eQTL identification and mapping in the population isolate of Norfolk Island

Miles Benton

Genomics Research Centre, Griffith Health Institute

Biomarker Discovery Conference
Shoal Bay, NSW, Australia

3-6th December 2012

Outline

1. Background
2. eQTL Mapping
3. Other Findings
4. Future Directions
5. Acknowledgements

Rationale

- eQTL's and eQTL mapping
  - genomic loci that regulate expression levels of transcripts (mRNAs)

Rationale

- eQTL's and eQTL mapping
  - genomic loci that regulate expression levels of transcripts (mRNAs)
- Why are we doing this study?

Rationale

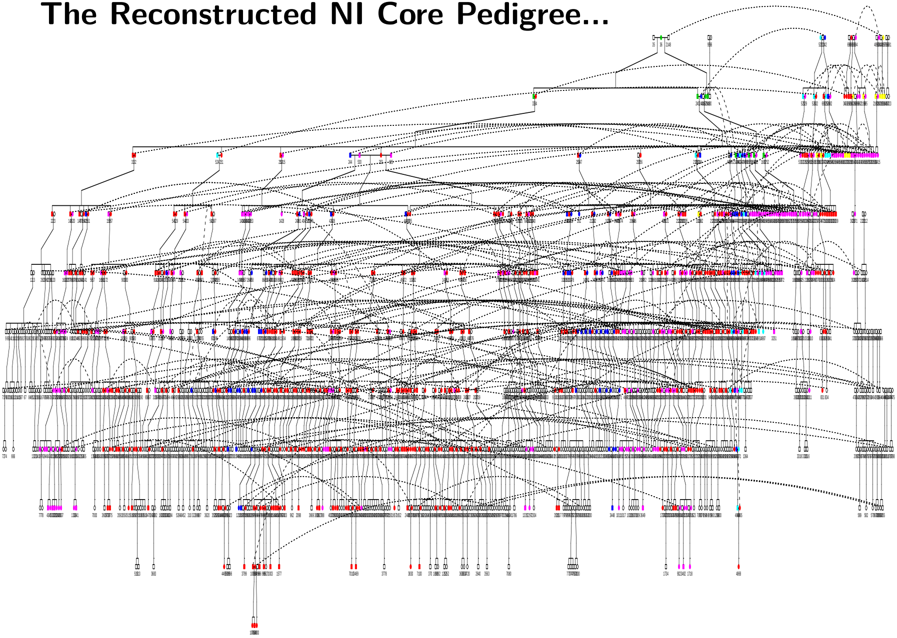
- eQTL's and eQTL mapping
  - genomic loci that regulate expression levels of transcripts (mRNAs)
- Why are we doing this study?
- Use the unique genetic isolate of Norfolk Island to hone in on functionally relevant loci

Rationale

- eQTL's and eQTL mapping
  - genomic loci that regulate expression levels of transcripts (mRNAs)
- Why are we doing this study?
- Use the unique genetic isolate of Norfolk Island to hone in on functionally relevant loci
- using gene expression and SNP assocication

A brief history lesson...

The Reconstructed NI Core Pedigree...



Study Design

- Participants - 330 NIHS individuals
    1. Samples - Blood (circulating lymphocytes)
    2. mRNA extracted > cDNA > expression analysis

Study Design

- Participants - 330 NIHS individuals
  1. Samples - Blood (circulating lymphocytes)
  2. mRNA extracted > cDNA > expression analysis
- Platforms:
  1. Expression: Illumina HT-12 beadarray
  2. SNPs: Illumina 610quad

Study Design

- Participants - 330 NIHS individuals
  1. Samples - Blood (circulating lymphocytes)
  2. mRNA extracted > cDNA > expression analysis
- Platforms:
  1. Expression: Illumina HT-12 beadarray
  2. SNPs: Illumina 610quad

$\sim$ 48000 mRNA probes & $\sim$ 590000 SNPs for 330 participants

Computational Genomics

- Normalised gene expression data (23000 transcripts)[1]

---

[1] Göring et al., (2007) *Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes.* Nature Genetics

Computational Genomics

- Normalised gene expression data (23000 transcripts)[1]
- Heritability analysis (batched using GenABEL/R: Polygenic Model)

---

[1] Göring et al., (2007) *Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes.* Nature Genetics

