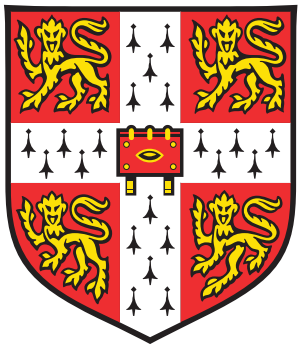
# DNA polymerase mutations as drivers of genome instability and cancer



**Mareike Herzog**

Wellcome Trust Sanger Institute

University of Cambridge

This dissertation is submitted for the degree of

*Doctor of Philosophy*

Christ's College

2016

“Think, think, think.”
- Winnie-the-Pooh

# Declaration

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

Mareike Herzog
2016

# Acknowledgements

First I would like to thank my supervisors Dave Adams, Steve Jackson and Thomas Keane for the opportunity to be a member of their teams, learn from them and earn my PhD with them. I want to thank them for the warm welcome I received, their patience, their kindness and their support. I am grateful to Dave for his many thoughts and advice on my work, for all his encouragement in stressful times and his trust in me, to Steve for never being too busy to talk to me and for the independence in exploring my projects and to Thomas for putting himself through the gruelling process of introducing me to bioinformatics.

I would like to thank all members of the Experimental Cancer Genetics Team at the Sanger Institute, for their friendship and their enthusiastic willingness to help me with anything I need: Daniela, James, Manu, Sofia, Mamun, Vivek, Martín, Clara, Stefan, Marco, Nicky, Chi, Louise, Rebecca, Marcela, Richard and Alistair. I want to especially thank James for supporting my with everything I could imagine: from getting me a desk, supplying me with reagents to getting me started in the world of tissue culture. Clara Alsinet Armengol and Manu Supper have my gratitude for teaching me how to take care of embryonic stem cells and for being great "babysitters" when I had to be away from the institute. My mouse work was greatly aided by Louise van der Weyden, who is the goddess of mouse work and a super friendly one at that. I am indebted to Mamun Rashid for his help with the EMu signature extraction and to Vivek Iyer for helping me manage import and storage of my mouse sequencing data. Martin Del Castillo Velasco Herrera and Stefan Dentro were providers of great advice for all things statistics and bioinformatics. And I cannot thank Daniela Robels Espinoza enough. Not only was she the patient and kind with all my questions, taught me endless programming tricks and helped me fix some tricky bugs, she also put up with me as a roommate. She has been an endless source of motivation, reassurance and encouragement for me.

I would also like to thank all of the members of the Jackson Group at the Gurdon Institute: Fabio, Israel, Josep, Yaron, Gabi, Rimma, Serena, Delphine, Paco, Paul, Will, Matt, Donna, Natasha, Pallavi, Chrstine, Andy, Sati, David, Carlos, Ryotaro, Jessica, Jon, Abdul, Linda, Nicola, Julia, Helen, Kate, Gopal, Matylda, Siyue, Muku and Ana. To highlight but a few, I will always remember Rimma Belotserkovskaya, because she was the first person from the lab

I met when she lectured me in my undergraduate days and got me interested in DNA repair, Julia Coates, because she shared a "bench" with me for years and we suffered the institute's temperamental air conditioning together and because she helped me move house no questions asked, Serena Bologna, for involving me in her project, being a great friend and speaking Italian to me, Linda Baskcomb, for being such a competent lab manager, and Nicola Geisler for supervising me during my rotation project, which convinced me to do my PhD in this group. Special thanks goes to Josep Forment, who invited me to collaborate on a rewarding project with him. He is a joy to work with and I will miss him greatly. I am also deeply grateful to Israel Salguero Corbacho, who helped me a lot with designing my plasmids and my cloning for the yeast strain construction. He is kind and funny and always happy to help me with anything. My deepest gratitude goes to Fabio Puddu, who, more than anyone, has been my daily mentor on this journey. From teaching me lab techniques (pulse-field gels are fiddly) to discussing and planning experiments (why would we run this analysis?), he has tried his best to make me a better scientist and I can only hope that somewhere along the way I was able to teach him something in return. I am grateful for all the criticism, encouragement, laughter, patience, praise and support.

I would like to acknowledge the Wellcome Trust for funding my PhD. Without their generous support this work would not have been possible and their funding has allowed me the privilege of not having to worry about feeding myself, paying the bills and having a roof over my head.

My friends have filled my days with joy and were always there with advice when needed and ultimately contributed much, that is intangible, to this project. I am very happy to count Chrissey, Böcki, Stephie, Monica and Julia as some of my best friends. Finally, I want to thank my family. My parents, Martina and Stefan, and my "baby" brother Gunnar, for their never ending support, motivation and their infinite love. And to Löffelchen I would just like to say with all my love "I couldn't have done it without you. Ich habe dich über alles lieb."

# Abstract

Genomic stability is essential to preserve the genetic information encoded in DNA, and many biochemical pathways are devoted to repair DNA damaged by external factors, or during the course of essential cellular processes such as transcription and DNA replication. Malfunctioning of these processes may alter the DNA, leading to abnormal cellular behaviour or cell death, which in multicellular organisms may be associated with disease. For this reason, the machineries that safeguard the integrity of eukaryotic genomes are of prime interest to research in the areas of ageing, rare disease and cancer. Every time a cell divides, duplication of the genome is principally carried out by two DNA polymerases — Pol $\delta$ and Pol $\varepsilon$ — which are highly processive and accurate. Together with polymerase gamma, which is active in mitochondria, these are the only human polymerases known to possess "proofreading" activity, making them extremely accurate. In parallel, cells have also evolved a repair system for base mismatches, to identify and correct mispaired bases occasionally produced by DNA polymerases. While it has been known that defects in mismatch repair promote carcinogenesis, mutations in replicative DNA polymerases driving tumorigenesis in mismatch repair proficient cells have only been recently identified. Here, I report the interrogation of twelve such DNA polymerase mutations for their potential to alter genetic information and contribute to genomic instability using the budding yeast *Saccharomyces cerevisiae* as model system. Of all the polymerase mutations tested, a subset caused significant increases in mutation accrual, and a shift in the observed mutation patterns/signatures. Most intriguingly, I observed that these increases are more severe than those caused by mutations disrupting the proofreading activity of the corresponding DNA polymerase, with my results further indicating that in some cases the high mutagenic potential depends on the proofreading activity. These strong increases in mutation rates do not likely result from inhibition of mismatch repair, as combination of these mutations with loss of mismatch repair factors results in synthetic sickness or lethality. My results point to these DNA polymerase mutations as driving extensive alterations of the genetic information, and are consistent with them being drivers of colorectal and endometrial cancer. Future work will be required to determine the exact mechanisms by which these mutations impair the fidelity of DNA replication.

# Contents

# List of Figures

# List of Tables