Chapter 1

Introduction

1.1 From peas to man: A (very) brief history of genetics

The study of genetics started more than a century ago, when the Austrian monk Gregor Mendel performed his classic experiments with common pea plants and deduced the basic principles of heredity, which he described in a paper written in German in 1866 [1]. Mendel, after careful characterisation of several generations of pea plants, hypothesised that physical characteristics were transmitted from parents to offspring in packets of information he termed 'elements'. The material responsible for carrying these elements was identified about a decade later, albeit unwittingly, by the German biologist Walther Flemming while he was studying cell division [2]. This fibrous material, which he termed 'chromatin', would later be shown to be formed of chromosomes [2], while Mendel's elements would form the basis for the definition of gene [3]. These discoveries, whose importance was unrecognised at the time, would launch a scientific revolution that would completely change the way researchers studied the laws of heredity over the coming century.

The identification of chromosomes as the carriers of heredity in the turn of the nineteenth century by Walter Sutton [4] and Theodor Boveri [5], along with the observation that chromosomes contain nucleic acids by Oskar Hertwig and Albert von Kölliker (reviewed in [6]), sparked huge interest in the scientific community to identify the mechanism by which chromosomes could preserve and transmit information from generation to generation. In 1953, James Watson and Francis Crick proposed that specific nucleotide base pairing in deoxyribonucleic acid (DNA) could provide a mechanism

for copying and transmitting information, making it the most likely gene carrier [7]. A few years later, scientists had managed to decipher the genetic code and efforts began to sequence genes and whole genomes.

Between 1977 and 1982, researchers were able to determine the complete genome sequences of the ΦX174 and λ bacteriophages [8, 9], the simian virus SV40 [10] and the human mitochondrion [11]. The success with these relatively small genomes prompted scientists to initiate efforts to sequence the entire human genome, and by 1990 the Human Genome Project had been launched with the participation of numerous research institutions around the globe [12]. This tremendous enterprise yielded the first draft of the human genome (covering about 94% of the total sequence) in 2001 [12], followed by the publication of the gold standard reference covering ~99% of the euchromatic sequence in 2004 [13].

In the few years since the publication of the human genome, sequencing technologies have evolved at an accelerated pace, bringing down the cost and the amount of human resources necessary for generating high-quality genome sequences [14]. As such, the last decade has seen a dramatic increase in the number of organisms that have had their whole genome sequenced, as well as the number of projects attempting to catalogue and understand all existing natural human variation. Examples of these are The 1000 Genomes Project [15], the UK10K Project [16] and the National Heart, Lung and Blood Institute Grand Opportunity Exome Sequencing Project (NHLBI GO ESP) [17]. Other endeavours have focused on charting the different elements encoded by and influencing regulation of the human genome, examples of which are The Encyclopaedia Of DNA Elements (ENCODE) [18], The National Institutes of Health Roadmap Epigenomics Mapping Consortium [19] and BLUEPRINT [20]. All of these interdisciplinary, multinational collaborations have allowed a much deeper understanding of the relationship between the human genome and the regulation of important biological processes, as well as the essential role it plays in health and disease.

1.2 The human genome: Its structure and gene content

In 1956, the long-standing quest to determine the number of chromosomes in human diploid cells was finally settled by two studies authored by the laboratories of Albert Levan and John Hamerton [21, 22]. These researchers were able to improve the karyotyping

techniques of the time by adding steps to treat cells with colchicine and an hypotonic solution [23], and finally determined that diploid cells possess 46 chromosomes including one sex-determining pair. Other pre-sequencing era genome measurements included the estimation of the size of the whole human genome, which was approximated to be between 3000 and 3200 megabases (Mb) by physical maps [24], and guesses on the number of genes it contains, which were generally in the range of 50,000-100,000 [25].

The delivery and initial analysis of the human genome helped refine these measurements and allowed an unprecedented description of our genomic landscape. The analysis revealed that the human genome showed marked variation in the distribution of features such as CpG islands and transposable elements, that it had many more segmental duplications than other previously sequenced eukaryotes, and that it seemed to contain only between 30,000 and 40,000 genes, among other remarks [12]. Many of these observations were confirmed with subsequent analyses and studies, whereas some estimates were further refined with the release of the gold standard reference genome and advances in assembly techniques and bioinformatic algorithms. For example, the amount of segmental duplication was estimated at ~5.3% of the euchromatic sequence as compared to a lower bound of 3.6% in the draft sequence, and the number of protein-coding genes was brought even lower to 20,000-25,000 [13].

The new estimated upper bound on the number of genes left many scientists perplexed. It was comparable to that of the simple nematode *Caenorhabditis elegans* [26], an organism composed of about 1,000 cells, and much lower than those in the single-celled pathogen *Trichomonas vaginalis* [27] and in common crops such as maize [28] and wheat [29]. Where did the perceived complexity in humans come from, if not from our protein-coding genes? Part of the explanation came from experiments that showed that the great majority of human genes are capable of generating more than one protein via alternative splicing [30], which less complex organisms, such as *T. vaginalis*, cannot do as their genes lack introns [27]. Proteins generated by human genes were also estimated to participate in many more interactions with each other than those of organisms of lower perceived complexity, such as *C. elegans* and *Drosophila melanogaster*, with the size of each species' interactome seeming to correlate better with complexity [31]. Other factors that may contribute to explain organism complexity relate to the amount of gene duplication in species such as *T. vaginalis*, maize and wheat [27, 28, 32], the intricate patterns of post-translational protein modification in humans [33], and gene expression regulatory mechanisms [34, 35].

However, protein-coding genes, their regulation, interactions and modifications offer

only a small part of the answer. The realisation that the number of genes in an organism and its perceived complexity did not seem to correlate pointed to additional factors, hidden in the non-protein-coding genome, contributing at least as importantly as proteins to biological function. In fact, John Mattick soon noticed that the proportion of protein-coding DNA in an organism decreases as a function of its complexity, with prokaryotes having less than 25% of non-coding sequences and humans having approximately 98.5% [36]. Additionally, it was estimated that the amount of genome that is transcribed into ribonucleic acid (RNA) greatly exceeds the amount that is translated into proteins, significantly adding weight to this hypothesis [37]. Given these observations, efforts began to identify and study the diverse repertoire of biological elements encoded by our genome, beyond protein-coding genes, that participate in structural, regulatory and functional tasks.

Undoubtedly, the ENCODE project has been the largest collaboration, launched in 2004, with the aim of providing a complete catalogue of all the functional components encoded by the human genome [38]. In 2012, the Consortium published their extensive set of analyses in a series of 30 papers published in *Nature*, *Genome Biology* and *Genome Research* (for an overview of these publications, see the Nature ENCODE explorer at `http://www.nature.com/encode/`). By using different approaches such as chromatin immunoprecipitation (ChIP-seq), RNA sequencing (RNA-seq) and mass spectrometry in more than 140 human cell lines, ENCODE was able to assign a biochemical function to the vast majority of the human genome [18]. About 62% of the genome was found to be covered by different RNA-encoding elements, the majority in intronic regions, and included long non-coding (lnc)RNAs and small RNAs, comprising small nuclear (sn)RNAs, small nucleolar (sno)RNAs, micro (mi)RNAs and transfer (t)RNAs [18, 39]. Additionally, about 56% of the genome was found to be enriched for histone modifications, and smaller regions were classified as open chromatin (15%) or transcription-factor-binding sites (8%) [18]. ENCODE also described regions of DNA methylation, long-range physical genomic interactions, and gave an estimate of the amount of human genome under purifying selection (3-8%), among other remarks [18].

Other projects have been initiated with the aim of understanding and cataloguing natural human variation. The 1000 Genomes Project was the first such undertaking, launched in 2008, that sequenced large numbers of individuals in order to discover and haplotype all forms of DNA polymorphisms present in humans [40]. The 1000 Genomes Project Consortium published the results of their pilot phase in 2010, followed two years later by a detailed description of the findings in the genomes of 1,092 individuals from

14 populations [15, 40]. So, although humans are thought to share about 99.9% of their genetic sequence with each other (reviewed in [41]), the Consortium reported around 38 million single nucleotide polymorphisms (SNPs), 1.4 million short insertions or deletions (indels) and more than 14,000 larger structural variants which showed marked differences in allele distribution across populations [15]. In 2012, the NHLBI GO ESP published an analysis of more than 15,000 genes in 6,515 individuals, in which they reported that individuals carry an excess of rare variants, thought to have arisen recently and attributable to explosive population growth [17]. These efforts have provided us with a deeper understanding of the history and migration patterns of human populations, the burden of rare variants typically carried by any single individual, regions of the genome that are essential in determining phenotypic characteristics, and genes that might play an important role in susceptibility to disease.

With other projects currently on-going with the aim of cataloguing all human variation and elucidating the biological function of these elements [16, 19, 20], our understanding of the human genome is likely to increase dramatically in the next years. However, at this moment, the ease of rapidly generating whole-genome or targeted sequencing of hundreds of individuals, coupled with the continued development of bioinformatic tools and the ability to search tens of in-depth catalogues of human variation and diverse encoded DNA elements, has substantially helped the identification of the genomic regions that play an important role in health and disease.

The present dissertation principally deals with the use of the above-mentioned sequencing methodologies, variation catalogues and bioinformatic tools in order to search for melanoma-predisposing genome regions in a cohort of high-density melanoma families from the United Kingdom (UK), The Netherlands and Australia. In the next section, I discuss the basis of cancer aetiology in humans, with special emphasis on the known genetic components of cancer development as it is the main focus of my work. The remaining sections provide a detailed description of melanoma and its risk factors, as it is the phenotype I have studied in the course of the last four years. Finally, I conclude this introduction with a discussion of the unanswered questions in melanoma genetics and the methodology I followed in an effort to answer them.

1.3 The basis of cancer aetiology in humans

Cancer is a disease of the genome [42]. As early as 1890, David von Hansemann examined cancer cells under a microscope and noticed striking aberrations in their chromosomes,

such as multipolar mitoses, breakage, asymmetry and altered chromosomal dosage [43, 44]. This observation made him postulate, contrary to beliefs at the time, that cancer was a disease of the internal hereditary material of the cell [44]. Following on these observations, Boveri made yet another landmark contribution to genetics: By studying sea urchin eggs, he hypothesised that tumours might originate as a consequence of cells passing abnormal numbers of chromosomes to their daughters [45, 46]. With their work, von Hansemann and Boveri laid the theory of the genetic origin of cancer.

Other experiments followed to show that cancer was caused by the progressive accumulation of somatic alterations in genomes [47]. By 1930, Katsusaburo Yamagiwa and Koichi Ichikawa had reported that exposing rabbits to coal tar could produce carcinomas with metastatic potential [48], and Ernest Kennaway had identified individual chemical compounds that generated tumours in mice and rabbits [49]. These chemical compounds were later shown to bind covalently to DNA and induce genetic mutations [50–53], thus providing the molecular link between exposure to certain chemicals and cancer development.

This hypothesis could explain well an observation that had been made in the 1770s. The British surgeon Percivall Pott had described that chimney sweeps, who were exposed to coal tar when working, were particularly likely to develop cancer of the scrotum, and ascribed it to be "a disease brought on them by their occupation" [54]. Since then, other industries in which workers face increased cancer risks have been described, such as mining, chemical manufacturing and iron and steel founding [55]. These observations established an important role for environmental risk factors in cancer development.

1.3.1 Environmental risk factors and their role in cancer development

Environmental risk factors, which include all non-genetic aspects such as diet, lifestyle and infectious agents, cause the majority of human cancers [56]. More than 100 agents have been classified as 'carcinogenic to humans' by the International Agency for Research on Cancer (IARC), which belongs to the World Health Organisation [57]. These agents, which are known as carcinogens, can have genotoxic or non-genotoxic mechanisms of action, or both. Genotoxic carcinogens can directly cause DNA damage, such as the chemicals in coal tar, whereas non-genotoxic carcinogens can have diverse modes of action and can alter diverse processes such as immune suppression, endocrine modification and inflammatory responses [58]. In this subsection, I review several of the most

established or commonly found carcinogens, their modes of action and the malignancies with which they have been causally associated.

1.3.1.1 Asbestos

Asbestos encompasses a series of naturally occurring silicate mineral fibres with desirable industrial properties such as tensile strength and resistance to heat, and that therefore are of commercial interest [59]. Asbestos fibres can be classified in two types: chrysotile, which is the most commonly used type and is flexible and easily breakable, and the amphiboles, which are rigid, sharp and durable [59]. The negative health effects of asbestos exposure have been known for more than a century, when in 1899, Montague Murray diagnosed the first fatal case of asbestosis arising from occupational exposure [60]. However, proper legislation regarding dust control in North American mines was not enacted until 1971 [61]. Because of this, asbestos is still present in buildings, and given the long latency period between initial exposure and disease manifestation (15 to 40 years), it still continues to pose a major health challenge [59]. Distinct modes of action have been described, with some sources noting its genotoxic effects, causing DNA base oxidation, breakage, mutations and deletions and chromosomal aneuploidy (reviewed in [62]) that alter processes such as cell proliferation, cell death and inflammation [63], and other sources describing non-genotoxic effects such as the generation of reactive oxygen and nitrogen species (ROS and RNS, respectively) and also mitogenic and cytotoxic consequences [62, 63]. Given these different effects, which might depend on the studied fibre type, lung clearance, genetics and other characteristics [59], the IARC has included both mechanisms of action in its classification of asbestos as carcinogenic to humans.

Exposure to asbestos can be occupational, as in the case of factory workers, or environmental, in the case of individuals living or working in communities near asbestos mining plants or buildings with a high asbestos content. Workers are mainly at risk of inhaling asbestos fibres when processing materials such as talc or vermiculite that might be contaminated [61], and studies have established that they have a much higher incidence of malignancies such as mesothelioma and lung cancers, among other diseases [61, 64, 65]. The risk for lung cancer can be modified by other factors such as smoking and persistent inflammation (reviewed in [62]). There is also evidence that environmentally-exposed individuals can develop mesothelioma, albeit at much lower frequencies [64, 66, 67].

1.3.1.2 Alcoholic beverages

Alcoholic beverages are pervasive in our society. In surveys conducted in 2012, more than half of people aged 18 or over reported drinking alcohol in the last month in the United States (US) [68], and more than half of adults reported consuming alcohol in the last week in the UK [69]. More than a century ago, French pathologists noticed an increase in the incidence of oesophageal cancer in absinthe drinkers [70], and since then, numerous studies have been conducted that show an association between alcohol consumption and cancers of the oral cavity, pharynx, larynx and oesophagus (reviewed in [71]). Finally, in 1988, the IARC concluded there was sufficient evidence for the carcinogenicity of alcoholic beverages in humans, and that malignant tumours of the oral cavity, larynx, pharynx, oesophagus and liver can be causally associated with their consumption [72].

Ethanol metabolism has been identified as the major mechanism by which alcoholic beverages can cause cancer. Ethanol is metabolised to acetaldehyde by the alcohol dehydrogenase (ADH) enzymes and cytochrome P450, family 2, subfamily E, polypeptide 1 (CYP2E1), and is further converted to acetate by aldehyde dehydrogenases [73]. This in turn causes genotoxic and non-genotoxic effects: Acetaldehyde can bind to DNA and form stable adducts [74], interfering with DNA synthesis and repair, and its production can generate ROS and an activation of other carcinogens present in the environment such as tobacco smoke (reviewed in [73]). Individuals that carry a certain polymorphism in one of the aldehyde dehydrogenase genes, *ALDH2*, have a much less efficient enzyme and thus accumulate higher levels of acetaldehyde, and have therefore been identified as a high-risk group for the development of oesophageal cancer (reviewed in [73]). Other mechanisms by which alcohol may cause cancer relate to an increase in the levels of hormones such as oestrogen, which increase breast cancer risk [75], an increase in the permeability of mouth and throat cells to other carcinogens [76], and a decrease in folate levels [77].

The proportion of all cancer cases and deaths attributable to alcohol intake worldwide has been estimated to be 3.6% and 3.5%, respectively [78]. Therefore, diminishing alcohol consumption has been highlighted as an important and generally underemphasised strategy for cancer prevention when compared to other preventive programs focusing on tobacco usage and genetic screening, among others [79].

1.3.1.3 Coal tar and soot

As discussed previously, coal tar and soot present in chimneys were the first occupational agents to be associated with an increase in cancer risk. In the years since Pott described the link between occupational exposure to coal and scrotal cancer [54], many other reports describing an association between skin cancers and soot were subsequently published (reviewed in [80]). As a result of his observations, Pott recommended that chimney sweeps take a daily bath, and his suggestion was so successful that it managed to greatly reduce the incidence of scrotal cancer in workers that followed his advice compared to those that did not [81, 82]. This evidence, and numerous studies that were conducted subsequently, led the IARC to conclude in 1973 that coal and soot are carcinogenic to humans.

The mechanism of carcinogenicity of coal has been studied extensively. In 1936, Alfred Winterstein reported that he could isolate benzo[*a*]pyrene (BP) from coal tar boiling at high temperature (Winterstein A. Festschrift. Basel: Emil Barell; 1936; quoted by [83]), and subsequent experiments showed that BP applied on mouse skin, even at low doses, was highly carcinogenic and generally produced squamous cell carcinomas [84]. Other researchers were later able to dissect the metabolism of BP: Upon exposure, BP is converted to BP-7,8 epoxide and then further hydrolysed and metabolised, in steps involving the microsomal epoxide hydrolase and cytochrome P450 enzymes, to the carcinogenic diol-epoxide 2 (DE 2) [85]. This species is highly reactive and can bind DNA [85, 86], and it has been shown by several studies that it can cause genetic mutations in both prokaryotes and eukaryotes (reviewed in [87]).

Exposure to coal tar and soot has been associated principally with higher incidences of lung and skin cancer, but limited evidence also exists linking it to bladder cancer [57].

1.3.1.4 Tobacco smoke

As early as 1930, tobacco smoking was proposed to be the underlying cause of the phenomenal rise in lung cancer that was seen after the end of the First World War, a suspicion that was confirmed over the next 20 years as more studies were carried out investigating lung cancer aetiology [88–90]. In numerous experimental systems in which mice, rats, Syrian hamsters, rabbits and dogs were exposed to mixtures of cigarette smoke and air, significant increases in lung tumours as well as emphysema and other cancers were observed (reviewed in [91]). Additionally, cigarette-smoke condensate was found to increase the incidence of tumours when applied to mouse and rabbit skin, as

well as acting as a potential co-carcinogen when used in conjunction with other agents (reviewed in [91]). Some of these studies, in conjunction with others, were considered by the IARC when it classified tobacco smoking as a confirmed carcinogen in 1986 [57].

