Chapter 5

Discussion

5.1 Summary of the work described in this dissertation

In this dissertation, I have described the methodology we followed to search for genes that predispose, with high penetrance, to the development of melanoma. This search started with whole-exome sequencing from several individuals from families with a high clustering of melanoma cases. We started an analysis of UK and Dutch families (European phase), but during the course of that study, a large collection of Australian pedigrees became available for analysis (integrative phase). For both phases, we implemented different gene prioritisation strategies, and then attempted to experimentally validate our findings.

The European phase was further divided into two parts, the discovery and replication phases. The discovery phase involved sequencing 24 pedigrees from the UK and The Netherlands and implementing a gene prioritisation strategy to define candidate melanoma susceptibility genes. This strategy took into account diverse criteria such as the number, position and consequence of gene variants at the protein level, the probability that these are shared among members of a given pedigree, their allelic frequency in catalogues of common human genetic variation, the likelihood of finding variants in these genes in a population matched by ancestry, and the biological function of the gene. We reasoned that a high score in these attributes would increase the likelihood that these are involved in melanoma predisposition.

Having a list of candidate melanoma susceptibility genes, we then sequenced these in an additional 92 pedigrees (replication phase). We implemented a novel prioritisation strategy taking into account results from both phases of the study, and validated by PCR variants in the highest-ranked genes. From this part of the study, detailed in Chapter

2, the highest-ranked candidate gene was *SMG1*.

For the integrative phase, we extended our dataset to include a large collection of Australian melanoma pedigrees, bringing the total number of pedigrees that had been whole exome- or genome-sequenced to 105. As this set included a large proportion of families with multiple affected members sequenced, we performed a new prioritisation strategy relying mainly on co-segregation of rare, potentially deleterious variants with melanoma. From this strategy, we prioritised two genes that contribute to telomere maintenance, *POT1* and *ACD*. I cover the analyses we performed in this phase in Chapter 3.

Although the experimental evidence we could gather for the involvement of *SMG1* and *ACD* in melanoma predisposition was inconclusive, we were able to establish a role for the variants within *POT1* in the aetiology of this disease. We found that *POT1* variants in these families either interfered with mRNA splicing or rendered the protein unable to bind to ssDNA, and therefore, affected its ability to mediate telomere protection and regulate telomere length. Accordingly, carrier individuals had telomere lengths that were significantly longer than *POT1* wild-type melanoma cases and non-carrier members of their own families. Variants affecting the ability of *POT1* to bind to telomeres seem to be extremely rare in the populations examined. I explain these analyses and biological assays in Chapter 4.

In summary, this work describes the methodology we followed to identify candidate melanoma susceptibility genes starting with exome sequencing data. It discusses the difficulties that are faced by large-scale sequencing projects investigating genetic diseases, namely, the diverse attributes that may or may not be indicative of the involvement of a gene in disease risk, problems with sequencing errors, the need for availability of control datasets and the suitability of biological assays to test genetic hypotheses. It also touches on complications arising from the genetic heterogeneity of the disease and the occurrence of phenocopies in familial studies. However, this dissertation also shows a successful example of the use of exome sequencing to pinpoint causal variants in a gene, and moreover, directly implicate a biological pathway in melanoma risk.

5.2 The identification of telomere dysregulation as an important contributor to familial melanoma

In this dissertation, I describe the identification of germline variants in *POT1* in more than 4% of familial melanoma pedigrees that do not carry variants in *CDKN2A* and *CDK4*, and in 2 out of 34 pedigrees (almost 6%) with 5 or more cases, making *POT1* the second most frequently mutated high-penetrance melanoma susceptibility gene reported thus far. At the same time we reported our results, another group based at the National Cancer Institute in the US independently reported French, US and Italian melanoma-prone families with rare germline variants in *POT1* [484]. They found that in families from Romagna, Italy, the frequency of rare *POT1* variants is comparable to that of *CDKN2A* variants, and identified a founder mutation occurring in five unrelated melanoma-prone families from this region. Carriers of this variant not only had longer telomeres than controls, but also a significant increase in the number of fragile telomeres, which indicates that these disruptive *POT1* variants alter not only telomere length maintenance but also telomere integrity. In accordance with these observations, susceptibility to cutaneous malignant melanoma due to disruptive variants in *POT1* has recently been included in the Mendelian Inheritance in Man catalogue of human genetic diseases (MIM #615848).

