126**, 1660–1663 (2006).

[662] Davies, M. A. *et al.* A novel *AKT3* mutation in melanoma tumours and cell lines. *British Journal of Cancer* **99**, 1265–1268 (2008).

[663] Chang, M. T. *et al.* Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nature Biotechnology* **34**, 155–163 (2016).

[664] Chang, M. T. *et al.* Accelerating Discovery of Functional Mutant Alleles in Cancer. *Cancer Discovery* **8**, 174–183 (2018).

[665] Huang, C.-H. *et al.* The Structure of a Human p110a/p85a Complex Elucidates the Effects of Oncogenic PI3ka Mutations. *Science* **318**, 1744–1748 (2007).

[666] Tsao, H., Goel, V., Wu, H., Yang, G. & Haluska, F. G. Genetic Interaction Between NRAS and BRAF Mutations and PTEN/MMAC1 Inactivation in Melanoma. *The Journal of Investigative Dermatology* **122**, 337–341 (2004).

[667] Goel, V. K., Lazar, A. J. F., Warneke, C. L., Redston, M. S. & Haluska, F. G. Examination of mutations in BRAF, NRAS, and PTEN in primary cutaneous melanoma. *The Journal of Investigative Dermatology* **126**, 154–160 (2006).

[668] Dankort, D. *et al.* BrafV600E cooperates with Pten loss to induce metastatic melanoma. *Nature Genetics* **41**, 544–552 (2009).

[669] Castel, P. *et al.* PDK1-SGK1 Signaling Sustains AKT-Independent mTORC1 Activation and Confers Resistance to PI3ka Inhibition. *Cancer Cell* **30**, 229–242 (2016).

[670] Rodrik-Outmezguine, V. S. *et al.* Overcoming mTOR resistance mutations with a new-generation mTOR inhibitor. *Nature* **534**, 272–276 (2016).

[671] Courtney, K. D., Corcoran, R. B. & Engelman, J. A. The PI3k pathway as drug target in human cancer. *Journal of Clinical Oncology* **28**, 1075–1083 (2010).

[672] Flores, J. F. *et al.* Loss of the p16ink4a and p15ink4b Genes, as well as Neighboring 9p21 Markers, in Sporadic Melanoma. *Cancer Research* **56**, 5023–5032 (1996).

[673] Coleman, K. G. *et al.* Identification of CDK4 Sequences Involved in Cyclin D1 and p16 Binding. *Journal of Biological Chemistry* **272**, 18869–18874 (1997).

[674] Shennan, M. G. *et al.* Lack of germline ***cdk6*** mutations in familial melanoma. *Oncogene* **19**, 1849–1852 (2000).

[675] Hocker, T. & Tsao, H. Ultraviolet radiation and melanoma: a systematic review and analysis of reported sequence variants. *Human Mutation* **28**, 578–588 (2007).

[676] Gembarska, A. *et al.* MDM4 is a key therapeutic target in cutaneous melanoma. *Nature Medicine* **18**, 1239–1247 (2012).

[677] Dang, C. V. MYC on the Path to Cancer. *Cell* **149**, 22 (2012).

[678] Sears, R. C. The life cycle of C-myc: from synthesis to degradation. *Cell Cycle* **3**, 1133–1137 (2004).

[679] The Cancer Genome Atlas Network. The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905 (2010).

[680] Parrella, P., Caballero, O. L., Sidransky, D. & Merbs, S. L. Detection of c-myc amplification in uveal melanoma by fluorescent in situ hybridization. *Investigative Ophthalmology & Visual Science* **42**, 1679–1684 (2001).

[681] Zhuang, D. *et al.* C-MYC overexpression is required for continuous suppression of oncogene-induced senescence in melanoma cells. *Oncogene* **27** (2008).

[682] Glatz-Krieger, K. *et al.* Anatomic site-specific patterns of gene copy number gains in skin, mucosal, and uveal melanomas detected by fluorescence in situ hybridization. *Virchows Archiv* **449**, 328–333 (2006).

[683] Hartman, M. L. & Czyz, M. MITF in melanoma: mechanisms behind its expression and activity. *Cellular and Molecular Life Sciences* **72**, 1249–1260 (2015).

[684] Cronin, J. C. *et al.* Frequent Mutations in the MITF Pathway in Melanoma. *Pigment Cell & Melanoma Research* **22**, 435–444 (2009).

[685] JÃ¶nsson, G. *et al.* Gene Expression Profiling-Based Identification of Molecular Subtypes in Stage IV Melanomas with Different Clinical Outcome. *Clinical Cancer Research* **16**, 3356–3367 (2010).

[686] Harbst, K. *et al.* Molecular Profiling Reveals Low- and High-Grade Forms of Primary Melanoma. *Clinical Cancer Research* **18**, 4026–4036 (2012).

[687] Ugurel, S. *et al.* Microphthalmia-Associated Transcription Factor Gene Amplification in Metastatic Melanoma Is a Prognostic Marker for Patient Survival, But Not a Predictive Marker for Chemosensitivity and Chemotherapy Response. *Clinical Cancer Research* **13**, 6344–6350 (2007).

[688] Guen, V. J. *et al.* CDK10/cyclin M is a protein kinase that controls ETS2 degradation and is deficient in STAR syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 19525–19530 (2013).

[689] Mao, J.-H. *et al.* Fbxw7/Cdc4 is a p53-dependent, haploinsufficient tumour suppressor gene. *Nature* **432**, 775–779 (2004).

[690] Abbate, F. *et al.* FBXW7 regulates a mitochondrial transcription program by modulating MITF. *Pigment Cell & Melanoma Research* **31**, 636–640 (2018).

[691] Cheng, Y. & Li, G. Role of the ubiquitin ligase Fbw7 in cancer progression. *Cancer and Metastasis Reviews* **31**, 75–87 (2012).

[692] Yada, M. *et al.* Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7. *The EMBO Journal* **23**, 2116–2125 (2004).

[693] Mao, J.-H. *et al.* FBXW7 targets mTOR for degradation and cooperates with PTEN in tumor suppression. *Science* **321**, 1499–1502 (2008).

[694] Kim, W. Y. & Kaelin, W. G. Role of VHL gene mutation in human cancer. *Journal of Clinical Oncology* **22**, 4991–5004 (2004).

[695] Khanna, K. K. & Jackson, S. P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genetics* **27**, 247–254 (2001).

[696] Bartek, J. & Lukas, J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell* **3**, 421–429 (2003).

[697] Lauss, M. *et al.* Mutational and putative neoantigen load predict clinical benefit of adoptive T cell therapy in melanoma. *Nature Communications* **8**, 1738 (2017).

[698] Hugo, W. *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **165**, 35–44 (2016).

[699] Allen, E. M. V. *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).

[700] Hoof, I. *et al.* NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* **61**, 1–13 (2009).

[701] Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**, eaaf8399 (2017).

[702] Knijnenburg, T. A. *et al.* Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Reports* **23**, 239–254.e6 (2018).

[703] Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).

[704] Sato, E. *et al.* Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian

cancer. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18538–18543 (2005).

[705] Schumacher, K., Haensch, W., Röefzaad, C. & Schlag, P. M. Prognostic significance of activated CD8(+) T cell infiltrations within esophageal carcinomas. *Cancer Research* **61**, 3932–3936 (2001).

[706] Garrido, F. *et al.* Implications for immunosurveillance of altered HLA class I phenotypes in human tumours. *Immunology Today* **18**, 89–95 (1997).

[707] Sade-Feldman, M. *et al.* Resistance to checkpoint blockade therapy through inactivation of antigen presentation. *Nature Communications* **8**, 1136 (2017).

[708] Ikeda, H., Old, L. J. & Schreiber, R. D. The roles of IFNγ in protection against tumor development and cancer immunoediting. *Cytokine & Growth Factor Reviews* **13**, 95–109 (2002).

[709] Detjen, K. M., Farwig, K., Welzel, M., Wiedenmann, B. & Rosewicz, S. Interferon γ inhibits growth of human pancreatic carcinoma cells via caspase-1 dependent induction of apoptosis. *Gut* **49**, 251–262 (2001).

[710] Del Castillo Velasco-Herrera, M. *et al.* Comparative genomics reveals that loss of lunatic fringe (*LFNG*) promotes melanoma metastasis. *Molecular Oncology* **12**, 239–255 (2018).

[711] Erard, N., Knott, S. R. & Hannon, G. J. A CRISPR Resource for Individual, Combinatorial, or Multiplexed Gene Knockout. *Molecular Cell* **67**, 348–354.e3 (2017).

[712] Wang, D. *et al.* Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning. *Nature Communications* **10** (2019).

[713] Dutton-Regester, K. *et al.* Melanomas of unknown primary have a mutation profile consistent with cutaneous sun-exposed melanoma. *Pigment Cell & Melanoma Research* **26**, 852–860 (2013).

[714] Hart, T., Brown, K. R., Sircoulomb, F., Rottapel, R. & Moffat, J. Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. *Molecular Systems Biology* **10**, 733 (2014).

[715] Shi, J. *et al.* Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nature Biotechnology* **33**, 661–667 (2015).

[716] Ong, S. H., Li, Y., Koike-Yusa, H. & Yusa, K. Optimised metrics for CRISPR-KO screens with second-generation gRNA libraries. *Scientific Reports* **7**, 7384 (2017).

[717] Sharma, S., Bartholdson, S. J., Couch, A. C., Yusa, K. & Wright, G. J. Genome-scale identification of cellular pathways required for cell surface recognition. *Genome Research* gr.231183.117 (2018).

[718] Pettitt, S. J. *et al.* Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance. *Nature Communications* **9**, 1849 (2018).

[719] Behan, F. M. *et al.* Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. *Nature* **568**, 511–516 (2019).

[720] Fields, B. N., Knipe, D. M. & Howley, P. M. *Fields Virology* (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2007). OCLC: 71812790.

[721] Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature Protocols* **12**, 828–863 (2017).

[722] Ruas, M. & Peters, G. The p16ink4a/CDKN2a tumor suppressor and its relatives. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1378**, F115–F177 (1998).

[723] Joo, Y. J. *et al.* Downstream promoter interactions of TFIID TAFs facilitate transcription reinitiation. *Genes & Development* **31**, 2162–2174 (2017).

[724] Müller, F. & Tora, L. The multicoloured world of promoter recognition complexes. *The EMBO Journal* **23**, 2–8 (2004).

[725] Albright, S. R. & Tjian, R. TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* **242**, 1–13 (2000).

[726] Lee, T. I. & Young, R. A. Transcription of Eukaryotic Protein-Coding Genes. *Annual Review of Genetics* **34**, 77–137 (2000).

[727] Cler, E., Papai, G., Schultz, P. & Davidson, I. Recent advances in understanding the structure and function of general transcription factor TFIID. *Cellular and Molecular Life Sciences* **66**, 2123–2134 (2009).

[728] Patel, A. B. *et al.* Structure of human TFIID and mechanism of TBP loading onto promoter DNA. *Science* **362**, eaau8872 (2018).

[729] Müller, F., Zaucker, A. & Tora, L. Developmental regulation of transcription initiation: more than just changing the actors. *Current Opinion in Genetics & Development* **20**, 533–540 (2010).

[730] Ribeiro, J. R., Lovasco, L. A., Vanderhyden, B. C. & Freiman, R. N. Targeting TBP-Associated Factors in Ovarian Cancer. *Frontiers in Oncology* **4**, 45 (2014).

[731] Hilton, T. L., Li, Y., Dunphy, E. L. & Wang, E. H. TAF1 histone acetyltransferase activity in Sp1 activation of the cyclin D1 promoter. *Molecular and Cellular Biology* **25**, 4321–4332 (2005).

[732] Kimura, J. *et al.* A functional genome-wide RNAi screen identifies TAF1 as a regulator for apoptosis in response to genotoxic stress. *Nucleic Acids Research* **36**, 5250–5259 (2008).

[733] Deato, M. D. E. & Tjian, R. An Unexpected Role of TAFs and TRFs in Skeletal Muscle Differentiation: Switching Core Promoter Complexes. *Cold Spring Harbor Symposia on Quantitative Biology* **73**, 217–225 (2008).

[734] Martinez, E. *et al.* Human STAGA Complex Is a Chromatin-Acetylating Transcription Coactivator That Interacts with Pre-mRNA Splicing and DNA Damage-Binding Factors In Vivo. *Molecular and Cellular Biology* **21**, 6782–6795 (2001).

[735] Ogryzko, V. V. *et al.* Histone-like TAFs within the PCAF Histone Acetylase Complex. *Cell* **94**, 35–44 (1998).

[736] Kolesnikova, O. *et al.* Molecular structure of promoter-bound yeast TFIID. *Nature Communications* **9**, 4666 (2018).

[737] Trowitzsch, S. *et al.* Cytoplasmic TAF2–TAF8–TAF10 complex provides evidence for nuclear holo–TFIID assembly from preformed submodules. *Nature Communications* **6**, 6011 (2015).

[738] Louder, R. K. *et al.* Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature* **531**, 604–609 (2016).

[739] Wang, H., Curran, E. C., Hinds, T. R., Wang, E. H. & Zheng, N. Crystal structure of a TAF1-TAF7 complex in human transcription factor IID reveals a promoter binding module. *Cell Research* **24**, 1433–1444 (2014).

[740] Curran, E. C., Wang, H., Hinds, T. R., Zheng, N. & Wang, E. H. Zinc knuckle of TAF1 is a DNA binding module critical for TFIID promoter occupancy. *Scientific Reports* **8**, 4630 (2018).

[741] Dikstein, R., Ruppert, S. & Tjian, R. TAFII250 Is a Bipartite Protein Kinase That Phosphorylates the Basal Transcription Factor RAP74. *Cell* **84**, 781–790 (1996).

[742] O'Brien, T. & Tjian, R. Functional Analysis of the Human TAFII250 N-Terminal Kinase Domain. *Molecular Cell* **1**, 905–911 (1998).

[743] Mizzen, C. A. *et al.* The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* **87**, 1261–1270 (1996).

[744] Pham, A.-D. & Sauer, F. Ubiquitin-Activating/Conjugating Activity of TAFII250, a Mediator of Activation of Gene Expression in Drosophila. *Science* **289**, 2357–2360 (2000).

[745] Boutet, S. C., Biressi, S., Iori, K., Natu, V. & Rando, T. A. Taf1 regulates Pax3 protein by monoubiquitination in skeletal muscle progenitors. *Molecular Cell* **40**, 749–761 (2010).

[746] Tavassoli, P. *et al.* TAF1 Differentially Enhances Androgen Receptor Transcriptional Activity via Its N-Terminal Kinase and Ubiquitin-Activating and -Conjugating Domains. *Molecular Endocrinology* **24**, 696–708 (2010).

[747] Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and Function of a Human TAFII250 Double Bromodomain Module. *Science* **288**, 1422–1425 (2000).

[748] Dunphy, E. L., Johnson, T., Auerbach, S. S. & Wang, E. H. Requirement for TAFII250 Acetyltransferase Activity in Cell Cycle Progression. *Molecular and Cellular Biology* **20**, 1134–1139 (2000).

[749] Sekiguchi, T. *et al.* Apoptosis Is Induced in BHK Cells by the tsBN462/13 Mutation in the CCG1/TAFII250 Subunit of the TFIID Basal Transcription Factor. *Experimental Cell Research* **218**, 490–498 (1995).

[750] Li, H.-H., Li, A. G., Sheppard, H. M. & Liu, X. Phosphorylation on Thr-55 by TAF1 Mediates Degradation of p53: A Role for TAF1 in Cell G1 Progression. *Molecular Cell* **13**, 867–878 (2004).