Computational Genomics

- Normalised gene expression data (23000 transcripts)[1]
- Heritability analysis (batched using GenABEL/R: Polygenic Model)
- Heritable transcripts GWAS using SNP set
  1. mmscore function - pedigree structure analysis
  2. study-wide significance for NI pop $= 1.84x10^{-7}$
  3. suggestive significance threshold $= 1.0x10^{-5}$

---

[1] Göring et al., (2007) *Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes.* Nature Genetics

Computational Genomics

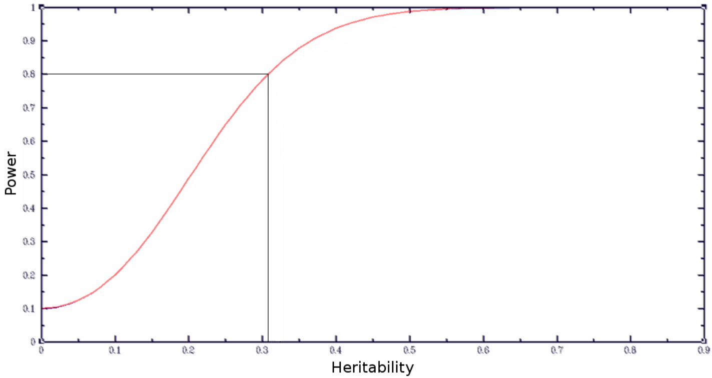
- Normalised gene expression data (23000 transcripts)[1]
- Heritability analysis (batched using GenABEL/R: Polygenic Model)
- Heritable transcripts GWAS using SNP set
  1. mmscore function - pedigree structure analysis
  2. study-wide significance for NI pop $= 1.84x10^{-7}$
  3. suggestive significance threshold $= 1.0x10^{-5}$
- A series of filters were designed to identify cis/trans eQTL's
  1. SNP/CHR location, Chromosome quadrants
  2. Graphical Filter - Modified Manhattan Plots with kern smoothing to facilitate peak identification (David Eccles)

---

[1] Göring et al., (2007) *Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes.* Nature Genetics

Expression Power Calculation

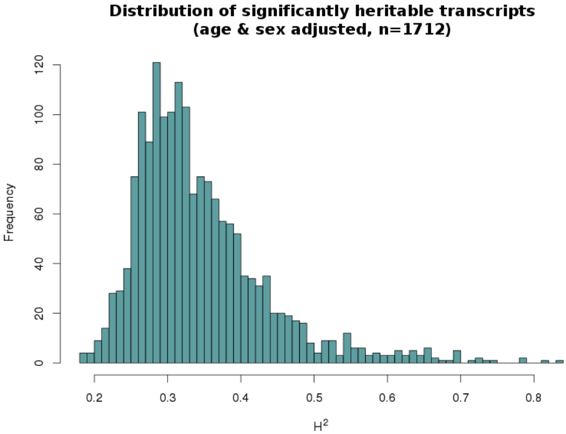
Estimation of power to detect significantly heritable transcripts in Norfolk Island pedigree was run in SOLAR



80% power to detect heritable transcripts above $H^2 = 0.3$

Heritable eQTL's

- H^2 analysis of n=23000 transcripts[2]