Cigarette smoke contains more than 60 well-known carcinogens, most of which require metabolic activation upon exposure, a step that generally involves the cytochrome P450 enzymes [92]. In addition to BP, which has also been found as a carcinogenic agent in coal tar, other compounds found in tobacco smoke that can form DNA adducts are N-nitrosodimethylamine, N´-nitrosonornicotine and ethylene oxide, among others [92]. Accordingly, in human cell line experiments, hotspots for the formation of BP-DNA adducts have been found in genes important in cancer progression such as Kirsten rat sarcoma (RAS) viral oncogene homolog (*KRAS*) and tumour protein p53 (*TP53*) [93, 94]. In the case of *KRAS*, adducts were found to preferentially affect important codons for kinase activation, and in the case of *TP53*, they matched the mutational signature observed in human lung cancers, implicating the genotoxicity of these carcinogens as a highly likely contributor to their aetiology. Other studies reporting higher levels of adduct formation resulting from exposure to other tobacco carcinogens support the causal relationship between smoking and cancer (reviewed in [92]).

Not only lung cancer has been causally associated with tobacco smoking. According to the IARC, there is sufficient evidence to support a causal role for smoking in the aetiology of many different types of malignancies, including myeloid leukaemia, and colorectal, liver, oesophageal, stomach and bladder cancers, among others. Second-hand smoking has also been linked to lung, laryngeal and pharyngeal cancers (reviewed in [57]).

Tobacco smoking has been identified as the main cause of cancer-related deaths worldwide, being responsible for approximately 16% of all cancers in developed countries (Stewart BW, Kleihues P. World Cancer Report. Lyon, France: IARCPress; 2003; quoted by [95]), and thus has influenced policymaking (for examples, see [96, 97]) and sparked numerous campaigns around the world to reduce its prevalence (for examples, see refs. [98, 99]).

1.3.1.5 Solar radiation

The link between exposure to solar radiation and skin cancer was suspected as far back as 1894, when Paul Unna described what he referred to as ‘seaman’s skin’: chronically sun-exposed skin that presented hyperkeratosis and squamous cell carcinomas (Unna, PG. Die Histopathologie der Hautkrankheiten. Berlin: August Hirschwald; 1894; quoted

by [100, 101]). Experiments in diverse model organisms over the next 70 years helped clarify the carcinogenic effects of ultraviolet (UV) light, one of the radiation types that the Sun emits (reviewed in [100]). In 1992, the IARC analysed data from numerous patients and animal studies and concluded that there was sufficient evidence to support a causal role for solar radiation in the development of cutaneous malignant melanoma and non-melanocytic skin cancer in humans [102].

Solar radiation is the combination of UV radiation and visible light that reaches the Earth's surface [103]. Solar UV radiation that reaches Earth is mainly composed of two wavelengths: UVA, which comprises about 95%, and UVB, which contributes the rest [103]. UV radiation seems to have both genotoxic and non-genotoxic effects: it can create cyclobutane pyrimidine dimers (CPDs) and (6,4)-photoproducts (6-4PPs) between adjacent pyrimidine bases in the same DNA strand [104] (Fig. 1.1), and it can also cause other systemic effects such as activation of the cell survival nuclear factor (NF)-κB signalling pathway [105] and generation of ROS [106], which appear to be independent of DNA damage. Although thymine bases in CPDs are not especially mutagenic [107], cytosine bases in these structures are chemically unstable and usually deaminate to generate uracil bases, which may then be incorrectly paired with adenine during replication inducing a C-to-T transition. The same outcome happens if the CPD occurs in a 5-methylcytosine (mC)-containing pyrimidine site, as is frequently the case for solar-induced DNA lesions, because they can rapidly deaminate to form thymine (reviewed in [108]). Therefore, the induction of C-to-T transitions at dipyrimidine sites is the signature consequence resulting from UV radiation.

The incidence rate for malignant melanoma, one of the cancer types for which UV radiation is an important risk factor, has more than quadrupled in the last 30 years in the UK [110]. This dramatic increase has been attributed to lifestyle factors such as the popularity of holidays in lower latitudes and the surge in sunbed usage [111], and has led to increasing numbers of campaigns targeting unsafe exposure to UV radiation (for examples, see [112–114]).

UV radiation, as one of the major causes of malignant melanoma, is revisited in Section 1.6.

1.3.1.6 Other carcinogenic agents

In addition to the carcinogens discussed above, there are about 100 other agents that have been classified in this category by the IARC. These can be chemicals (*e.g.*, aflatoxins or benzene), occupations (*e.g.* aluminium production or painting), metals (*e.g.* arsenic

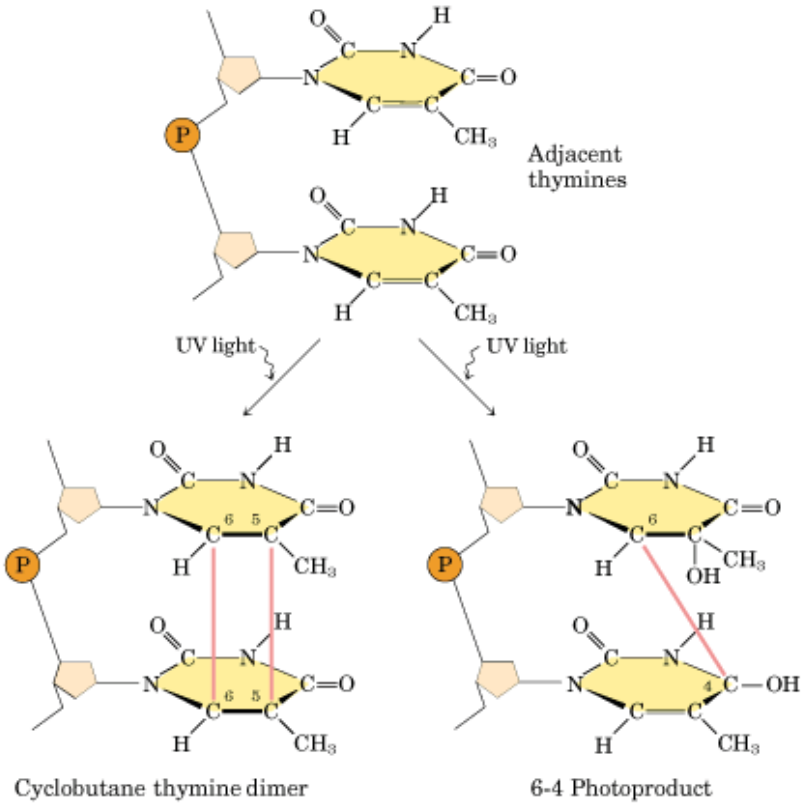

Figure 1.1: **Types of DNA damage caused by UV radiation.** An overview of the structure changes induced by UV radiation in adjacent pyrimidine bases, in this example, thymines are shown. Phosphate and sugar groups forming the DNA backbone are shown in orange and light orange, respectively. In the case of CPDs, the double bonds between carbons 5 and 6 become saturated, forming a four-membered ring. In 6-4PPs, a bond is formed between carbons 6 and 4 of two adjacent pyrimidines. Figure taken from ref. [109]. Copyright 2013 by W.H. Freeman and Company. Used with permission of the publisher.

or nickel compounds), dusts and fibres (*e.g.* leather or wood dust), radiation sources (*e.g.* X or gamma radiation), biological agents (*e.g.* hepatitis B or C or HIV viruses), personal habits (*e.g.* consuming areca nuts or Chinese-style salted fish), or drugs (*e.g.* cyclosporine or tamoxifen) [57]. The value of identifying these carcinogenic agents lies in actionable measures, such as educational campaigns and prevention programmes, aimed at diminishing exposure to them. For example, encouraging results can be seen from efforts of tobacco control campaigns and consequential reductions in lung cancer incidence and mortality in the US [115, 116].

Exposure to carcinogens is an established cause of malignancy, but genetic predisposition also plays an important role in cancer aetiology. In the next subsection, I discuss

the genetic factors that have been identified in cancer predisposition, the importance of familial studies and the biological insight we have gained from such analyses.

1.3.2 Genetic risk factors and their influence on cancer development

In 1866, the same year in which Mendel published the basic principles of heredity, French surgeon Paul Broca published perhaps what is the first report on the existence of familial predisposition to malignancy. He constructed the pedigree of his wife's family, who suffered from early-onset breast cancer, and found fifteen cases of breast, liver and uterine cancer spread across four generations of women (Broca P. Traité des tumeurs. Paris: P. Asselin; 1866; pedigree reproduced by [117]) (Fig. 1.2). Then, in 1913, the American pathologist Aldred Warthin published a study of 3,600 cancer cases that had been examined in the University of Michigan in the period from 1895 to 1913. Of all the carcinoma cases where detailed family history was available, he determined that as many as 15% had a familial history of the disease, and thus supported the idea that predisposition to malignancy could be inherited [118]. One of the families in his original study has been followed now for more than a century, and the causes for its cancer-prone phenotype have now been defined molecularly, contributing enormously to our knowledge of the biological processes underlying cancer development [119]. With his work, Warthin provided solid evidence for the importance of genetic factors in cancer aetiology, and thus has been referred to as "the father of cancer genetics" [120].

Studies of other families, in which many different types of cancer seemed to cluster, followed Broca's and Warthin's reports (reviewed in [121]), and epidemiological studies carried out several decades later supported an important role for genetic predisposition to cancer [122]. However, with the exception of these rare malignancies, the origin of common cancers was still considered to have predominantly environmental causes throughout almost all of the twentieth century [123].

Germline transmission of predisposition to certain malignancies, such as the childhood cancer retinoblastoma and syndromes of multiple endocrine neoplasms, had been recognised in the 1980s [124–126]. Several years earlier, scientists had started to regard cancer as a multi-stage process, having a fixed number of rate-limiting steps that should be overcome before malignancy could progress [121]. The American geneticist Alfred Knudson, after meticulously studying 48 cases of retinoblastoma, formulated in 1971 his groundbreaking two-hit hypothesis: he postulated that in the hereditary form of the

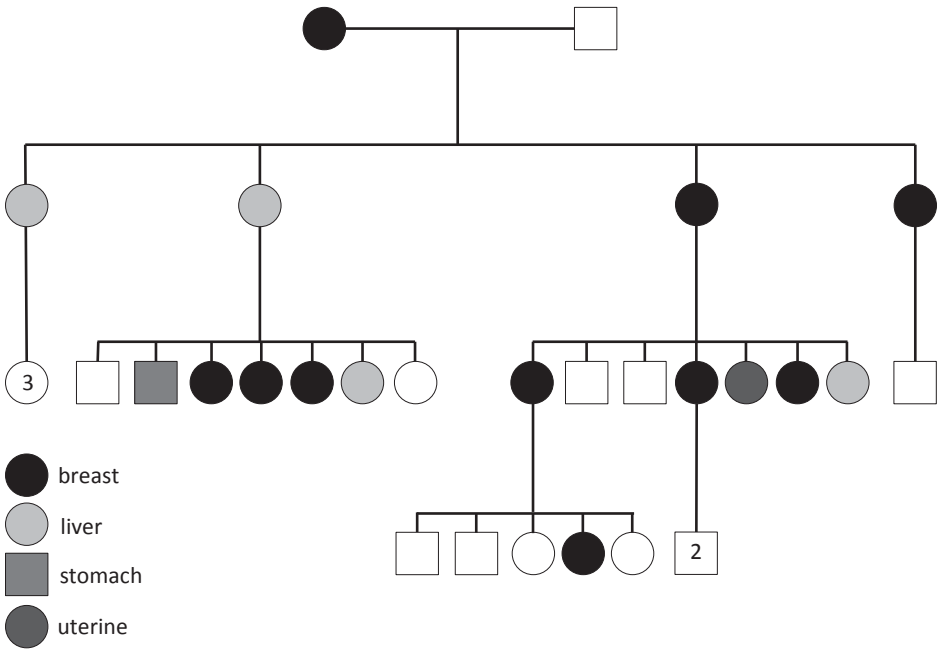

Figure 1.2: **Paul Broca's pedigree evidences familial predisposition to cancer.** This pedigree, possibly the first report on familial cancer, shows fifteen cases of different malignancies spread across four generations of women. Circles represent females, squares represent males. Numbers in symbols refer to numbers of unaffected offspring. Pedigree adapted from ref. [117].

disease, one mutation was inherited while the other happened somatically, whereas in the sporadic form both happened somatically [127]. Knudson's hypothesis was confirmed in 1986 with the successful isolation of the gene responsible for retinoblastoma [128], now known as *RB1*, and the characterisation of its aberrations in retinoblastoma DNA samples.

Since these landmark discoveries, many other genes have been discovered that, when mutated, predispose their carrier to the development of malignancies. Almost all of these have been categorised as tumour suppressor genes, and echo Knudson's two-hit hypothesis in the mechanism by which they contribute to cancer development [129]. Tumour-suppressor genes normally have functions inhibiting cellular growth or genomic instability, and thus inactivation of both copies of the gene in a cell leads to the loss of these protective functions. Genes that contribute positively to cell proliferation can also be affected by germline mutations, although this happens much less frequently. These genes are referred to as proto-oncogenes, and inherited mutations render the encoded proteins constitutively active [129], so a single copy of the gene harbouring the mutation

is sufficient to promote tumourigenesis. Because of their modes of action, tumour suppressor and proto-oncogenes show distinct patterns of germline and somatic alterations, with tumour suppressors displaying truncating or disruptive mutations distributed throughout the encoded protein and proto-oncogenes displaying clusters of mutations in key residues that activate them.

The discoveries mentioned above, along with other genetic models in epidemiological studies of cancer clustering, supported the hypothesis that a substantial proportion of cancer incidence might be accounted for by genetic effects [123], and thus obligated researchers to revisit the theory for the origins of common cancer. The importance of genetic effects in the aetiology of common cancers was further supported, in 1990, by the discovery of a genomic region linked to early-onset familial breast cancer [130]. Four years later, breast cancer 1 (*BRCA1*), the gene responsible for the disease, was identified by positional cloning methods [131]. It is now recognised that between 5 and 10% of all cancers are inherited, and that the interaction between genetic factors, or gene-environment interactions, might play an important role in a further 15% [132]. These numbers give us an idea of the importance of genetic make-up in cancer aetiology.

However, identifying cancer susceptibility genes is not an easy task. Although a strong genetic component can be suspected in families where there are multiple affected individuals across generations or individuals that present with multiple primary cancers or with an early age of onset, various confounding factors can greatly complicate its isolation. For example, the allele might show incomplete penetrance, in which case some carriers will not be affected by the disease. It can also be the case that sporadic forms of the disease, termed phenocopies, arise in non-carrier family members. In these situations, a genotype comparison between affected and unaffected family members will not be informative [133]. Despite these and other difficulties, such as small family size or uncertain family history when assessing affected pedigrees [133], hundreds of genomic regions have been found to influence cancer susceptibility over the last 30 years.

The advent of next-generation sequencing (NGS) technologies and other analysis methodologies has facilitated the discovery of cancer-predisposing genes and their relative contributions to disease development. Over 100 high-penetrance genes have been identified to date [129], and hundreds of other genomic regions have been associated with cancer risk [134], deepening our understanding of the biological processes that underlie normal cell proliferation and maintenance. In this subsection, I review the different models of cancer risk inheritance that have been proposed, the most established tumour suppressor genes and proto-oncogenes and the technologies that have been developed in

their identification.

1.3.2.1 Mendelian inheritance of elevated cancer risk

Germline variants with strong cancer-predisposing effects can be inherited in a simple Mendelian fashion, and generally underlie an early-onset of cancer and the appearance of multiple primaries [135]. High-penetrance mutations in cancer-predisposing genes can show different modes of transmission, with the majority displaying autosomal dominant inheritance but some others requiring germinal bi-allelic inactivation to promote cancer progression [136]. Syndromes that show autosomal dominant inheritance in families are generally recessive at the molecular level, with one inactivated allele being transmitted germinally from generation to generation and the other one being inactivated by subsequent somatic mutations [135].

In general, rare germline variants with large effects (for example, underlying lethal childhood cancers) tend to be purged rapidly by natural selection, because the carriers generally do not reproduce [135]. In contrast, variants that predispose to later-onset cancer, such as those in *BRCA1*, are removed more slowly from the population's gene pool. Therefore, rare alleles with large effects will tend to be younger than common alleles with smaller effects, and will generally not be associated with any polymorphisms in the genomic vicinity [135]. This distinction is important because it impacts the methodologies used for gene identification, with whole-genome or exome sequencing (WGS and WES, respectively) being better suited for the identification of rare alleles with large effects and genome-wide association studies (GWAS) being more useful in the identification of common alleles with weak effects. The genes described in this Subsection belong to the first category, and have been identified by linkage and candidate gene studies.

1.3.2.1.1 *RB1* and retinoblastoma: The prototypical example of an inherited cancer-predisposition syndrome The gene underlying almost all cases of childhood retinoblastoma, *RB1,* was identified in 1986 by Stephen Friend and colleagues [128] following Knudson's hypothesis [127]. Accordingly, the majority of inherited cases present with bilateral retinoblastoma, whereas this proportion is much lower among sporadic cases [137], evidencing the different numbers of rate-limiting mutational steps required for cancer development in mutation carriers versus non-carriers. *RB1* is the prototypical tumour suppressor gene, displaying about 90% penetrance and showing autosomal dominant inheritance in families [137]. It has since been found mutated in

other cancers, such as oesteosarcoma and small-cell lung carcinoma [138].

The biology of *RB1* has been studied extensively in the three decades since its discovery. Shortly after its isolation, two research teams found that its gene product was the target of two virus-encoded oncogenes, adenovirus E1A and SV40 large T antigen [139, 140], implicating that tumour-promoting and suppressor genes could act in the same biological pathway. Further studies established that the RB1 protein acted as an inhibitor of cell proliferation given that its over-expression caused cell-cycle arrest at the G1 phase, and, conversely, its deficiency caused fast G1 progression (reviewed in [138]). In the following years, researchers reported that RB1 was able to exert its molecular functions by inhibiting members of the E2F family of transcription factors and cyclin-dependent kinases (CDKs), which promote cell cycle entry (reviewed in [141]).

Retinoblastoma is a rare cancer, with an incidence of only about 12 cases per million young children in the US, similar to several European countries [142]. Although more than 500 distinct mutations in the *RB1* gene have been described [143], novel mutations are constantly being reported as most of the germline mutations underlying hereditary retinoblastoma arise *de novo* [144, 145].

1.3.2.1.2 *BRCA1* and *BRCA2*: High risk for breast and ovarian cancer

The first report of breast cancer might be a 1500 BC Egyptian papyrus which described tumours of the breast and their palliative treatment, although this is disputed [146, 147]. In any case, it is clear that breast cancer has been recognised for centuries. It is by far the most common cancer in women, and indeed, the most prevalent cancer in the majority of countries worldwide [148]. Broca's pedigree (Fig. 1.2) showed breast cancer segregating as an autosomal dominant trait, and accordingly, modern analyses have estimated a 96% chance that Mrs. Broca harboured a *BRCA1* or *BRCA2* mutation, and that the other cancers seen in the family represent metastatic disease [117].

