**Acknowledgement**

*You can't connect the dots looking forward; you can only connect them looking backwards. So you have to trust that the dots will somehow connect in your future. You have to trust in something - your gut, destiny, life, karma, whatever.*

[Steve Jobs' 2005 Stanford Commencement Address]

This PhD thesis gives me the opportunity to reflect on my past and recognise the books, the events and people who have helped me to become who I am.

As a child, I was initially drawn to physicists with their acumen and ability to describe part of Nature with mathematics and later, I was inspired like many others to study the software of life and the manifestation of that software after reading *What Is Life* by Erwin Schrödinger. Three other books (*Genentech: The Beginnings of Biotech* by Sally Smith Hughes, *Life at the Speed of Light: from the Double Helix to the Digital Life* by J. Craig Venter and *The Billion-Dollar Molecule: The Quest for the Perfect Drug* by Barry Werth) also springs to my mind when I am asked which books inspired me to become a scientist. I don’t know why, but I must have always loved the idea of a group of people working towards a shared goal to not only improve their understanding of the world, but to positively transform the lives of other people.

As an undergraduate studying biochemistry at Imperial College London, starting and finish a PhD degree was a distant dream and countless number of people have helped me achieve what I thought was impossible. My words cannot fully express my gratitude towards people who have helped me on my journey.

First, I would like to thank my parents. They have always believed in me. They have invested in my education. They have showered me with their care and attention. What I appreciate the most is that they did not ask me to conform to the social norms and instead they cultivated fierce independence to say no when it was necessary and to challenge and verify what I was taught and to have a healthy scepticism for everything. I could not have asked for a better family.

Second, I would like to thank Anny King, Rebecca Sawalmeh and Veronica McDouall for their care and warmth during my graduate studies at Churchill College. I still fondly remember weekly teak breaks with Anny, and light-hearted conversations with Rebecca. I absolutely could not have completed the MPhil in Computational Biology without their support. In the past, I dreaded waking up and I mightily struggled to complete the computational assignments. Now, I relish at the opportunity to design and implement new methods to explore the unexplored biological phenomena. How the tables have turned!

Third, I would like to thank Professor Jeong-Sun Seo, Chairman of Macrogen, for providing the opportunity to participate in the Korean Genome Project as part of my national service. I had no prior experience in sequence analysis, but he took a chance on me. I had the immense fortune to use the latest sequencing and genome mapping technologies to assemble chromosome-length scaffolds of the Korean reference genome. I cannot emphasize enough how important this research experience has been in increasing both my breadth and depth of knowledge and influencing the direction of research.

Fourth, I would like to thank University of Cambridge and Wellcome Sanger Institute for the generous PhD studentship, creating an environment where I can be dedicated to research and providing the infrastructure to ask and answer original scientific questions. When I stroll through Cambridge, I am always in awe of the architecture and the fact I could breathe the same air and walk the same grounds as other great scientists who laid the foundation for human genomics.

Fifth, I have nothing but sincere gratitude towards my three supervisors Peter Campbell, Richard Durbin and Raheleh Rahbari for the opportunity to ask and answer original scientific questions. I had the unbelievable fortune to tackle three amazing questions: is genome-wide single molecule somatic single-base-substitution detection possible? If single molecule somatic mutation detection is possible, is single molecule structural rearrangement detection possible as well? What is the germline and somatic mutational process across the Tree of Life? I still cannot fathom the sequence of events that led me to this fortunate circumstance. I was the only PhD student in my year who was interested in exploring the capabilities and applications of PacBio circular consensus sequencing and Peter had the brilliant idea to assess the possibility of single molecule somatic mutation detection with PacBio CCS reads with samples with single ongoing somatic mutational process. An amazing opportunity presented itself and I was the only person who wanted to pursue it. I might not have another opportunity to work with such great supervisors and I wanted to record what I learnt and what I appreciated from them for perpetuity. I think they believed more in me than I believed in myself and their confidence in me in turn motivated me to push myself and to burn the midnight oil. I cannot count the number of times I wondered if someone else might have been better suited to complete the projects. What I appreciated the most is that they had the courage to ask and attack the important questions and had the patience for me to make the mistakes and learn from mistakes such that I have ownership of my projects. I have been to many labs and I could not have had a better PhD and supervision elsewhere.

Sixth, I would like to thank my mentor for his wisdom and friends from high school (Anuran Makur, Gaurav Kankanhali, Jinseok Lee, Jisoo Kim, Kok Weng Chan and Victor Trisna), Imperial College (Claire Rebello, Euikon Jeong, Jiyae Kang, Jiyoon Kim, Jongseok Ahn, Quentin Godefroi, Rebecca Yu, Seonwook Park, Soo Young Yoon, William Gao, Woochan Hwang and Yunsung Na) and University of Cambridge (Dongseok Kim, Emily Sellman, Haerin Jang, Hans Werner, Hyesoo Lee, Ioana Olan, Ju An Park, Juyeon Heo, Kwon Juneyoung, Layla Hosseini-Gerami, Michal Tykac, Omid, Rob Henderson, So Yeon Kim, Sul Ki Park, Sunwoo Lee) for their continued friendship. Anuran and Gaurav have already completed their PhD and have started their assistant professorship at Purdue University and University of Pittsburgh, respectively. Jinseok just started his PhD at University of North Carolina at Chapel Hill and I have no doubt he will graduate with flying colours.

Seventh, I would also like to thank colleagues from Macrogen (Junsoo Kim, Chang-Uk Kim) and Wellcome Sanger Institute (Chiara Bortoluzzi, Chloe Pacyna, Emily Mitchell, Haynes Heaton, Hyunchul Jung, Jongeun Park, Jun Sung Park, Kenichi Yoshida, Lori Kregar, Matthew Young, Mike Spencer Chapman, Rashesh Sanghvi, Sigurgeir Olafsson, Thomas Mitchell, Thomas Oliver and Yichen Wang) for the stimulating conversations. A special mention goes to Mike Spencer Chapman and Heaton Haynes who were instrumental in maintaining my physical and mental health through regular afternoon runs and pair programming, respectively.

I will dearly miss my time at the University of Cambridge and Wellcome Sanger Institute.

If I have forgotten anyone in haste, you have my sincere apologies.