Decoding telomere gene variation through saturation genome editing: from functional genomics to clinical interpretation

Obolenski, S.

English Summary

Telomeres and their associated protein complexes are essential for maintaining chromosome integrity. Variants that disrupt this system can predispose to cancer and to telomere biology disorders. In clinical practice, genes such as *POT1* and *TINF2* are now increasingly included on hereditary cancer gene testing panels. However, many identified variants lack sufficient functional and epidemiological evidence and are therefore classified as variants of uncertain significance (VUS). This limits the clinical utility of genetic testing by preventing definitive diagnosis, constraining risk stratification, and leaving clinical management and surveillance decisions unresolved for patients and their families.

This thesis investigates whether saturation genome editing (SGE) can provide systematic functional evidence for variants in telomere biology genes and whether such data can be integrated into existing clinical frameworks. SGE is a pooled genome-editing approach that introduces large numbers of variants into a gene at its endogenous genomic locus and measures their effects in parallel. Using *POT1* and *TINF2* as model genes, dense variant-effect maps were generated and integrated with telomere phenotypes, clinical observations, and population-level cancer risk.

Chapter 2 reviews available functional genomics approaches for variant interpretation and positions SGE within the broader landscape of CRISPR-based endogenous multiplexed assays of variant effect (MAVE), outlining the motivation for using SGE. **Chapter 3** describes the development of an SGE workflow in haploid HAP1-A5 cells, including variant library design, editing strategy and quantitative fitness-based readouts enabling reproducible scoring of large numbers of coding variants. In **Chapter 4**, this platform is applied to *POT1*, a key shelterin component with an established role in inherited melanoma and other cancers. Approximately 97% of all possible coding sequence single-nucleotide variants were assayed, resulting in a functional map showing that *POT1* variants form a continuum of functional effects, with strong and weak loss-of-function alleles clustering in the DNA-binding protein domains. SGE functional scores were validated using telomere length and DNA damage response phenotyping in HAP1-A5 cells and were associated with cancer diagnoses in UK BioBank and clinically ascertained patient cohorts, supporting a relationship between functional *POT1* impairment, telomere dysfunction, and cancer risk. **Chapter 5** extends the SGE approach to *TINF2*, focusing on exons encoding core interaction domains that harbour the majority of known cancer-associated variants. The SGE functional map confirms that established pathogenic and newly identified variants described in this thesis significantly

impair cellular fitness. Structural mapping indicates that disruption of TRF1, TRF2 and TPP1 interaction surfaces is particularly deleterious. Integration of functional data with available clinical and pedigree information supports a broader *TINF2*-associated tumour spectrum, with melanoma and thyroid carcinoma as core malignancies, as well as recurrent occurrence of rarer malignancies, including sarcoma, renal cell carcinoma and multiple brain tumour subtypes. **Chapters 6** and **7** assess the clinical application of SGE. In a population-based melanoma case-control study, functional classification distinguished rare *POT1* variants that disrupt telomere binding from numerous rare missense variants without measurable functional effect, showing that pathogenic *POT1* variants are uncommon in the general population and occur more frequently among melanoma cases. In Swedish melanoma families, SGE scores were used as supplementary evidence to aid interpretation of *POT1* variants detected through routine clinical testing, particularly in complex pedigrees with early-onset disease, multiple primaries and phenocopies.

Chapter 8 discusses all the findings of this thesis in the context of clinical variant interpretation and hereditary cancer genetics, demonstrating how calibrated SGE data can be integrated into the ACMG/AMP framework, reducing the burden of VUS and improving interpretation across populations. It also discusses limitations of the experimental model and outlines future directions, including expansion to non-coding variation, additional phenotypic readouts and integration with predictive approaches.

Overall, this thesis establishes SGE as a clinically relevant functional genomics framework for telomere biology genes. For the cancer predisposition genes *POT1* and *TINF2*, this thesis provides high-resolution functional maps that reflect underlying biology and identify domain-specific vulnerabilities, supporting interpretation of rare germline variants in clinical cohorts. By linking variant-level functional effects to telomere phenotypes and cancer risk, this work advances the clinical interpretation of *POT1* and *TINF2* variants and demonstrates how functional genomics can move beyond variant annotation to inform practical applications in hereditary cancer predisposition.