Familial studies in the mid-1990s led to the isolation of *BRCA1* and *BRCA2* [131, 149], the two genes underlying the great majority of multi-case breast cancer families [150]. A meta-analysis considering several estimates of the penetrance of mutations in these genes reported that the breast and ovarian cancer risk was around 60% and 40% for *BRCA1*, and 50% and 18% for *BRCA2* mutation carriers, respectively [151]. However, although these genes confer high cancer risks and might explain a large proportion of multiple-case breast cancer families, they only account for about 2-6% of all breast cancer cases [152]. This is the reason why other genes and susceptibility regions have been intensively searched for using other methods, such as GWAS and WES.

The BRCA proteins interact with several regulatory proteins, and have an important role in the transcriptional regulation of the DNA damage response and thus in its associated cell cycle checkpoints. Upon DNA damage, BRCA1 is phosphorylated by several kinases, including ataxia telangiectasia-mutated (ATM) and ATM-related (ATR), and participates in double-strand (ds) DNA break (DSB) repair (reviewed in [153]). BRCA1 has also been shown to transcriptionally co-activate NF-κB-regulated genes, which might contribute to maintain genomic stability [154]. It is thought that *BRCA1*-haploinsufficient cells need to accumulate further genomic damage before inactivating the wild-type allele, because although tumours display BRCA1 loss-of-function, its bi-allelic inactivation in normal cells results in cellular lethality [155]. The mechanism by which *BRCA1* haploinsufficiency predisposes to breast carcinogenesis might therefore represent a modification of the two-hit hypothesis, with an additional third hit (genomic instability) required before the inactivation of the wild-type allele [155].

BRCA2 may also play an important role in DSB repair, as shown by the chromosomal breakage and abnormal mitotic exchanges observed in *BRCA2*-deficient cells (reviewed in [153]). Through its binding to the RAD51 recombinase, its synergistic interactions with P53 and its association with CDKs, BRCA2 promotes genomic stability and the response to radiation-induced DNA damage [156, 157]. Loss of the *BRCA2* wild-type allele is commonly seen in tumours from germline mutation carriers [158, 159].

Other methodologies have been used to find important genes in breast cancer susceptibility. For example, WES has identified germline mutations in the DNA repair genes Fanconi anaemia complementation group C (*FANCC*) and Bloom syndrome (*BLM*) [160], and GWAS has found a few other low-penetrance loci influencing breast cancer risk (reviewed in [161]).

1.3.2.1.3 DNA mismatch repair genes, *APC* and familial colon cancer syndromes One of the families in Warthin's original report, Family G, had several members that presented with colon, stomach and abdominal cancers across two generations [118]. This family has been followed regularly, initially by Warthin and his colleagues and subsequently by other researchers, for more than a century now (reviewed in [162]). Their cancer-prone phenotype is now known as Lynch syndrome (in honour of one of the physicians that has been studying it for several decades) or hereditary nonpolyposis colon cancer [132]. The molecular basis for their phenotype remained elusive throughout the twentieth century, but in 2000, Hai Yan and colleagues tested one member of this family and identified a T to G transversion at a splice acceptor site in the mutS homolog 2

(*MSH2*) gene [163], encoding a protein important for DNA mismatch repair [164]. Since then, mutations in any of the five DNA mismatch repair genes (*MSH2, MSH6*, mutL homolog 1 [*MLH1*], PMS1 postmeiotic segregation increased 1 [*PMS1*] and *PMS2*) have been found to underlie this syndrome [132]. Because of these defects, a type of DNA replication error called microsatellite instability is the mutational hallmark of colorectal tumours in Lynch syndrome [132]. It presents as an autosomal dominant disease with early onset, and has a typically high penetrance (~85%) [132, 164]. The development of cancer in these individuals follows the two-hit hypothesis, with the wild-type allele often found inactivated by deletion, somatic mutations or epigenetic alterations [165].

The other major subtype of hereditary colorectal cancer is known as familial adenomatous polyposis (FAP). Individuals with this disease, in contrast to Lynch syndrome, present at a young age with hundreds to thousands of adenomatous polyps in the colon and rectum, and have a greatly increased risk of developing colorectal cancer [132]. The gene underlying this syndrome was discovered by linkage analysis in 1987 [166], and is now known as adenomatous polyposis coli (*APC*). The APC protein behaves as a tumour suppressor, having a myriad of important cellular functions in cell cycle control, migration, differentiation and apoptosis that it exerts through its modulation of the integration/wingless (Wnt) signalling pathway [132]. Abrogation of APC function is not only important in tumours from individuals with hereditary disease, as mutations in *APC* are the earliest genetic alteration found in sporadic cases as well [167, 168]. Malignancy then ensues with the acquisition of additional somatic mutations in genes such as *KRAS* and *TP53* [168].

The susceptibility to colon cancer has also been linked to other genes such as mutY homolog (*MUTYH*), axin 2 (*AXIN2*), mothers against decapentaplegic homolog 4 (*SMAD4*), bone morphogenetic protein receptor, type IA (*BMPR1A*) and phosphatase and tensin homolog (*PTEN*), among others (reviewed in [169]). GWAS have also uncovered several different loci associated with the risk of colorectal cancer and adenomas (reviewed in [161]).

1.3.2.1.4 *TP53* and the multi-cancer Li-Fraumeni syndrome In 1969, Frederick Li and Joseph Fraumeni reported four families in which children and young adults presented with a high frequency of different types of malignancies, including soft-tissue sarcomas and breast cancer, and often with multiple primaries [170]. Some years later, two groups identified germline mutations in the tumour suppressor *TP53* in six families presenting with this syndrome [171, 172]. Since then, many other mutations

in *TP53*, both in coding and non-coding regions, have been described. These mutations are sometimes associated with inhibition of growth arrest, apoptosis, or transcriptional activation, and their activity can be influenced by genetic background, accumulation of further genetic alterations or other epigenetic or environmental factors (reviewed in [173]). Although the germline alterations result in *TP53* loss of function, experiments made *in vitro* suggest that missense mutations harbour oncogenic potential (reviewed in [173]). This observation could help explain why only a fraction of tumours present loss of the wild type allele (following Knudson's hypothesis) [174].

TP53 encodes an important tumour suppressor gene that integrates signals from multiple pathways and controls cell cycle progression and fate (reviewed in [175]). Mutations in this gene have been found in a fraction of virtually every sporadically-occurring malignancy [173], and given its pivotal role in maintaining genomic stability, has been referred to as "the guardian of the genome" [176].

Because not all families with Li-Fraumeni syndrome have segregating mutations in *TP53*, other genes have been searched in connection with the phenotype. Genes in the P53 pathway, such as checkpoint kinase 1 (*CHEK1*), *CHEK2,* cyclin-dependent kinase inhibitor 2A (*CDKN2A*) and *PTEN* have all been considered as candidates, however, their involvement in the aetiology of the disease remains unconvincing (reviewed in [176]).

1.3.2.1.5 *RET*: A proto-oncogene activated by germline mutations Multiple endocrine neoplasia (MEN) syndromes are a group of diseases in which endocrine glands present with two or more tumours, and in which ectopic hormone production is common (reviewed in [124]). MEN syndromes are generally divided into three main types: MEN1, MEN2 and MEN4, which are characterised by different combinations of endocrine tumours [177]. MEN2 is further divided in MEN2A, MEN2B and familial medullary thyroid carcinoma (FMTC), that have medullary thyroid carcinoma as their principal clinical characteristic [178]. In 1993, two groups studying MEN2A and FMTC families discovered germline mutations in a proto-oncogene called rearranged during transfection (*RET*), and most of these mutations affected the same cysteine residue in the protein [179, 180]. A year later, MEN2B was also attributed to germline mutations in the same gene [181]. Since then, numerous studies have shown that >95% of MEN2A patients carry mutations affecting any one of 6 different cysteine residues that result in constitutive RET signalling, and that the MEN2B phenotype is almost exclusively associated with two different methionine and alanine substitutions that result in protein

conformational changes (reviewed in [182]). In contrast to tumour suppressors, which normally require inactivation of both copies of the gene before tumourigenesis can progress, proto-oncogenes require mutation of only one copy, as it is enough to render them constitutively active. Since the discovery that certain cancer-prone families carry activating mutations in *RET*, only a handful of other genes have been found that predispose to cancer through gain-of-function mutations [129].

RET encodes a single-pass transmembrane receptor tyrosine kinase (RTK) that is highly evolutionarily conserved and participates in spermatogonial stem cell maintenance, kidney induction, neural crest cell migration among other biological processes [182]. Upon co-receptor- and ligand-binding, RET autophosphorylates and activates the mitogen-activated protein kinase (MAPK)-RAS-rapidly accelerated fibrosarcoma (RAF) and the phosphatidylinositol-3 kinase (PI3K) pathways [183], providing a biological explanation for the observed cancer-prone phenotype upon its constitutive activation. *RET* is notorious not only for being the first proto-oncogene discovered to be implicated in a cancer predisposition syndrome, but also because variable penetrance has been observed depending on the amino acid affected in germline mutation carriers (reviewed in [132]). Interestingly, *RET* germline loss-of-function mutations have also been observed in patients with Hirschsprung's disease, which is a common congenital malformation characterised by functional intestinal blockage [132]. A small number of patients have been observed with both Hirschsprung's and MEN2, as they seem to carry mutations that lead both to a decrease in *RET* expression and constitutive kinase activation (reviewed in [184]).

RET has also been implicated more broadly in diverse sporadic tumour types, such as pancreatic ductal carcinoma, invasive breast cancer and acute myeloid leukaemia (AML), among others (reviewed in [182]). Because of its mode of action and its possible implication in several cancer types, RET constitutes an important therapeutic target. Several small-molecule RTK inhibitors that target RET, such as vandetanib and cabozantinib, have been tested on patients with thyroid cancers, with varying results depending on *RET* germline status and the identity of the mutations (reviewed in [182]).

1.3.2.1.6 *CDKN2A* and *CDK4*: Melanoma risk increased by mutations in both a tumour suppressor and a proto-oncogene By 1994, the tumour suppressor CDK4 inhibitor (INK4A, also called p16, one of the proteins encoded by the *CDKN2A* locus) had emerged as an important player in human cancer development. INK4A is an inhibitor of CDK4 and CDK6, and thus prevents G1/S transition by

inhibiting RB1 hyperphosphorylation (reviewed in [185]) (Fig. 1.3a,b). Deletions in this gene had been originally reported at high frequencies in a number of cell lines, including those derived from melanoma, lung, bladder, leukaemia and brain cancers, among other malignancies [186, 187]. Only two years before, a potential melanoma predisposition locus, named MLM, had been mapped via linkage studies in melanoma-prone kindreds to the same chromosomal location, 9p21 [188]. Several studies then ensued that demonstrated that MLM and INK4A were the same gene and delineated some of the predisposing mutations, which appeared to be missense in its majority [189–192]. These mutations were later shown to impair the ability of INK4A to inhibit CDK4 and CDK6, thus providing a rationale for the cancer-predisposing syndrome observed in carriers [193] (Fig. 1.3c).

However, not all the families that showed disease linkage to chromosome 9p21 had mutations in INK4A, and furthermore, not all melanoma-prone families showed linkage to 9p21. *CDKN2A* not only encodes INK4A, but also the structurally unrelated alternate reading frame (ARF) (Fig. 1.3a). ARF stabilises P53 signalling and thus suppresses growth by blocking mouse double minute 2 homolog (MDM2), an important negative regulator of P53 (reviewed in [185]) (Fig. 1.3b). Although almost half of the *CDKN2A* germline alterations in melanoma-prone kindreds affect the shared exon 2 and thus disrupt both gene products (Fig. 1.3c), mutations affecting only ARF have been described [196, 197], and a role for this tumour suppressor in melanoma susceptibility is further supported by studies in mice (reviewed in [185]). Both INK4A and ARF behave like classic tumour suppressors, with loss of the wild type allele frequently seen in tumour cell lines and, some times, in uncultured tumours (reviewed in [198]). These observations evidence the importance of both products of the *CDKN2A* locus in melanoma susceptibility.

In some of the families that did not show linkage to the *CDKN2A* locus, germline mutations were found in the gene encoding one of the INK4A-binding partners, *CDK4*, two years later [199]. Mutations were found to cluster in the arginine at position 24, displaying the properties of a proto-oncogene. Arginine 24 was shown to be important for INK4A binding, but not for its kinase activity [199], and thus this mutation results in constitutional CDK4 activation by preventing INK4A-mediated inhibition (Fig 1.3d). These observations established that the functions of both a tumour suppressor and a proto-oncogene, that cooperate in the same pathway, can be altered by germline mutations that result in the same phenotype.

Genetic susceptibility to familial melanoma, as the subject of this dissertation, is

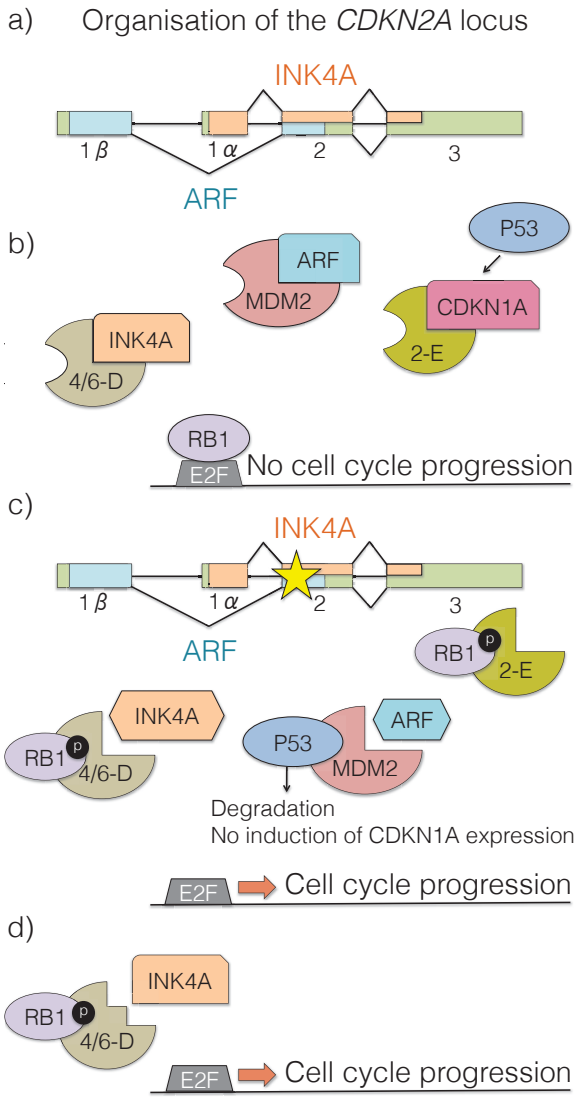

Figure 1.3: **Known melanoma susceptibility genes and their functions.** a) Organisation of the *CDKN2A* locus and its alternatively spliced products INK4A and ARF. Both proteins are structurally unrelated, but they share exon 2. The green rectangles depict untranslated regions (UTRs). b) Normal physiological function of the INK4A and ARF tumour suppressors. INK4A prevents the Cyclin D-CDK4/6 (depicted as '4/6-D') complex from phosphorylating RB1, and thus E2F-mediated cell cycle progression is inhibited. CDKN1A (also known as p21), whose expression is induced by P53, also inhibits RB1 phosphorylation by binding the Cyclin E-CDK2 complex (shown as '2-E'). ARF prevents MDM2-mediated ubiquitylation and degradation of P53, and thus contributes to its stability. c) Effect on the cell cycle of a disruptive mutation affecting INK4A and ARF. Mutations that affect exon 2 (depicted by a star) alter the function of both INK4A and ARF. If these proteins are incapable of binding their targets, P53 degradation ensues and hyperphosphorylation of RB1 (symbolised by 'p') occurs via direct interaction with the cyclin complexes, allowing progression of the cell cycle. d) CDK4 is a proto-oncogene that has been found mutated in a subset of melanoma-prone families. Mutations affect the INK4A binding site, but not its kinase activity, so CDK4 constitutively phosphorylates RB1. Adapted from refs. [194, 195].

further discussed in Section 1.6.

1.3.2.1.7 Other high-penetrance cancer-predisposing mutations and autosomal recessive inheritance of cancer predisposition Many other cancer-predisposing genes have been discovered by diverse methodologies such as linkage studies, candidate gene analyses and genome-wide mutation analyses (reviewed in [129]). These genes have different inheritance patterns, penetrances, mechanisms of disease (requiring biallelic inactivation or showing haploinsufficiency or dominant-negative effects, or constitutive activation in the case of proto-oncogenes) and participate in a diverse array of biological processes, such as cell cycle regulation and DNA repair (reviewed in [129]). Other examples of autosomal-dominant inheritance of cancer predisposition are Cowden syndrome, caused by germline mutations in *PTEN*, von Hippel-Lindau disease, caused by alterations in the eponymous *VHL*, and juvenile polyposis, caused by mutations in the *SMAD4* and *BMPR1A* genes (reviewed in [132]).

There are other cancer predisposition syndromes that are inherited in an autosomal-recessive manner. Examples of these are ataxia telangiectasia (A-T), caused by germline mutations in *ATM*, Nijmegen breakage syndrome (NBS), caused by mutations in nibrin (*NBN*), and the Bloom (BS) and Werner (WS) syndromes, caused by mutations in the DNA helicase genes *BLM* and *WRN*, respectively. All of these syndromes are characterised by telomere abnormalities and genomic instability (reviewed in [200]).

1.3.2.2 Polygenic model of cancer risk inheritance

Studies of cancer-prone families carrying deleterious alleles of some of the genes mentioned in Section 1.3.2.1 suggested that additional genetic components might be involved in the aetiology of their cancers, albeit with lower or modifier effects. For example, although specific mutations in *RET* can distinguish MEN2A and FMTC from MEN2B, they cannot distinguish between MEN2A and FMTC, suggesting that other genes might play a role in predisposition to these syndromes [133, 182]. Additionally, only about a fifth of the total clustering of familial breast cancer is observed in families carrying a predisposing *BRCA1* or *BRCA2* mutation, but diverse twin and familial studies suggest that genetic components are significant in the aetiology of the remaining families [123]. In fact, a polygenic risk model was found to best describe the distribution of breast cancer not due to *BRCA1* or *BRCA2* mutations in a population-based series, with the quintile most at risk showing 40-fold higher risk than the quintile least at risk [201]. Similar conclusions have been reached in other studies (reviewed in [123]) (Fig. 1.4).

