Transcriptome Profiling in Population, Community, and Field Studies

USC/UCLA Center of Biodemography and Population Health
UCLA Social Genomics Core Laboratory
Geneome-wide transcriptional profiling

1. Proximal determinant of health
2. Socially/behaviorally/environmentally modifiable (potential mediator)
3. Large effect sizes
4. Bioinformatically extensible

Question: Can we do this in field-friendly DBS samples?
RNA Comparison Study

59 generally healthy adults
Sampled for diversity in chronic inflammation

Age: 30 young (21-31 yo), 29 older (51-98 yo)
BMI: 17-39  (50% > 25, 14% > 30)

Female = 81%
Male = 19%

White = 44%
Asian = 24%
Hispanic = 14%
African American = 14%

Rx Med = 51%
OTC Med = 17%

Extract RNA in parallel
Test for RNA mass (amount)
Test for RNA integrity (degradation)

Assay RNA in parallel
2012: qRT-PCR
2013: genome-wide microarray
Aims for 2013 analyses

1. Establish feasibility of RNA amplification to overcome DBS mass limitations
   
   Goal: transcriptome-wide profiling & bioinformatics from DBS

2. Catalogue all human genes for DBS/PBMC correspondence

   Goal: a list of “Reliably ascertainable genes”
RNA Comparison Study: RNA Yield

![Graph showing RNA yield comparison and cost]

- **Affymetrix**
  - PBMC: $150
  - Ora: $150

- **Illumina TotalPrep**
  - PAX: $220
  - DBS: $220

- **NuGEN Ovation Pico SL**
  - PBMC: $280
  - Ora: $280

- **Epicentre TargetAmp Pico**
  - PAX: $220
  - DBS: $220

The graph depicts the total RNA mass in ng for different samples and methodologies.
RNA Comparison Study: Nugen assay performance

Representative DBS results

Transcriptome profiling by Illumina HT-12 v4 Bead Arrays
**RNA Comparison Study: NuGEN assay performance**

### Genes detectable (among PBMC-present > 80%*)

- PBMC: 1,800 genes
- PAXgene: 1,600 genes
- Oragene: 1,400 genes
- DBS: 1,200 genes

### PBMC validation \( r \) (cross-gene**)

- PAXgene: 0.9
- Oragene: 0.8
- DBS: 0.7

### PBMC validation \( r \) (cross-subject***)

- PAXgene: 80
- Oragene: 60
- DBS: 50

---

* Detectable = \( p < .05 \) above background in \( \geq 80\% \) of PBMC samples

** Computed over all assayed genes

*** Computed over genes >80% present in PBMC
RNA Comparison Study: Epicentre assay performance

Genes detectable (among PBMC-present > 80%)

PBMC > 5,000
PAXgene > 4,000
Oragene > 3,000
DBS > 2,000

PBMC validation $r$ (cross-gene)

PAXgene > 1.0
Oragene > 0.9
DBS > 0.8

PBMC validation $r$ (cross-subject)

PAXgene > 70
Oragene > 60
DBS > 50

Number genes with cross-subject $r \geq 0.40$
RNA Comparison Study: Epicentre vs NuGEN

Genes detectable

PBMC validation $r$ (cross-gene)

PBMC validation $r$ (cross-subject)
RNA Comparison Study: Reliably ascertainable genes

NuGEN

Type I IFN-related antiviral genes: MX1, GBP2, IFI6, IFITM2/3
Myeloid cell genes: CD16/FCGR3A, CD68
Inflammatory genes: IL8, IL8RB, PTGS2/COX2, S100s, SOD2, CCR1, FOS, JUND, CREB1, CXCL1/3
Regulatory cytokines: IL10, IL18
Missed relative to PBMC: DC markers (CD83, CD1D), IL1B

Epicentre

Missed relative to PBMC: Pretty much everything, except…
Type I IFN-related antiviral genes
RNA Comparison Study: Take-home points

1. Genome-wide transcriptional profiling is feasible in DBS.

   Reliably detect ~75% of PBMC-expressed genes (~1300)

   **Caveat:** DBS is a challenging context for genome-scale analyses
   
   Limiting dilution statistical issues (low SNR)
   
   Cross-gene $r = $ good     Cross-subject $r = $ variable
   
   Balance: still $>100x$ yield from qRT-PCR or microarray of Oragene

2. DBS “gold standard” convergent validity varies across genes

   **Highest consistency:** Type I IFN antiviral genes, Monocyte/macrophage,
   Inflammatory ($IL8$, $PTGS2/COX2$, $SOD2$), Regulatory cytokines ($IL10$)
   
   Variability driven largely by cell type & “true transcript variance”

3. Specific approaches

   Similar in abstract: reliably detect ~75% of PBMC-expressed genes

   **NuGEN:** slightly better quantitative performance
   
   MUCH better yield of immunologically interpretable genes