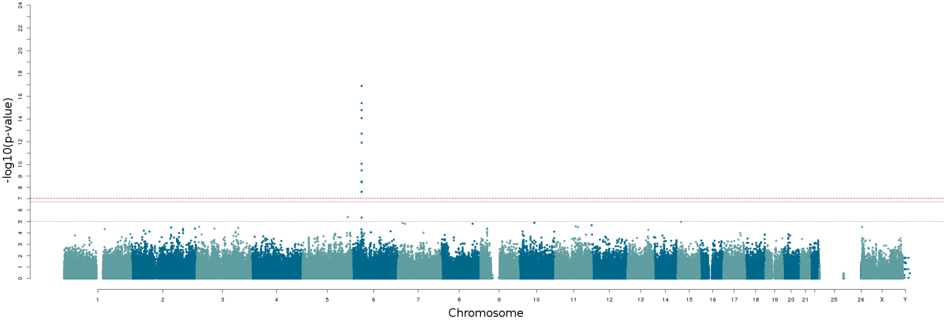


---

[2] Max sig. $H^2 = 0.84$ Min sig. $H^2 = 0.15$

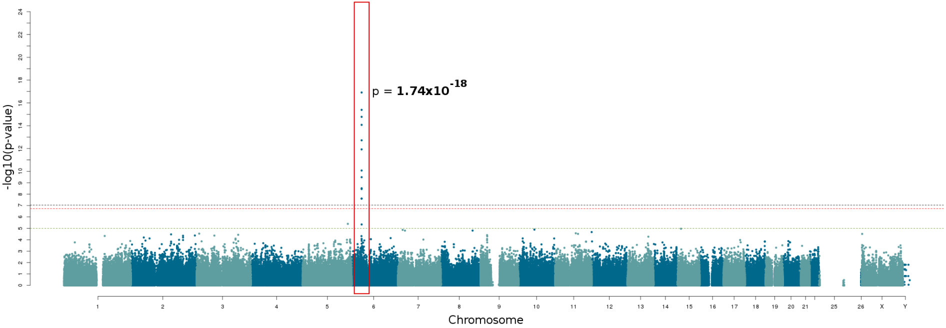
Heritable eQTL's

Overview of eQTL Manhattan Plots



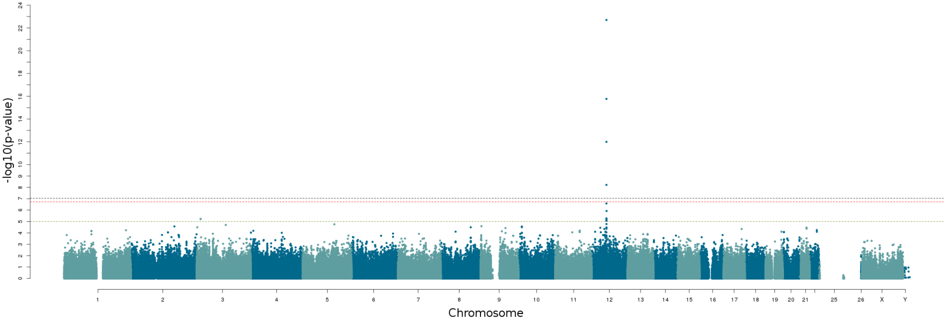
Heritable eQTL's

Overview of eQTL Manhattan Plots



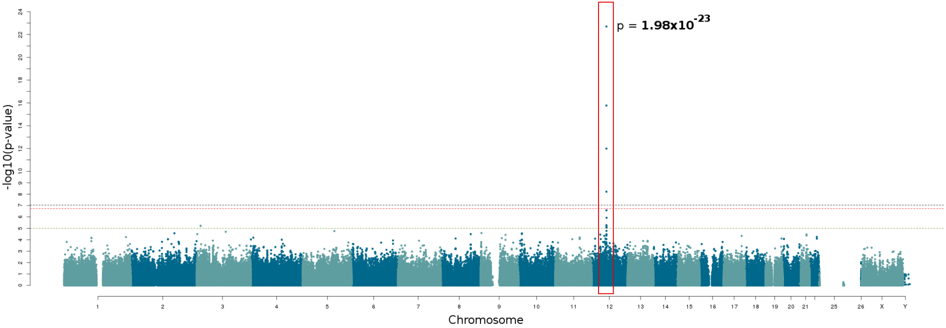
Heritable eQTL's

Overview of eQTL Manhattan Plots



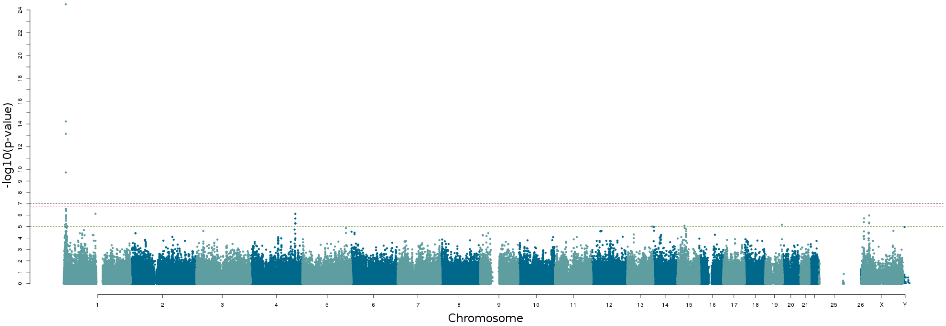
Heritable eQTL's