[751] Buchmann, A. M., Skaar, J. R. & DeCaprio, J. A. Activation of a DNA Damage Checkpoint Response in a TAF1-Defective Cell Line. *Molecular and Cellular Biology* **24**, 5332–5339 (2004).

[752] Talavera, A. & Basilico, C. Temperature sensitive mutants of BHK cells affected in cell cycle progression. *Journal of Cellular Physiology* **92**, 425–436 (1977).

[753] Verrijzer, C. P., Yokomori, K., Chen, J. L. & Tjian, R. Drosophila TAFII150: similarity to yeast gene TSM-1 and specific binding to core promoter DNA. *Science* **264**, 933–941 (1994).

[754] Chalkley, G. E. & Verrijzer, C. P. DNA binding site selection by RNA polymerase II TAFs: a TAFII250–TAFII150 complex recognizes the Initiator. *The EMBO Journal* **18**, 4835–4845 (1999).

[755] Walker, S. S., Shen, W.-C., Reese, J. C., Apone, L. M. & Green, M. R. Yeast TAFII145 Required for Transcription of G1/S Cyclin Genes and Regulated by the Cellular Growth State. *Cell* **90**, 607–614 (1997).

[756] Gegonne, A., Weissman, J. D., Zhou, M., Brady, J. N. & Singer, D. S. TAF7: A possible transcription initiation check-point regulator. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 602–607 (2006).

[757] Devaiah, B. N. *et al.* Novel Functions for TAF7, a Regulator of TAF1-independent Transcription. *Journal of Biological Chemistry* **285**, 38772–38780 (2010).

[758] Munz, C. *et al.* TAF7 (TAFII55) Plays a Role in the Transcription Activation by c-Jun. *Journal of Biological Chemistry* **278**, 21510–21516 (2003).

[759] Fukuchi, J. *et al.* TATA-binding Protein-associated Factor 7 Regulates Polyamine Transport Activity and Polyamine Analog-induced Apoptosis. *Journal of Biological Chemistry* **279**, 29921–29929 (2004).

[760] Thut, C. J., Chen, J. L., Klemm, R. & Tjian, R. p53 transcriptional activation mediated by coactivators TAFII40 and TAFII60. *Science* **267**, 100–104 (1995).

[761] Bell, B., Scheer, E. & Tora, L. Identification of hTAFII80δ Links Apoptotic Signaling Pathways to Transcription Factor TFIID Function. *Molecular Cell* **8**, 591–600 (2001).

[762] Wang, W., Nahta, R., Huper, G. & Marks, J. R. TAFII70 Isoform-Specific Growth Suppression Correlates With Its Ability to Complex With the GADD45a Protein11nih grants CA84955 and CA73802 (J.R. Marks).Note: W. Wang and R. Nahta contributed equally to this work. R. Nahta is currently at Breast Medical Oncology, University of Texas M.D. Anderson Cancer Center, Houston, TX. *Molecular Cancer Research* **2**, 442–452 (2004).

[763] Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annual Review of Biochemistry* **73**, 1019–1049 (2004).

[764] Skropeta, D. The effect of individual N-glycans on enzyme activity. *Bioorganic & Medicinal Chemistry* **17**, 2645–2653 (2009).

[765] Stanley, P., Taniguchi, N. & Aebi, M. N-Glycans. In Varki, A. *et al.* (eds.) *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 2015), 3rd edn.

[766] Ning, B. & Elbein, A. D. Cloning, expression and characterization of the pig liver GDP-mannose pyrophosphorylase. *European Journal of Biochemistry* **267**, 6866–6874 (2000).

[767] Kelleher, D. J. & Gilmore, R. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**, 47R–62R (2006).

[768] Graham, D. B. *et al.* TMEM258 Is a Component of the Oligosaccharyltransferase Complex Controlling ER Stress and Intestinal Inflammation. *Cell Reports* **17**, 2955–2965 (2016).

[769] Sitia, R. & Braakman, I. Quality control in the endoplasmic reticulum protein factory. *Nature* **426**, 891 (2003).

[770] Meusser, B., Hirsch, C., Jarosch, E. & Sommer, T. ERAD: the long road to destruction. *Nature Cell Biology* **7**, 766 (2005).

[771] Varki, A., Kannagi, R., Toole, B. & Stanley, P. Glycosylation Changes in Cancer. In Varki, A. *et al.* (eds.) *Essentials of Glycobiology* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 2015), 3rd edn.

[772] Pinho, S. S. & Reis, C. A. Glycosylation in cancer: mechanisms and clinical implications. *Nature Reviews Cancer* **15**, 540–555 (2015).

[773] Lau, K. S. & Dennis, J. W. N-Glycans in cancer progression. *Glycobiology* **18**, 750–760 (2008).

[774] Nath, S. & Mukherjee, P. Muc1: a multifaceted oncoprotein with a key role in cancer progression. *Trends in Molecular Medicine* **20**, 332–342 (2014).

[775] Milde-Langosch, K. *et al.* Prognostic relevance of glycosylation-associated genes in breast cancer. *Breast Cancer Research and Treatment* **145**, 295–305 (2014).

[776] Ono, M. *et al.* The expression and clinical significance of ribophorin II (RPN2) in human breast cancer. *Pathology International* **65**, 301–308 (2015).

[777] Zhang, J. *et al.* Integrated transcriptional profiling and genomic analyses reveal RPN2 and HMGB1 as promising biomarkers in colorectal cancer. *Cell & Bioscience* **5**, 53 (2015).

[778] Sengupta, P. K., Bouchie, M. P., Nita-Lazar, M., Yang, H.-Y. & Kukuruzinska, M. A. Coordinate regulation of N-glycosylation gene DPAGT1, canonical Wnt signaling and E-cadherin adhesion. *Journal of Cell Science* **126**, 484–496 (2013).

[779] Powell, L. D., Sgroi, D., Sjoberg, E. R., Stamenkovic, I. & Varki, A. Natural ligands of the B cell adhesion molecule CD22 beta carry N-linked oligosaccharides with alpha-2,6-linked sialic acids that are required for recognition. *Journal of Biological Chemistry* **268**, 7019–7027 (1993).

[780] Wolfert, M. A. & Boons, G.-J. Adaptive immune activation: glycosylation does matter. *Nature Chemical Biology* **9**, 776–784 (2013).

[781] Honma, K. *et al. RPN2* gene confers docetaxel resistance in breast cancer. *Nature Medicine* **14**, 939–948 (2008).

[782] Fujiwara, T. *et al.* RPN2 Gene Confers Osteosarcoma Cell Malignant Phenotypes and Determines Clinical Prognosis. *Molecular Therapy - Nucleic Acids* **3**, e189 (2014).

[783] Li, C.-W. *et al.* Eradication of Triple-Negative Breast Cancer Cells by Targeting Glycosylated PD-L1. *Cancer Cell* **33**, 187–201.e10 (2018).

[784] Barrowman, J., Bhandari, D., Reinisch, K. & Ferro-Novick, S. TRAPP complexes in membrane traffic: convergence through a common Rab. *Nature Reviews Molecular Cell Biology* **11**, 759–763 (2010).

[785] Kim, J. J., Lipatova, Z. & Segev, N. TRAPP Complexes in Secretion and Autophagy. *Frontiers in Cell and Developmental Biology* **4** (2016).

[786] Kong, X. *et al.* Synbindin in Extracellular Signal-Regulated Protein Kinase Spatial Regulation and Gastric Cancer Aggressiveness. *Journal of the National Cancer Institute* **105**, 1738–1749 (2013).

[787] Weng, Y.-R. *et al.* The role of ERK2 in colorectal carcinogenesis is partly regulated by TRAPPC4. *Molecular Carcinogenesis* **53**, E72–E84 (2014).

[788] Zhao, S.-L. *et al.* TRAPPC4-ERK2 Interaction Activates ERK1/2, Modulates Its Nuclear Localization and Regulates Proliferation and Apoptosis of Colorectal Cancer Cells. *PLOS ONE* **6**, e23262 (2011).

[789] Witkos, T. M. & Lowe, M. Recognition and tethering of transport vesicles at the Golgi apparatus. *Current Opinion in Cell Biology* **47**, 16–23 (2017).

[790] Oka, T., Ungar, D., Hughson, F. M. & Krieger, M. The COG and COPI Complexes Interact to Control the Abundance of GEARs, a Subset of Golgi Integral Membrane Proteins. *Molecular Biology of the Cell* **15**, 2423–2435 (2004).

[791] Ungar, D., Oka, T., Krieger, M. & Hughson, F. M. Retrograde transport on the COG railway. *Trends in Cell Biology* **16**, 113–120 (2006).

[792] Spang, A. Membrane Tethering Complexes in the Endosomal System. *Frontiers in Cell and Developmental Biology* **4** (2016).

[793] van der Beek, J., Jonker, C., van der Welle, R., Liv, N. & Klumperman, J. CORVET, CHEVI and HOPS – multisubunit tethers of the endo-lysosomal system in health and disease. *Journal of Cell Science* **132**, jcs189134 (2019).

[794] Lindmo, K. & Stenmark, H. Regulation of membrane traffic by phosphoinositide 3-kinases. *Journal of Cell Science* **119**, 605–614 (2006).

[795] Bonifacino, J. S. & Hierro, A. Transport according to GARP: receiving retrograde cargo at the trans-Golgi network. *Trends in Cell Biology* **21**, 159–167 (2011).

[796] Seaman, M. N. J. Cargo-selective endosomal sorting for retrieval to the Golgi requires retromer. *The Journal of Cell Biology* **165**, 111–122 (2004).

[797] Kweon, Y., Rothe, A., Conibear, E. & Stevens, T. H. Ykt6p Is a Multifunctional Yeast R-SNARE That Is Required for Multiple Membrane Transport Pathways to the Vacuole. *Molecular Biology of the Cell* **14**, 1868–1881 (2003).

[798] Alonso-Curbelo, D. *et al.* RAB7 controls melanoma progression by exploiting a lineage-specific wiring of the endolysosomal pathway. *Cancer Cell* **26**, 61–76 (2014).

[799] Alonso-Curbelo, D. *et al.* RAB7 counteracts PI3k-driven macropinocytosis activated at early stages of melanoma development. *Oncotarget* **6**, 11848–11862 (2015).

[800] Mellman, I. & Yarden, Y. Endocytosis and Cancer. *Cold Spring Harbor Perspectives in Biology* **5** (2013).

[801] Li, S. *et al.* Transcriptional regulation of autophagy-lysosomal function in BRAF-driven melanoma progression and chemoresistance. *Nature Communications* **10**, 1–18 (2019).

[802] Munson, M. Tip20p reaches out to Dsl1p to tether membranes. *Nature Structural & Molecular Biology* **16**, 100–102 (2009).

[803] Guo, Z. *et al.* Subunit Organisation of In Vitro Reconstituted HOPS and CORVET Multisubunit Membrane Tethering Complexes. *PLOS ONE* **8**, e81534 (2013).

[804] Falkenberg, K. J. & Johnstone, R. W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nature Reviews Drug Discovery* **13**, 673–691 (2014).

[805] Deng, S. *et al.* HDAC3 Inhibition Upregulates PD-L1 Expression in B-Cell Lymphomas and Augments the Efficacy of Anti–PD-L1 Therapy. *Molecular Cancer Therapeutics* **18**, 900–908 (2019).

[806] Hu, G. *et al.* HDAC3 modulates cancer immunity via increasing PD-L1 expression in pancreatic cancer. *Pancreatology* **19**, 383–389 (2019).

[807] Zhu, H. *et al.* BET Bromodomain Inhibition Promotes Anti-tumor Immunity by Suppressing PD-L1 Expression. *Cell Reports* **16**, 2829–2837 (2016).

[808] West, A. C. & Johnstone, R. W. New and emerging HDAC inhibitors for cancer treatment. *The Journal of Clinical Investigation* **124**, 30–39 (2014).

[809] Insinga, A. *et al.* Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. *Nature Medicine* **11**, 71–76 (2005).

[810] Wilting, R. H. *et al.* Overlapping functions of Hdac1 and Hdac2 in cell cycle regulation and haematopoiesis. *The EMBO Journal* **29**, 2586–2597 (2010).

[811] Chen, X. *et al.* Requirement for the histone deacetylase Hdac3 for the inflammatory gene expression program in macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E2865–2874 (2012).

[812] Shakespear, M. R., Halili, M. A., Irvine, K. M., Fairlie, D. P. & Sweet, M. J. Histone deacetylases as regulators of inflammation and immunity. *Trends in Immunology* **32**, 335–343 (2011).

[813] Banik, D., Moufarrij, S. & Villagra, A. Immunoepigenetics Combination Therapies: An Overview of the Role of HDACs in Cancer Immunotherapy. *International Journal of Molecular Sciences* **20** (2019).

[814] Zhao, Y., Lin, H., Jiang, J., Ge, M. & Liang, X. TBL1xr1 as a potential therapeutic target that promotes epithelial-mesenchymal transition in lung squamous cell carcinoma. *Experimental and Therapeutic Medicine* **17**, 91–98 (2019).

[815] Liu, H. *et al.* Correlations between TBL1xr1 and recurrence of colorectal cancer. *Scientific Reports* **7** (2017).

[816] Li, J. Y., Daniels, G., Wang, J. & Zhang, X. TBL1xr1 in physiological and pathological states. *American Journal of Clinical and Experimental Urology* **3**, 13–23 (2015).

[817] Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology* **10**, 445–457 (2009).

[818] Bednarczyk, R. B. *et al.* Macrophage inflammatory factors promote epithelial-mesenchymal transition in breast cancer. *Oncotarget* **9**, 24272–24282 (2018).

[819] Marck, V. L. V. & Bracke, M. E. *Epithelial-Mesenchymal Transitions in Human Cancer* (Landes Bioscience, 2013).

[820] Abbink, P. *et al.* Rapid Cloning of Novel Rhesus Adenoviral Vaccine Vectors. *Journal of Virology* **92**, e01924–17 (2018).

[821] Irwin, C. R., Farmer, A., Willer, D. O. & Evans, D. H. In-fusion® cloning with vaccinia virus DNA polymerase. *Methods in Molecular Biology* **890**, 23–35 (2012).

[822] Ma, H. *et al.* Pol III Promoters to Express Small RNAs: Delineation of Transcription Initiation. *Molecular Therapy - Nucleic Acids* **3**, e161 (2014).

[823] Sasaki, T. *et al.* Autolysosome biogenesis and developmental senescence are regulated by both Spns1 and v-ATPase. *Autophagy* **13**, 386–403 (2017).

[824] Yanagisawa, H. *et al.* L-leucine and SPNS1 coordinately ameliorate dysfunction of autophagy in mouse and human Niemann-Pick type C disease. *Scientific Reports* **7**, 1–9 (2017).

[825] Rong, Y. *et al.* Spinster is required for autophagic lysosome reformation and mTOR reactivation following starvation. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7826–7831 (2011).

[826] Yanagisawa, H., Miyashita, T., Nakano, Y. & Yamamoto, D. HSpin1, a transmembrane protein interacting with Bcl-2/Bcl-x L , induces a caspase-independent autophagic cell death. *Cell Death & Differentiation* **10**, 798–807 (2003).

[827] Nakano, Y. *et al.* Mutations in the Novel Membrane Protein Spinster Interfere with Programmed Cell Death and Cause Neural Degeneration inDrosophila melanogaster. *Molecular and Cellular Biology* **21**, 3775–3788 (2001).

[828] Spns1 MGI Mouse Gene Detail - MGI:1920908 - spinster homolog 1. Mouse Genome Database (MGD), Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine. URL: http://www.informatics.jax.org/marker/MGI:1920908.