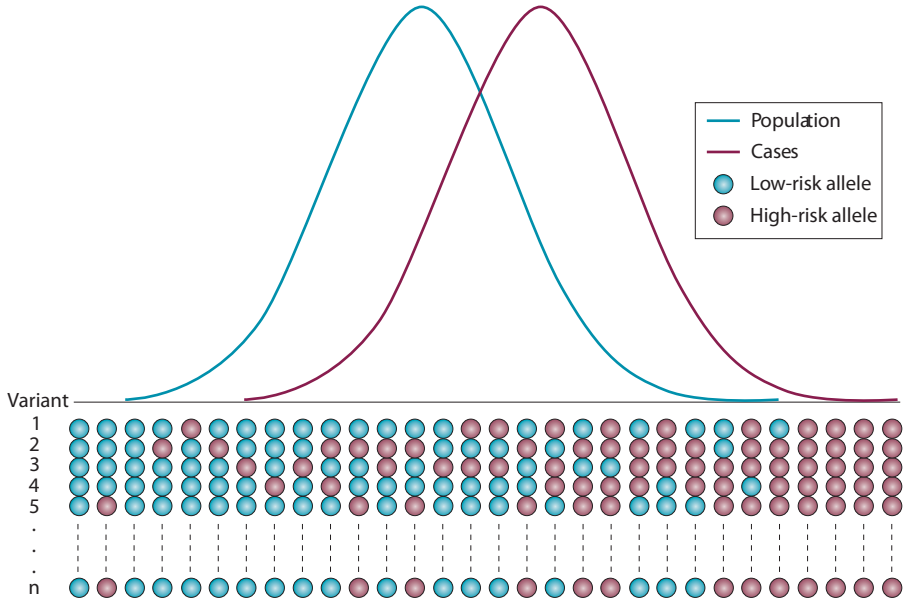

Figure 1.4: **Polygenic model of the distribution of cancer risk in the population and in individual cases.** Models of cancer risk that take into account several predisposition loci with small effects have been found to fit observed population-based data. Cases show an enrichment of high-risk alleles when compared to the rest of the population. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Cancer ([122]), copyright (2010).

Two models for polygenic predisposition to cancer have been formulated. The first of these, comprising common low-penetrance alleles predisposing to common cancers, is practically impossible to explore using linkage or candidate gene approaches, and thus GWAS have been carried out to compare allele frequencies in cases and controls. These common alleles are thought to have arisen once during evolution, and thus can be marked and genotyped in large numbers of cases and controls with high-density SNP arrays [123]. GWAS have been successful at pinpointing more than 200 low-penetrance genomic regions that influence cancer risk (reviewed in [134]), but identifying the causal alterations within these regions remains a daunting task. The second model involves rare alleles with moderate effects that are thought to have arisen recently or multiple times during evolution, and thus cannot be marked by SNPs in the vicinity. Several examples of these alleles have been found principally in breast cancer, such as variants in *CHEK2*, *BRCA1*-interacting protein C-terminal helicase 1 (*BRIP1*) and partner and localiser of BRCA2 (*PALB2*) (reviewed in [122]).

In conclusion, human cancers originate from complex interactions between environ-

mental and genetic components. Environmental factors can be behavioural in nature, such as alcohol drinking and smoking, occupational, such as exposure to asbestos or coal tar, or naturally prevalent, such as solar radiation and outdoor air pollution. Although the majority of cancers arise from environmental effects, it is estimated that up to a quarter of all human cancers have a major genetic component. Genetic predisposition to cancer can display Mendelian inheritance, when single genes account for the majority of the inherited cancer risk. Autosomal dominant conditions require only one mutated allele to confer an elevated cancer risk, and these can be caused by mutations in tumour suppressor genes (*e.g.*, *TP53* or *CDKN2A*) or in proto-oncogenes (*e.g.*, *RET* and *CDK4*). Autosomal recessive cancer-predisposing conditions also exist, such as A-T, NBS, BS and WS. However, the majority of the genetic risk for cancer is thought to arise from the additive or multiplicative effects of many genes with small effects, and in support of this hypothesis, polygenic models of cancer predisposition have been found to explain observations made in population-based series.

In the remaining sections of this introduction, I focus on the particular cancer I studied during my PhD, melanoma. I cover its biology, classification and presentation, its risk factors, the melanoma genome, and unanswered questions in melanoma genetics.

1.4 Melanoma: Facts, origin and biology

Melanoma is a malignancy arising from melanocytes, the pigment-producing cells in our skin. Even though it accounts for less than 2% of all dermatological cancer cases, it causes more than 75% of skin cancer-related deaths [202]. Over the last thirty years, its incidence has increased more rapidly than that of any other common cancer in the UK (Fig. 1.5) [110]. Some of this effect is perhaps due to better surveillance and improved detection methods and diagnosis criteria, but the majority is thought to be real and linked to changes in sun-related behaviour, and has been reflected, for example, in the increase in popularity of holiday packages in lower latitudes and sunbed usage [110]. This increase means that malignant melanoma is now the fifth most commonly diagnosed cancer in the UK, and is projected to become as common in the US by 2030 [203, 204].

The great majority of melanomas are diagnosed at an early stage, in which case the 5-year survival rates are high (~95% for young adults and ~90% for older people) [206, 207]. However, melanoma diagnosed at a metastatic stage is highly aggressive and resistant to chemotherapeutic treatment [208, 209], and the 5-year survival rates decline

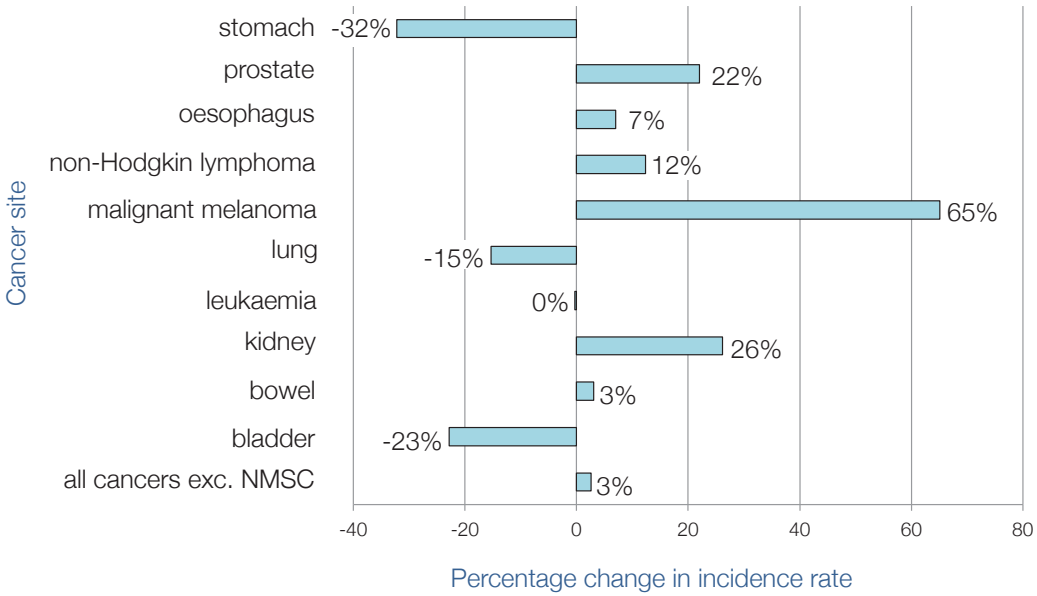

Figure 1.5: **Percentage change in European age-standardised three-year average incidence rates, in the UK, for the ten most common cancers.** European age-standardised three-year average incidence rates, UK, between 1999-2001 and 2008-2010. The information shown displays only incidence in males, but is representative of both sexes (malignant melanoma is, in both sexes, the fastest increasing common cancer in incidence). Information from ref. [205].

to ~60% and ~15% for regional and distant metastatic disease, respectively [202]. These effects are thought to arise from particular melanocyte biology properties, which I review in the next subsection.

1.4.1 Melanocytes: How they originate, where they reside, what they do

Melanocytes are specialised cells that synthesise melanin, a biopolymer that has various roles in cellular protection against damaging stimuli. Almost all melanocytes are derived from the neural crest, which is a transient embryonic cell population that can migrate extensively and give rise to numerous cell types, including much of the peripheral nervous system (reviewed in [210]). The exception are those melanocytes that constitute the retinal pigment epithelium (RPE), which are derived from the neuroepithelium [211]. The neural crest is formed during the embryonic process of neurulation, when the

neuroectoderm (at the outermost of the three germ cell layers) is transformed into the neural tube (Fig. 1.6a). Cells at the border between the neural and non-neural ectoderm dissociate and induce the epithelial-mesenchymal transition (EMT), allowing them to migrate out of the neuroepithelium [210]. These migrating cells are initially multipotent but they become gradually lineage-restricted depending on anatomical location (reviewed in [212]). The moment at which neural crest cells become committed to the melanocytic lineage (after which point they are called melanoblasts) is open for discussion [213], but they do so before reaching their final destination in the developing embryo [214]. Melanoblasts migrate mainly dorsolaterally, and settle principally in the epidermis, dermis, hair follicles and the inner ear cochlea [215].

During their migration process, melanoblasts rely on the Notch and Wnt signaling pathways for lineage commitment [212], and express several genes essential for their survival, such as paired box 3 (*PAX3*), sex determining region Y-box 10 (*SOX10*), microphthalmia-associated transcription factor (*MITF*) and v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (*KIT*) [217, 218] (Fig. 1.6b). In line with their functions in the developing melanocyte, defects in these genes underlie piebaldism and the Waardenburg and Tietz syndromes, which have in common pigmentation defects (reviewed in [219]).

Having reached their target site, melanoblasts differentiate into melanocytes by up-regulating genes important for the production of melanin, such as tyrosinase (*TYR*), tyrosinase-related protein 1 (*TYRP1*) and dopachrome tautomerase (*DCT*), mainly driven by *MITF* (reviewed in [214]) (Fig. 1.6b). Although the production of melanin (known as melanogenesis) is widely considered to be the most important function of melanocytes, it is by no means the only one, as these cells also have important sensory and immunological roles [220, 221]. They also participate in eye organogenesis and vision, hearing and possibly cardiac functions, depending on their anatomical location [211]. As such, melanocytes can be described as "classical", which are those found in the skin and contribute to its pigmentation, and "non-classical", which are those found in all other parts of the body [211].

1.4.1.1 Classical melanocytes

The skin is divided into three layers: The hypodermis, which consists of fatty tissue and connects the skin with underlying tissues, the dermis, which consists of connective tissue and fibroblasts and houses the lymphatic, neural, vascular and secretory systems in the skin, and the epidermis, the outermost layer [222]. The epidermis can be further

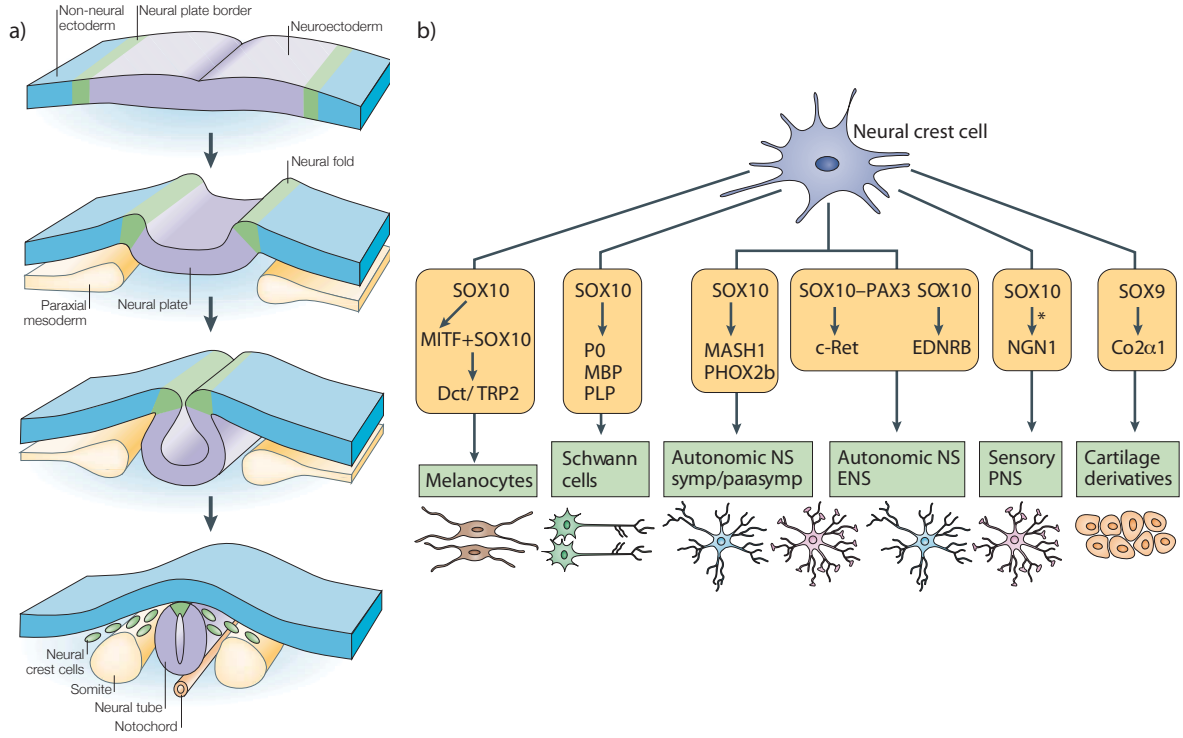

Figure 1.6: **Origin of the melanocytic lineage.** a) The neurulation process. In this embryonic stage, the neural plate rolls into the neural tube, and neural crest cells, localised at the border between the neural and non-neural ectoderm (green), delaminate from the neural folds and induce EMT, at which point they are able to migrate out of the neuroepithelium. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Neuroscience [216], copyright (2003). b) Neural crest cells can differentiate into numerous cell types, including much of the peripheral nervous system. Coordinated expression of *SOX10* and *MITF*, in conjunction with genes important for melanin synthesis such as *DCT*, determine the melanocytic fate. Important genes for other neural crest lineages are indicated. NS: autonomic nervous system; ENS: enteric nervous system; PNS: peripheral nervous system. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology [210], copyright (2008).

divided into four layers, which, in order from innermost to outermost, are known as the basal, spinous, granular and cornified layers [222]. Classical melanocytes can be found in the dermis or in the basal layer of the epidermis (Fig. 1.7).

Epidermal melanocytes are dynamic cells, with dendritic projections that they use to communicate with keratinocytes and Langerhans cells, thus mediating their pigmentation and immune functions [211, 224]. The interactions between epidermal melanocytes and keratinocytes constitute the basis of the epidermal melanin unit, a complex that coordinates the production and distribution of melanin and is capable of responding

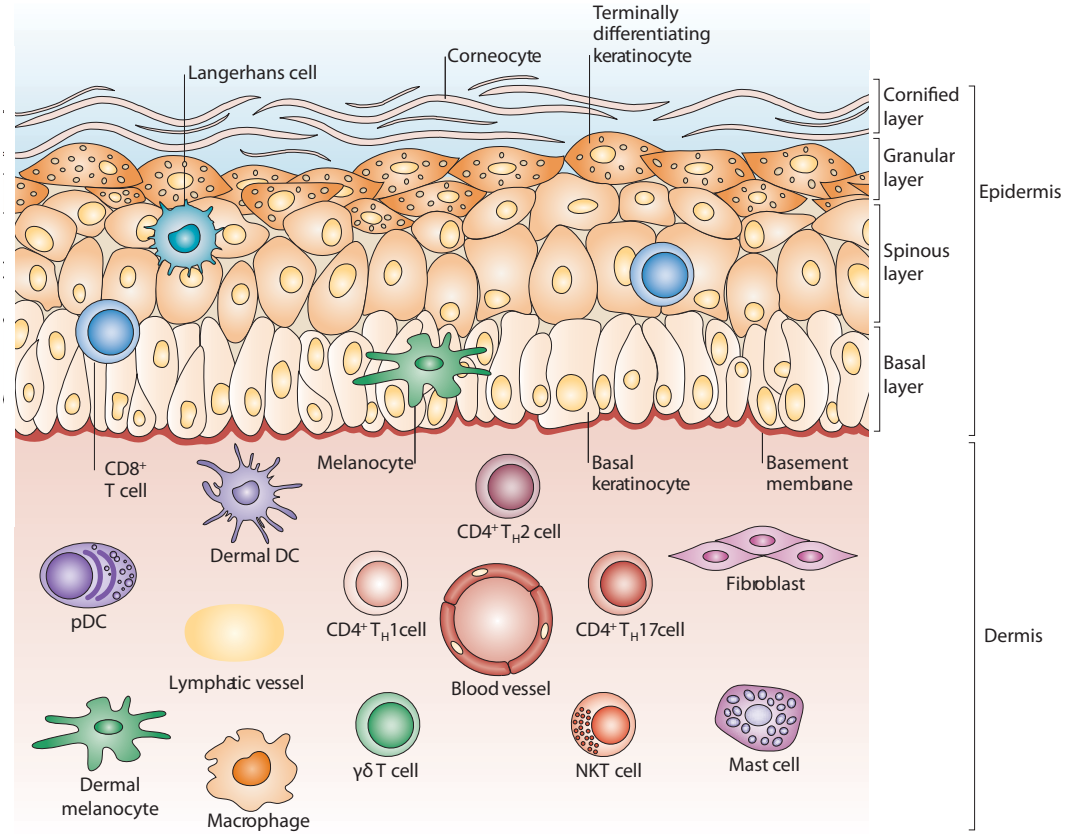

Figure 1.7: **Skin structure and melanocyte location.** The two outermost layers of the skin, dermis and epidermis, are shown. The four layers of the epidermis are indicated. Melanocytes (in green) are located in the dermis and the basal layer of the epidermis. Other types of cells they cohabitate and interact with, such as keratinocytes, Langerhans cells and fibroblasts, are depicted. The dendritic nature of melanocytes is also depicted. DC: dendritic cell, pDC: plasmacytoid DC, NKT: natural killer T, T_H: T helper. Modified and reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Immunology [223], copyright (2009).

rapidly to a large array of environmental stimuli [222]. It has been calculated that each epidermal melanocyte is in contact with about 30 keratinocytes [211], thus being able to distribute melanin upon a large area when required.

Within the epidermis, melanocytes can be located in the papillary bulbs of hair follicles or can be distributed in inter-follicular space. Inter-follicular melanocytes rarely undergo cell division [225], but those in hair follicles are able to proliferate and are involved in hair pigmentation and in the purging of toxic byproducts of melanin production [211].

Dermal melanocytes do not interact with keratinocytes, as they are surrounded by

fibroblasts. They are present in very small numbers, and although they are capable of producing melanin, little is known about their function [211].

