CCHMC Biobank
“Better Outcomes for Our Children”

Co-PI: John Harley, MD, PhD
Co-PI: David Witte, MD
PI: Michael Barnes, PhD

Biobanking: Navigating the Practical, Ethical and Regulatory Pathways

Cincinnati Children’s Hospital Medical Center
9:00 to 11:00 am, Thursday, January 27, 2011
CCHMC BioBank Components

• **Tissue Repository** (underway Pathology)
  – Excess frozen tissues from surgeries
  – 20,000 samples for de-identified research
  – “opt-in” consent for future

• **DNA Repository** (February)
  – From excess CBC samples
  – “opt in” consent – de-identified research
  – 25,000 per year (100 per day)

• **Trio Repository** (FY 2012 or 2013)
  – 5000 parents, 2500 children per year
  – DNA, RNA, full consent
AutoGenFlex STAR

QIAGEN FlexiGene chemistry on AutoGenFlex STAR:

- Human whole blood (fresh and frozen)
- Buffy Coat (fresh and frozen)

Throughput capabilities:

- 0.5 - 5 ml blood:
  - 40 samples per batch
  - 80 samples per day
- 6 - 10 ml whole blood:
  - 20 samples per batch
  - 40 samples per day
CCHMC BioBank – other details

• Service individual investigators and teams of investigators.
• Warehouse tissues, DNA and other samples.
  – Central subsidized service
  – Automated equipment
  – Large storage inventory
• Inventory system.
• Bioinformatics.
Review

The Lupus Family Registry and Repository

Astrid Rasmussen1, Sydney Sevier1,2, Jennifer A. Kelly1, Stuart B. Glenn1, Teresa Aberle1, Carisa M. Cooney3, Anya Grether1, Ellen James1, Jared Ning1, Joanne Tesiram1, Jean Morrisey1, Tiny Powe1, Mark Drexel1, Wes Daniel1, Bahram Namjou1, Joshua O. Ojwang1, Kim L. Nguyen1, Joshua W. Cavett1, Jeannie L. Te1, Judith A. James2,4, R. Hal Scofield1,2,5, Kathy Moser1, Gary S. Gilkeson6, Diane L. Kamen6, Craig W. Carson7, Ana I. Quintero-del-Rio8,9, Maria del Carmen Ballesteros8,9, Marilynn G. Punaro10,11, David R. Karp11, Daniel J. Wallace12, Michael Weisman13, Joan T. Merrill14, Roberto Rivera15, Michelle A. Petri16, Daniel A. Albert17, Luis R. Espinoza18, Tammy O. Utset19, Timothy S. Shaver20, Eugene Arthur21, Juan-Manuel Anaya22, Gail R. Bruner1 and John B. Harley1,2,5
11,482 received samples from consented subjects
213,758 filled serum tubes at -20°C
55,823 filled serum tubes at -80°C
269,606 filled plasma tubes at -20°C
233,267 filled DNA aliquots
4,528 filled PBMC tubes
19,286 filled TLC tubes
1,126 clinical data fields on average per subject
15,624,049 clinical data points collected
2,722,3442,219 genotype data points for LFRR samples
## Genotyping Capacity in SLE

<table>
<thead>
<tr>
<th>Year Period</th>
<th>Total Genotypes</th>
<th>Genotypes/day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992 to 2006</td>
<td>1,500,000</td>
<td>293</td>
</tr>
<tr>
<td>2006</td>
<td>10,000,000</td>
<td>27,700</td>
</tr>
<tr>
<td>2007-2009</td>
<td>&gt;400,000,000</td>
<td>&gt;1,000,000</td>
</tr>
<tr>
<td>2010</td>
<td>&gt;4,000,000,000</td>
<td>&gt;100,000,000</td>
</tr>
</tbody>
</table>
iScan (…>100,000,000 genotypes /day)