[829] Dermaut, B. *et al.* Aberrant lysosomal carbohydrate storage accompanies endocytic defects and neurodegeneration in Drosophila benchwarmer. *The Journal of Cell Biology* **170**, 127–139 (2005).

[830] Sasaki, T. *et al.* Aberrant Autolysosomal Regulation Is Linked to The Induction of Embryonic Senescence: Differential Roles of Beclin 1 and p53 in Vertebrate Spns1 Deficiency. *PLOS Genetics* **10**, e1004409 (2014).

[831] Harris, A. L. Hypoxia — a key regulatory factor in tumour growth. *Nature Reviews Cancer* **2**, 38–47 (2002).

[832] Vaupel, P., Kallinowski, F. & Okunieff, P. Blood flow, oxygen and nutrient supply, and metabolic microenvironment of human tumors: a review. *Cancer Research* **49**, 6449–6465 (1989).

[833] Nsengimana, J. *et al.* β-Catenin–mediated immune evasion pathway frequently operates in primary cutaneous melanomas. *The Journal of Clinical Investigation* **128**, 2048–2063 (2018).

[834] Goldman, M. *et al.* The UCSC Xena platform for public and private cancer genomics data visualization and interpretation. *bioRxiv* 326470 (2019).

[835] Klionsky, D. J. *et al.* Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition). *Autophagy* **12**, 1–222 (2016).

[836] Rebecca, V. W. *et al.* A Unified Approach to Targeting the Lysosome's Degradative and Growth Signaling Roles. *Cancer Discovery* **7**, 1266–1283 (2017).

[837] Brandes, N., Linial, N. & Linial, M. Modeling Functional Genetic Alteration in Cancer Reveals New Candidate Driver Genes. *bioRxiv* 242354 (2018).

[838] Newton-Bishop, J. A. *et al.* 25-Hydroxyvitamin D2 /D3 levels and factors associated with systemic inflammation and melanoma survival in the Leeds Melanoma Cohort. *International Journal of Cancer* **136**, 2890–2899 (2015).

Appendix A

Supplementary data

A.1 Details of software parameters

Analysis	Thesis section	Software	Version	Main command	Optional parameters
Read mapping	2.2.3	Bwa	0.7.15	mem	-t 14 -p -Y -K 100000000
Mark duplicates	2.2.3	biobambam	2.0.72	bamsormadup	-
Generate statistics of sequencing run	2.3.1	samtools	1.5	stats	-c 5,100,10 -t
Generate coverage information	2.3.1	samtools	1.5	bedcov	-
Genotype check	2.3.4	bcftools	1.3	gtcheck	-
Generate pileup for bam files	2.3.7	samtools	1.3	mpileup	t AD,INFO/AD -C50 -pm3 -F0.2 -d2000 -L500 -g -f
Mutation calling	2.3.8	bcftools	1.9	call	-m -f GQ -O z
Driver gene discovery	3.4	R: dndscv	0.0.0.9	dndscv	-
Pairwise mutation pattern analysis	3.5	R: DISCOVER	0.9.2	pairwise.discover.test	-
Groupwise mutation pattern analysis	3.5	R: DISCOVER	0.9.2	pairwise.discover.test	-
Mutational subtyping	3.1.6	R: SAMBAR	0.2	sambar	k = 23, kmin=1, kmax=250, signatureset c2.cp.v6.2.symbols.gmt, esize = genes_cds_length, cangenes = genes_incl
Neoantigen prediction	4.8	pvactools	1.0.5	Pvacseq run	NetMHC --top-score-metric=lowest -d full -e 8,9,10,11
CRISPR screen ROC curve analysis	5.4.1	Mageck	0.5.6	test	-k counts_table -c plasmid_sample -t ctr_rep_1, ctr_rep_2, ctr_rep3, ctr_rep4, ctr_rep5, ctr_rep6 --norm-method total
CRISPR screen analysis	5.4.2 and 5.6.2	Mageck	0.5.6	test	-k counts_table -c ctr_rep_1, ctr_rep_2, ctr_rep3, ctr_rep4, ctr_rep5, ctr_rep6 -t sort_rep_1, sort_rep2, sort_rep3, sort_rep4, sort_rep5, sort_rep6 --norm-method total

A.2 Targeted capture bait design information

Probe group	Information
1	Genes: *ABL1, ACAN, ACD, ACVR1B, ACVR2A, ADAM29, ADAMTS18, ADAR, AFF4, AHCTF1, AHNAK, AJUBA, AKAP9, AKT1, AKT2, AKT3, ALK, ALPK1, AMER1, ANK3, ANKRD35, AOX1, APC, AQR, AR, ARFGEF2, ARHGAP21, ARHGAP29, ARHGAP35, ARHGEF2, ARID1A, ARID1B, ARID2, ARIH2OS, ASPM, ASXL1, ASXL2, ATF6B, ATM, ATP1B4, ATP2A2, ATR, ATRX, AXIN1, AXIN2, B2M, BAP1, BAZ2B, BCL11A, BCL2L12, BCL3, BCL9L, BCLAF1, BCOR, BLM, BMPR2, BPTF, BRAF, BRCA1, BRCA2, BRD7, BRWD1, BUB1B, C1orf116, CAMTA1, CASP8, CBFB, CBLB, CCDC28A, CCND1, CCND2, CD1D, CD274, CD48, CD58, CDC27, CDH1, CDH2, CDK12, CDK4, CDK6, CDK8, CDKN1A, CDKN1B, CDKN2A, CDKN2B, CDKN2C, CDX2, CELA3B, CEP290, CHD3, CHD6, CHD9, CHEK2, CIC, CLCC1, CNOT3, CNTFR, COL11A1, COL1A1, COL2A1, COL6A3, COL9A2, CREB1, CREB3, CREBBP, CRLF2, CRTC3, CRYM, CSF2RA, CSF3R, CSMD1, CTCF, CTNNB1, CUX1, CXCL10, CXCL12, CXCR4, CYLD, DAXX, DCC, DDR2, DDX3X, DDX58, DICER1, DLG2, DNAH17, DNAJC11, DNMT3A, DOCK2, EBI3, ECT2L, EEF1A1, EGFR, EHHADH, EIF2AK3, EIF4G1, EIF4G3, ELAVL4, ELF3, EMG1, EP300, EPHA3, ERBB2, ERBB3, ERBB4, ERCC2, ERCC4, ESR1, EYS, EZH2, F3, F5, FAM58A, FANCA, FANCC, FANCD2, FANCE, FANCF, FANCG, FANCI, FAS, FAT1, FAT2, FBXW7, FCRL1, FCRL4, FGFR1, FGFR2, FGFR3, FGFR4, FLT3, FN1, FOXA1, FOXA2, FOXL2, FOXP1, FRYL, FSIP1, FUBP1, GABRA6, GATA3, GATA6, GATAD1, GNA11, GNAI2, GNAQ, GNAS, GPS2, GRB2, GRIN2A, H3F3A, H3F3B, H3F3C, HBD, HDAC9, HDLBP, HIST1H2BG, HMCN1, HNF1A, HNF4A, HRAS, HRNR, HYDIN, IDH1, IDH2, IFI16, IFI27, IFI30, IFI35, IFI44, IFI44L, IFI6, IFIH1, IFIT1, IFIT1B, IFIT2, IFIT3, IFITM1, IFITM2, IFNA1, IFNA14, IFNA2, IFNA21, IFNA4, IFNA5, IFNA6, IFNA8, IFNAR1, IFNAR2, IFNB1, IFNE, IFNG, IFNGR1, IFNGR2, IFNK, IFNL1, IFNL2, IFNLR1, IFNW1, IFRD1, IFRD2, IFT80, IKBKB, IKBKG, IL10RA, IL10RB, IL11RA, IL12A, IL12B, IL13RA1, IL15, IL20RA, IL20RB, IL21R, IL22RA2, IL2RB, IL2RG, IL31RA, IL36A, IL3RA, IL4R, IL5RA, IL6, IL6R, IL7R, IL9R, IREB2, IRF1, IRF2, IRF2BP1, IRF2BP2, IRF3, IRF4, IRF5, IRF6, IRF7, IRF8, IRF9, IRGM, IRS2, ISG15, ITGA4, ITGA9, ITSN1, JAK1, JAK2, JUN, KAT6A, KCNQ3, KDM5C, KDM6A, KDR, KEAP1, KEL, KIT, KLF4, KLF6, KMT2C, KMT2C, KMT2D, KMT2E, KNSTRN, KRAS, LEPR, LFNG, LNPEP, LRP6, LRPPRC, LRRC37A3, LSAMP, LZTR1, MAGEC3, MAGI2, MAP2K1, MAP2K2, MAP2K4, MAP3K1, MAP3K13, MAP3K4, MAP3K5, MAP3K9, MAX, MBD1, MDM2, MDM4, MECOM, MED12, MED23, MEN1, MET, MFNG, MITF, MKL1, MLH1, MLLT11, MLLT4, MPL, MRPL30, MRPL33, MRPS31, MRPS5, MSH2, MSH3, MSR1, MTOR, MUC17, MX1, MYB, MYC, MYCL, MYCN, NBPF10, NCOR1, NDUFB9, NF1, NF2, NFATC4, NFE2L2, NFIB, NIPBL, NKX2-1, NKX3-1, NOTCH1, NOTCH2, NOTCH2NL, NOTCH3, NOTCH4, NPAS3, NRAS, NTN4, NTRK2, NUGGC, NUP210L, NUP98, NUTM1, OAS1, OLFML2B, OR2Y1, OXA1L, OXNAD1, PAK3, PALB2, PAX5, PBRM1, PCBP1, PCDH18, PCDHA2, PCDHB12, PCDHGA1, PCED1B, PDGFRA, PER1, PHF6, PHOX2B, PIAS1, PIAS4, PIK3C2B, PIK3C2G, PIK3CA, PIK3CB, PIK3R1, PLCB1, PLCG1, PMS2, POLE, POLQ,*

	POLR2A, POLR2B, POM121, POT1, PPIAL4G, PPM1B, PPP2R1A, PPP6C, PRDM1, PRDM2, PREX2, PRG4, PRKAR1A, PRKCD, PSG4, PSME1, PTCH1, PTEN, PTGS1, PTPN11, PTPRB, PTPRK, PYHIN1, QKI, RAC1, RAD21, RAD51C, RAF1, RANBP17, RAPGEF5, RASA1, RASA2, RASSF2, RB1, RBM10, RCAN2, RELN, RET, RGS3, RHOA, RIPK1, RNF43, ROBO1, ROBO2, ROS1, RPL5, RPS27, RQCD1, RUNX1, SEC24D, SERPINB3, SETBP1, SETD2, SF3B1, SIRPB1, SLC27A3, SLIT2, SMAD2, SMAD3, SMAD4, SMARCA4, SMARCB1, SMARCD1, SMO, SNX31, SOCS1, SOCS3, SOS1, SOS2, SOX10, SOX2, SOX9, SP110, SPEN, SPOP, SPRED1, SPTA1, SQSTM1, STAG1, STAG2, STAT1, STAT2, STAT3, STAT4, STK11, STK19, SUFU, SVEP1, SYK, SYNE1, TACC1, TAOK1, TBC1D3B, TBL1XR1, TBX21, TBX3, TCF12, TCF4, TCF7L2, TDRD9, TET2, TGFBR2, TGIF1, TJP2, TLR4, TM2D1, TMEM216, TNFRSF1A, TP53, TP53BP1, TPTE, TPTE2, TRAF6, TRAF7, TRAIP, TRERF1, TRPA1, TRRAP, TSC1, TSC2, TTN, TYK2 , U2AF1, UBC, UBR5, UGGT2, USP6, USP9X, VDR, VHL, VIM, WASF3, WDR12, WNK1, WT1, XBP1, XRN1, ZBTB20, ZFP36L2, ZFX, ZNF638, ZNF831
2	Genes: *BNC2, CASP1, NUP107, PSIP1, SETDB1, TERT*
3	1934 probes targeting 960 reference SNPs, with approximately 1 SNP per 3MB of genome
4	Genes: *AHCTF1, BNC2, BRAF, CASP1, CCND1, CD274, CDK4, CDKN2A , CTLA4, FOXP1, GAB2, JAK2, KIT, KRAS, MDM2, MITF, MYC, NF1, NOTCH2, NRAS, NUP107, PAX5, PDCD1, PSIP1, PTEN, SETDB1, SMARCA4, TERT*
5	*DPH3* positions chr3:16306504, chr3:16306505 and chr3:16306508, *NDUFB9* position chr8:125551344, *NFKBIE* position chr6:44233400, *SDHD* positions chr11:111957523, ch11:111957541 and chr11:111957544, *TERT* positions chr5:1295161, chr5:1295191, chr5:1295228, chr5:1295242, chr5:1295243 and chr5:1295250
6	*ALK* intron 16, 18, 19, 20 and 21, *BRAF* intron 7, 8, 9 and 10 *NFIB* intron 7, 8, 9, *NTRK1* intron 7, 8, 9, 10, 11, 12 and 13, *RET* intron 7, 8, 9, 10, 11 and 12, *ROS1* intron 30, 31, 32, 33, 34 and 35
7	16351 probes targeting MHC alleles from the IMGT/HLA database including *HLA-A*, *HLA-B*, *HLA-C*, *HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRB1* [433]
8	1116 probes targeting 134 SNPs as part of a spiradenocarcinoma validation cohort

A.3 Caveman filters

Filter	Description
DTH	Less than 1/3 mutant alleles were >= 25 base quality
RP	Coverage was less than 8 and no mutant alleles were found in the first 2/3 of a read (shifted 0.08 from the start and extended 0.08 more than 2/3 of the read length)
MN	More than 0.03 of mutant alleles that were >= 15 base quality found in the matched normal
PT	Mutant alleles all on one direction of read (1rd allowed on opposite strand) and in second half of the read. Second half of read contains the motif GGC[AT]G in sequenced orientation and the mean base read contains the motif GGC[AT]G in sequenced orientation and the mean base quality of all bases after the motif was less than 20
MQ	Mean mapping quality of the mutant allele reads was < 21
SR	Position falls within a simple repeat using the supplied bed file
CR	Position falls within a centromeric repeat using the supplied bed file
PH	Mutant reads were on one strand (permitted proportion on other strand: 0.04
TI	More than 10 percent of reads covering this position contained an indel according to mapping
SRP	More than 80 percent of reads contain the mutant allele at the same read position
HSD	Position falls within a high sequencing depth region using the supplied bed file
AN	Position could not be annotated against a transcript using the supplied bed file
VUM	Position has >= 3 mutant allele present in at least 1 percent unmatched normal samples in the unmatched VCF.
SE	Coverage is >= 10 on each strand but mutant allele is only present on one strand
MNP	Tumour sample mutant allele proportion - normal sample mutant allele proportion < 0.2
DTH	Less than 1/3 mutant alleles were >= 25 base quality