1.4.1.2 Non-classical melanocytes

Non-classical melanocytes are present in many different locations, such as the eye, the inner ear, the heart, the brain and in adipose tissues, and do not generally contain large amounts of melanin, although some exceptions exist [211]. Melanocytes in the eye can be part of the RPE, in which case they play an essential role in retinal function and visual acuity, or be located in the uvea, where they contribute to its development and are responsible for iris coloration [211]. Melanin produced by the RPE is crucial for neural retina development, and can protect it from ROS, thus preventing age-related macular degeneration, whereas that produced by uveal melanocytes contributes to protection from oxidative damage related to their location in a densely vascularised section of the eye [211]. Otic melanocytes can be found in the cochlea or the vestibular organ, in which they are necessary for normal hearing functions and might also participate in balance perception [211, 226]. Cardiac melanocytes have only recently been described [227]; they do not appear to transfer melanin to surrounding cells [228] and are not essential for normal cardiac function, so their contribution in the heart remains unclear [211]. Brain melanocytes are mainly located in the leptomeninges, which are the innermost tissue layers that cover the brain [229]. It has been suggested that these melanocytes may aid in sequestering toxic compounds from the circulation and participate in neuroendocrine functions, although their physiological role has not been clarified [211, 226]. Melanocytes that reside in adipose tissue have been found to synthesise melanin at much higher levels in obese people than in controls, and it is thought that they might help to neutralise ROS and cellular fat deposition present in these patients [211].

In summary, the majority of melanocytes originate from the neural crest, a multipotent embryonic cell population that also gives rise to much of the peripheral nervous system; however, melanocytes in the RPE differ in that they originate from the neuroepithelium. Melanocytes can be classified according to their anatomical location, with “classical” melanocytes being those that populate the skin (and thus have been extensively studied) and “nonclassical” melanocytes being those residing in all other parts of the body. Melanocytes contribute importantly to organ development and function, but melanin biosynthesis is widely considered to be their most important task. In the next paragraphs, I review the melanin biosynthetic pathway, as well as the specialised organelles in which it takes place, the melanosomes.

1.4.1.3 Melanosomes and melanin biosynthesis

Melanins are important biopolymers that determine the most obvious phenotypic characteristic in human and other vertebrates, skin colour, and have an essential role in protecting the body against harmful UV radiation and other environmental challenges [222]. Melanins can be divided in two main groups: eumelanins, which are dark brown or black, and phaeomelanins, which are lighter or yellowish. Eumelanins are insoluble and participate in various protective functions, being able to oxidise and reduce other molecules, bind diverse metal ions, and absorb the most hazardous components of solar radiation (reviewed in [230]). Phaeomelanins have poor protective properties as they are photolabile at physiological conditions, and thus might even increase phototoxicity [231].

The pathway for melanin synthesis is the same for all melanocytes, although some types of melanocytes display particularities. For example, although melanin synthesis generally occurs in specialised organelles called melanosomes, it happens in the cytosol of melanocytes that reside in adipose tissue [211]. Additionally, dopaminergic neurons are able to synthesise a pigment related to melanin, termed neuromelanin, that appears to have similar properties but may arise from a different non-enzymatic synthesis pathway [232]. Also, the timing of melanosome generation and melanin biosynthesis can differ significantly, for example, skin melanocytes produce melanosomes throughout life but the RPE synthesises melanin only during embryonic and early post-natal life [211]. Nonetheless, the majority of melanocytes carry out the synthesis of melanin in melanosomes and express a common set of catalysing enzymes.

Melanosomes are membrane-bound, lysosome-related organelles that possess the conditions and proteins required for melanin synthesis, and that prevent toxic byproducts associated with its production from harming other cellular components [233]. Melanosome development is generally divided into four stages, termed I-IV (Fig. 1.8). Stage I melanosomes probably originate from the endoplasmic reticulum, and possess an amorphous matrix and internal vesicles. Premelanosome protein (encoded by *PMEL*) is then transferred to stage I melanosomes and aids in transforming them into fibrillar, elongated organelles with TYR activity, now termed stage II melanosomes. Pigment synthesis starts in stage III melanosomes, which accumulate melanin in their fibrillar matrix. Finally, when they become fully pigmented they are termed stage IV melanosomes, at which point they are ready for transfer to neighbouring cells (reviewed in [226, 233, 234]).

Proteins necessary for melanosome structure and function are delivered to these

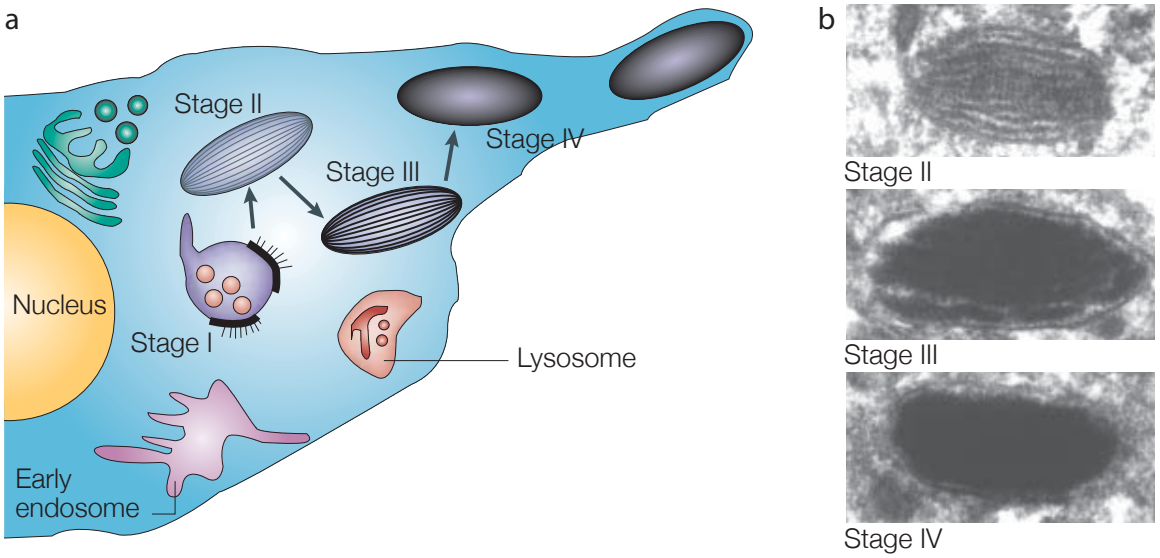

Figure 1.8: **Stages of melanosome maturation.** a) Diagram of a portion of a melanocyte body and dendrite. All stages of melanosome maturation are shown, alongside other relevant organelles. The degree of melanisation is indicated by black. b) Electron micrograph of melanosomal stages II, III and IV are shown. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Molecular Cell Biology [235], copyright (2001).

organelles throughout their maturation process, PMEL early and TYR, TYRP1 and DCT during the late stages [233]. The last three enzymes are required for melanin production, and therefore their malfunction or absence underlies distinct oculocutaneous albinism syndromes [222].

Biosynthesis of both eumelanins and phaeomelanins start with the hydroxylation of L-tyrosine to form L-3,4-dihydroxyphenylalanine (L-dopa), a reaction catalysed by TYR, and which is the rate-limiting step in melanin synthesis. The same enzyme then is able to oxidise L-dopa to L-dopaquinone, the last common step in the synthesis of dark and light pigments (see proximal phase, Fig. 1.9).

In eumelanin synthesis, L-dopaquinone yields L-dopachrome in a series of spontaneous chemical reactions, which then can form different precursors of dark pigments, such as 5,6-dihydroxyindole (DHI) and DHI-2-carboxylic acid (DHICA). These compounds then undergo spontaneous polymerisation reactions to form black and insoluble eumelanin (DHI-melanin) and brown, somewhat soluble eumelanin (DHICA-melanin). DCT and TYRP1 participate in the processing of L-dopachrome to these eumelanin precursors (reviewed in [236]) (see distal phase, Fig. 1.9).

Phaeomelanin synthesis is preferred if compounds such as L-cysteine or reduced glutathione are present. In this case, L-dopaquinone reacts with the thiol group in these compounds to form a complex mixture of intermediates, of which 5-cysteinyldopa is thought to be the most abundant. This compound then is hypothesised to undergo dehydration and several structural rearrangements to yield an alanyl-hydroxy-benzothiazine monomer, which is thought to be the monomeric subunit of phaeomelanin (Fig. 1.9). However, the phaeomelanin synthesis pathway is poorly understood, and thus could proceed via different and more complex intermediaries [236].

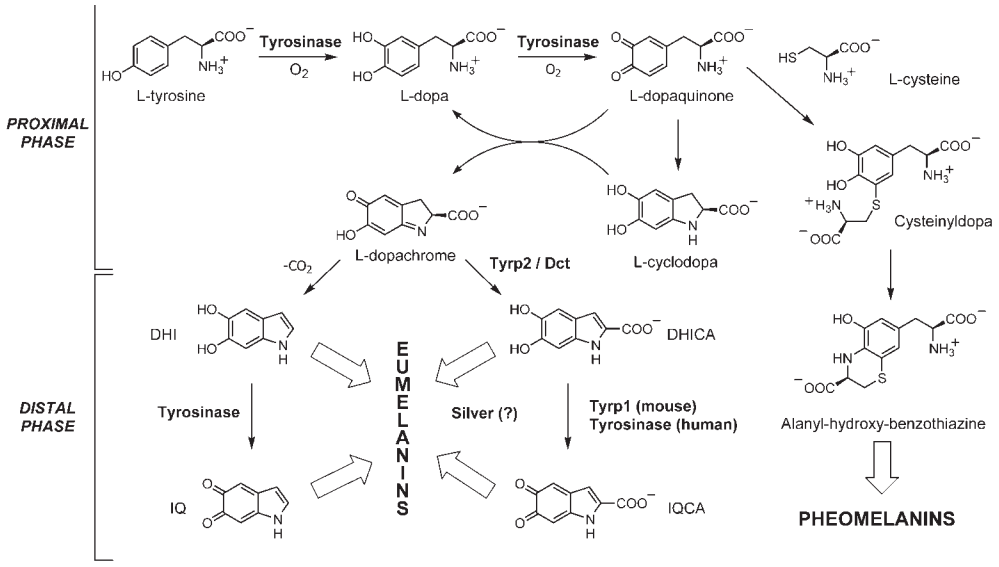

Figure 1.9: **Melanin biosynthesis pathway.** For simplicity, the pathway can be divided into the proximal phase, that contains the tyrosinase (TYR)-catalysed reactions, and the distal phase, that contains the subsequent reactions. In eumelanin synthesis, L-dopaquinone may form L-cyclodopa before this compound is spontaneously oxidised to L-dopachrome. IQ: 5,6-indolequinone and IQCA: indole-2-carboxylic acid-5,6-quinone are oxidation products from DHI and DHICA, respectively, and can be precursors for eumelanins. The silver locus, named after the mouse phenotype, encodes in humans the PMEL protein. Reproduced by permission from John Wiley & Sons, ref. [236], copyright (2011).

Melanocytes respond to different environmental challenges by producing melanin. Upon UV radiation, both melanocytes and keratinocytes increase their expression of the proopiomelanocortin (POMC) protein, which is then processed to different peptides, including α-melanocyte stimulating hormone (α-MSH) and adrenocorticotropin (ACTH). These peptides can bind the melanocortin 1 receptor (MC1R) on the surface of melanocytes and increase MITF expression through a pathway involving cyclic adenosine

monophosphate (cAMP), protein kinase A (PKA) and the cAMP response element binding protein (CREB) [222, 237]. The epidermal melanin unit can also respond to other environmental stimuli such as physical pressure, growth factors or cytokines [222, 238]. Keratinocytes also produce and secrete other factors that stimulate melanocyte dendricity and melanin synthesis, such as endothelin 1 (EDN1), colony-stimulating factor 2 (CSF2), prostaglandins E2 and F2α, leukaemia inhibitory factor (LIF), KIT ligand (KITLG), fibroblast growth factor 2 (FGF2) and the hepatocyte and nerve growth factors (HGF and NGF respectively) (reviewed in [222, 225]). These factors bind to different receptors in the surface of melanocytes and signal down different but cross-talking biological pathways, stimulating melanogenesis, dendrite formation, proliferation and differentiation [222] (Fig. 1.10).

1.4.1.4 Implications of the melanocyte lineage in melanoma treatment

The embryonic origin of melanocytes and their function offer some clues as to the aggressiveness and resistance of melanoma. For example, it has been shown that human primary melanocytes become metastatic when transfected with a specific set of genes, but not other cell types, indicating that at least a part of the characteristic aggressiveness of melanoma can be attributed to melanocyte lineage-specific factors [240]. In human benign naevi, the transcription factor snail family zinc finger 2 (SNAI2), that strongly contributes to EMT during neural crest migration [241], has been shown to be expressed along with other neural crest cell migration-associated genes [240]. Additionally, other studies have shown that melanoma cells exploit lineage-specific RTKs or transcription factors to promote their plasticity and metastatic potential [242, 243]. These observations might indicate that melanocytes are predisposed to acquiring invasive properties, thus necessitating fewer alterations to metastasise than other tissue types.

Additionally, melanocyte function may help explain why melanoma is highly resistant to chemotherapeutic treatment. Many chemotherapeutic drugs approved or being tested for treating melanoma, such as dacarbazine [244], temozolomide [245] and cisplatin [246], exert at least part of their function by inducing DNA damage and thus triggering cell cycle arrest or cell death [247, 248]. Given the crucial role that melanocytes have in protecting our skin against damage caused by UV radiation, they have developed powerful anti-apoptotic mechanisms that contribute to their intrinsic resistance to DNA damage and cell death [215]. Lineage-specific factors, such as MITF and KIT, are thought to play a role in activating anti-apoptotic genes such as B-cell chronic lymphocytic leukaemia (CLL)/lymphoma 2 (*BCL2*), although the mechanisms of melanoma chemoresistance

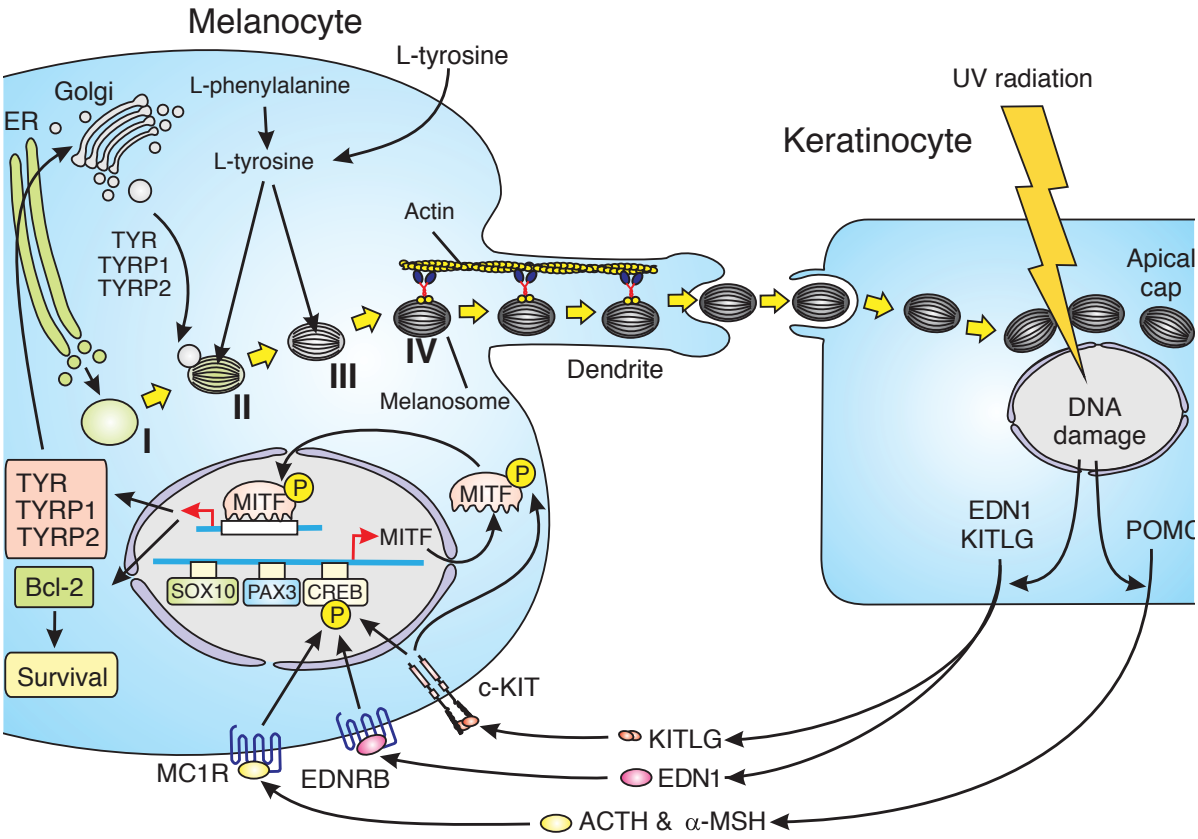

Figure 1.10: **Interactions between keratinocytes and melanocytes.** Schematic diagram of a melanocyte (left) and a keratinocyte. Upon UV radiation, keratinocytes produce factors that stimulate melanin production in melanocytes. These factors can, via signalling pathways, stimulate the trascription of genes important for survival and melanin synthesis. These gene products are then shipped to melanosomes, which upon maturation are transferred to keratinocytes where they contribute to DNA protection. Figure reproduced, with some modifications, from ref. [239].

remain controversial [215, 249]. Some authors have described melanocytes as cells that are "born to survive" [249, 250], and indeed, this seems to be the case given the low success rate of treatments for metastatic melanoma, making it one of the most challenging cancers to treat [251].

But how can a benign melanocyte give rise to one of the most aggressive and resistant human cancers? In the following paragraphs, I address the different models and theories that attempt to explain the transition from a benign melanocyte to malignant melanoma.

1.4.2 From melanocytes to melanoma: Different models of disease progression

It has been estimated that about three quarters of all melanomas arise from *de novo* melanocyte transformation, and the rest from pre-existing naevi [215, 252]. In both of these distinct paths to malignancy, changes in skin morphology are associated with a series of underlying molecular events.

1.4.2.1 The Clark model: Naevi become malignant

In 1984, Wallace Clark Jr. and colleagues described a step-wise model for the development of malignant melanoma from common naevi [253]. Naevi in the skin normally have an activating mutation in either v-raf murine sarcoma viral oncogene homolog B (*BRAF*) or neuroblastoma RAS viral (*NRAS*), which promotes growth by hyperactivation of the MAPK signalling pathway [254, 255] (Fig. 1.11). However, naevi rarely progress towards cancer, as oncogene-induced senescence is triggered and tumour suppressors such as INK4A (Fig. 1.3) and PTEN are up-regulated [256, 257]. However, this oncogene-induced senescence can be overridden by additional mutational events, sometimes in INK4A or PTEN themselves, and thus benign naevi can progress towards a dysplastic state (Fig. 1.11). As discussed in Paragraph 1.3.2.1.6, families with inactivating mutations in INK4A are predisposed to familial melanoma, perhaps because naevi progress more easily to the dysplastic state, whereas in sporadic cases the loss of PTEN is a more common alteration [258].