These results implicate telomere dysregulation as an important factor in the aetiology of familial melanoma. It had been recognised by previous studies that both short and long telomeres can be risk factors for cancer development. Some examples of telomere shortening syndromes due to deleterious mutations in telomere-associated proteins are dyskeratosis congenita, ataxia telangiectasia and the Bloom and Werner syndromes (reviewed in [200]). Individuals with these conditions have higher frequencies of cancer than the general population [485–487], possibly due to chromosomal breakage and telomere fusions.

Nonetheless, individuals with long telomeres are also cancer-prone. For example, it has been observed that both breast cancer cases and women at high genetic risk for developing the disease have longer telomeres than controls, with telomere length displaying a positive correlation with risk [488, 489]. In fact, *BRCA1* and *BRCA2* mutation carriers have recently been found to have longer telomeres than controls [490], and longer telomeres have also been associated with a worse prognosis in a subset of breast cancer patients with advanced disease [489]. Other cancers in which longer telomeres have been associated with increased risk are non-Hodgkin lymphoma [491],

and lung cancer in both smokers [492] and non-smokers [493]. Interestingly, a higher risk of developing melanoma, but not squamous or basal cell carcinomas, has also been associated with longer telomeres [494]. This risk is not only seen in sporadic cases, but in *CDKN2A*-negative familial cases as well [495], and in both studies, longer telomeres were associated with a higher naevus count. In accordance with these results, shorter telomeres have also been found to be associated with decreased melanoma risk [496].

The studies cited above report comparisons of telomere lengths, usually measured in whole blood or leukocytes, between groups of cancer cases and controls. Here I have described what is, to the best of our knowledge, the first hereditary mechanism underlying telomere lengthening in humans. It is known that cancer cells activate telomerase or the alternative lengthening of telomeres (ALT) pathway to bypass replicative senescence (reviewed in [497]), and thus cells from individuals with inactivating *POT1* variants might be behaving in a similar way: If *POT1* cannot inhibit telomerase, then cells might have a longer lifespan, allowing the accumulation of somatic mutations. A non-functional *POT1* might lead to progressive lengthening of telomeres, despite the telomere erosion that ensues every cell division.

In accordance with this idea, previous experiments have shown that in *POT1*-deficient cells, telomeres get progressively longer with each cell division [460, 462]. In fact, although we have limited evidence, it is tempting to speculate that families with *POT1* variants in both our study and that one from the US might be showing genetic anticipation, in the form of a higher number of primaries or an earlier age of onset with each successive generation. It has long been known that families with progressive telomere shortening show this effect due to shortening of telomeres in germ cells (reviewed in [498]), and it might be the case that *POT1* families display the same effect. However, assessment of a higher number of families and molecular studies will be necessary in order to address this question.

Moreover, telomere lengthening might not be the only mechanism by which *POT1* loss-of-function variants might contribute to tumourigenesis, but telomere dysfunction is likely to play an important role as well. As noted by previous publications [462, 484, 499], cells with defects in POT1 display loss of telomeric overhangs, chromosomal fusions and breaks, multitelomeric signals and fragile telomeres. Even so, this effect might not be severe enough to arrest cell division or lead to cell death, thus rendering *POT1* loss-of-function variants compatible with life. In accordance with this hypothesis, it has been shown that *POT1* knockdown in primary fibroblasts reduced their proliferative potential, but they nonetheless could continue dividing [459]. Interestingly, this effect

could be rescued by the abrogation of the P53 or INK4A/RB1 pathways, which could offer some clues as to the additional molecular events that are necessary for *POT1*-deficient cells to progress to malignancy. The effect of a *POT1* knockdown is better tolerated than that of *TERF2* (which leads to severe telomere fusions [459]), which might also explain why we did not find any deleterious variants in *TERF1* and *TERF2* when we examined 510 melanoma-prone families from around the world (discussed in Chapter 4).

An interesting observation arises from the examination of pedigree UF23 (Fig. 3.2d). Whereas the individual that was sequenced, III-3, is a carrier of the POT1 R273L variant, her half sister (III-1), who also developed melanoma, is not. However, these two individuals presented with very different clinical characteristics: The carrier of R273L presented with three primaries, the first one at 45 years of age, whereas the non-carrier presented with a single melanoma *in situ* at 62 years of age. Additionally, the POT1 R273L variant was found in one case part of the Leeds Melanoma Case-Control study, who presented with MPM and an early age of onset, similar to the phenotype of the familial cases. This fact could point to other cancer-predisposing genetic or non-genetic aspects being shared by the two individuals of pedigree UF23, or just to the occurrence of phenocopies within the same family. Notably, individual III-1 did not seem to have longer telomeres when compared to *POT1* wild-type melanoma cases (Fig. 4.10b), further implicating the involvement of telomere dysregulation in the development of a more severe phenotype.