RP	Coverage was less than 8 and no mutant alleles were found in the first 2/3 of a read (shifted 0.08 from the start and extended 0.08 more than 2/3 of the read length)
MN	More than 0.03 of mutant alleles that were >= 15 base quality found in the matched normal
PT	Mutant alleles all on one direction of read (1rd allowed on opposite strand) and in second half of the read. Second half of read contains the motif GGC[AT]G in sequenced orientation and the mean base read contains the motif GGC[AT]G in sequenced orientation and the mean base quality of all bases after the motif was less than 20
MQ	Mean mapping quality of the mutant allele reads was < 21
SR	Position falls within a simple repeat using the supplied bed file
CR	Position falls within a centromeric repeat using the supplied bed file
PH	Mutant reads were on one strand (permitted proportion on other strand: 0.04
TI	More than 10 percent of reads covering this position contained an indel according to mapping
SRP	More than 80 percent of reads contain the mutant allele at the same read position
HSD	Position falls within a high sequencing depth region using the supplied bed file
AN	Position could not be annotated against a transcript using the supplied bed file
VUM	Position has >= 3 mutant allele present in at least 1 percent unmatched normal samples in the unmatched VCF
SE	Coverage is >= 10 on each strand but mutant allele is only present on one strand
MNP	Tumour sample mutant allele proportion - normal sample mutant allele proportion < 0.2

A.4 Pindel filters

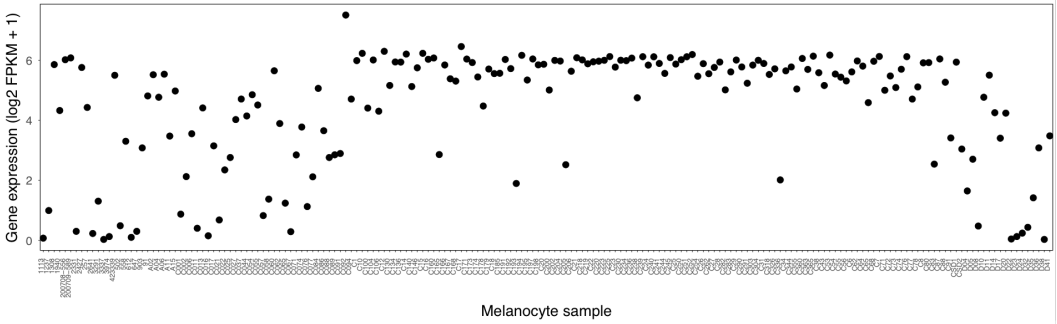
Filter	Description
F001	Pass if Mt > Wt Reads: Likely GERMLINE
F002	No Wt calls in variants over 4bp in length: Likely GERMLINE
F003	Tum low call count strand bias check
F004	Tum medium read depth strand bias check: Calls In 8% Reads Bt Depth 10 And 200 (inclusive)
F005	Tum high read depth strand bias check: Calls In 4% Reads > Depth 200
F006	Small call excessive repeat check: Fail if Length <= 4 and Repeats > 9
F007	Sufficient Normal Depth: If Mt Depth > 5 then Wt > 8% tum depth
F008	Wildtype contamination: Fail when wt reads > 5% mt reads
F009	Is coding: Pass when in gene footprint
F010	Variant must not exist within the Unmatched Normal Panel

A.5 DISCOVER: Analysis of mutational patterns

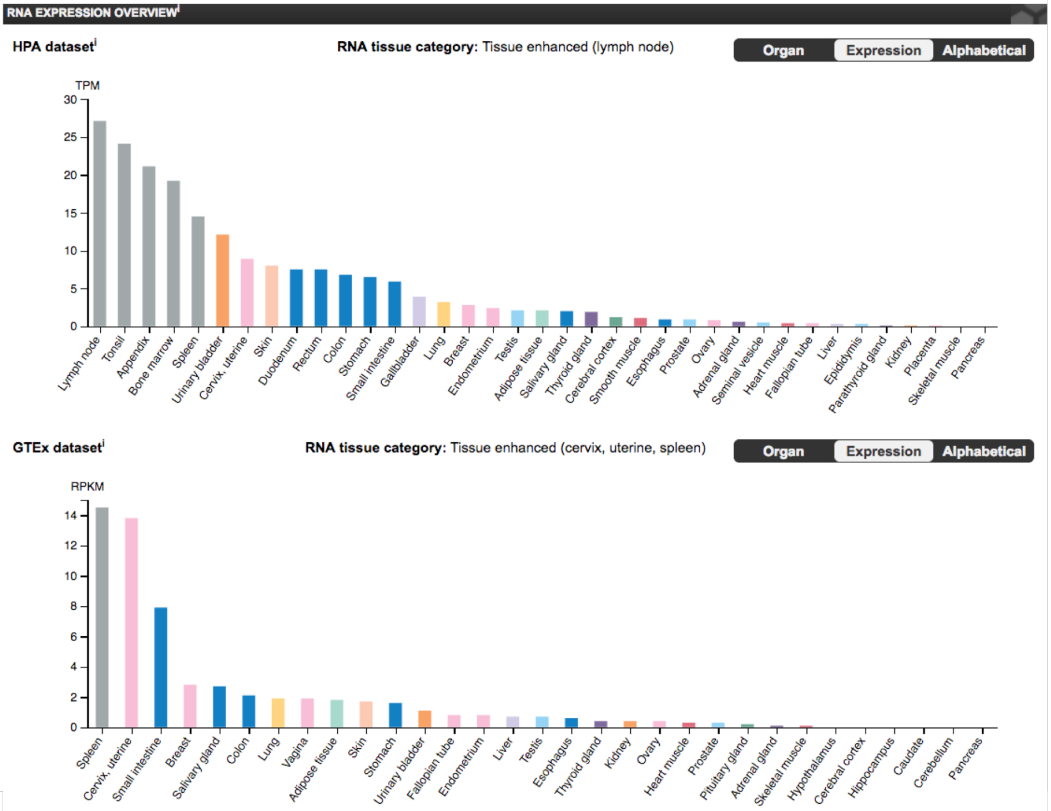
Analysis of mutual exclusivity were done for all 524 melanoma patients, including SNPs and indels in all genes of the panel with a minimum 5% mutation recurrence. To provide a cleaner output, only nonsynonymous coding mutations and highly recurrent promoter mutations were included. Furthermore, only known hotspot variants in *BRAF*, *NRAS* and *TERT* were considered. A 5% FDR threshold was applied, followed by an additional filter to address a weakness in the algorithm. The method was found to discover several mutually exclusive gene pairs where a notable fraction of patients showed a substantial number of co-occurring mutations in both genes. Therefore, such gene pairs were removed by applying a filter that excluded gene pairs where either gene had $\geq$15% of all its mutations in tumors showing co-occurring mutations with the second gene. By applying this filter, the number of gene pairs were reduced from 78 to 8. Notably, genes with a high alteration frequency such as *BRAF* had a tendency to show mutual exclusivity with an unrealistically high number of other genes, but a large proportion of such hits were successfully removed using the 15% overlap filter.

A.6 Expression of *IRF4* in normal tissue

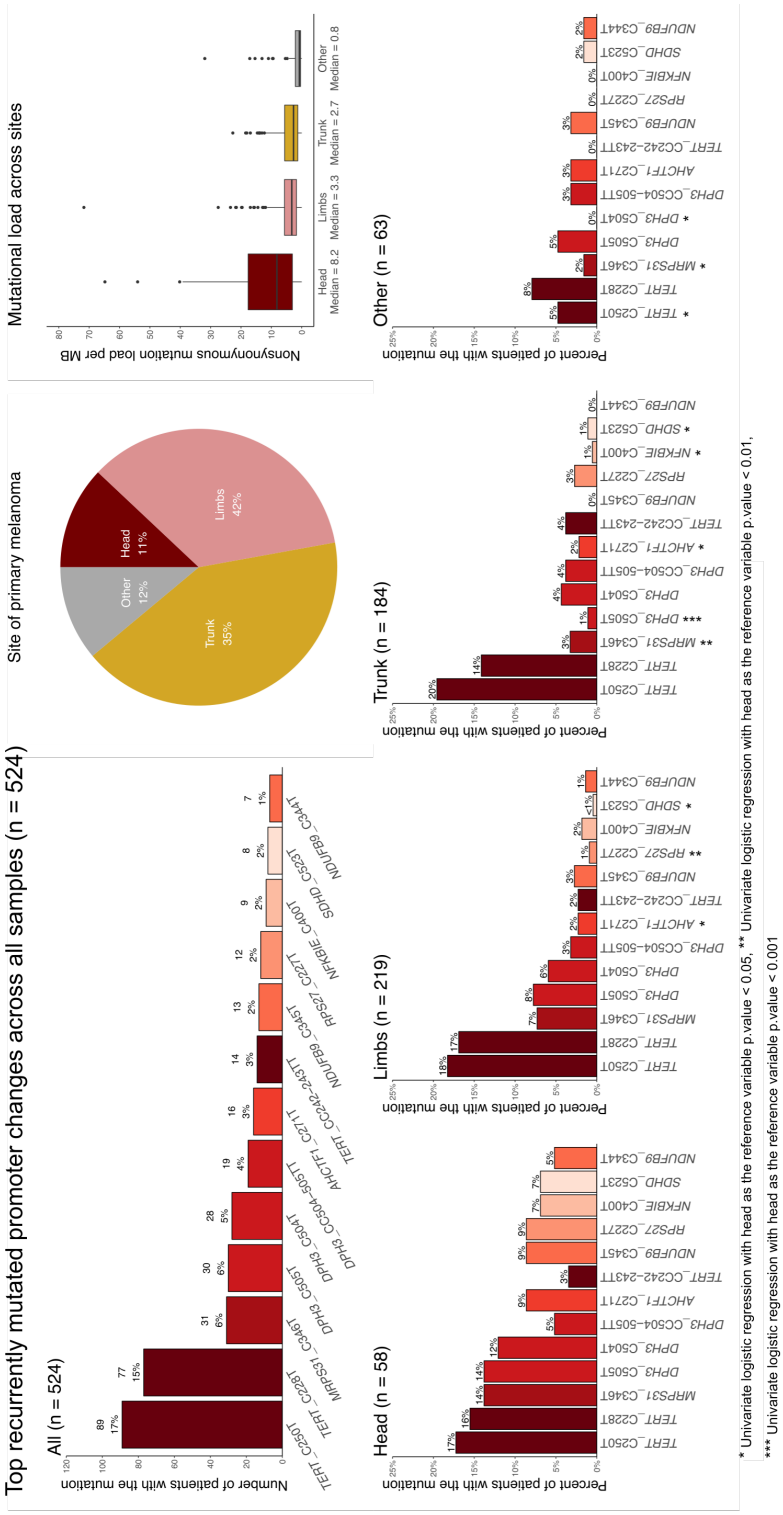
A. Expression of *IRF4* in primary melanocytes



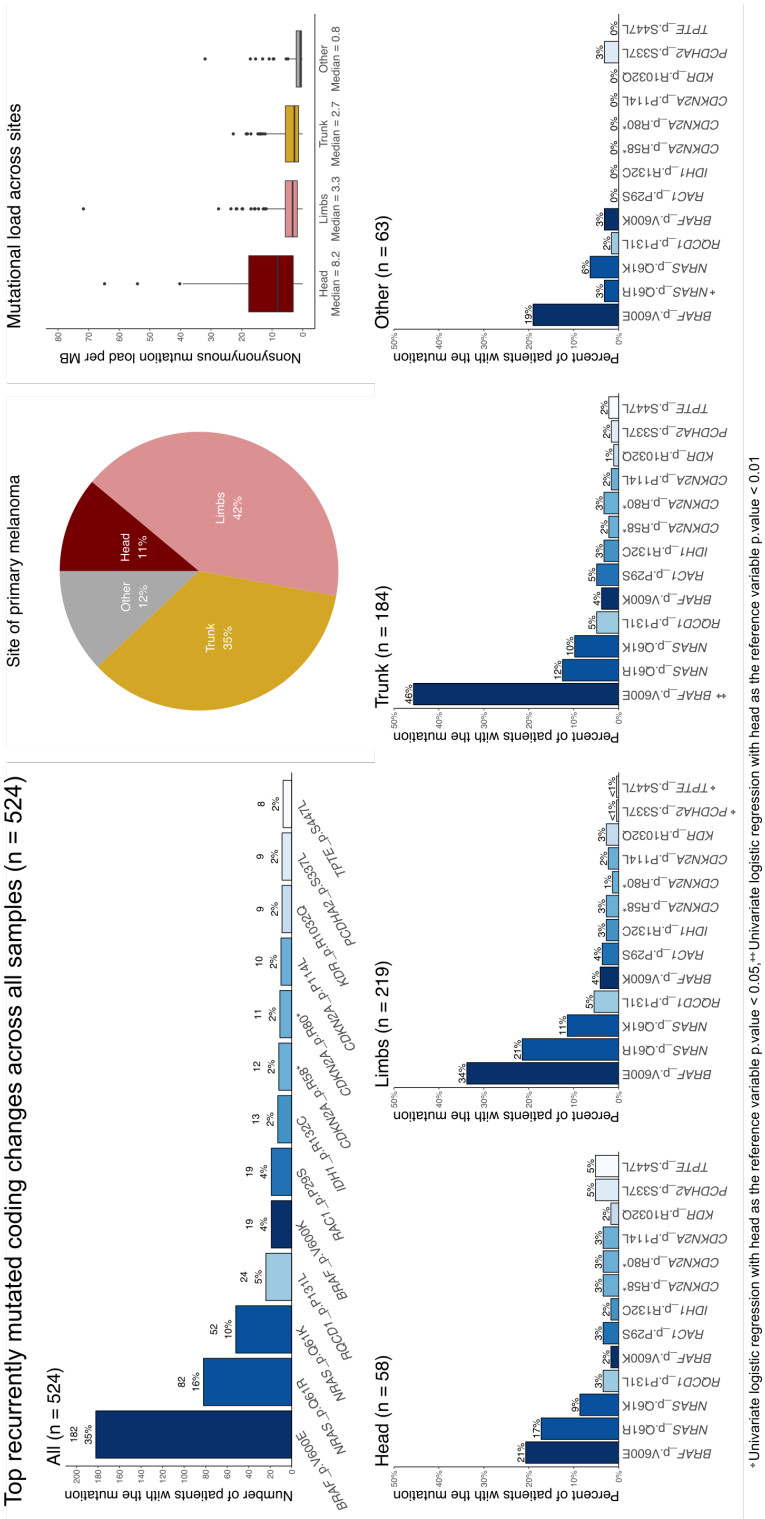
B. Expression of *IRF4,* The human protein atlas, accessed 18 February 2019



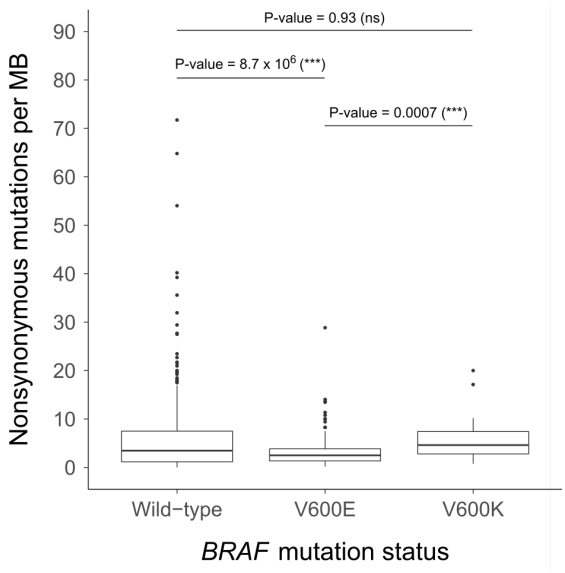
A.7 Top promoter mutations across body sites