Additional mutational events might then affect MITF, although its role in melanoma progression is varied. In 10-20% of samples, MITF is found amplified, an event that is associated with worse prognosis. Although MITF is associated with melanocyte differentiation, it is thought that it confers a growth advantage to cells by cooperating with *BRAF* activation to transform them, thus being able to function as an oncogene [259]. Other samples have been found where MITF and its targets are down-regulated, suggesting that there are different subsets of melanoma with different MITF activities (reviewed in [260]). Amplification of cyclin D1 (encoded by *CCND1*) is also a marker of some acral melanomas, in which its inhibition has been shown to cause apoptosis [258] (Fig. 1.11). This phase, referred to as the radial growth phase, is characterised by an ability of melanocytes, now immortalised, to proliferate intradermally. The transition to the vertical growth phase, in which melanocytes gain the ability to invade the dermis and form tumours, is characterised by a loss of cell adhesion markers and the expression

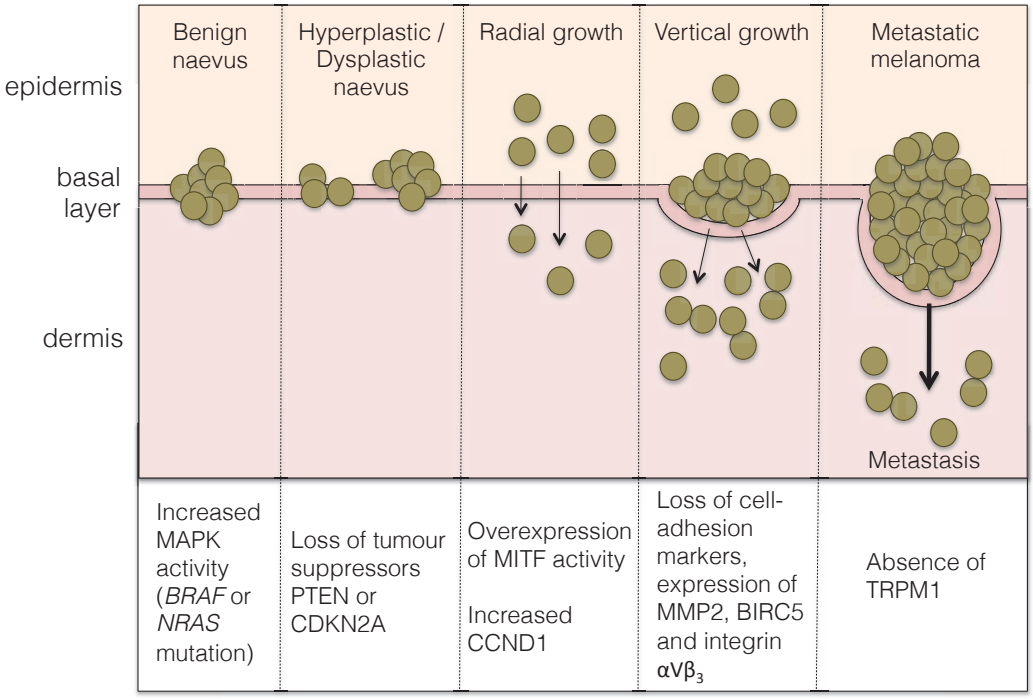

Figure 1.11: **The Clark model for the progression from naevi to malignant melanoma.** The hyperplastic and dysplastic phases (originally described as separate stages) are drawn together in this figure. Molecular events associated with each stage are in the bottom panel. Figure adapted from ref. [258].

of integrin αVβ3, baculoviral inhibitor of apoptosis domain repeat containing 5 (BIRC5) and matrix metallopeptidase 2 (MMP2), which contribute to survival and degradation of the collagen in the basal layer of the epidermis (reviewed in [258]). It has been also observed that the level of telomerase activity correlates well with tumour clinical stage, with higher levels of tumour cell penetration having higher telomerase activity [261]. Finally, melanocytes that have gained the ability to colonise other tissues, thus giving rise to malignant melanoma, commonly have a reduction or absence of expression of transient receptor potential cation channel, subfamily M, member 1 (TRPM1), a target of MITF. Although its function is not completely clear, this gene has been hypothesised to function as a tumour suppressor by comparison to other members of its family [262] (Fig. 1.11).

Many models of melanoma progression have been based on Clark's initial description of the melanoma stages, and thus is widespread in melanoma research community [263].

However, this view has been challenged by the observation that melanocytes within a naevus are not monoclonal as the Clark model would suggest, and the realisation that most melanomas arise in normal skin, not in association with an existing naevus [263]. Thus, an alternative model for the origin of malignant melanoma, aiming to explain how the majority of these cancers arise, has recently been put forward.

1.4.2.2 *De novo* progression from melanocytes to melanoma

Recently, Minoru Takata and colleagues proposed an alternative model for melanoma progression that does not pass through the dysplastic naevus stage and thus is different from Clark's [263]. They proposed that a yet unidentified hit might cooperate with the inactivation of INK4A or over-expression of cyclin D1 or CDK4, therefore allowing the cell to bypass oncogene-induced senescence. The acquisition of an activating mutation in *BRAF* would come after this first hit, resulting in clonal proliferation and thus yielding the *BRAF* mutation heterogeneity observed in naevi. Activation of telomerase and additional hits in late stages might be similar to those in cancers arising from benign naevi. Therefore, the order of mutational events might determine whether a melanoma arises from a pre-existing naevus or *de novo* [263]. This hypothesis is supported by observations made in acral and mucosal melanoma, in which cyclin D1 amplification arises frequently as a founder event, followed by oncogene activation, normally *KIT* [263].

At least one mouse model has been developed in an effort to recapitulate *de novo* disease progression. Mayuko Kumasaka and colleagues achieved *de novo* malignant melanoma formation in endothelin receptor B (Ednrb)-heterozygous mice constitutively expressing the RET oncoprotein [264]. This model could reproduce some characteristics of the human disease, such as the late age of onset, poor prognosis and high percentage of metastases, and, coupled with the observation that melanoma risk is increased in patients with *EDNRB* loss-of-function mutations [265], could provide a molecular rationale for the development of *de novo* malignant melanoma.

In summary, there seem to exist different paths to malignancy, with the majority of melanomas arising *de novo* but some arising from pre-existing naevi. This distinction might arise from the order in which important players for melanoma progression are somatically altered, but the lesions involved in both cases (*e.g.* senescence bypass, oncogene and telomerase activation and down-regulation of adhesion markers) are thought to be similar.

1.4.3 Clinical classification and staging systems

The models described above attempt to describe how melanocytes progress from a benign to a malignant state. Researchers have also classified melanoma in different types depending on the microscopic growth patterns of these lesions and have created diverse staging systems to aid treatment choice when diagnosing patients. In this Subsection, I review the recognised types of melanoma and the diverse systems used in diagnostic settings.

1.4.3.1 Types of melanoma

Clark not only defined the stages that a benign naevus had to progress through to become a malignant tumour, but also classified melanoma according to the microscopic growth patterns of these lesions [266]. He and his team divided it into four broad main types: superficial spreading melanoma (SSM), lentigo maligna melanoma (LMM), nodular melanoma (NM) and acral lentiginous melanoma (ALM) [263, 267] (Fig. 1.12).

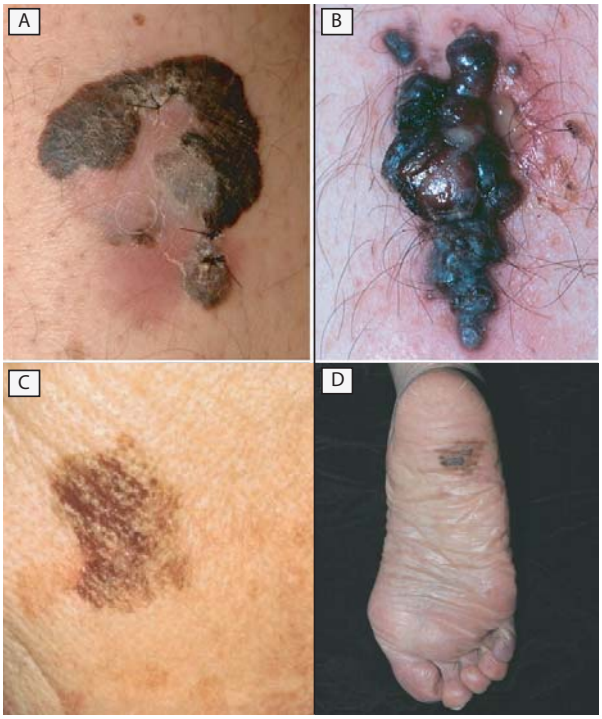

Figure 1.12: **Types of melanoma.** a) Superficial spreading melanoma, b) Nodular melanoma, c) Lentigo maligna melanoma and d) Acral lentiginous melanoma. Panels a, b and d reproduced from ref. [209] under a Creative Commons CC BY-NC 4.0 license. Panel c taken from the Skin Cancer Foundation, and reproduced by [268].

SSMs constitute about 70% of all invasive melanomas, and are characterised by melanocytes expanding upward within the epidermis and radially in its basal layer. They display variation in pigmentation and the majority occur *de novo* [267, 269]. LMMs evolve from lentigo maligna, a lesion that occurs on sun-exposed skin of older patients, although the progression is rare and slow. They are considered to be at the earliest stage of melanoma (see Paragraph 1.4.3.2.3), and are characterised by a proliferation of abnormal melanocytes along the basal layer of the epidermis [267, 270]. NMs are the second most common subtype of melanoma, comprising 10-15% of cutaneous melanomas, and are more aggressive than SSMs. They occur more commonly in sun-exposed areas of the skin, and present as expanding, darkly pigmented nodular lesions [271, 272]. ALMs present in palms and soles, and are the most common manifestation of melanoma in people of Black ancestry. They are highly aggressive, and present as darkly pigmented patches with varying degrees of pigmentation that expand rapidly [267, 273, 274].

Other types of melanoma have been identified that do not conform to the typical classification system. The IARC recognises, apart from the types mentioned above, desmoplastic melanoma (DM), melanoma arising from blue naevi (BN), melanoma arising from giant congenital naevi (GCN), childhood melanoma, naevoid melanoma and persistent melanoma [275]. DMs are superficial melanocytic lesions in which malignant cells are separated by fibrous stroma or collagen fibres, and can present cellular abnormalities. Melanomas arising from BN, which owe their colour to their deep position within the epidermis, are exceedingly rare and seem to be highly metastatic. GCN are large lesions that commonly cover more than 2% of the body surface, and are direct precursors to melanoma. Indeed, about 6% of GCN progress to malignancy, and the tumours are commonly sharply demarcated and characterised by asymmetry. Childhood melanoma comprises melanomas that develop in individuals before they reach puberty. Their incidence is quite low, and their clinical features are similar to melanomas arising in adults. Naevoid melanomas are also rare, and are distinctive because they resemble common intradermal naevi, but with the potential to metastasise. Persistent melanomas are those that grow out of primary lesions that have been excised, and normally have the same epidemiological and etiological characteristics as the primary tumour they are derived from [275].

Because of the heterogeneity in melanoma lesions and the existence of tumours that do not conform to the above classification system, John Curtin and colleagues proposed in 2005 an alternative classification in which melanoma types could potentially be distinguished not by histological characteristics but by molecular signatures [276]. As

such, characteristics such as the number of DNA copies, gains in the *CCND1* locus or mutations in *BRAF* or *NRAS* in a sample were able to accurately distinguish among acral melanomas, mucosal melanomas or melanomas on skin with or without chronic sun-induced damage [276]. Melanomas with and without chronic sun-induced damage roughly correspond to LMM and SSM, respectively, and acral melanoma to ALM in Clark's classification [263].

These distinct types of melanoma and their associated molecular events reveal a glimpse of the complexity of melanoma aetiology, indicating that diverse biological pathways can be altered in different ways in melanoma development, depending on characteristics such as anatomical site, amount of sun exposure, age and ancestry.

1.4.3.2 Melanoma staging systems

Different staging systems have been developed to describe tumour depth and the amount of spreading to other parts of the body. These are routinely utilised by medical doctors and researchers when diagnosing a patient and recommending treatment procedures.

1.4.3.2.1 Clark levels Tumour depth had been known to inversely correlate with prognosis, which inspired Clark to propose, in 1969, a 5-level classification system to score tumour depth [277] (Table 1.1).

Table 1.1: **Clark levels used to score tumour depth.**

Clark level	Description
Level I	All tumour cells are above the basal layer of the epidermis (melanoma *in situ*)
Level II	Tumour cells have broken into the papillary dermis
Level III	Tumour cells have reached the junction between the papillary and reticular dermis
Level IV	Tumour cells have invaded the reticular dermis
Level V	Tumour cells have spread into the subcutaneous tissue

Clark levels were used as the main tumour staging determinant in the American Joint Committee on Cancer (AJCC) staging system of 1997, but following studies that demonstrated that they only held independent predictive value for thin melanoma lesions (T1 stage, see Table 1.2), the AJCC revised their classification system in 2001 to restrict their assessment to this type. This recommendation has stayed the same for the last release of their classification system, in 2009 [278, 279].

1.4.3.2.2 Breslow thickness Around the same time that Clark published his classification system, the pathologist Alexander Breslow also published a paper arguing that melanoma lesions had to be scored on tumour thickness as this factor, along with Clark's level, seemed to correlate well with prognosis [280]. Tumour thickness is measured by determining the depth of invasion from the skin surface to the farthest malignant cell, usually by a pathologist examining a lesion under the microscope. Breslow thickness is still used as the main determinant of tumour staging, according to the 2009 AJCC melanoma staging system [279] (Table 1.2).

Table 1.2: **Breslow thickness measurements used by the 2009 AJCC melanoma tumour staging system.** Tis stands for tumour *in situ*, where cells are confined to the top layer of the epidermis. Modified from ref. [279].

Tumour stage	Breslow thickness (millimetres)
Tis	Not applicable
T1	≤ 1.00
T2	1.01-2.00
T3	2.01-4.00
T4	> 4.00

1.4.3.2.3 TNM scale The TNM classification is the main cancer staging system in use, and the most recent guidelines for melanoma were published by the AJCC in 2009 [279]. T stands for tumour size, N for the number of metastatic nodes, and M for the number of distant metastases. Each of these can be further subdivided to indicate characteristics such as whether ulceration is present, the amount of nodal metastatic burden and the site where metastases are present (Table 1.3). As such, this system utilises the Clark levels and Breslow thickness to aid in defining tumour stage and prognosis. The TNM stages include the most important parameters that retain individual prognostic value, and thus are the most important indicator of probability of survival [279] (Fig. 1.13).

1.4.3.2.4 Melanoma diagnosis: The ABCDE acronym In 1985, the mnemonic "ABCD" was created and published by three medical doctors to aid both healthcare professionals and the public in the early diagnosis of malignant melanoma [281]. Its main idea is that naevi that present the ABCD criteria should be examined closely; A stands for asymmetry, B for border irregularity, C for colour variegation and D for diameter larger than 6 millimetres (mm). More recently, a fifth criterium, E (evolving),

Table 1.3: **Most recent pathologic stage categories for cutaneous melanoma as determined by the AJCC.** Tumour thickness (T) stages are as described in Table 1.2, with the letters a and b indicating the absence and presence of ulceration, respectively. The number of metastatic nodes (N) are scored as follows: N0, not applicable, N1, 1 metastatic node, N2, 2 or 3 metastatic nodes, and N3, 4 or more. The letters a, b and c indicate micrometastases, macrometastases and in transit metastases, respectively. M0 and M1 refer to the absence and presence of distant metastases, respectively. Reprinted with permission from ref. [279].

Stage	T	N	M
0	Tis	N0	M0
IA	T1a	N0	M0
IB	T1b	N0	M0
	T2a	N0	M0
IIA	T2b	N0	M0
	T3a	N0	M0
IIB	T3b	N0	M0
	T4a	N0	M0
IIC	T4b	N0	M0
IIIA	T1-4a	N1a	M0
	T1-4a	N2a	M0
IIIB	T1-4b	N1a	M0
	T1-4b	N2a	M0
	T1-4a	N1b	M0
	T1-4a	N2b	M0
	T1-4a	N2c	M0
IIIC	T1-4b	N1b	M0
	T1-4b	N2b	M0
	T1-4b	N2c	M0
IV	Any T	Any N	M1

was added to recognise the dynamic nature of melanoma and to aid in identifying lesions that might not conform to the ABCD rule, such as NMs [282]. The ABCD rule has been found useful in recognising early lesions [283, 284], and due to its simplicity, continues to be widely used and promoted.

As described in this Section, the different types of melanoma, their progression through different disease stages and their resistance mechanisms are associated with genomic lesions such as activation of *KIT*, *NRAS* or *BRAF*, or inactivation of *CDKN2A* or *PTEN*, among others. In the next section, I review the current knowledge regarding the human melanoma genome, including known driver mutations (*i.e.*, those that confer

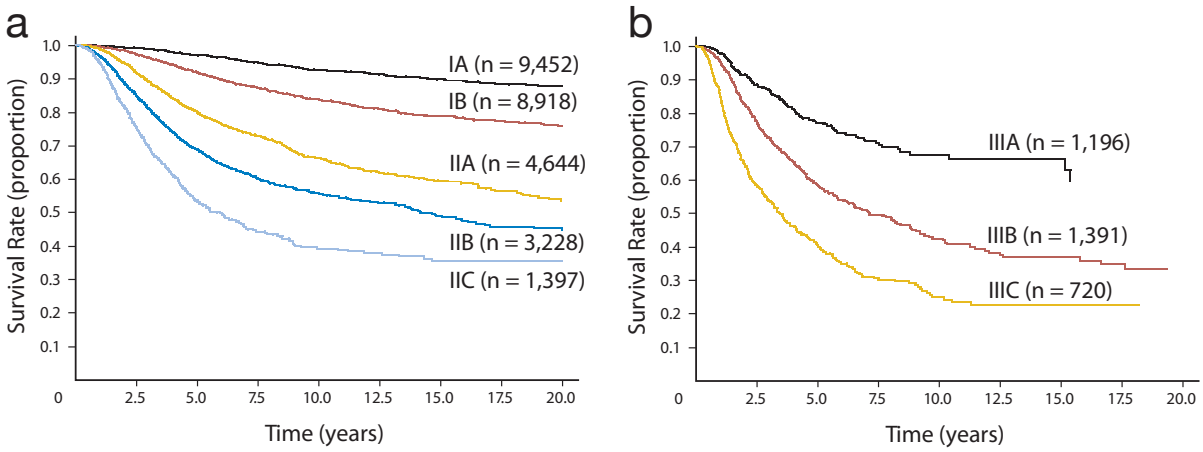

Figure 1.13: **Survival curves from the AJCC Melanoma Staging Database grouped by TNM stage.** The AJCC analysed data from 30,946 patients with Stage I-III melanoma, survival curves grouped by TNM stage over a 20-year period are shown. a) Stages I and II. b) Stage III. Reprinted with permission from ref. [279]. © (2009) American Society of Clinical Oncology. All rights reserved.

a fitness advantage to the carrier cell) and mutational signatures.