Overview of eQTL Manhattan Plots



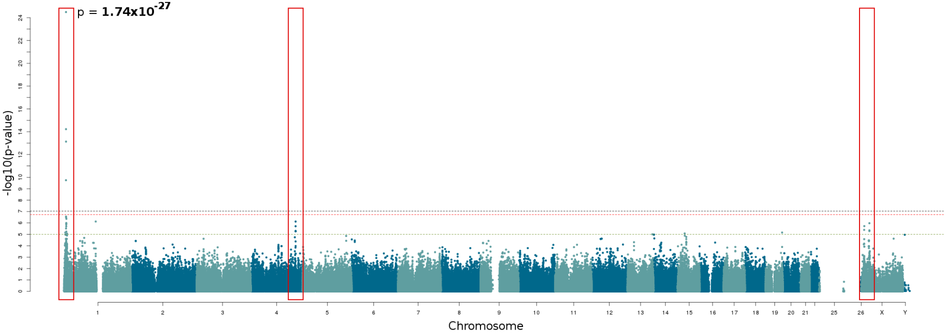
Heritable eQTL's

Overview of eQTL Manhattan Plots

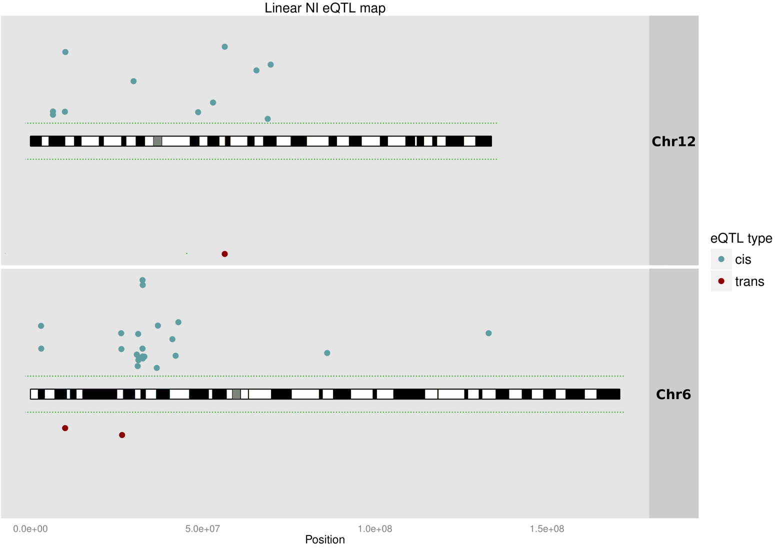


Heritable eQTL's

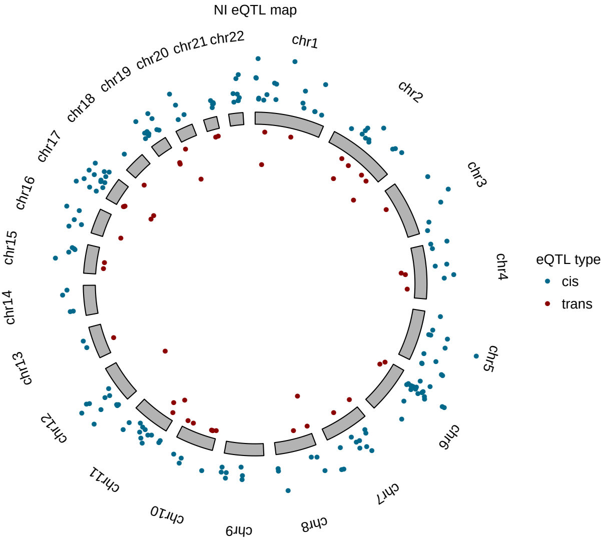
Overview of eQTL Manhattan Plots



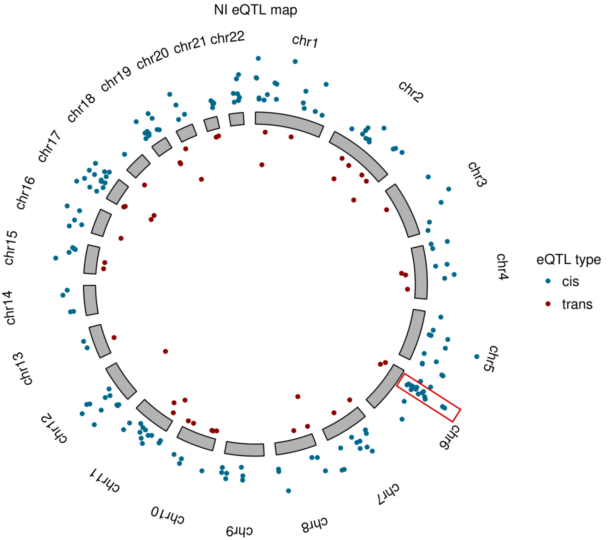
The NI eQTL map



The NI eQTL map



The NI eQTL map



Heritable eQTL's: Comparison

200 cis & 70 trans eQTL's identified at study wide sig.

