Regulation of gene expression in macrophage immune response

Kaur Alasoo

Wellcome Trust Sanger Institute
Hughes Hall

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy
September 2016

Abstract

Regulation of gene expression in macrophage immune response

Kaur Alasoo

Gene expression quantitative trait loci (eQTL) mapping studies can provide mechanistic insights into the functions of disease-associated variants. However, many eQTLs are cell type and context specific. This is particularly relevant for immune cells, whose cellular function and behaviour can be substantially altered by external cues. Furthermore, understanding mechanisms behind eQTLs is hindered by the difficulty of identifying causal variants. We differentiated macrophages from induced pluripotent stem cells from 86 unrelated, healthy individuals derived as part of the Human Induced Pluripotent Stem Cells Initiative. We generated RNA-seq data from these cells in four experimental conditions: naïve, interferon-gamma (IFNɣ) treatment (18h), *Salmonella* infection (5h), and IFNγ treatment followed by *Salmonella* infection. We also measured chromatin accessibility with ATAC-seq in 31-42 individuals in the same four conditions. We detected gene expression QTLs (eQTLs) for 4326 genes, over 900 of which were condition-specific. We also detected a similar number of transcript ratio QTLs (trQTLs) that influenced mRNA processing and alternative splicing. Macrophage eQTLs and trQTLs were enriched for variants associated with Alzheimer's disease, multiple autoimmune disorders and lipid traits. We also detected chromatin accessibility QTLs (caQTLs) for 14,602 accessible regions, including hundreds of long-range interactions. Joint analysis of eQTLs with caQTLs allowed us to greatly reduce the set of credible causal variants, often pinpointing to a single most likely variant. We found that caQTLs were less condition-specific than eQTLs and ~50% of the stimulation-specific eQTLs manifested on the chromatin level already in the naive cells. These observations might help to explain the discrepancy between strong enrichment of diseases associations in regulatory elements but only modest overlap with current eQTL studies, suggesting that many regulatory elements are in a 'primed' state waiting for an appropriate environmental signal before regulating gene expression.

Declaration of Originality

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the beginning of each chapter. t is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution. This dissertation does not exceed the word limit set by the Degree Committee for the Faculty of Biology.

Signature:

Date:

Kaur Alasoo
September 2016

Acknowledgements

First, I am grateful to many people who have inspired and nurtured my passion for science. I thank Hedi Peterson for drawing me into the field before I had even finished my first semester at the university. I thank Phaedra Agius for introducing me into machine learning and teaching me that it is completely normal in research not to know beforehand what works and what does not. I thank Prof. Jaak Vilo for supporting and guiding me during my undergraduate years, Harri Lähdesmäki for letting me play with new and exciting data, and Isabel Sá-Correia for giving me an independent experimental project with no prior experience. All of these experiences gave me the necessary skills and confidence to help me realize that research was what I wanted to do.

Many people have made the four years of my PhD studies such a rewarding and stimulating experience. Over the past four years, fair and constructive feedback from my supervisor Daniel Gaffney has greatly improved the clarity and focus of my research and writing. I also greatly value the independence that I had in shaping the projects that I was working on. I thank my co-supervisor Gordon Dougan for accepting me to his lab even though I had never seen neither human nor a bacterial cell before. Subhankar Mukhopadhyay was very patient in explaining both intricacies of immunology as well as basics of cell culture. Finally, most of the experimental work presented in this thesis would not have been possible without the dedication and hard work of Julia Rodrigues. Thank you, Julia, for not giving up even when cells were dying and differentiations failing. I will always remember the fruitful discussions with other members of the Gaffney on a wide range of experimental on computational topics.

I have received great advice from many other people on the campus. Leopold Parts has been a great mentor by listening to my ideas, thinking about them hard and asking important questions. Conversation with Leo have made me to reconsider my experimental design more than once, ultimately allowing me to answer more important questions.

Last but not least, I am most grateful to my family. Aitäh isale nende õhtuste matemaatika ja füüsika olümpiaadiülesannete lahendamiste eest, mis olid küll väga põnevad, aga millega ma ise alati esimesel katsel hakkama ei saanud. Aitäh vennale, õele, vanaemadele, vanaisadele, tädile ja teistele suurtele ja väikestele sugulastele pidevalt meelde tuletamast, kui hea on kodus olla. Aitäh, Maret, et oled minu kõrval olnud ja mind toetanud kogu selle aja.

Table of Contents