It is still not completely clear whether the biological mechanism by which these *POT1* variants predispose to familial melanoma is the same as that one by which mutations in the promoter of *TERT* do so, but it is unlikely. Telomere length measurements for germline *TERT* variant carriers have not been reported in the literature, but our own bioinformatic measurements in the Leeds family that was found to harbour the *TERT* promoter variant did not indicate that they had telomeres that differed significantly from controls (data not shown). Although similar somatically-acquired variants in *TERT* do lead to its increased expression [500], *TERT* has many other roles outside telomere maintenance that could be contributing to the elevated cancer risk. For example, it can act as a transcriptional modulator of the Wnt-β-catenin signalling pathway and physically occupy Wnt-dependent promoters [501], and bind the RNA component of mitochondrial RNA processing endoribonuclease (*RMRP*), [502] possibly to regulate gene expression. In addition to this, individuals with activating variants in *TERT* would not, in principle, show the fragile telomere syndrome displayed by *POT1*-deficient

cells. More experiments will be necessary to address how *TERT* over-expression affects telomere maintenance, if indeed it does.

But why would variants in *POT1* lead to the development of melanoma? What is so special about the relationship between melanocytes and telomeres? This is indeed something we do not know at the moment, not unlike the question of why germline defects in the major tumour suppressor *CDKN2A* lead primarily to the development of melanoma, with other cancers manifesting at lower frequencies in carriers. It may be the case that defects in *POT1* also predispose to other cancers, as we can see in the carrier families (Fig. 3.2), and pedigrees with other cancers and germline variants in *POT1* will be described in the future. However, as case-control studies analysing melanoma risk factors have shown [494–496], longer telomere length seems to be associated with melanoma, but not squamous or basal cell carcinomas. It can also be that longer telomere length is indicative of a defect in the telomere maintenance machinery that also leads to fragile telomeres and a higher tendency to suffer telomere breaks and fusions, just as has been shown in *POT1* variant carriers. This indeed points to the idea that there is a special relationship between melanocytes and their telomeres.

Some clues as to the above question might come from experiments that have shown that DNA-damaging agents, including UV irradiation, cause damage throughout the genome but cause proportionately more damage in telomeric sequences [503]. Telomeres have been shown to be hypersensitive to UV-induced DNA damage, and to be refractory to DNA repair [504, 505]. Perhaps this effect could be exacerbated by the longer and unprotected telomeres in *POT1* variant carriers. This hypothesis would point to the involvement of other genes that contribute to UV-induced melanomagenesis, such as *TERT* and *TP53*, in melanoma tumours initiated by telomere dysfunction. However, future experiments should be performed to address the biological pathway, or pathways, by which these cells progress toward malignancy.

In conclusion, our study of the genomes of melanoma-prone pedigrees has led us to describe what is, to the best of our knowledge, the first mechanism underlying hereditary telomere lengthening in humans. This telomere lengthening is associated with high melanoma risk, and possibly other cancers. Many questions are posed by this discovery, for example, whether families with this defect display genetic anticipation, what other genetic hits are necessary for melanoma progression, and whether there is a reason why melanocytes would be more susceptible to telomere dysfunction or this is just an effect of sample ascertainment bias. The investigation of these questions might have the potential to facilitate better clinical management of families with *POT1*, and potentially variants

of a similar consequence in other genes, in the future.

5.3 Future directions

As mentioned above, there are many questions that need to be addressed relating to the manner by which individuals with germline variants leading to telomere dysfunction develop cancer. Initially, to help translate this discovery to clinical practice, we need to establish the true penetrance of the variants described in this study and the one from the US. In the work presented here, we did not find any *POT1* variant carrier that did not develop melanoma, whereas the US group determined that the founder variant they identified showed dominant inheritance with incomplete penetrance [484]. Whether this effect is related to the positions of the variants themselves (*i.e.*, the variants identified in our study all affect DNA-binding residues whereas the founder mutation identified by the US group does not) or whether we need a larger sample size to see the same effect remains to be determined. We also need to determine the penetrance and biological mode of action of the splice acceptor variant. Although we only have one family with this type of variant (Fig. 3.2b), its affected members seem to have presented with melanoma later in life than the missense carriers. This observation might arise from the small sample size, but it can also mean that a variant affecting splicing does not generate a protein able to compete with wild-type POT1 and perturb protein interactions at the telomere [465]. This could imply that the mechanism by which this variant predisposes to melanoma is not dominant-negative, but haploinsufficiency. We will be able to answer these questions, hopefully, as we gather additional data from families with these types of variants. We are in the process of doing so as part of the Melanoma Genetics Consortium (GenoMEL, `http://www.genomel.org/`).