1.5 The human melanoma genome

Since the discovery of *BRAF*-activating mutations in almost 70% of human melanomas [285], somatically-acquired aberrations in the MAPK signalling pathway have been intensely studied. Overall, the great majority of melanomas have MAPK pathway dysregulation, with 15% of melanomas acquiring activating *RAS* mutations in a *BRAF* mutually-exclusive manner [258, 286]. Curtin and colleagues found that activating mutations of the MAPK pathway seemed to be less common in melanomas with chronic sun damage and non-cutaneous melanomas, and in general, that distinct subsets of genetic alterations were present depending on the type of melanoma [276]. Additionally, traditional *BRAF* or *NRAS* activating alterations are not C-to-T transitions, which are typically induced by UV radiation (see Subsection 1.3.1.5) [287]. These observations indicate that UV exposure significantly affects melanoma progression, possibly through the mutation of other oncogenic pathways.

In the last five years, the human melanoma genome has been intensely investigated using WGS, WES and exon capture methodologies. Two large studies analysing the somatic mutation frequency in more than 25 cancer types agreed that melanoma has the highest mutational burden across all of them, principally due to the ubiquitous

presence of C-to-T transitions [288, 289] (Fig. 1.14). This elevated mutational rate therefore can act as a confounding factor when attempting to uncover melanoma driver mutations.

Despite these difficulties, several studies have described many genomic events contributing to melanomagenesis or tumour progression. The majority of these mutations, that have been classified as drivers according to different criteria (*e.g.*, mutation frequency across samples, reduction of tumour fitness upon gene silencing or tumour progression upon ectopic gene expression), affect signalling pathways important for cell proliferation and migration. As such, somatically-aquired mutations resulting in MAPK pathway activation other than *BRAF* or *RAS* have been identified, such as gain-of-function events in glutamate receptor metabotropic 3 (*GRM3*) [290], MAPK kinase 1 (*MAP2K1*), *MAP2K2* [291], and v-erb-b2 avian erythroblastic leukaemia viral oncogene homolog 4 (*ERBB4*) [292]. Additionally, loss-of-function events in MAPK kinase kinase 9 (*MAP3K9*) and *MAP3K5* have been found to decrease MAPK signalling but increase drug resistance [293].

The PI3K pathway also plays a key role in melanoma progression. Activating mutations in *NRAS* and loss of function on *PTEN*, which result in activation of this signalling pathway, have been found in 15% and 20-30% of melanomas, respectively, in a mutually-exclusive manner [294]. Other significantly mutated loci might also activate the PI3K pathway, such as phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2 (*PREX2*) [295], phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha (*PIK3CA*) [287], v-akt murine thymoma viral oncogene homolog 3 (*AKT3*) [294], and *KIT* (which also signals down the MAPK pathway) [287, 296].

Expectedly, other common somatic alterations involve pathways important for cell cycle control and genome maintenance. Copy-number gains or somatic activation of the telomerase reverse transcriptase (*TERT*) gene are a frequent event in melanoma, as point mutations in its core promoter leading to increased telomerase expression were found in about 70% of examined tumours [297, 298]. These mutations were all found to be C-to-T transitions, providing a mechanism for UV radiation-induced tumourigenicity. Additionally, copy-number gains in *CCND1* are common in the acral subtype of cutaneous melanoma [299, 300], where its encoded product cyclin D1 contributes to tumour development by binding the CDK4/6 complex and by phosphorylating RB1 (see Paragraph 1.3.2.1.6). Recently, two independent studies identified protein phosphatase 6, catalytic subunit (*PPP6C*) as frequently mutated exclusively in melanoma samples with *BRAF* or *RAS* mutations [287, 301]. *PPP6C* encodes a protein able to block entry

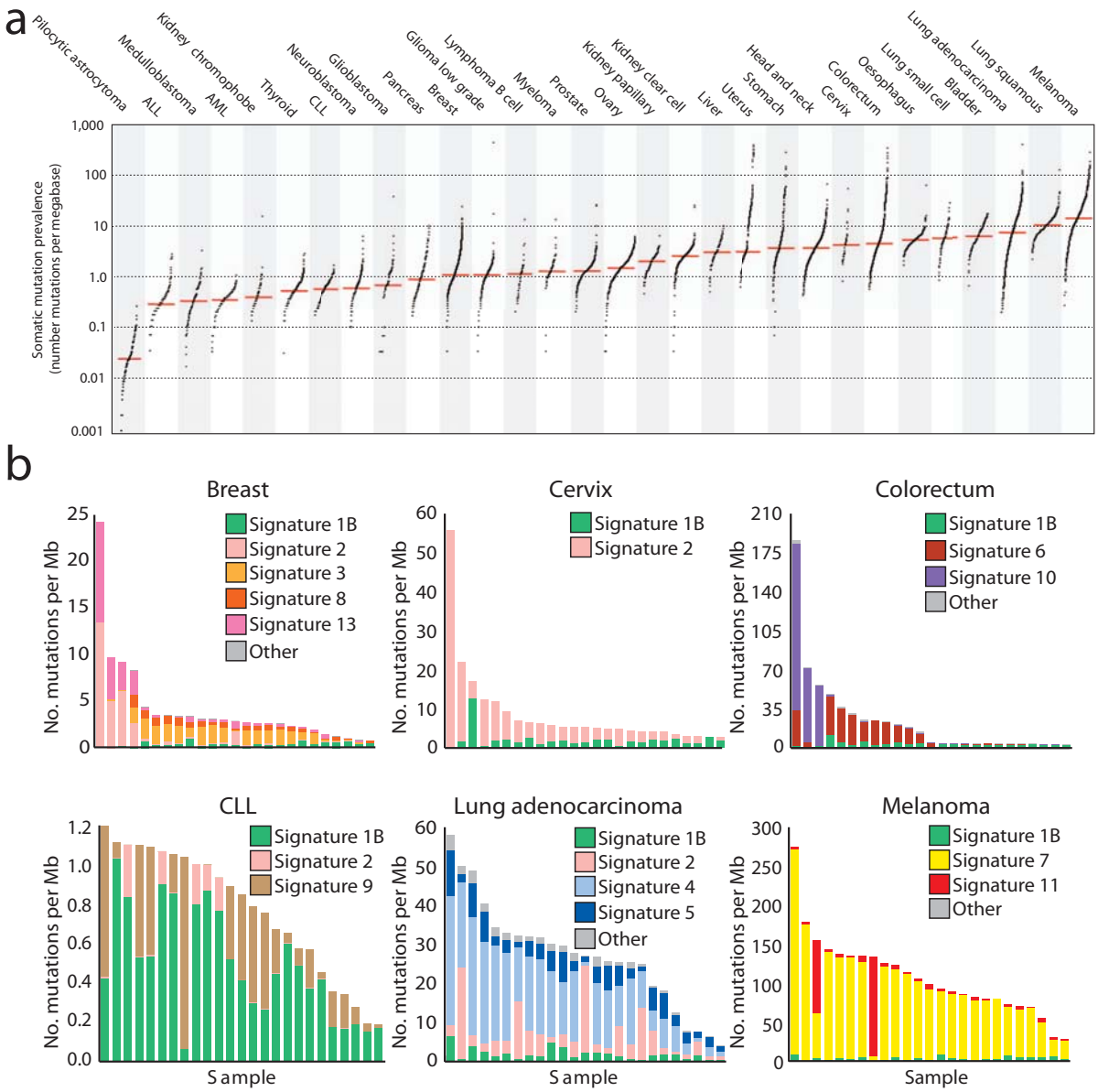

Figure 1.14: **Frequency and type of somatically-acquired mutations in melanoma genomes.** a) Prevalence of somatic mutations in human cancers. Each dot represents one sample, and the red line represents the median number of mutations per cancer type. The vertical axis is log-scaled and shows the number of somatic mutations per megabase (Mb), whereas the horizontal axis has the cancer types ordered by the number of mutational events. Melanoma is the type of cancer with the highest mutational burden overall. b) The contributions of somatic signatures to individual cancer types. Bars represent selected samples from each cancer type whereas the vertical axis denotes number of mutations per Mb. Whereas other cancer types have distinct contributions from different mutational processes, melanoma tumours are dominated by C-to-T transitions (Signature 7) and temozolomide-induced mutations in treated cases (Signature 11). Signature 1B has been linked to ageing, and is present across different cancer types. Reprinted by permission from Macmillan Publishers Ltd: Nature [288], copyright (2013).

into S phase, suppress the levels of cyclin D1 and reduce RB1 phosphorylation [302]. However, the two studies did not agree on the nature of the mutations: One identified probable loss-of-function events [301], whereas the other one identified clusters of variants in highly conserved residues that might represent oncogenic mutations [287].

Although the majority of cancers inactivate the P53 pathway at the level of *TP53* itself, mutations in *TP53* are not too frequent in melanoma, generally found in less than 20% of samples [209, 287, 301]. However, loss of this pathway's functions seems to be achieved, in a mutually-exclusive manner, by mutating *RB1* or the INK4A-coding region of *CDKN2A*, or amplifying *CDK4* or *MDM2* [209, 287, 303] (see Paragraph 1.3.2.1.6). Interestingly, more than half of the mutations in *TP53* in one large systematic review could be attributed to UV stress, implicating this locus in UV-induced disease progression [304]. Another study identified a synonymous C-to-T somatic mutation in the BCL2-like 12 (*BCL2L12*) gene, which caused accumulation of its messenger (m)RNA and protein product due to differential miRNA targeting [305]. This accumulation was found to cause increased binding to P53, decreased induction of UV-induced apoptosis, and a reduction in the transcription of endogenous P53 target genes.

Other pathways that might be commonly targeted for somatic alterations are those important in melanocyte development. The MITF pathway is altered in about 20% of metastatic tumours, with *MITF* amplification being a frequent event [259, 306]. Copy-number gains in *MITF*, in conjunction with oncogenic *BRAF*, were shown to be able to transform melanocytes [259], and the *MITF* regulator *SOX10* was found to be mutated in about 10% of primary melanomas and to be inactivated in some metastatic tumours [306].

Recently, the glutamate receptor pathway has also been found to be somatically deregulated in melanoma cells. Mutations in *GRM3* not only activate the MAPK signalling pathway, but also alter glutamate receptor function [290]. Additionally, mutations in the glutamate receptor, ionotropic, N-methyl D-aspartate 2A (*GRIN2A*) have been found in about a quarter of melanomas, and phospholipase C, beta 4 (*PLCB4*), a downstream effector from GRIN2A, has also been found to be recurrently mutated [307]. Glutamate receptor dysfunction has been noted in neuronal tumours, with gliomas that have high glutamate release being more aggressive [308]. Given the neural crest origin of melanocytes, it might indicate that shared molecular mechanisms between neurons and melanocytes allow melanoma cells to take advantage of neurophysiological receptors [209]. Although it is unclear from the pattern of mutations if GRIN2A acts as a tumour suppressor or an oncogene, mutant *GRM3* was found to harbour a mutation hotspot and

to confer increased anchorage-independent growth and migration, possibly indicating that it acts as an oncogene [290, 307].

Many other significantly mutated loci have been described. Recently, ras-related C3 botulinum toxin substrate 1 (*RAC1*) was described as a new oncogene in two independent studies, as both found UV-induced recurrent mutations affecting the same codon [287, 301]. RAC1 is an important player in cell proliferation and cell migration, and when mutated, it could provide a migratory advantage to melanocytes through ERK activation [301]. Other candidate genes include deleted in colorectal carcinoma (*DCC*), dynein cytoplasmic 1 intermediate chain 1 (*DYNCI1I*), RPTOR independent companion of MTOR, complex 2 (*RICTOR*), AT rich interactive domain 2 (*ARID2*), serine/threonine kinase 19 (*STK19*), sorting nexin 31 (*SNX31*) and transforming acidic coiled-coil containing protein 1 (*TACC1*) [287, 301].

Although somatic mutations are required to drive melanomagenesis, the acquisition of mutations is facilitated by certain factors, such as induction of C-to-T transitions upon UV radiation exposure, or having a preponderance of phaeomelanin over eumelanin, which can generate oxidation-induced DNA damage. In addition, germline mutations in some of the genes mentioned above have been found to predispose to melanoma with high penetrance. In the next section, I discuss these and other risk factors for melanoma development.

1.6 Risk factors for melanoma development

Several risk factors for the development of melanoma have been identified. These can be environmental, such as exposure to UV radiation, physical characteristics, such as having fair skin, eyes or hair, or genetic, such as carrying mutations in genes that control melanogenesis or more general cell integrity processes. In this Section, I review established risk factors for melanoma and their modes of action.

1.6.1 Environmental factors

1.6.1.1 UV radiation

It has been estimated that about 86% of all melanomas can be attributed to exposure to UV radiation, whether from the sun or from UV-emitting devices such as sunbeds [110], and as such, it is considered the main risk factor for developing melanoma (for mutagenic mode of action, see Subsection 1.3.1.5). However, different patterns of sun exposure carry

different risks. For example, a meta-analysis of 57 studies and a total of 38,671 patients identified that intermittent and intense sun exposure significantly increased melanoma risk when compared to chronic sun exposure [309]. Although a history of sunburn might appear to increase melanoma risk, this effect may arise because people are more likely to remember it, but a causal relationship cannot be excluded [309].

Additionally, the relationship between the use of sunbeds and melanoma has been analysed in several studies, and at least two meta-analyses concluded that melanoma risk was increased in individuals that had ever used sunbeds, with risk increasing if the exposure was at a young age [310, 311]. As such, the IARC decided to raise the carcinogenicity level of sunbed usage to the top category in 2009 [312]. In 2011, legislation came into force in England and Wales that made it illegal for people under 18 years of age to use sunbeds [313], and other countries, such as France, Brazil, Austria and Germany, have also taken steps to limit the exposure of younger individuals [314].

1.6.1.2 Other environmental factors

Diverse studies have identified other factors that might play a role, albeit probably limited, in melanoma susceptibility. Melanomagenesis has been linked to arsenic exposure [315], regular swimming (probably water chlorination) [316], exposure to residential magnetic fields [317], and industries such as electronics, metal and transport and communications [318]. However, these studies analysed only a limited number of people and reported small odds ratios (ORs) with wide confidence intervals, so further research is necessary to establish whether a causal relationship exists between these environmental factors and melanoma formation.

1.6.2 Physical characteristics and medical history

1.6.2.1 Skin, hair and eye colour

Skin and hair colour are primarily determined by the amount and types of melanins produced, although other pigments such as haemoglobin and dietary carotenoids also contribute to them [319]. The absence of both eumelanin and phaeomelanin results in white hair (albinism), higher eumelanin production results in dark hair and skin, and higher phaeomelanin production results in blond or red hair and fair skin [319]. Therefore, given the crucial role of eumelanin in skin protection against UV radiation, it is expected that individuals with darker skin and hair would be better protected than

blond or red-haired individuals when exposed to UV radiation [226] (see Subsection 1.4.1.3).

Although the molecular details underlying hair and skin colour are still incompletely understood, some important players have been elucidated. The switch between eumelanin and phaeomelanin production within a melanocyte is controlled by *MC1R*, which is highly polymorphic in the human population [319]. Variants that impair MC1R binding to α-MSH and/or subsequent cAMP production, and therefore affect eumelanin production, are over-represented in redheads and fair-skinned people [320]. Furthermore, *MC1R* mutation carriers are not only more susceptible to UV radiation-induced damage, but may also have increased oxidation-induced DNA damage arising from phaeomelanin synthesis, contributing to melanomagenesis [321].

MC1R variants play a role not only in red-haired individuals, but also account for inability to tan in individuals without red hair [322], and increase the risk of developing freckles independently of skin and hair colour [323]. As such, an inability to tan and the presence of freckles represent independent melanoma risk factors [324, 325]. More recently, a mutation altering *KITLG* expression levels was found to underlie classic blond hair colour in Europeans [326], and given that this pathway seems to play an important role in melanocyte number, size and dendricity [327], it will be interesting to see whether this variant contributes to melanoma susceptibility in this population, and if so, whether the mechanism involves the KIT pathway.

Eye colour is a polygenic trait, but two main contributing loci have been identified: oculocutaneous albinism II (*OCA2*) and homologous to the E6-AP carboxyl terminus (HECT) and regulator of chromosome condensation 1-like domain (RLD) containing E3 ubiquitin protein ligase 2 (*HERC2*) [328]. OCA2 is involved in melanosome transport and maturation, and thus variations in this gene affect the type and quantity of melanin produced [328]. Polymorphisms within *HERC2* probably act in the same pathway, as they have been found to affect *OCA2* expression levels [329]. Therefore, eye colour reflects genetic variants that affect melanin distribution, and therefore has been identified as an independent melanoma risk factor [324, 330].

1.6.2.2 Presence of multiple naevi

The presence of an elevated number of naevi is considered one of the most important risk factors for melanoma development. Naevi are benign melanocytic tumours, and although the majority are stable and will not progress to melanoma, some will become malignant (see Subsection 1.4.2.1). A meta-analysis carried out in 2005 identified a 7-

fold higher melanoma risk in people with more than 100 common naevi compared with those that have fewer than 15, and individuals with five dysplastic naevi presented a 6-fold higher risk compared with individuals with no dysplastic naevi [331].

Although the majority of melanomas arise *de novo* (see Subsection 1.4.2.2), it has been calculated that around 20% of all melanomas originate from a dysplastic naevus (reviewed in [332]), and, at least within *CDKN2A* mutation-carrier families, individuals with dysplastic naevi are more likely to carry the mutant gene compared with family members without dysplastic naevi [333]. Therefore, the presence of multiple common or dysplastic naevi might indicate the presence of genetic or environmental factors that favour melanoma progression, such as carrying a mutation that accelerates melanocyte proliferation, or having a history of sun exposure.

1.6.2.3 Presence of a previous melanoma

Individuals with a previous diagnosis of melanoma have a 8-12-fold increased risk of developing a subsequent melanoma if they do not have a family history of melanoma, and 300-500-fold increased risk if they do and have dysplastic naevi when compared to population incidence rates [334, 335]. As is the case with dysplastic naevi, the development of a single melanoma might be indicative of the presence of underlying risk factors, but it can also happen that cells from a primary tumour that failed to be surgically excised proliferate and originate a secondary tumour. This event is referred to as local recurrence or persistent melanoma (discussed in Subsection 1.4.3.1).