	n	**cis**	**trans**	**significance threshold**
NIHS	330	200	70	1.84E-007
BSGS	852	1529	256	5.25E-012
San Antonio	1240	750	1072	LOD score >3

Heritable eQTL's: Comparison

- Overlap of several top hits with other studies:
  1. BSGS[3]: overlap 7 of there top 12 cis-eQTL results
     (genes: HLA-DRB1, HLA-DQB1, ERAP2, RPS26, CLEC12A, TUBB2A, PAM)
  2. some overlap with San Antonio Family Heart Study[4]
     An overlap of 7 of the top 20 cis-eQTL results:
     (genes: UTS2, RPS26, TIMM10, LGALS2, RPL14, HLA-DRB3, HLA-DRB5)

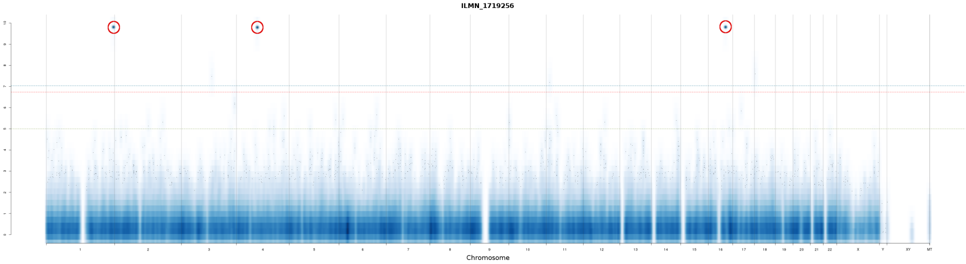
---

[3] Powell et al., (2012) *The Brisbane Systems Genetics Study: Genetical Genomics Meets Complex Trait Genetics.* PLoS ONE

[4] Göring et al., (2007) *Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes.* Nature Genetics

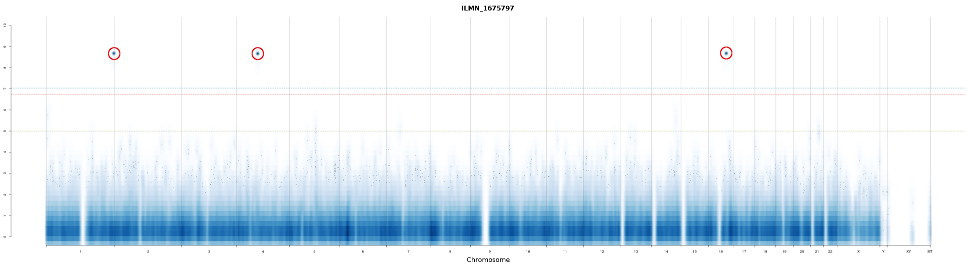
eQTL with Genomewide SNP signature?

- Another interesting finding... a group of trans eQTL that were originally filtered out appear to form a genome-wide 'signature' associated with 20 separate transcripts:



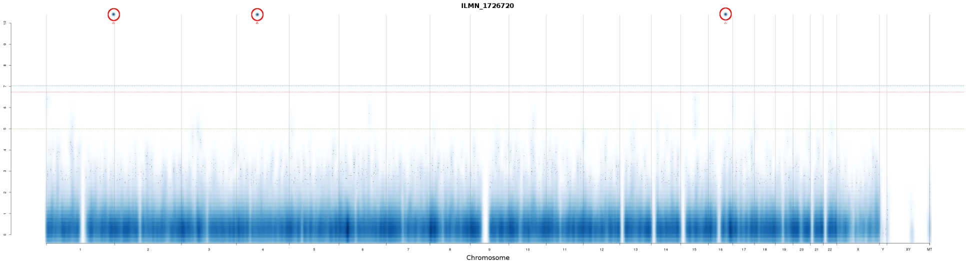
eQTL with Genomewide SNP signature?

