Somatic mutations in the pancreatic islets

Supervisor: Iñigo Martincorena

Pantelis Andreas Nicola

King's College

University of Cambridge

July 2018

This dissertation is submitted for the degree of
Master of Philosophy

Preface

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.

It does not exceed the prescribed word limit prescribed by the Degree Committee for the Faculty of Biology. The word count is 19,956 exclusive of tables, footnotes, bibliography and appendices.

Abstract

The endocrine pancreas is composed of the islets of Langerhans. These micro-organs play a crucial role in glucose homeostasis by producing and regulating insulin and glucagon secretion. Dysfunction of these islets is part of the pathogenesis of diabetes mellitus. The global health burden of diabetes mellitus is growing, with an estimated one in eleven adults affected worldwide. A better understanding of the development and maintenance of the pancreatic islets could prove crucial to reversing this trend.

Whilst somatic mutations have been studied extensively in tumours, the exploration of normal tissue is still in its infancy. Here I present a novel workflow, using laser capture microdissection, whole-genome sequencing and innovative bioinformatics to accurately identify somatic mutations in healthy pancreatic islets. By sequencing 32 islets, all from a single individual, this work reveals islets to be polyclonal units formed by multiple embryonic founding lineages that often come to be dominated by one or two major lineages. The very low mutational burden observed here suggests these islets expand very early in embryogenesis and do not undergo large clonal expansions later in life. This is consistent with the islets being a slowly dividing tissue during adulthood, making it less likely that islet neogenesis or replenishment by a small number of progenitors, occurs in adults.

The 32 islets sequenced here all share a single common ancestor, one that also gives rise to the bladder urothelium. This is presumably the first cell that gave rise to all adult tissues. Spatially, nearby islets are more genetically similar than distant islets. This pattern demonstrates that different embryonic lineages contribute disproportionately to the islets in different areas of the pancreas, the implication being that this emerges during embryonic pancreatic development, or in adulthood through hypothetical islet fission events.

The translational potential of this line of work is substantial. Using this approach to understand the maintenance of islets in diabetes could yield a greater understanding of the pathogenesis, and whether somatic mutations could play a role in the disease. Applications to other normal tissues could similarly refine our knowledge of their development, maintenance and disease, with exciting prospects for clinical application.

Acknowledgements

This work could not have been possible without the kind generosity of the transplant donor, who so altruistically donated their tissues for research. The surgical team, led by Kourosh Saeb-Parsy at Addenbrooke's Hospital, Cambridge, collected the donor specimens and Yvette Hooks processed and sectioned them. The low-input sequencing pipeline was designed by Pete Ellis while the laser capture microdissection pipeline was developed by Luiza Moore, who also gave an expert histological opinion on all samples used in this project. Laura O'Neill was responsible for managing the sequencing submissions, which was undertaken by the Sanger Sequencing Core. The Cancer, Ageing and Somatic Mutation (CASM) IT department were central to the data analysis used here, as was Mathijs Sanders, who crafted the in-house variant filter used in this work.

CASM, and the Wellcome Sanger Institute, has been a stimulating environment full of friendly colleagues and I thank all those to whom I've spoken. Alex Cagan and Andrew Lawson have both been instrumental throughout this work and have shown themselves to not just be exceptional scientists, but natural-born mentors. Federico Abascal proved to be an encouraging teacher and I thank him for his help with the clustering analysis (section 4.10 and 8.4). Tim Coorens has been the backbone of this project and his computational input is omnipresent throughout this work, particularly in the design of the beta-binomial filter (section 3.10) and the production of the phylogenetic trees (section 3.15 and 4.13).

My thesis committee, made up of Peter Campbell and Kourosh Saeb-Parsy, provided critical insights for which I'm grateful. Additionally, I thank Peter Campbell for writing the clustering algorithm used here (section 3.15 and 8.1). My supervisor, Iñigo Martincorena, deserves the highest praise for taking me on with such passion and nurturing my development as a budding clinician-scientist. The lessons I've learnt from him, not just in research, will stay with me through the years.

Finally, I am humbled by the generous support that King's College, Cambridge and the Isaac Newton Trust have provided me with this year. To my family, friends and loved ones, you have all been unparalleled in your unconditional love and support and I dedicate this work to you all, with the hope that one day I can repay the sentiment.

Highlights

1. The whole genomes of 32 pancreatic islets were sequenced from a single human donor.

2. An unmatched analysis proved efficient in the removal of germline variants and artefacts, as well as accurately calling somatic mutations, including those occurring in early embryonic development.

3. The observed somatic mutation burden in the pancreatic islets is low and is driven by intrinsic mutational processes.

4. Almost all somatic mutations identified appeared to have no functional impact.

5. Islets are polyclonal units.

6. Pancreatic islets and bladder urothelium share the same most recent common ancestor.

7. Multiple embryonic founders establish each pancreatic islet but islets develop major and minor lineages.

8. Pancreatic islets do not appear to be maintained by a rapidly-dividing stem cell population, but whether there are multiple stem cell populations, or self-duplicating islet cells, needs further study.

9. The spatial distribution of islets and their embryonic lineages reveals their founding cells are non-randomly distributed.

10. There are potential applications of this work to the fields of tissue development, maintenance and disease. The possible role of somatic mutations in diabetes mellitus is a target of future research.

Contents