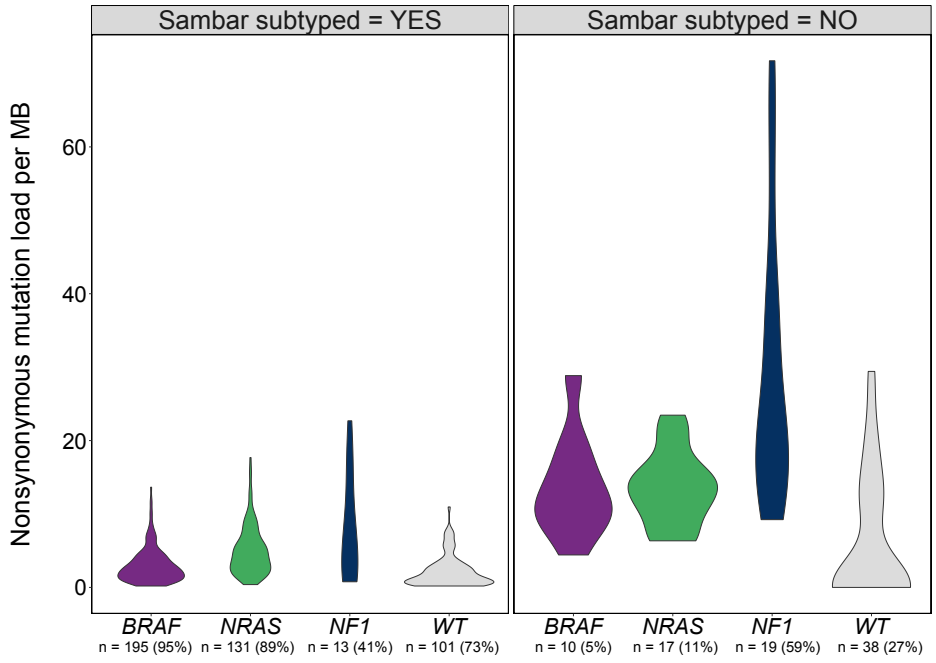
A.8 Top coding mutations across body sites



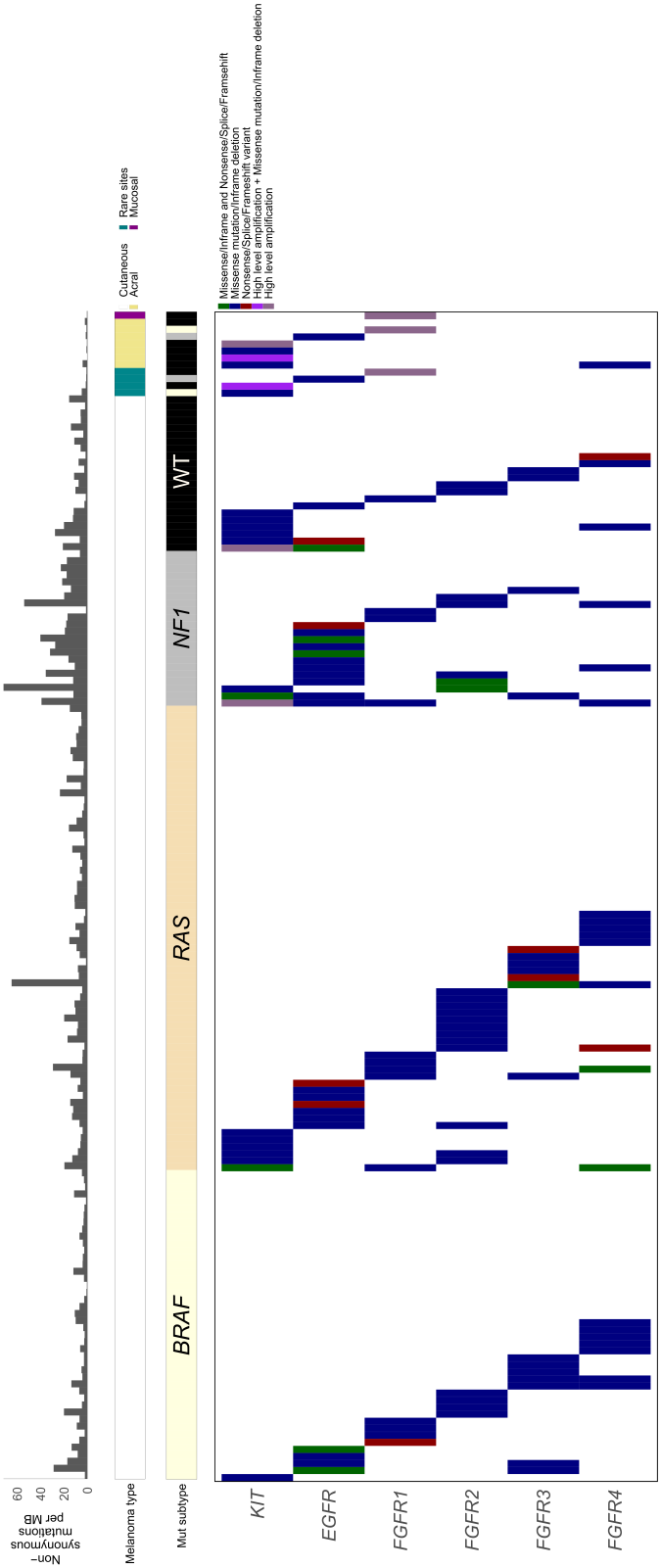
A.9 Comparison of mutation load between *BRAF* variants



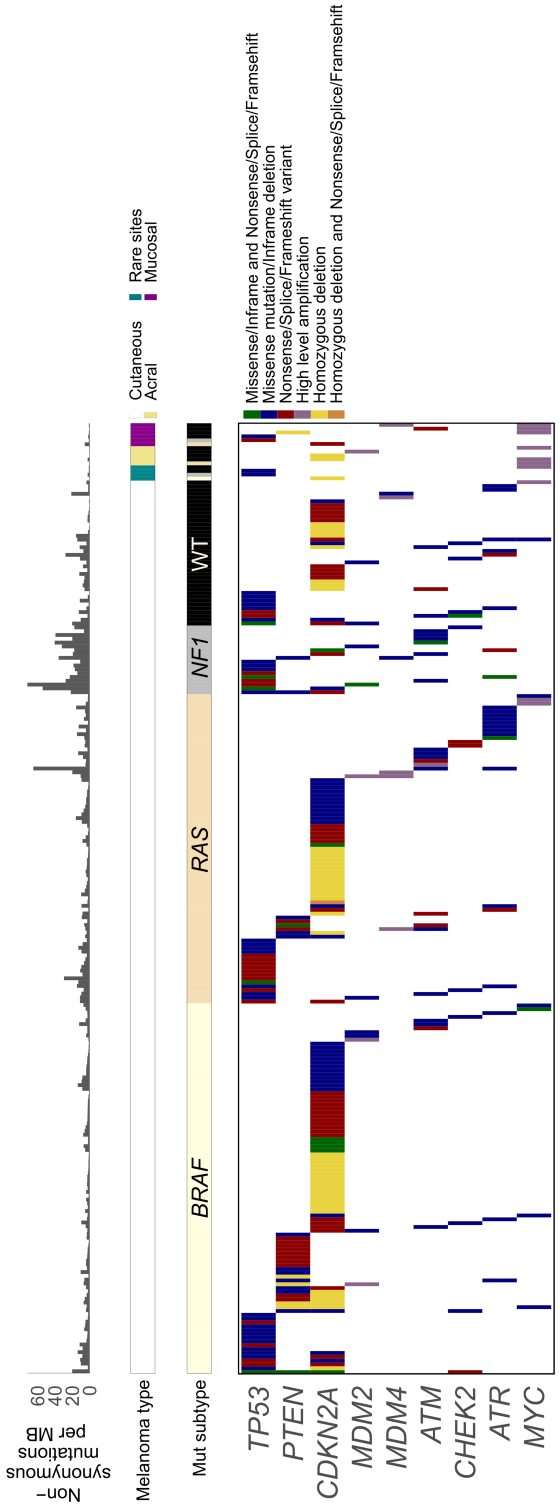
A.10 Distribution of mutation load in Sambar subtyped versus non-subtyped samples



A.11 Patterns of genetic alterations in RTKs

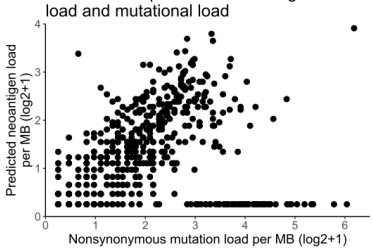


A.12 Patterns of genetic alterations in *TP53*–associated pathways in primary melanoma

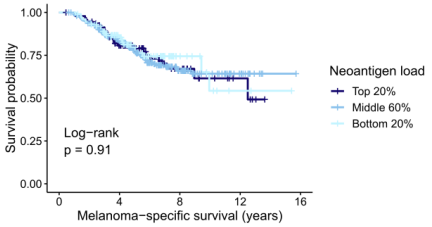


A.13 The effect of neoantigen load on survival and its correlation with mutation load

A. Relationship between neoantigen load and mutational load



B. Kaplan−Meier survival analysis: 20% / 60% / 20% groups

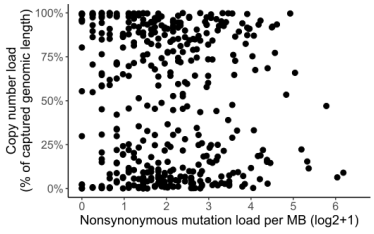


C. Cox proportional hazards regression analysis with neoantigen load

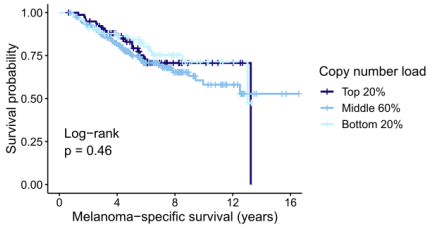
	Multivariate analysis		
Variable	**HR**	**95% CI**	**P-value**
Neoantigen load per MB	1.00	0.92-1.09	0.973
Age at diagnosis	1.02	1.001-1.04	0.043
Sex, male	1.62	1.09-2.42	0.017
Ulceration, present	1.89	1.21-2.97	0.005
Breslow thickness	1.14	1.04-1.23	0.003
Stage			
II vs I	0.76	0.41-1.41	0.387
III vs I	1.96	1.04-3.70	0.038

A.14 The effect of copy number load on survival and its correlation with mutation load

A. Relationship between copy number load and mutational load



B. Kaplan−Meier survival analysis: 20% / 60% / 20% groups



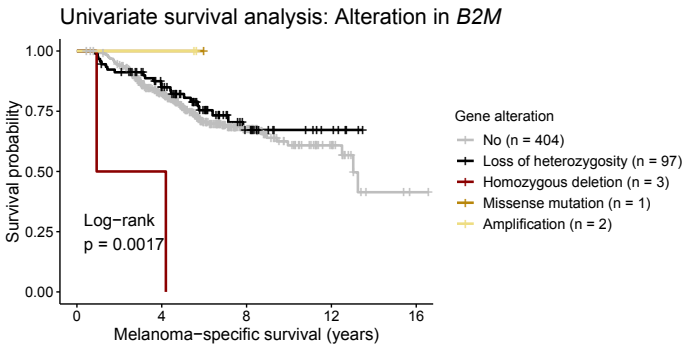
C. Cox proportional hazards regression analysis with copy number load

	Multivariate analysis		
Variable	**HR**	**95% CI**	**P-value**
Copy number load (%)	1.00	0.995-1.01	0.744
Age at diagnosis	1.03	1.01-1.05	0.011
Sex, male	1.71	1.11-2.61	0.014
Ulceration, present	1.96	1.19-3.22	0.008
Breslow thickness	1.12	1.03-1.21	0.005
Stage			
II vs I	0.65	0.34-1.25	0.197
III vs I	1.55	0.77-3.01	0.219

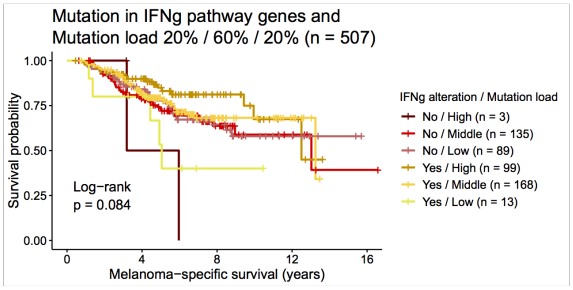
A.15 Comparison of key clinical variables between datasets

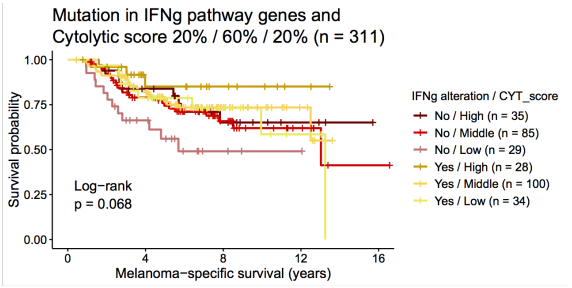
	Overall (n=524)	Mutation and transcriptome (n=319)	Mutation only (n=205)
Sex			
Female	263 (50 %)	168 (53 %)	95 (46 %)
Male	261 (50 %)	151 (47 %)	110 (54 %)
Age (years)			
Mean (SD)	57 (± 12)	56 (± 12)	58 (± 12)
Stage			
I	167 (32 %)	88 (28 %)	79 (39 %)
II	253 (48 %)	173 (54 %)	80 (39 %)
III	97 (19 %)	52 (16 %)	45 (22 %)
Unknown	7 (1 %)	6 (2 %)	1 (0 %)
Breslow thickness (mm)			
Mean (SD)	3.0 (± 2.4)	3.3 (± 2.7)	2.6 (± 1.9)
Missing	20 (3.8%)	9 (2.8%)	11 (5.4%)
Ulceration			
No	289 (55 %)	165 (52 %)	124 (60 %)
Yes	169 (32 %)	119 (37 %)	50 (24 %)
Unknown	66 (13 %)	35 (11 %)	31 (15 %)
Mitotic rate (mitoses/ per mm2)			
<1	66 (13 %)	39 (12 %)	27 (13 %)
>=1	402 (77 %)	241 (76 %)	161 (79 %)
Unknown	56 (11 %)	39 (12 %)	17 (8 %)
Tumour-infiltrating lymphocytes			
Absent	83 (16 %)	41 (13 %)	42 (20 %)
Yes (Unclassified)	47 (9 %)	31 (10 %)	16 (8 %)
Non-brisk	215 (41 %)	136 (43 %)	79 (39 %)
Brisk	77 (15 %)	41 (13 %)	36 (18 %)
Unknown	102 (19 %)	70 (22 %)	32 (16 %)
Ns mutation load per MB			
Mean (SD)	5.1 (± 7.2)	4.9 (± 7.4)	5.5 (± 6.9)
Relapse			
No	333 (64 %)	201 (63 %)	132 (64 %)
Yes	191 (36 %)	118 (37 %)	73 (36 %)
Type of melanoma			
Acral	24 (5 %)	18 (6 %)	6 (3 %)
Cutaneous	468 (89 %)	284 (89 %)	184 (90 %)
Mucosal	7 (1 %)	4 (1 %)	3 (1 %)
Other rare sites	13 (2 %)	12 (4 %)	1 (0 %)
Unknown	12 (2 %)	1 (0 %)	11 (5 %)

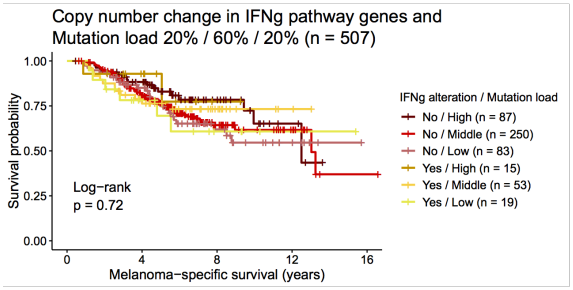
A.16 The effect of genetic alterations in *B2M* on survival