- Another interesting finding... a group of trans eQTL that were originally filtered out appear to form a genome-wide 'signature' associated with 20 separate transcripts:



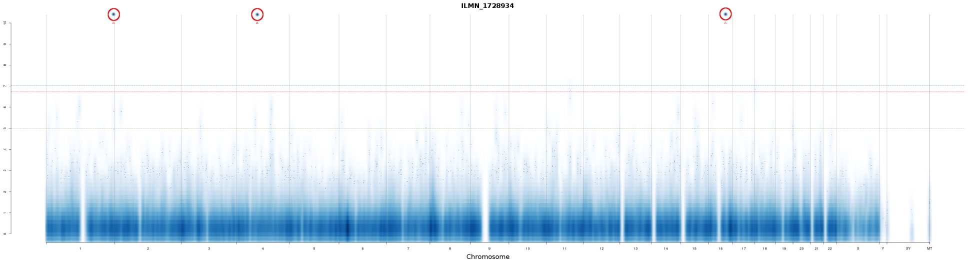
eQTL with Genomewide SNP signature?

- Another interesting finding... a group of trans eQTL that were originally filtered out appear to form a genome-wide 'signature' associated with 20 separate transcripts:



eQTL with Genomewide SNP signature?

- Another interesting finding... a group of trans eQTL that were originally filtered out appear to form a genome-wide 'signature' associated with 20 separate transcripts:



eQTL with Genomewide SNP signature?

- No SNP peaks, but potential genomewide SNP signature:
    1. 3 locus SNP signature
    2. 20 transcripts, 20 genes from 12 different chromosomes

Probe_ID	Gene	Chromosome	Top SNP p-value	eQTL
ILMN_1719256	CKS1B	1	1.65e-10	trans
ILMN_1675797	EPDR1	7	2.05e-09	trans
ILMN_1726720	NUSAP1	15	8.81e-11	trans
ILMN_1709634	CMBL	5	7.74e-09	trans
ILMN_1741133	NME1	17	1.28e-09	trans
ILMN_1786125	CCNA2	4	3.54e-06	trans
ILMN_1800197	MRPL36	5	3.64e-07	trans
ILMN_1728934	PRC1	15	2.39e-11	trans
ILMN_1663390	CDC20	1	4.34e-12	trans

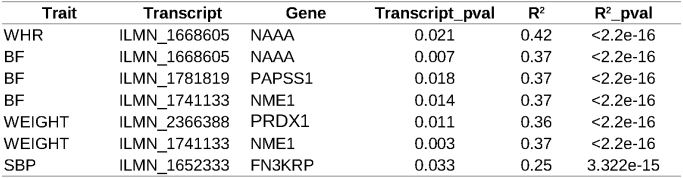
- GATHER analysis suggests all 20 genes involved in possible cell division/mitosis pathway...

Transcript/Trait/SNP associations

- Smaller, targeted association analysis
  - all SNPs from eQTL peaks (n= 2200)

Transcript/Trait/SNP associations

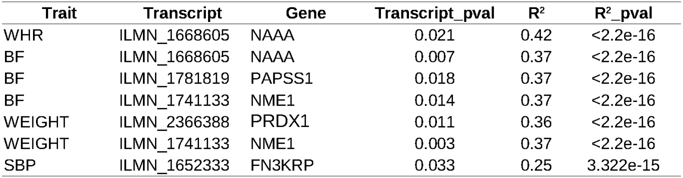
- Smaller, targeted association analysis
  - all SNPs from eQTL peaks (n= 2200)
- stepwise regression modeling trait transcript
  - included age|sex|genetic structure



Trait	Transcript	Gene	Transcript_pval	R^2	R^2_pval
WHR	ILMN_1668605	NAAA	0.021	0.42	<2.2e-16
BF	ILMN_1668605	NAAA	0.007	0.37	<2.2e-16
BF	ILMN_1781819	PAPSS1	0.018	0.37	<2.2e-16
BF	ILMN_1741133	NME1	0.014	0.37	<2.2e-16
WEIGHT	ILMN_2366388	PRDX1	0.011	0.36	<2.2e-16
WEIGHT	ILMN_1741133	NME1	0.003	0.37	<2.2e-16
SBP	ILMN_1652333	FN3KRP	0.033	0.25	3.322e-15

Transcript/Trait/SNP associations

- Smaller, targeted association analysis
  - all SNPs from eQTL peaks (n= 2200)
- stepwise regression modeling trait transcript
  - included age|sex|genetic structure