1.6.2.4 Other medical conditions

Immunocompromised patients, such as organ transplant recipients, are at an increased risk of developing many types of cancer, with skin cancer including melanoma being one of the most common [336, 337]. Additionally, Parkinson's disease patients also seem to have a significant increase in melanoma incidence, although the reason remains unknown [338]. Perhaps not surprisingly, individuals with other types of cancer, such as non-Hodgkin's lymphoma, leukaemia and renal and gastrointestinal tumours also have increased melanoma risk [339–342]. Additionally, several studies have reported that retinoblastoma patients that survive their disease are also at high risk of developing melanoma [209, 343, 344].

1.6.3 Genetic risk factors

Another important melanoma risk factor is having a family history of the disease, which could be indicative of underlying genetic alterations that predispose to its formation. Meta-analyses have shown a ~2-fold increase in risk for individuals with a family history compared to those without [345, 346]. Several loci have been discovered, through methodologies such as linkage studies, WGS, WES or GWAS, to influence cutaneous melanoma risk (Fig. 1.15). These genes participate in cell cycle control or genome stability, and have varying degrees of penetrance.

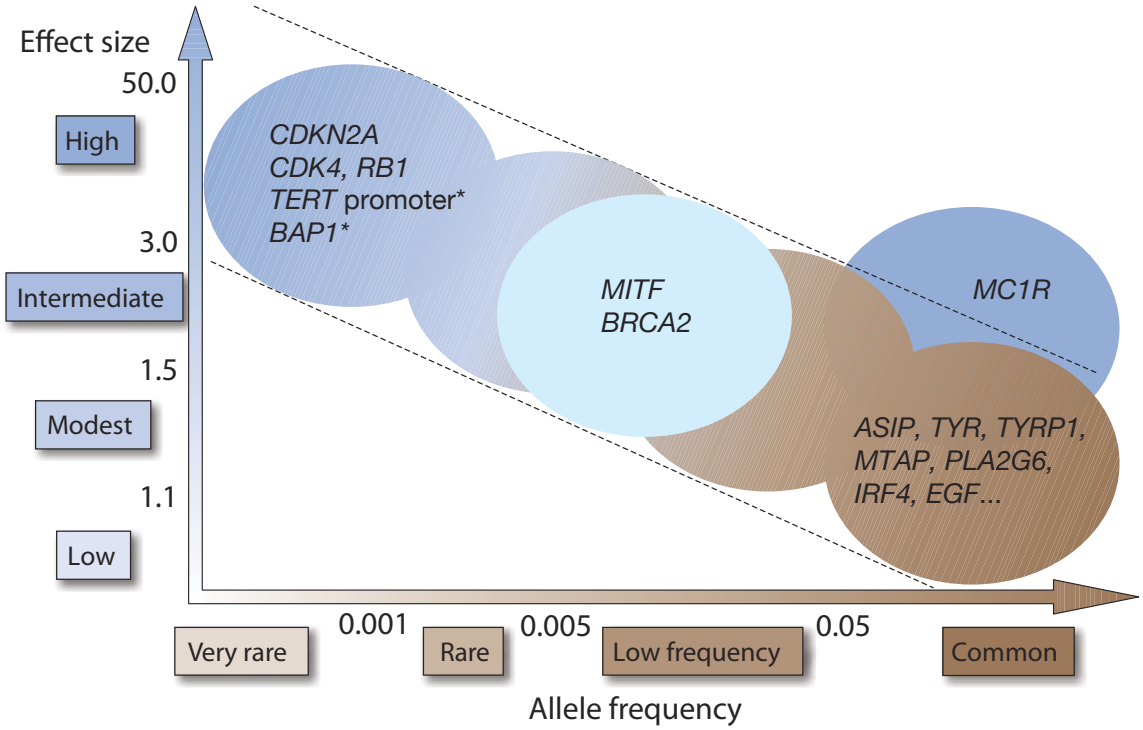

Figure 1.15: **Effect sizes and allele frequencies of loci influencing melanoma risk.** The dotted line indicates where the highest emphasis is when searching for predisposition loci. Loci with asterisks have been estimated, as very few families have been studied to ascertain their true effect size and allele frequency. The position of *BRCA2* has been estimated from refs. [347, 348]. Remaining loci positions are from ref. [215]. Modified and reprinted by permission from Macmillan Publishers Ltd: Nature [349], copyright (2009).

1.6.3.1 High-penetrance loci

The biological function of the two classical high-penetrance melanoma susceptibility loci, *CDKN2A* and *CDK4*, is discussed in Paragraph 1.3.2.1.6. About 40% of melanoma-prone families carry mutations in *CDKN2A*, whereas very few kindreds with pathogenetic *CDK4* mutations have been described worldwide [350, 351]. For familial cases, the overall penetrance of *CDKN2A* mutations has been estimated at 30% by 50 years of age and 67% by 80 years of age, although it can be modified by residence location [352]. For sporadic cases, it has been calculated at 14% by 50 years of age and 28% by 80 years of age [353], indicating that other shared environmental or genetic risk factors might influence familial melanoma risk.

RB1 mutation carriers that survive retinoblastoma (see Paragraph 1.3.2.1.1) are at a 4-80-fold risk of developing melanoma when compared to population incidence levels [209]. Other genes have been reported recently that might be highly penetrant, although estimating their penetrance has been hampered by the small numbers of carrier individuals identified. In 2011, Thomas Wiesner and his team identified deleterious mutations in BRCA1-associated protein 1 (*BAP1*) co-segregating with melanocytic tumours in two families, with tumours showing loss of heterozygosity of the wild-type allele [354]. Two years later, Susanne Horn and colleagues identified a mutation in the promoter of *TERT* co-segregating with the disease in a large melanoma-prone German family [297]. This mutation was found to increase *TERT* expression levels, and since then, similar somatically-acquired mutations have been found in a variety of other cancers [355].

1.6.3.2 Medium-penetrance loci

As discussed in Subsection 1.6.2.1, individuals that carry variants in *MC1R* have a higher risk of developing melanoma. ORs for different variants within the gene range from 1.42 to 2.45 [356]. More recently, a single nucleotide variant (SNV) affecting codon 318 in MITF was found to disrupt a conserved small ubiquitin-like modifier (SUMO)-ylation site, thus altering the spectrum of MITF target genes and leading to melanomagenesis. The OR for this allele was calculated to be between 2.7 and 4.78, consistent with an intermediate penetrance effect, but its allele frequency is much rarer than that of *MC1R* variants [357, 358]. *BRCA2* mutation carriers have more than 2.5-fold risk of developing the disease when compared to population incidence levels [347].

1.6.3.3 Low-penetrance loci

Many low penetrance loci, with high allele frequencies, have been identified through GWAS. Many of these genes affect pigmentation functions or skin sensitivity to UV-induced DNA damage [215]. For example, variants in agouti signalling protein (*ASIP*), an antagonist of MC1R, *TYR* and *TYRP1* have been associated with cutaneous melanoma with ORs ranging from 1.15 to 1.45 [359], and similar ORs were obtained for variants in the naevus-associated genes methylthioadenosine phosphorylase (*MTAP*), phospholipase A2, group VI (*PLA2G6*) and interferon regulatory factor 4 (*IRF4*) [360]. Variants in epidermal growth factor (*EGF*) and its receptor (*EGFR*) have also been associated with melanoma and its progression [361, 362]. Other loci potentially implicated are caspase 8 (*CASP8*), *CCND1*, solute carrier family 45, member 2 (*SLC45A2*), *ATM*, *OCA2*, myxovirus resistance 2 (*MX2*), among others (reviewed in [194, 209, 215]).

As we can see from the above paragraphs, melanoma is a complex disease arising from the interaction of diverse environmental factors, such as UV radiation and potentially other carcinogens, with the underlying genetic make-up of an individual. It is possible that dozens or maybe hundreds of genes, contributing to pigmentation, skin sensitivity, cell cycle control and genome stability, influence the genetic risk of this disease. In the next section, I discuss melanoma clustering in families, as the subject of study of this dissertation.

1.7 Familial melanoma

The English general practitioner William Norris first reported a case of familial melanoma in 1820 [363]. He speculated, given that his patient's tumour had originated from a naevus, that his children and brothers had many naevi, and that his father had also died of the disease, that melanoma was hereditary. This observation is notable because it was made nearly half a century before Mendel published his treatise on genetics [1].

Familial melanoma can be characterised by multiple melanoma cases across several generations on one side of the family, or by the presence of multiple primary melanomas in a single individual, or by an early age of onset [209]. In 1978, Lynch coined the term "familial atypical multiple mole-melanoma syndrome" (FAMMM) for families with a clustering of multiple large, dysplastic naevi of variable colour with pigmentary leakage, one of the main risks for the development of melanoma [364] (see Subsection 1.6.2.2). However, tumours from individuals with FAMMM do not show any histopathologic

differences when compared to sporadic cases, and they might not show all typical FAMMM characteristics, so their analysis is not useful when diagnosing the familial condition [365]. Individuals with FAMMM caused by mutations in *CDKN2A* also have been found to have a higher incidence of other malignancies, especially pancreatic cancer [365, 366].

Since Norris's report, it has become clear that although the majority of melanomas can be attributed to UV radiation exposure [110], about 10% of all melanoma cases have a family history of the disease [194]. However, heritability of melanoma is high, as it has been estimated that about 55% of variation in liability to melanoma is due to genetic effects [367]. This discrepancy between high heritability and low familial melanoma rate might indicate that a large proportion of the risk arises from common, low-penetrance variation [368], although some rare high-penetrance loci might still remain undiscovered. This could be because individual variants in a locus might have a very low allele frequencies and/or might not be obviously disruptive, which precludes systematic identification of potential candidates.

The definition of familial melanoma varies depending on geographical location, as some regions, such as Australia, have a higher prevalence of melanoma and thus a higher probability of seeing clusters of melanoma within a family caused by non-genetic reasons [352, 369, 370]. In the UK and Europe, familial melanoma is generally defined as a cluster of two or more first- or second-degree relatives with melanoma [370, 371], whereas in Australia it constitutes a diagnosis of one invasive melanoma and two or more other cancer diagnoses among first- or second- degree relatives [370]. Additionally, patients might be referred for genetic testing if they live in a region with low melanoma incidence and present with at least two primary tumours, or if they present with three or more primary tumours and they live in a high melanoma incidence region [370].

1.8 Unanswered questions in melanoma genetics

Since the discovery of *CDKN2A* and *CDK4* about 20 years ago [192, 199], only a handful of other high-penetrance genes in very rare families have been described, namely *BAP1*, *RB1* and *TERT* (see Subsection 1.6.3.1). Collectively, these genes explain only about 50% of all familial melanoma cases. This fact begs the question: Are there any more high-penetrance genes remaining to be discovered, or can all unexplained familial risk be attributed to many lower-penetrance alleles?

Additionally, these known high-penetrance genes encode proteins that participate in

cell cycle control and genome maintenance pathways, with *CDKN2A*, *CDK4*, *RB1* and *BAP1* participating in G1/S cell cycle progression ([372], see Paragraph 1.3.2.1.6), and *TERT* having a paramount role in genome stability after each cell division [373]. If other high-penetrance loci remain to be discovered, do these participate in the same biological pathways, or are there any other processes important in melanoma predisposition?

During my PhD, I endeavoured to help answer these questions. The dramatic drop in the costs in sequencing costs over the past decade [14] has allowed the use of this technology to explore large numbers of affected individuals in the search for common affected loci or biological pathways. When planning this project, we, as a team, decided to use WES in multiple members of several melanoma-prone families for four main reasons:

- Unbiased projects (*i.e.*, assessing the whole genome or exome rather than sets of candidate genes) to date studying genetic susceptibility to melanoma are very few, and have been carried out in three or fewer individuals [354, 358]. This means that the possibility of rare, high-penetrance alleles in coding regions has not been properly addressed.

- The cost of sequencing an exome is ~20× less than that of sequencing a whole genome, which would allow us to study multiple individuals from a greater number of affected families. This is helpful when attempting to obtain statistical support for the co-segregation of candidate mutations.

- Variation in protein-coding regions can be directly interpretable, *e.g.* causing an amino acid change or introducing a premature stop codon in a protein, which then can be assessed by functional, structural or conservation scores.

- The possibility that deep catalogues of human genetic variation would be developed and made publicly available during the course of this project, which would allow a better estimate of the allele frequency of candidate variants in human populations.

1.8.1 Overview of whole-exome sequencing as a tool for providing answers

In order to analyse the protein-coding regions of the genome, the WES method involves, firstly, the enrichment of the DNA sample for exonic regions (or any other custom regions, in the case of candidate gene sequencing). Generally, this is achieved by solid- or liquid-phase hybridisation. In the first case, probes complimentary to the target regions are

fixed in a solid surface, such as a microarray, which can then be washed and captured regions eluted; in the second case they can be hybridised to biotinylated DNA or RNA probes and captured with streptavidin beads [374, 375] (Fig. 1.16). Captured DNA regions can then be massively sequenced, mapped onto the human genome reference assembly, and potential variants can be called and filtered in the search for candidate melanoma predisposition genes.

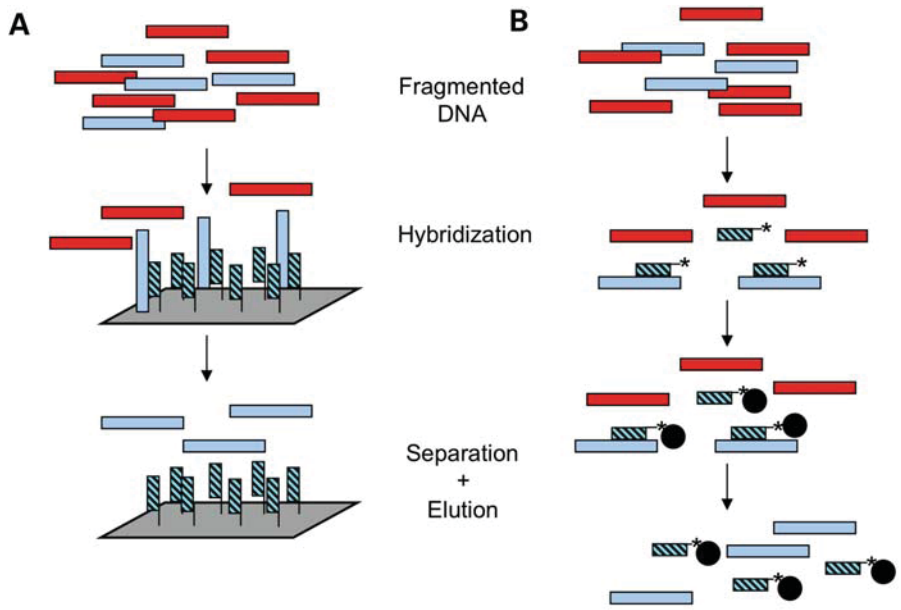

Figure 1.16: **Hybridisation methods for target DNA capture.** A) Solid-phase. Baits complementary to the targeted sequence (blue with black stripes) are synthesised on a microarray. Upon fragmentation, target DNA (blue) hybridises with the microarray baits, the array is washed and the captured DNA is eluted. B) Liquid-phase. Baits complementary to the targeted sequence (blue with black stripes) are synthesised and biotinylated, indicated by an asterisk. Fragmented DNA hybridises with the baits, and these are captured with streptavidin beads (black). The bead-bait complexes are washed and the target DNA eluted. Target DNA is then sequenced. Reprinted from ref. [374] by permission of Oxford University Press.

1.8.2 Overview of the methodology followed to study melanoma predisposition genes

I have organised the chapters in this dissertation roughly according to the sequential order in which I carried out the different steps in the study (Fig. 1.17). In order to

search for high-penetrance melanoma susceptibility genes, I used three main familial melanoma collections, originating from the UK, The Netherlands, and Australia. These datasets were not all available at the start of the project and thus the initial analyses were only performed with the UK and Dutch datasets (discovery and replication phases, covered in Chapter 2). Then, with the availability of the Australian cohort, I could perform an integrative analysis and identify candidate susceptibility genes (covered in Chapter 3). Finally, I cover the mechanistic investigation of the mode of action of these candidate genes in Chapter 4. Finally, in Chapter 5, I discuss the relevance of the results presented here, as well as the future directions of this project.

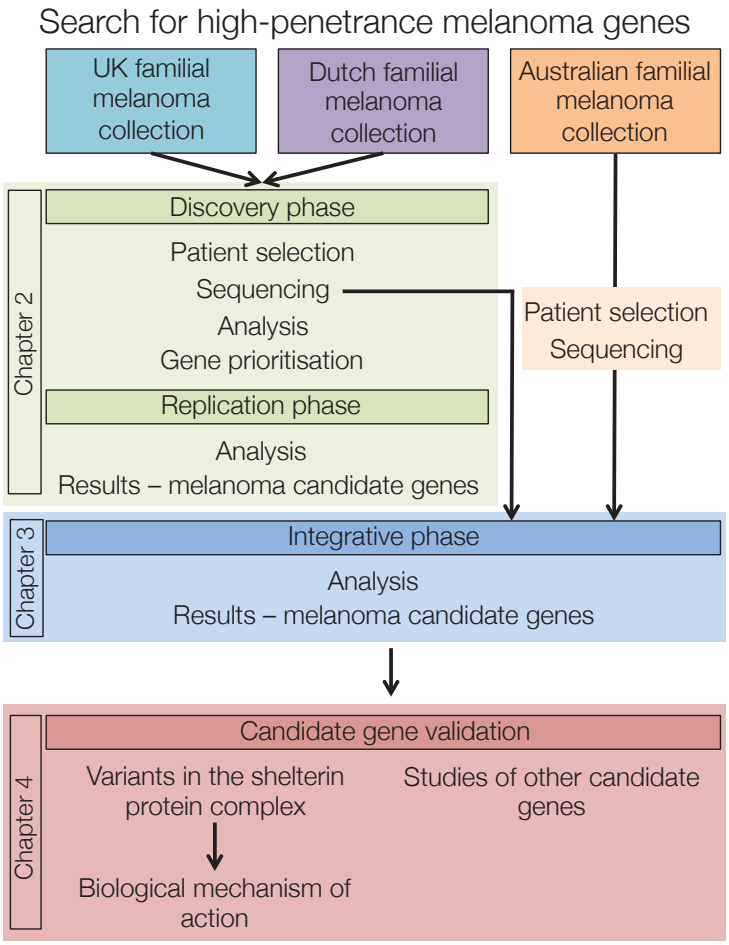

Figure 1.17: **Methodology followed during my PhD to study melanoma susceptibility genes.** The chapter in which each step is covered is indicated at the left. An arrow indicates a contribution of that dataset to the next analysis.