SNP Genotyping & CNV Analysis, Custom Genotyping, Cytogenetic Analysis, Focused Genotyping, Linkage Analysis, Whole-Genome Genotyping & Copy Number Analysis, Gene Regulation & Epigenetic Analysis, Array-Based Methylation Analysis, Gene Expression Analysis, Array-Based Transcriptome Analysis, FFPE Sample Analysis, Whole-Genome Gene Expression Analysis
Joris Veltman, Han Brunner and colleagues report results of a family based exome sequencing study of ten individuals with unexplained mental retardation. They identified and validated de novo mutations in nine genes, six of which are likely to be pathogenic based on functional criteria, suggesting an important role for de novo point mutations in the etiology of unexplained mental retardation.
Overview of variants detected per proband and impact of the prioritization steps for selecting candidate non-synonymous *de novo* mutations

<table>
<thead>
<tr>
<th>Trio</th>
<th>1</th>
<th>4</th>
<th>6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-confidence variant calls</td>
<td>20,810</td>
<td>22,647</td>
<td>22,333</td>
<td>21,755</td>
</tr>
<tr>
<td>After exclusion of nongenic, intronic and synonymous variants</td>
<td>5,556</td>
<td>5,991</td>
<td>5,567</td>
<td>5,640</td>
</tr>
<tr>
<td>After exclusion of known variants</td>
<td>165</td>
<td>155</td>
<td>136</td>
<td>143</td>
</tr>
<tr>
<td>After exclusion of inherited variants</td>
<td>4</td>
<td>7</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
### Overview of all *de novo* variants identified by exome sequencing in ten individuals with unexplained mental retardation

<table>
<thead>
<tr>
<th>Gene</th>
<th>Trio</th>
<th>cDNA</th>
<th>Protein</th>
<th>PhyloP</th>
<th>Probability in dbSNP&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Gene function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYNC1H1</td>
<td>1</td>
<td>c.11465A&gt;C</td>
<td>p.H3822P</td>
<td>5.5</td>
<td>77</td>
<td>Retrograde axonal transporter</td>
</tr>
<tr>
<td>ZNF599</td>
<td>1</td>
<td>c.532C&gt;T</td>
<td>p.L187F</td>
<td>-1.5</td>
<td>22</td>
<td>X-linked mental retardation</td>
</tr>
<tr>
<td>YY1</td>
<td>3</td>
<td>c.1138&gt;T</td>
<td>p.N380W</td>
<td>6.9</td>
<td>160</td>
<td>Ubiquitous transcription factor</td>
</tr>
<tr>
<td>CIC</td>
<td>6</td>
<td>c.1474C&gt;T</td>
<td>p.N492W</td>
<td>2.6</td>
<td>101</td>
<td>Neural granule cell</td>
</tr>
<tr>
<td>JARID1C</td>
<td>10</td>
<td>c.1919G&gt;A</td>
<td>p.C640W</td>
<td>5.1</td>
<td>194</td>
<td>X-linked mental retardation gene</td>
</tr>
</tbody>
</table>
Large Lupus Association Study #2 (LLAS2)

- 42 Investigators
- 18,252 attempted subjects (16,435 produced quality genotypes).
- 33,789 ordered SNPs (32,216 produced quality genotypes)
- 547 MILLION genotypes completed Jan 25, 2010
- Initial results produced February 17, 2010
SLE Sample Sources

SLEGEN

LFRR

UAB P01
Gaffney Moser

James
Bae

MERRILL

Alarcon

MUSC
HSS

LLAS2 = 18,288 samples
The End
Virginia Pascual, MD

Combined Immunology Seminar
8:00 am, Wednesday
January 12
MSB.7051

“A Genomic Approach to Pediatric Rheumatic Disease”
Diagnosis of infection with *Staph aureus*