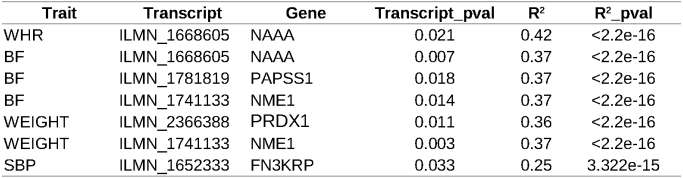
Trait	Transcript	Gene	Transcript_pval	R^2	R^2_pval
WHR	ILMN_1668605	NAAA	0.021	0.42	<2.2e-16
BF	ILMN_1668605	NAAA	0.007	0.37	<2.2e-16
BF	ILMN_1781819	PAPSS1	0.018	0.37	<2.2e-16
BF	ILMN_1741133	NME1	0.014	0.37	<2.2e-16
WEIGHT	ILMN_2366388	PRDX1	0.011	0.36	<2.2e-16
WEIGHT	ILMN_1741133	NME1	0.003	0.37	<2.2e-16
SBP	ILMN_1652333	FN3KRP	0.033	0.25	3.322e-15

- adding top rank eQTL SNP improves prediction scores

Transcript/Trait/SNP associations

- Smaller, targeted association analysis
  - all SNPs from eQTL peaks (n= 2200)

- stepwise regression modeling trait transcript
  - included age|sex|genetic structure



Trait	Transcript	Gene	Transcript_pval	R^2	R^2_pval
WHR	ILMN_1668605	NAAA	0.021	0.42	<2.2e-16
BF	ILMN_1668605	NAAA	0.007	0.37	<2.2e-16
BF	ILMN_1781819	PAPSS1	0.018	0.37	<2.2e-16
BF	ILMN_1741133	NME1	0.014	0.37	<2.2e-16
WEIGHT	ILMN_2366388	PRDX1	0.011	0.36	<2.2e-16
WEIGHT	ILMN_1741133	NME1	0.003	0.37	<2.2e-16
SBP	ILMN_1652333	FN3KRP	0.033	0.25	3.322e-15

- adding top rank eQTL SNP improves prediction scores

- have shown relationship between trait, transcipt and SNP

Conclusions & Future Directions

In conclusion...

- Identified a genomewide eQTL map in NI.
- 1712 expression transcripts were found to be significantly H^2
- GWAS identified 200 cis & 70 trans eQTL (study wide threshold).
- potentially novel SNP/eQTL signatures have been identified.

Conclusions & Future Directions

In conclusion...

- Identified a genomewide eQTL map in NI.
- 1712 expression transcripts were found to be significantly H^2
- GWAS identified 200 cis & 70 trans eQTL (study wide threshold).
- potentially novel SNP/eQTL signatures have been identified.

Future Directions...

- More comprehensive meta-analysis of current eQTL maps/databases
- **should facilitate the detection of novel (NI/population specific?) eQTL's**
- linkage analysis with SOLAR (STR & SNP??) for comparison
- exploration of clustering within heritable transcripts - potential detection of biological pathways and/or biomarkers

Conclusions & Future Directions

In conclusion...

- Identified a genomewide eQTL map in NI.
- 1712 expression transcripts were found to be significantly H^2
- GWAS identified 200 cis & 70 trans eQTL (study wide threshold).
- potentially novel SNP/eQTL signatures have been identified.

Future Directions...

- More comprehensive meta-analysis of current eQTL maps/databases
- **should facilitate the detection of novel (NI/population specific?) eQTL's**
- linkage analysis with SOLAR (STR & SNP??) for comparison
- exploration of clustering within heritable transcripts - potential detection of biological pathways and/or biomarkers

We've identified a trans-eQTL mapping to a gene which associates with an obesity related phenotype (COMP3) with associated kidney dysfunction.

Acknowledgements

Project Leaders: Prof Lyn Griffiths, Dr Rod Lea

Co Supervisors: Dr Donia Macartney-Coxson, Dr Geoff Chambers

Texas Biomedical Research Institute: Melanie Carless, Claire Bellis, Matt Johnson, Harald Göring, Thomas Dyer, Jo Curran, John Blangero

Genomics Research Centre: Michelle Hanna, Dr Bridget Maher, Dr David Eccles

Funding: NHMRC, John Corbett (Scholarship)

Ethics: Griffith University Ethics Committee