Centre for Genetic Epidemiology and Biostatistics

“Advancing global health through the application of statistical genetics, bioinformatics, functional genomics and human phenomics”

Senior Lecturer Phillip E. Melton
Statistical Genetics
Two universal principles of human genetics

1. Virtually all traits, including diseases, have a genetic component

2. There are no perfect humans - all of us carry a significant number of DNA variations
Centre Focus

- A major focus is the genetic basis of common complex human conditions such as cardiovascular and metabolic diseases, respiratory diseases, neuropsychiatric disorders, sleep disorders, pregnancy disorders, cancers....
A) Monogenic versus B) Complex Disease
Why should we try to identify the genetic risk factors?
Centre: Skills & Resources

- Statistical Genetics
- Bioinformatics
  - The Ark
- Molecular/Functional Genetics
  - WA DNA Bank
  - NGS
- Supercomputing
  - iVEC (UWA & Murdoch)
Current Centre Projects/Initiatives

- Family-based genomics/whole genome study for schizophrenia (with Prof Assen Jablensky and others)
- An adolescent cohort-based study on genetic risk factors for metabolic syndrome using lipidomic phenotypes (with Raine Study investigators)
- Chronic disease related to ethnicity & diet in Aboriginal and Torres Strait Islander people (with Assoc Prof David Whyatt from UWA and W/Prof Jenny Blackwell at TICHR)
- A family-based population genomics study for common disease on Manus Island/Peres village, Papua New Guinea (with USA investigators, John Blangero, Ellen Demerath & others)
- Case/control cohort study on genetic markers for melanoma severity and progression using transcriptomic phenotypes (with WAMHS, QIMR and TICHR investigators)
- Several projects around pre-eclampsia and pre-term birth (with Craig Pennell, Women’s and Infant’s Health, International collaborators in US, Scandinavia, Ireland)
- Family-based population genomics study for common disease in WA ‘Towards a Busselton Family Genome Project’ (with Busselton investigators)
- Meta-analysis of 150 publicly available genome-wide methylation datasets to look for signatures associated with cancer, aging, gender and ethnicity (with Lion’s Eye Institute investigators)
- GWAS that identified several regions/genes that appear to be associated with mesothelioma risk and currently preparing to resequence these regions using NGS (with GUARD investigators)
What is Next Generation Sequencing (NGS)?

- Approach to sequencing that has triggered numerous ground-breaking discoveries and ignited a revolution in genomic science.
- Small fragments of DNA are sequentially identified from signals emitted as each fragment is re-synthesized from a DNA template strand.
- NGS extends this process across millions of reactions in a massively parallel fashion, rather than being limited to a single or a few DNA fragments.
- Enables rapid sequencing of large stretches of DNA base pairs spanning entire genomes.
- Latest instruments capable of producing hundreds of gigabases of data in a single sequencing run.
What do we do with all these Data?

- NGS data output has increased at a rate that outpaces Moore’s law, more than doubling each year.
- In 2007, a single run produced a maximum of around one gigabase (Gb).
- By 2011, rate has reached one terabase (Tb) of data in a single run.
- 1000× increase in 4 years.
- Now sequence a human genome in one run, producing data in one week, for a cost of ~ $1,000 per genome.
- First human genome required 10 years to sequence and 3 years to analyse cost $3 billion.
What is Statistical Genetics?

- Development and application of mathematical models to identify the genetic structure of biological components influencing health and disease.
- Significant interest in the development of models for the understanding of genetic history.

My Research Interests

- Development of statistical models for joint and/or longitudinal analysis of genetic traits.
- Identification of population variation of genetic traits influencing cardiovascular disease and associated quantitative risk factors.
Diabetes, Heart Disease, and American Indians

- American Indians and Alaska Natives have the highest age-adjusted prevalence of diabetes among all U.S. ethnic groups (16.1%).*
- 95% of American Indians with diabetes have Type 2. Increase of 68% from 1994 to 2004 in youth aged 15-19.*
- Heart Disease is leading cause of death in American Indians and Alaska Natives.*
- Strong Heat Family Study
  - 3500 American Indian participants from three US regions
    - 1) Arizona;
    - 2) Oklahoma
    - 3) North and South Dakota.

*Source: American Diabetes Association
Bilirubin

- Endogenous antioxidant that suppresses lipid oxidation and retards atherosclerosis formation.
- Inverse relationship with heart disease.
- Levels vary with gender, ethnicity, and smoking status.
- Principal product of heme degradation and levels are regulated by five enzymes.
  - HMOX1
  - BLVRA and B
  - SLCOB1
  - UGT1A1
    - Number of TA repeats in UGT1A1 promoter is negatively associated with transcriptional activity.
**UGT1A1*28, American Indians and Bilirubin**

Mean serum bilirubin levels by *UGT1A1* genotype

---

**Graph:**
- **X-axis:** Chromosome Position (cM)
- **Y-axis:** LOD Score
- **Lines:**
  - Red: Combined LOD
  - Dashed Red: Combined LOD with UGT1A1
  - Blue: Oklahoma LOD
  - Dashed Blue: Oklahoma LOD with UGT1A1

---

**Legend:**
- *1/*1
- *1/*28
- *28/*28

---

**Figure Sources:**
- Melton et al. 2011

---

**Table:**
- **Bilirubin Levels (umol/L):**
  - **Arizona:**
  - **Dakotas (SD):**
  - **Oklahoma (SD):**
  - **Combined (SD):**

---

**Footnote:**
- The University of Western Australia
UGT1A1 and Pharmacogenetics

- While UGT1A1 promoter polymorphism is clearly associated with bilirubin there are other functional variants involved in American Indians.
- Promoter polymorphism is common in Europeans but less prevalent in Asians where different exonic variant is present.
- UGT1A1 involved in glucuronic pathway, so has important implications for pharmacogenetic research.
  - Promoter and exon variant recommended for genetic testing by US Food and Drug Administration for Irinotecan.
- Little is known about variants in the region for American Indians.
Exome Sequencing in American Indians

- Current Research
  - Identify functional variants in a subsample of SHFS American Indians from Oklahoma using exome sequencing.
  - Assess genetic association between potential functional variants for glucuronidation and bilirubin levels.
  - Investigate identified association levels with genome wide transcript levels from whole blood.

- Preliminary Results
  - Does not appear that the UGT1A1*6 variant is detected in American Indians or is very rare.
  - Have identified two rare functional variants: one known (rs57307513) and one unknown in UGT1A1 for American Indians.

- Lots and lots of work still needs to be done.
Acknowledgements

Centre for Genetic Epidemiology and Biostatistics

- W/Prof Eric Moses
- StatGen Team
  - Senior Lecturer Phil Melton (ex Texas Biomed)
  - Dr Nina McCarthy (ex RCSI, Ireland)
  - Dr Gemma Cadby (ex OICR, Canada; UWA graduate)
  - Dr John Blangero (visiting Scientist)
- Bioinformatics team
  - Paul White
  - Travis Endersby
  - Chris Ellis
  - Juan Manuel Peralta (ex Texas Biomed; PhD candidate)
- Molecular/Functional Genetics Team
  - W/Prof Eric Moses
  - W/Prof Lawrie Abraham
  - Dr Tegan McNab

Texas Biomedical

- Laura Almasy
- John Blangero
- Karin Haack
- Harald H. Goring
- Sandy Laston
- Jean W. MacCluer
- Shelley A. Cole

Strong Heart Study

- Jason G. Umans
- Elisa T. Lee
- Richard R. Fabsitz
- Richard B. Devereux
- Lyle G. Best