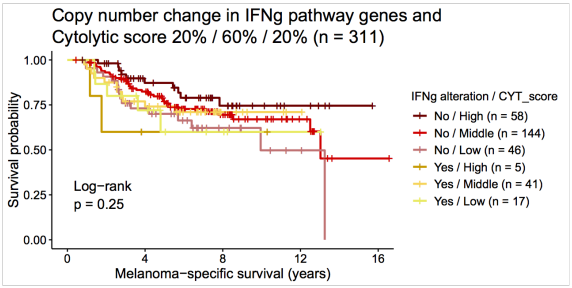


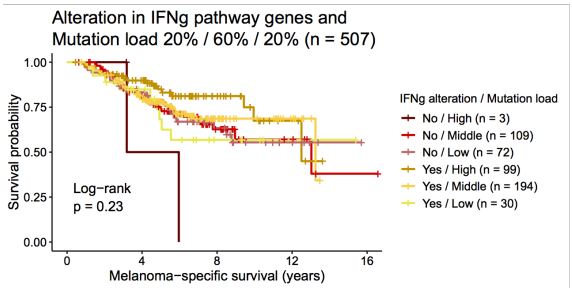
A.17 The effect of genetic alterations in IFN-γ pathway genes on survival

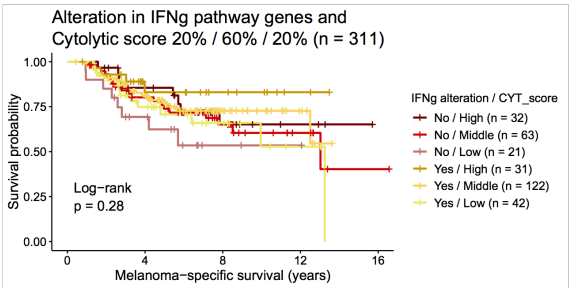




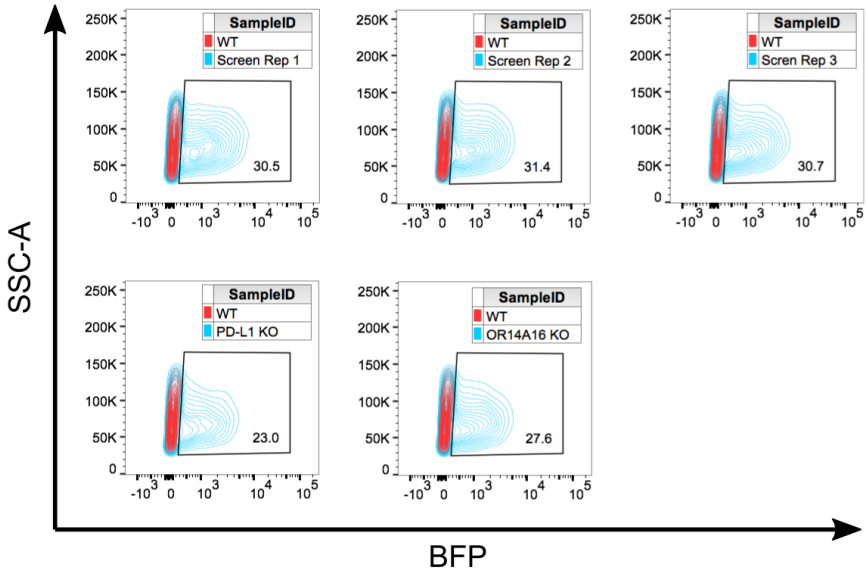




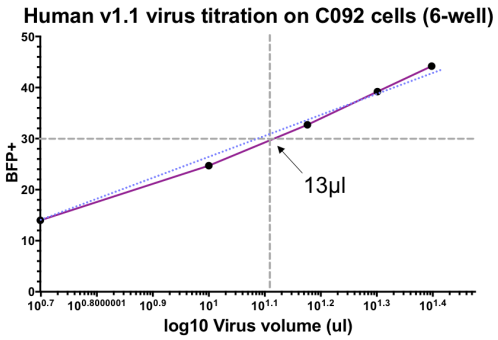


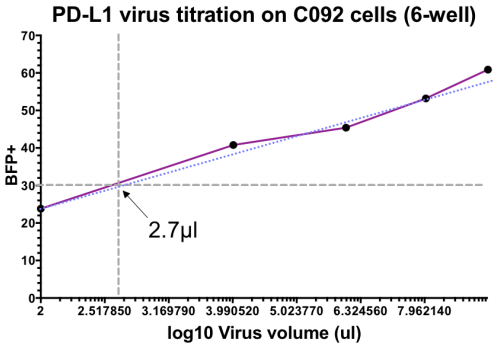


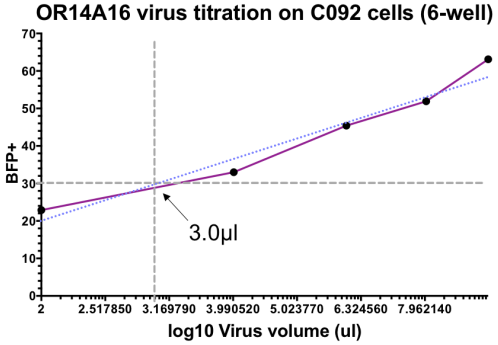
A.18 MOI for C092 screen replicates and controls



A.19 Titration of virus and calculations of MOIs for the CRISPR-Cas9 screen



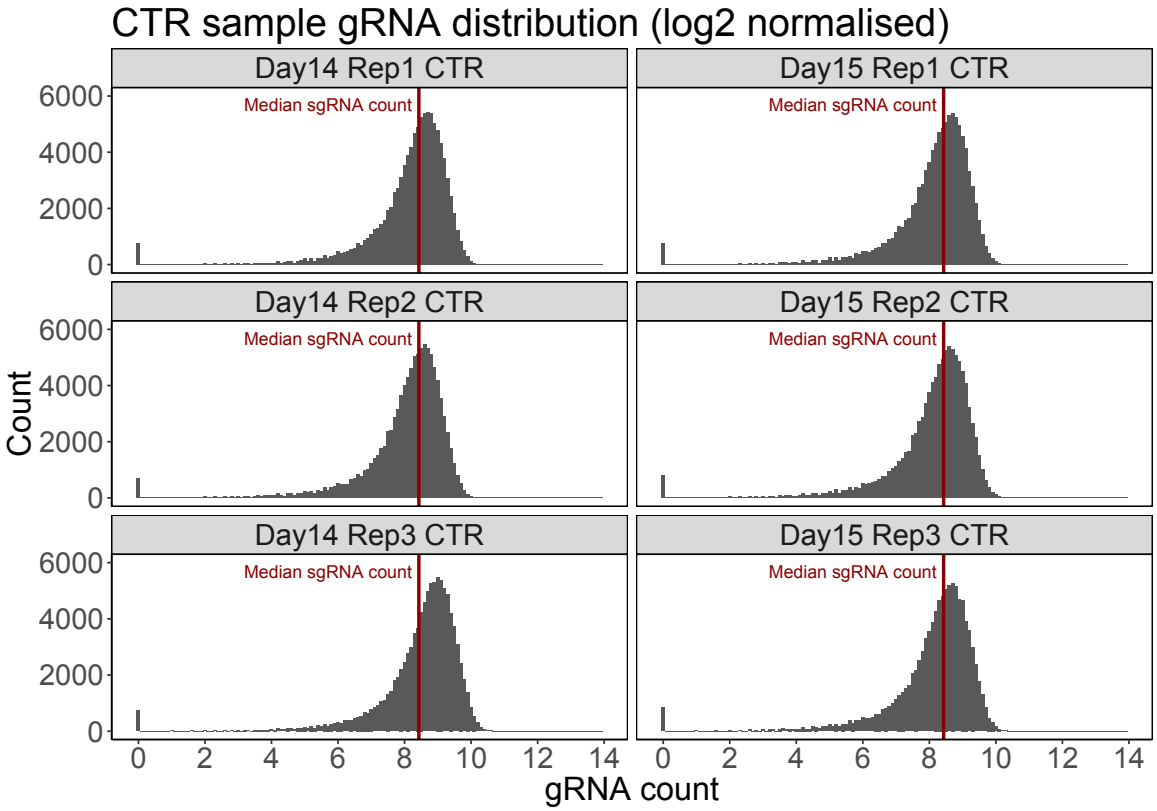




Calculation of virus volumes for the screen

Lentivirus	Volume to get MOI 0.3 in 6-well (µl)	Increase in surface area 6-well to T150	Estimated volume to get MOI 0.3 in T150 (µl)
Human v.1.1	13.0	15x	195
PD-L1	2.7	15x	40
OR14A16	3.0	15x	45

A.20 Distribution of gRNA counts in control samples of the CRISPR-Cas9 screen



A.21 Sort statistics for the C092 screen

Sample	Total sorted events	PD-L1 low fraction	PD-L1 dim fraction
Day 14 Rep 1	90 million	0.45 million	2.2 million
Day 14 Rep 2	90 million	0.30 million	1.63 million
Day 14 Rep 3	110 million	0.35 million	1.85 million
Day 15 Rep 1	70 million	0.18 million	0.53 million
Day 15 Rep 2	95.8 million	0.30 million	1 million
Day 15 Rep 3	87 million	0.29 million	1 million

A.22 Method to compute the ROC curve and AUC calculations

1. Perform Mageck analysis comparing plasmid counts with counts in the control samples (Section A.1)

2. Order the data frame to list genes in ascending order based on the Mageck negative rank (neg.rank)

3. Annotate in a new column whether each gene is part of the Bagel core reference genes of essential or non-essential genes [714]

4. At each position in the data frame, loop through the genes in ascending order and assign a true positive (TP), true negative (TN), false positive (FP) and false negative (FN) value based on:

   (a) TP: sum of the total number of bagel essential genes identified

   (b) TN: sum of the total number of bagel non-essential genes not yet identified

   (c) FP: sum of the total number of bagel non-essential genes identified

   (d) FN: sum of the total number of bagel essential genes not yet identified

5. Scale the negative rank to range from 0 to 1

6. Calculate the sensitivity, specificity and false positive rate (FPR) as follows:

   (a) $Sensitivity = \frac{TP}{TP+FN}$

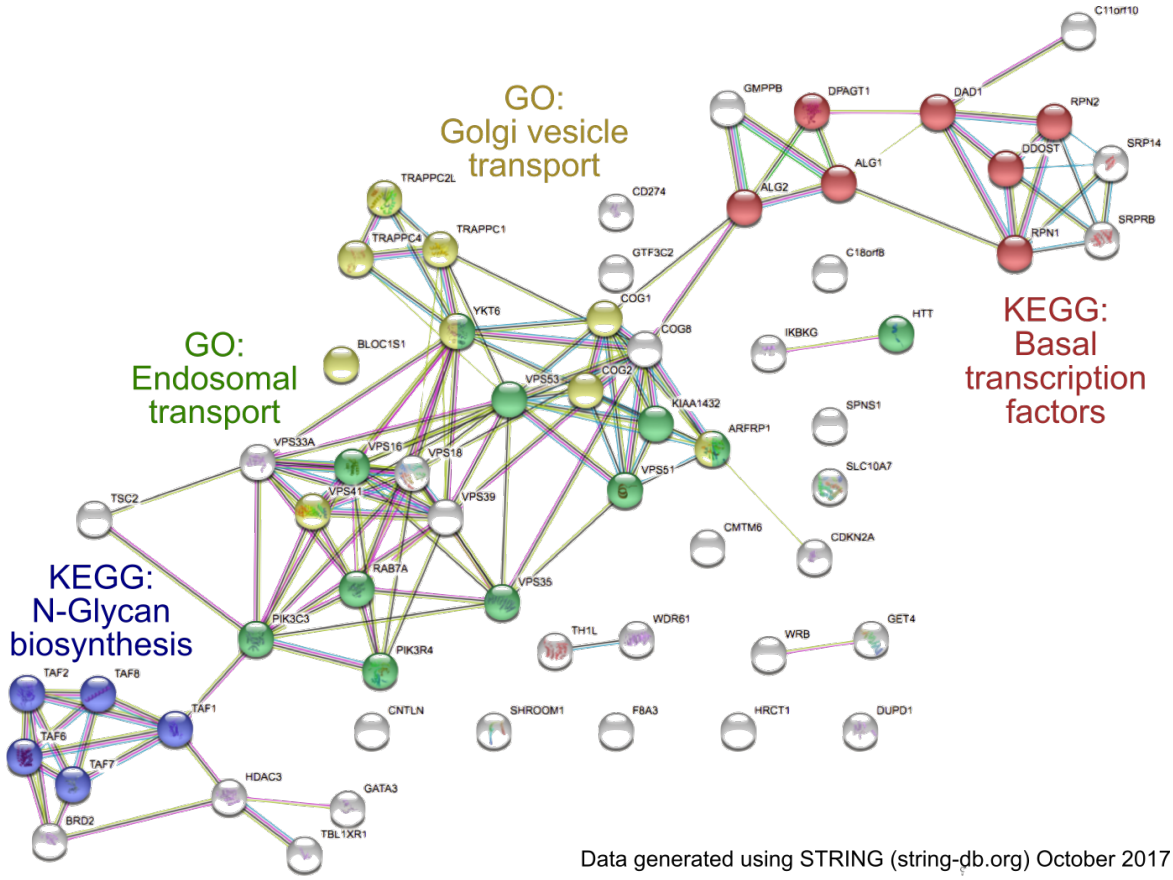
   (b) $Specificity = \frac{TN}{TN+FP}$

   (c) $FPR = 1 - Specificity$

7. AUC is calculated using the R package kulife using the FPR and sensitivity values

8. ROC curve is generated the R package ggplot2 using the FPR and sensitivity values

A.23 STRING analysis of PD-L1 CRISPR-Cas9 screen hits



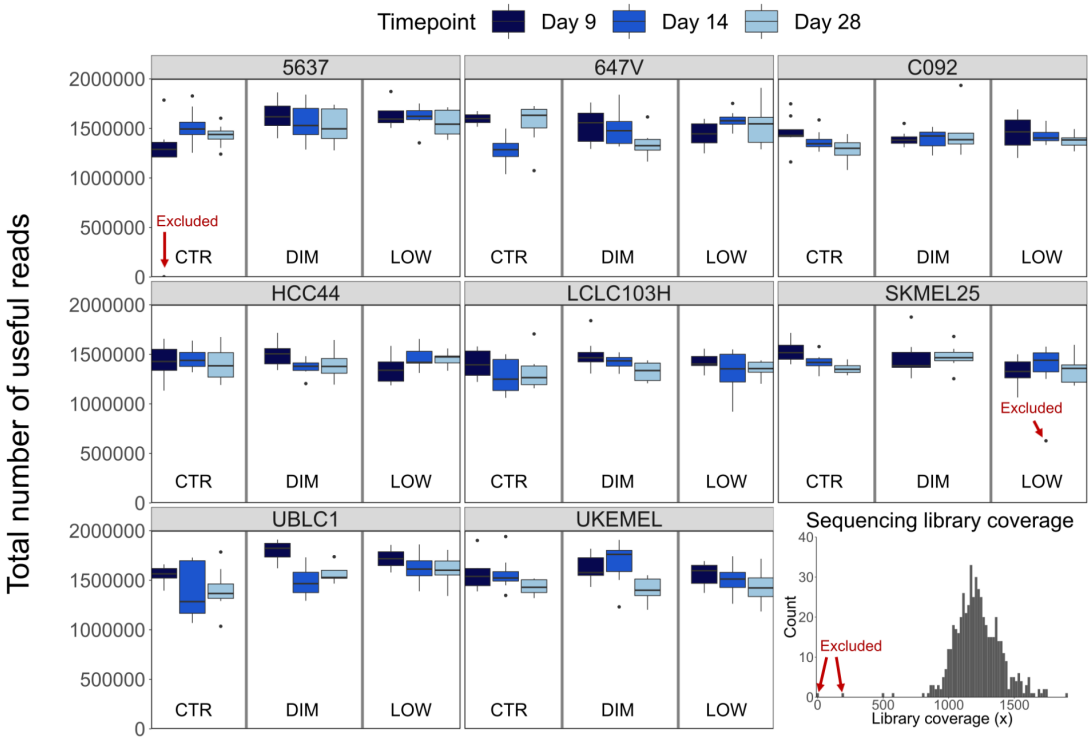
Data generated using STRING (string-db.org) October 2017

