THE CAMBRIDGE STATISTICS DISCUSSION GROUP

Monday 3rd April 2017 7:15 for 7:45

The Large Seminar Room, Institute of Public Health,
University Forvie Site, Robinson Way, Cambridge, CB2 0SR

Mapping regulatory variation in human cells

Daniel Gaffney
Sanger Institute

Abstract: Association mapping of cellular traits such as gene expression can provide powerful insights into the functions of human genetic variation. However, until recently mapping studies have been limited to a very restricted range of tissues or cell states. An additional limitation of association mapping is that, depending on the structure of linkage disequilibrium, the causal variant that drives an association can be extremely difficult to pinpoint. I will discuss work from our group that attempts to address some of these limitations. First I will discuss our work, as part of the Human Induced Pluripotent Stem Cells Initiative (HIPSCI: www.hipsci.org), to develop IPSCs as model systems for understanding the functions of human genetic variation. I will show how variation due to the cell line donor contributes to heterogeneity at the level of the epigenome, transcriptome and proteome of IPSCs, and illustrate how individual quantitative trait loci (QTLs) have confounded past studies of human pluripotent stem cells. Second, I will discuss a new statistical method (RASQUAL) that uses allele-specific signals to improve power and fine-mapping accuracy in association analysis of read-based phenotypes. I will show how RASQUAL handles a range of technical and biological biases in allele-specific signals without requiring data filtering or removal, and significantly outperforms all other existing methods. I will finish by discussing our application of RASQUAL to ATAC-seq data in human LCLs to pinpoint likely causal variants in disease, and to uncover long-range interactions between different regulatory regions .

Speaker: Daniel earned his PhD in evolutionary genetics from Edinburgh University in 2006 studying variation in the mutation rate and natural selection in noncoding DNA. From 2006 to 2008 he pursued a postdoc at McGill University and Genome Quebec Genome Centre, where he worked on the evolution of transcriptional regulation and alternative splicing in mammals. From 2008 until 2011 he worked on population genetic variation in gene expression and regulation at the University of Chicago. In 2011 Daniel started at the Wellcome Trust Sanger Institute becoming Group Leader in 2015. The goal of the group is to understand the molecular and cellular consequences of genetic changes in gene regulatory regions.

Directions: (From Central Cambridge) Turn right off Hills Road into the Addenbrooke's site then turn left at the hospital roundabout onto Robinson Way. Follow Robinson Way until you see an access road on the left signed 'Forvie Site' (but note that the sign is on your right). Turn into the access road and follow signs - first to the Institute and then to the Large Seminar Room. There is ample car parking. The front doors will be locked at 7:45. Arrivals after 7:45pm can gain admittance by contacting the secretary on 07761769436.

Provisional Next Meetings:
3rd May – Alun Bedding (Roche Products) on 'Innovative statistical approaches for studies in anti-infective drug combination development'.
10th October – Zoubin Ghahramani (Engineering).
20th November – Adam Kashlak (Statistical Laboratory).
5th February 2018 – Anthony Edwards (Gonville & Caius) on 'Cambridge Statistics from Venn to Fisher and Beyond'.
Supper: Some members eat regularly in the University Centre before each meeting at **5-45pm**. Feel free to join them.
Subscriptions: of 1 pound are now due for attending the 2016-2017 session.
Secretary: Peter Watson, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF;
telephone 01223 355294 Extension 801; E-mail peter.watson@mrc-cbu.cam.ac.uk

Slides and .mp3 files of old talks: http://www.mrc-cbu.cam.ac.uk/people/peter.watson/csdg.html