A mouse model might be useful in elucidating the additional somatic variants that must occur in *POT1* variant carriers to develop melanoma. Within our group, Chi Wong has successfully created, via CRISPR/Cas9 technology [506], mice carrying variants orthologous to the human *POT1* Y89C and Q94E variants. Analyses of their consequences might be complicated by the fact that mice have much longer telomeres than humans [507] and two *POT1* orthologues, *Pot1a* and *Pot1b* [454]. Nonetheless, we hope to be able to recapitulate the human cancer-prone phenotype, just as it was possible with shortening telomere syndromes due to variants in *TERC* [508]. We can then seek to answer questions such as whether telomeres get progressively longer throughout life, the spectrum of cancer types that these mice are predisposed to and the contribution of UV

irradiation toward melanoma development.

However, mouse studies should be compared to and supported by studies of human tissue whenever possible, as it has been shown that mouse models not always recapitulate human disease perfectly (reviewed in [509]). Undoubtedly, the study of tumours from patients with *POT1* variants will prove invaluable as we endeavour to establish the genomic events leading to malignancy. We can then assess the contribution of UV-induced damage in tumour DNA and find recurrently mutated oncogenes or tumour suppressors, for example. As more patients with rare *POT1* variants are being identified by dermatologists part of GenoMEL, we hope to be able to obtain this important biological material in the coming months.

Additionally, a more in-depth investigation into the biological consequences of rare variants in *ACD* and *TERF2IP* should hopefully help us complete the picture of how telomere dysfunction predisposes to the development of melanoma. Presumably, they function as haploinsufficient alleles and lead to a malformed shelterin complex, as the variants that segregate with melanoma in these pedigrees introduce premature termination codons into the affected proteins. We hope to be able to address these questions by performing protein-protein interaction assays and introducing the detected variants into cells *in vitro*, thus being able to assess their contribution to biological processes such as cell cycle progression and telomere maintenance.

Finally, although we have identified a novel biological pathway relevant to melanoma predisposition, we cannot disregard the bigger picture: We still have 100 families in the Leeds and Leiden collections for which we have not identified any genetic predisposition loci. It could be that some of these have high-penetrance variants in genes that we have sequenced, such as *SMG1*, and we just need to identify a suitable biological assay to test them. One of the genes prioritised in the integrative phase, NIMA-related kinase 10 (*NEK10*), which plays a role in G2/M cell cycle arrest upon UV irradiation [510], is currently being investigated by our Dutch colleagues.

Variants in genes with roles in biological processes previously found to be important in melanoma development could also represent interesting candidates. In this category, we have variants segregating with melanoma within a single family in discs, large homolog 1 (*DLG1*), proteasome 26S subunit, non-ATPase, 3 (*PSMD3*), *PSMD12*, *PSMD13*, *AKT1*, minichromosome maintenance complex component 5 (*MCM5*) and nuclear protein, ataxia-telangiectasia locus (*NPAT*). All these genes encode proteins that play a role in the G1/S checkpoint of the mitotic cell cycle, the same pathway in which CDKN2A and CDK4 participate. Genes encoding proteins that participate in other cell

cycle regulatory tasks or cell differentiation could represent interesting candidates, such as centrosomal protein 250kDa (*CEP250*), *CEP290*, eukaryotic translation initiation factor 4E binding protein 2 (*EIF4EBP2*) or *SOX3*. All these genes have variants segregating with melanoma in the families we studied, and could represent plausible melanoma susceptibility candidates.

Additionally, it could also be that these families have high-penetrance predisposition variants in non-exonic regions, in which case we would not have the data as we did not sequence the non-coding genome. A good example of this type of variant is the activating mutations in the promoter of *TERT* [297, 298, 500], which have been found subsequently in a large number of cancer types. In order to address this question, we are in the process of whole genome-sequencing 29 familial melanoma pedigrees from the Dutch cohort. Other possibilities that we need to consider relate to the occurrence of several low-penetrance alleles within families or the contribution of non-genetic or epigenetic effects.

In conclusion, there are many questions we are looking forward to answer relating to the role of telomere dysregulation in melanoma susceptibility, and possibly other cancers. We expect to be able to investigate these as we gather more families with rare variants in *POT1* and related genes, and with the generation of mouse models and the study of human tumours from carriers. Additionally, we need to investigate the involvement of other biological processes, and the contribution of non-genic effects, in the remaining families for which have not been able to identify any predisposition loci thus far. Hopefully, as we gather more data and exploit alternative technologies such as whole-genome or bisulphite sequencing, we will be able to help complete the description of the processes that influence genetic susceptibility to familial melanoma.