A.24 MOIs in the pooled validation screen

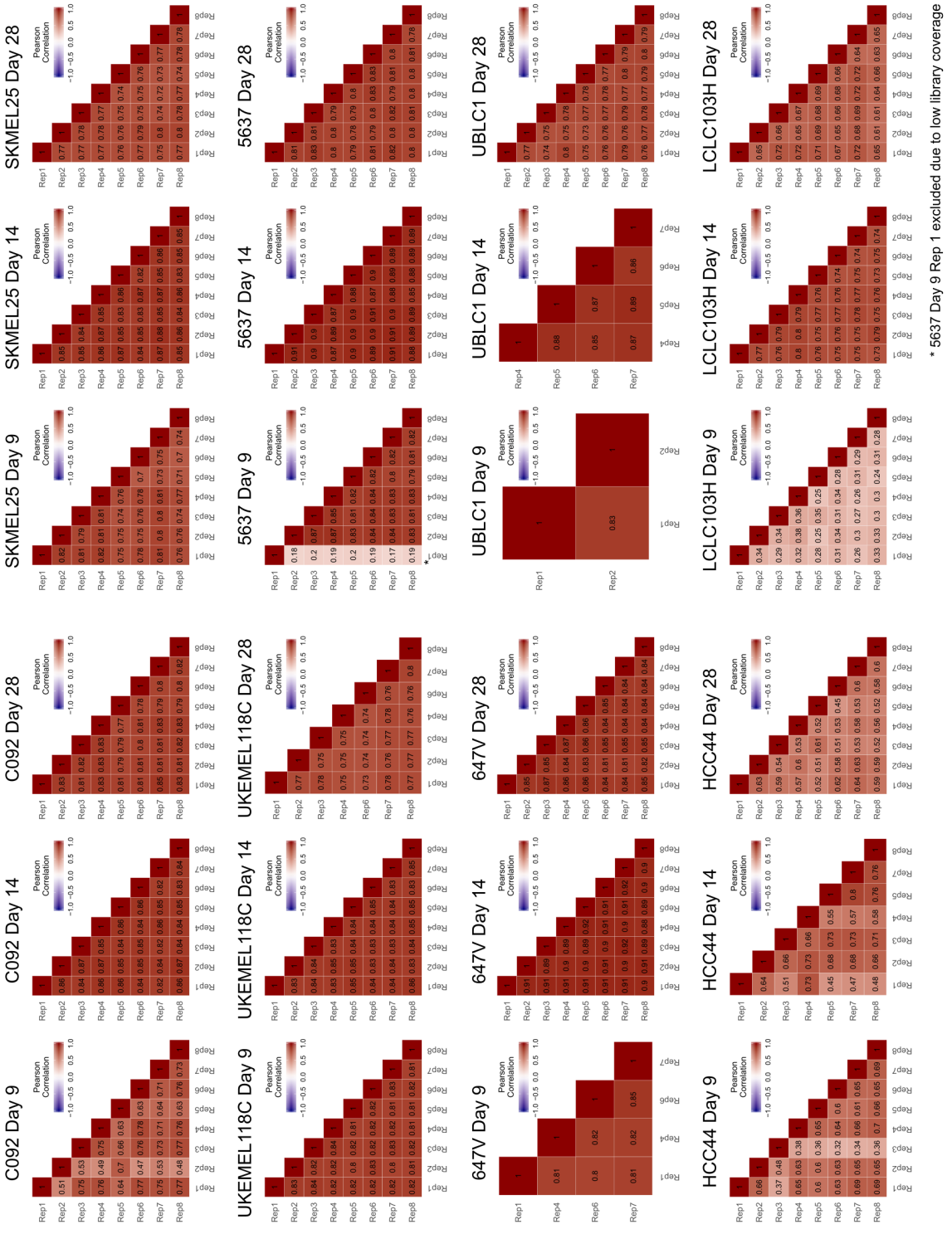
Cell line	Replicate	MOI (%)	Median MOI (%)
5637	1	27	
5637	2	29	
5637	3	25	
5637	4	21	23
5637	5	26	
5637	6	21	
5637	7	22	
5637	8	22	
647V	1	31	
647V	2	23	
647V	3	27	
647V	4	29	29
647V	5	29	
647V	6	32	
647V	7	29	
647V	8	34	
C092	1	31	
C092	2	26	
C092	3	29	
C092	4	28	29
C092	5	27	
C092	6	29	
C092	7	31	
C092	8	28	
HCC44	1	33	
HCC44	2	33	
HCC44	3	32	
HCC44	4	31	33
HCC44	5	31	
HCC44	6	33	
HCC44	7	34	
HCC44	8	31	

Cell line	Replicate	MOI (%)	Median MOI (%)
LCLC103H	1	31	
LCLC103H	2	27	
LCLC103H	3	29	
LCLC103H	4	27	28
LCLC103H	5	27	
LCLC103H	6	31	
LCLC103H	7	32	
LCLC103H	8	25	
SKMEL25	1	21	
SKMEL25	2	19	
SKMEL25	3	22	
SKMEL25	4	20	22
SKMEL25	5	22	
SKMEL25	6	23	
SKMEL25	7	22	
SKMEL25	8	21	
UBLC1	1	28	
UBLC1	2	27	
UBLC1	3	28	
UBLC1	4	26	28
UBLC1	5	29	
UBLC1	6	26	
UBLC1	7	30	
UBLC1	8	27	
UKEMEL118C	1	40	
UKEMEL118C	2	27	
UKEMEL118C	3	32	
UKEMEL118C	4	33	33
UKEMEL118C	5	34	
UKEMEL118C	6	33	
UKEMEL118C	7	39	
UKEMEL118C	8	31	

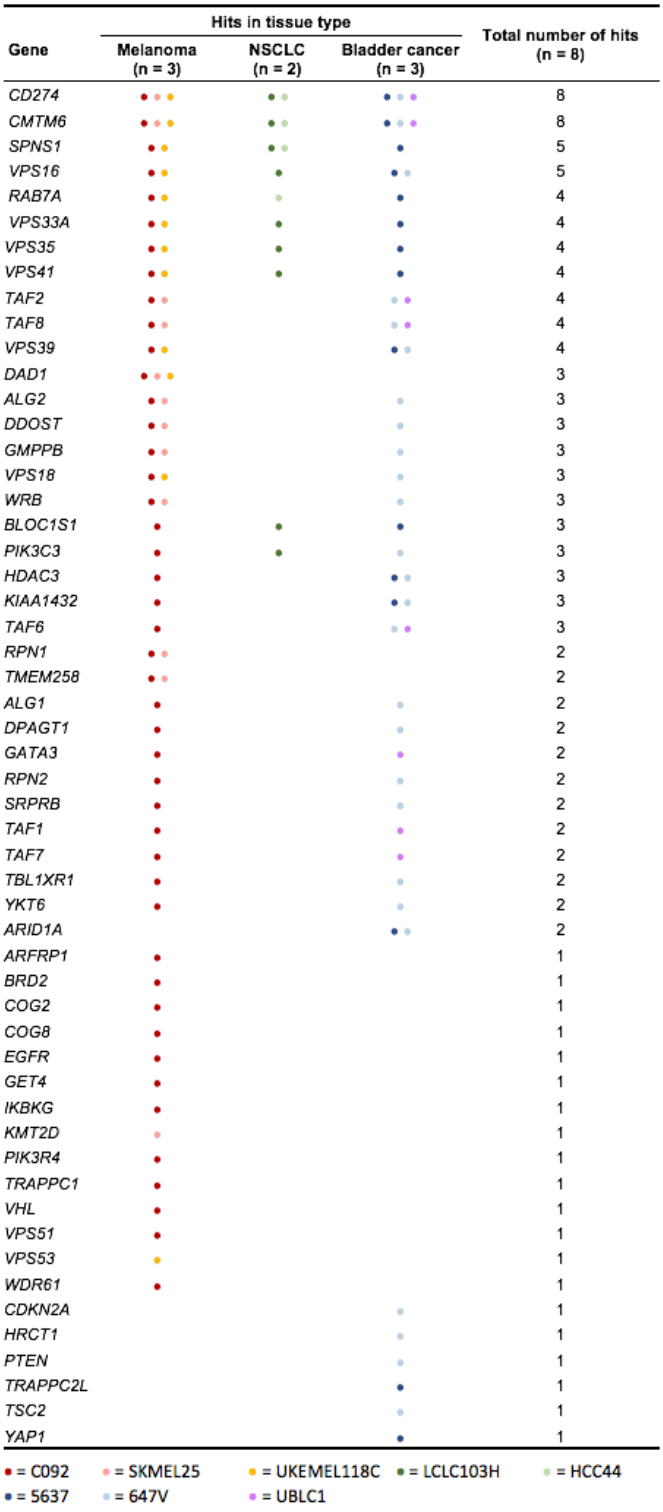
A.25 Validation screen QC



A.26 Correlation between gRNA counts in control samples in the validation screen



A.27 Validation pattern across cell lines (FDR <10%)

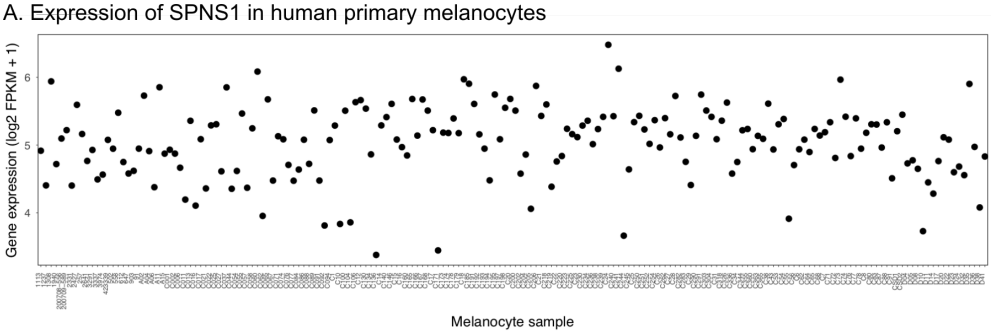


Gene	Hits in tissue type: Melanoma (n = 3)	NSCLC (n = 2)	Bladder cancer (n = 3)	Total number of hits (n = 8)
CD274	● ● ●	● ●	● ● ●	8
CMTM6	● ● ●	● ●	● ● ●	8
SPNS1	● ●	● ●	●	5
VPS16	● ●	●	● ●	5
RAB7A	● ●	●	●	4
VPS33A	● ●	●	●	4
VPS35	● ●	●	●	4
VPS41	● ●	●	●	4
TAF2	● ●		● ●	4
TAF8	● ●		● ●	4
VPS39	● ●		● ●	4
DAD1	● ● ●			3
ALG2	● ●		●	3
DDOST	● ●		●	3
GMPPB	● ●		●	3
VPS18	● ●		●	3
WRB	● ●		●	3
BLOC1S1	●	●	●	3
PIK3C3	●	●	●	3
HDAC3	●		● ●	3
KIAA1432	●		● ●	3
TAF6	●		● ●	3
RPN1	● ●			2
TMEM258	● ●			2
ALG1	●		●	2
DPAGT1	●		●	2
GATA3	●		●	2
RPN2	●		●	2
SRPRB	●		●	2
TAF1	●		●	2
TAF7	●		●	2
TBL1XR1	●		●	2
YKT6	●		●	2
ARID1A			● ●	2
ARFRP1	●			1
BRD2	●			1
COG2	●			1
COG8	●			1
EGFR	●			1
GET4	●			1
IKBKG	●			1
KMT2D	●			1
PIK3R4	●			1
TRAPPC1	●			1
VHL	●			1
VPS51	●			1
VPS53	●			1
WDR61	●			1
CDKN2A			●	1
HRCT1			●	1
PTEN			●	1
TRAPPC2L			●	1
TSC2			●	1
YAP1			●	1

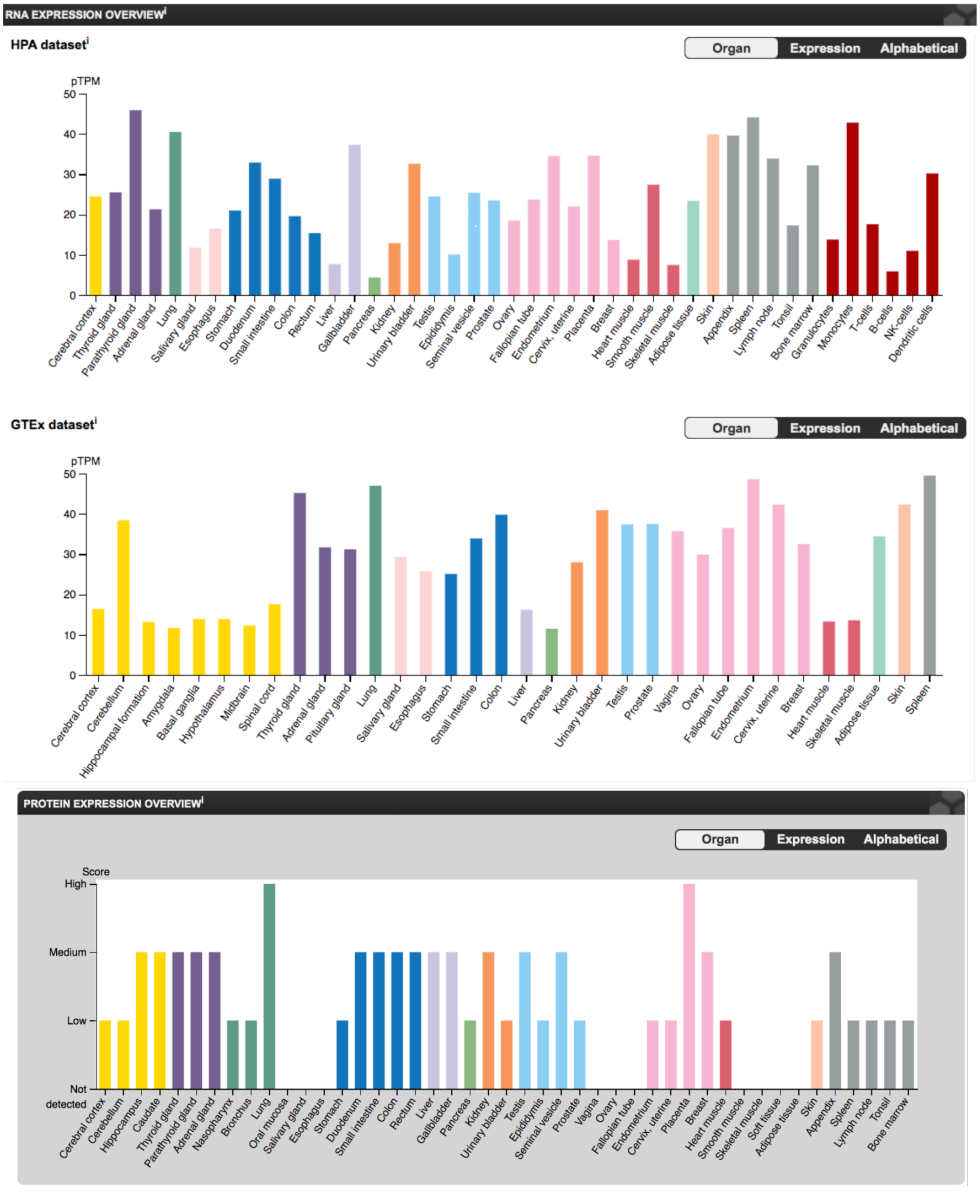
● = C092 ● = SKMEL25 ● = UKEMEL118C ● = LCLC103H ● = HCC44
● = 5637 ● = 647V ● = UBLC1

A.28 Expression of *SPNS1* in normal tissues

A. Expression of SPNS1 in human primary melanocytes

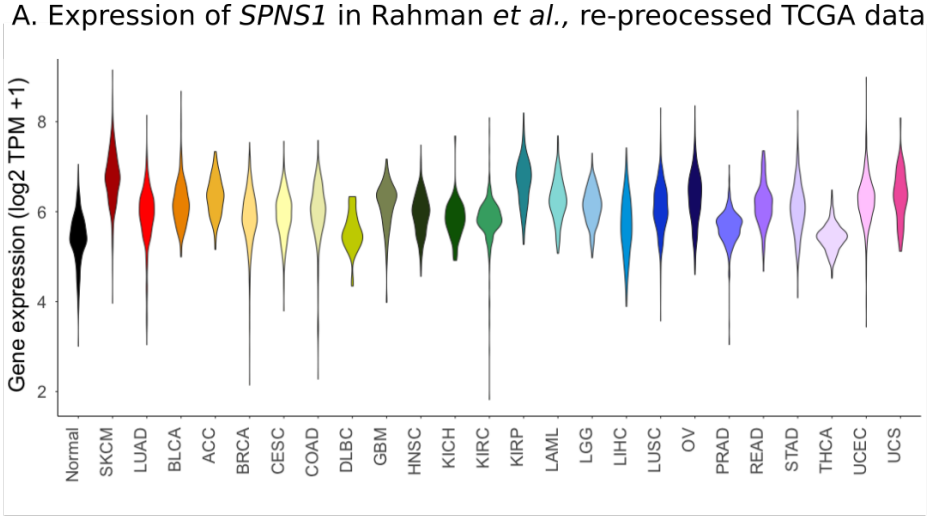


B. Expression of *SPNS1*, The human protein atlas, accessed 15 November 2019

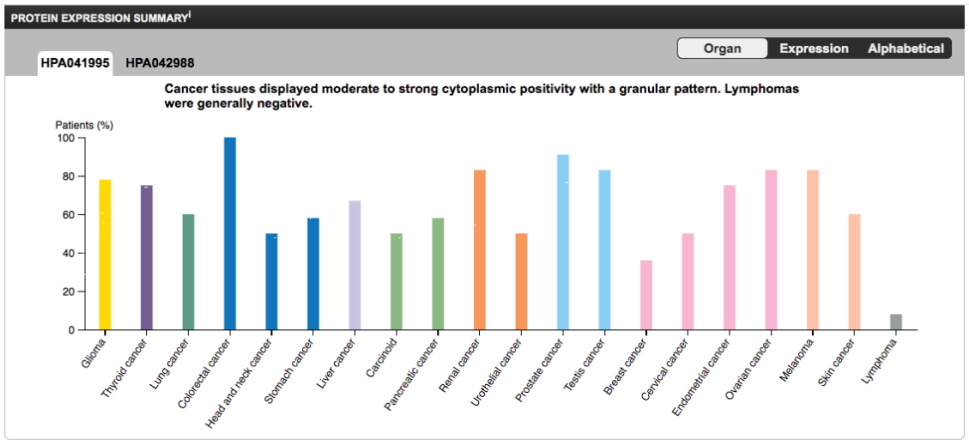


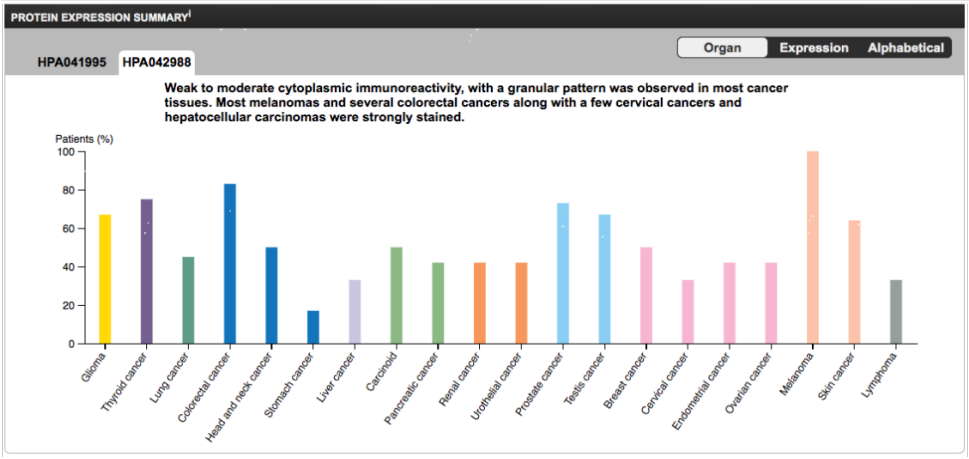
A.29 Expression of *SPNS1* in cancer tissues

A. Expression of *SPNS1* in Rahman *et al.*, re-preocessed TCGA data



B. Expression of SPNS1 in cancer tissues, The human protein atlas, accessed 15 November 2019





Appendix B

Supplementary data in electronic format

The following supplementary files are provided in electronic format as well as on a CD.

B.1 Targeted capture bait design

This file contains the genomic location of all positions aimed to be captured using my bait design, excluding probe group 7: HLA typing panel which has been published [433].

File name: targeted_panel_covered.bed
Link to file: `https://drive.google.com/open?id=1HylHmk2wZzRJH4k9ds-VFfDsTZQJKv0A`

B.2 ASCAT SNP distribution

This file shows the distribution of SNPs used by ASCAT, split by chromosome, as well as the genomic location of all genes in my targeted capture bait panel.

File name: Panel_SNPs_and_genes_across_chromosomes_detailed.pdf
Link to file: `https://drive.google.com/open?id=1LhfQOrpflMrUxhesZ71Q4SAEL8rBgzeC`

B.3 Survival curves IFN-γ pathway

This file contains melanoma-specific survival curves, where patients were stratified by the presence or absence of mutation in IFN-γ pathways genes as well as cytolytic score.

File name: Alt_IFNg_pathway_genes_by_mutation_load_and_cyt_score_KM.pdf
Link to file: `https://drive.google.com/open?id=1ysEQ-yhEDhIY5fVqXNrumGegM_2C6hGz`

B.4 Pooled validation screen library design

This file contains information about the gRNAs in the pooled validation screen.
• Total_pooled_library: This sheet provides an overview of the library design
• Master_pooled_sgRNA_sequences: This sheet contains detailed information about the gRNA sequences
• Gene_rationale: This sheet outlines the rationale for gene selection

File name: Validation_screen_suppl_info.xlsx
Link to file: `https://drive.google.com/open?id=12Cczl5cshILvuj3tx2zm3GK6wt558Gpk`

B.5 MOI figures for pooled validation screen

This file contains flow cytometry scatter plots showing the MOI of each replicate and cell line used in the pooled validation screen.

File name: Val_screen_MOIs.png
Link to file: `https://drive.google.com/open?id=1VINCsWKJB6GZRHLoOH65b-TYm_o2132k`

B.6 Pooled validation screen statistics

This file contains statistics of read counts and library coverage for each sample in the pooled validation screen.

File name: Stats_combined_PDL1_validation_screen_results.txt
Link to file: `https://drive.google.com/open?id=1QWJOFzIYqrkRZnHkZCyFJTh-XtB24YwI`

Appendix C

Collaborators and datasets

C.1 Collaborators contributing to my thesis

Martin Del Castillo Velasco-Herrera: Dr. Del Castillo Velasco-Herrera was a PhD student in the group of Dr. David Adams, and is currently a postdoctoral fellow in the group of Dr. Sam Behjati. He selected gRNAs against *CD274* and *OR14A16* from his design published in [710]. These were used in Chapter 5 of my thesis to generate stable knock-out cell lines as controls for my screen.

Kerstin Haase: Dr. Haase is a postdoctoral fellow in the group of Dr. Peter van Loo, at the Francis Crick Institute. She contributed to this thesis by running the ASCAT software to generate copy number data for the Leeds melanoma cohort (Section. 2.2.5). She provided the purity, ploidy and copy number data on a segment level, from which I generated gene level copy number estimates and did all further analysis. She also provided the original code and data to generate whole genome copy number overview figures, both for the Leeds melanoma cohort and the TCGA SKCM dataset (Section. 4.3.3).

Victoria Harle: Dr. Harle is a postdoctoral fellow in the group of Dr. David Adams. She helped out with one of the the small-scale validation experiments in LCLC103H at day 14 (Section 5.5). She also carried out the assembly and production of the pooled validation gRNA library used in Section 5.6 of my thesis. Additionally, she also performed the lentiviral production of the pooled validation gRNA library, Cas9-transductions for SKMEL25, UKEMEL118C, 5637, 647V and UBLC1 cell lines and antibiotic and library titrations for all cell lines used in the validation screen. Finally, the experimental work for the pooled validation screen were done

collaboratively with Dr. Harle.

Vivek Iyer: Dr. Iyer was a principal scientist in the group of Dr. David Adams, and is now leading the Human Genetics Informatics team at Sanger. He selected heterozygous SNPs against the 28 genes selected for detailed copy number estimation of probe group 4 (Section. 2.2.2). He also interpreted and performed the ABSOLUTE power calculation to fit my dataset (Section. 2.3.2).

Marieke Kuijjer: Dr. Kuijjer is a group leader at the Centre for Molecular Medicine Norway, University of Oslo. She developed the tool Sambar, and ran it for my dataset (Section. 4.2.1). I prepared the input data, including the required mutation matrix, from which Dr. Kuijjer ran the subtyping (Appendix. A.1), and returned the output to me for further analysis.

Rashid Mamunur: Dr. Mamunur was a senior computational biologist in the group of Dr. David Adams, and is currently employed by Cambridge Epigenetics. Dr. Mamunur ran the cake pipeline to generate Mutect mutation calls, used to compare my Caveman calls to in Section. 2.3.6. Dr. Mamunur also performed the DepMap CRISPR gene dependency analysis whereby he discovered 35 genes which were significantly associated with lethality in skin cancer cell lines (Section. 3.3.1). I intersected these results with my analysis and identified a melanoma-associated genetic vulnerability mediated by the *IRF4* gene.

Ultan McDermott: Dr. McDermott was a group leader at Sanger, and is currently employed by AstraZeneca as a Chief Scientist. He designed the targeted capture baits for probe group 3 and 6, and a solid tumour gene panel from which I selected 254 genes to include in probe group 1 (Section. 2.2.2).

Julia Newton-Bishop: Prof. Newton-Bishop is a Clinician Scientist and Professor of Dermatology at the University of Leeds, where she is leading the Melanoma Research Group. Prof. Newton-Bishop and her team recruited, biopsied, and followed up all the patients in the Leeds melanoma cohort (Section C.2). Jonathan Laye, Mark Harland, Tracey Mell and Timothy Bishop were involved in the sample processing to DNA extraction from tumour and normal samples (Section. 2.2.3). Measures of sun sensitivity metrics used in Chapter 4 were generated by John Davies. Mark Iles provided me with the SNP array data, used for the genotype concordance analysis in Section. 2.3.4.

Annette Paschen: Prof. Paschen is a group leader at Universitätsklinikum Essen. Sonia Leonardelli in her team obtained flow cytometry PD-L1 expression data of the patient-derived melanoma cell line UKEMEL118C, and sent the cell line. After I confirmed its high PD-L1 expression, I decided to include this cell line in the pooled validation screen in Chapter 5 of my thesis.

Marco Ranzani: Dr. Ranzani provided me with RNA-sequencing data in FPKM-format of his in-house collection of melanoma cell lines, from which C092 was selected for the screen (Section 5.3.1).

Kosuke Yusa: Prof. Yusa was a former group leader at Sanger, and is currently a Professor at Kyoto University. He developed and gifted the human genome-wide gRNA library human v.1.1, used in Section 5.4 of my thesis.

Sanger core facilities: The Sanger CGP and sequencing core facilities performed the sample processing from second round of PCR to sequencing of my CRISPR-Cas9 screen samples (Section 5.4). Additionally, they performed all steps from DNA extraction to sequencing of the samples part of the pooled validation screen (Section 5.6). The Sanger flow cytometry core facility helped with the sorting of samples for my CRISPR-Cas9 screen (Section 5.4), and performed all sorting steps for the pooled validation screen samples (Section 5.6).

Sanger pipelines teams: The pipelines teams at Sanger processed all DNA from the Leeds melanoma cohort, including targeted capture pull-down, library generation, sequencing and alignment (Section. 2.2.3).

C.2 Datasets used in my thesis

Leeds melanoma cohort: This is the main dataset used throughout part I of my thesis. It originally comprises 2182 patients recruited from the Northern England or Yorkshire region [838], from which 524 primary tumours were successfully sequenced as part of my project. Generally throughout my thesis these 524 primary tumours are referred to as the Leeds melanoma cohort. Additionally, transcriptomes from 700 primary tumours of this cohort were sequenced as part of another study [833], and this data was used in Chapter 5 for the *SPNS1* survival analysis. 318 tumours have been profiled to yield both mutation data and transcriptomic data, which was used in Section 4.8 of my thesis. This cohort and corresponding clinical data were collected

by Prof. Julia Newton-Bishop's group at Leeds University, and the patients have been followed up for over 15 years.

TCGA SKCM: The TCGA SKCM dataset, also known as the Skin Cutaneous Melanoma (TCGA, PanCancer Atlas) dataset is the largest and most widely used dataset comprising comprehensive genetic information about melanoma. It has been used throughout Part I of my thesis, as a comparison to my analysis of the Leeds melanoma dataset.

TCGA Pan-cancer: The TCGA Pan-cancer cohort are a collective dataset comprising 33 major tumour types. This dataset was used in Chapter 5 of my thesis to compare variants found in the Leeds melanoma cohort with those found in other cancer types as well as melanoma.

TCGA BLCA and NSCLC Pan-Lung: These TCGA datasets were used in Chapter 4 of my thesis, to look at alterations in the *SPNS1* gene, using the cBioPortal platform [480, 481]. They were also used for a survival analysis using the UCSC Xena platform [834].

Primary melanocyte expression dataset: This unpublished dataset consisting of gene expression data from primary melanocytes extracted from human foreskin, generated by Dr. Kevin Brown, Senior Investigator at the NIH. This dataset was used to assess *IRF4* expression in melanocytes in Section 3.3.1 of my thesis.

Broad DepMap dataset: This is an extensive CRISPR-Cas9 dropout screen dataset, where 342 cancer cell lines of varying tissue types have been studied. The data was downloaded and reprocessed at Sanger by Dr. Francesco Iorio, to generate lethality scores for each condition. This data was then used in Section 3.3.1 of my thesis.