Figure 5. Blood transcriptional fingerprints of patients with *Staphylococcus aureus* infection. Relative changes in transcript abundance in the blood of patients with *S. aureus* infection compared to that of healthy controls are recorded for a set of 28 transcriptional modules. Colored spots represent relative increase (red) or decrease (blue) in transcript abundance \((P < 0.05, \text{Mann Whitney})\) within a module. The legend shows functional interpretation for this set of modules. Fingerprints have been generated for two independent cohorts of subjects (divided into a training set used in the discovery phase, \(n = 30\), and an independent test set used in the validation phase, \(n = 32\)).
Ken M. Kaufman, PhD

Special Seminar
10:00 am Tuesday
January 18, 2011
S1.203/4

“15 Years of Lupus Genetics: Robust Results Despite Terabytes of Data!”
“The Genomics of Interferon in Systemic Lupus Erythematosus”
A Modular analysis framework for blood genomics studies: Application to systemic lupus erythematosus
Exome data of 10 mental retardation cases sequenced on SOLiD 3 Plus

Read mapping and variant calling
- Default mapping settings
- High-stringency variant calling
- Exclude low quality

Variant analysis
- Exclude nongenic, intronic and synonymous
- Exclude known SNPs and in-house database
- Exclude inherited

Validation
- Exclude non-validated
- Exclude inherited
- Test occurrence in control cohort

Interpretation
- Gene function
- Mutation impact
RNA sequencing shows no dosage compensation of the active X-chromosome

Yuanyan Xiong, Xiaoshu Chen, Zhidong Chen, Xunzhang Wang, Suhua Shi, Xueqin Wang, Jianzhi Zhang & Xionglei He
doi:10.1038/ng.711

Jianzhi Zhang and Xionglei He report analyses of published RNA sequencing data examining relative expression levels between genes located on the X chromosome and genes located on autosomes. Unlike previous reports of dosage compensation between the X chromosome and autosomes, their analyses detect an X:autosome expression ratio of ~0.5.

Nature Genetics Vol 42, pp1043-1047, 2010
Technology has taken us from the “Horse and Buggy” - circa 1900...

to marvel the V-8 Ferrari harnessing 490 hp at 8,500 rpm and 343 lbs - feet of torque.

In the last two years such a technical transition has occurred in genetic analysis... No human genetic problem remains beyond our reach, and only temporarily beyond our pocketbook...
ImmunoChip

- 184 p<5x10^-8 associations
- Crohn’s (63)
- Type 1 Diabetes (40)
- **SLE (36 with 12 shared)**
- Ulcerative Colitis (28)
- Celiac Disease (30)
- Multiple Sclerosis (26)
- Rheumatoid Arthritis (25)
- Psoriasis (19)
- Ankylosing Spondylitis (8)
- AITD (2)
- Primary biliary cirrhosis (1)
- IgA deficiency (1)

- 196,000 SNPs
- 1000 genomes data
- 5000 AIMs
- 6400 in HLA
- ~740 SNPs / assn.
- SNPs ~85% saturation
- 150,000 sold
- **$39 each**
- Ordered Feb 2010
- Available Apr 2010
Lupus Family Registry & Repository (LFRR) Participant Collection. September 2010

- 11326 participants useful for SLE genetic studies
  - 3086 SLE affecteds
  - 499 alleged SLE affecteds (still in progress)
  - 6171 unaffected family members
  - 1570 controls
A SNP Array (e.g., ImmunoChip) for your field?

Think about it…

- ~200,000 markers
- $39 each
- Process ~400 samples per day
- 80,000,000 genotypes per day

NEED:
- 150,000 arrays purchased, $6,000,000
- 30 collaborators with ~ $200,000 each
- Six months & 2 FTE to select content & administer consortium.
The 46 Human Chromosomes
SLE Genes
Alleged & Published <2007*

• HLA
• FCGR3A, F176V
• FCGR2A, H131R
• IRF5
• PTPN22
• SPP1 (osteopontin)
• PD-1

*\( p \sim < 5 \times 10^{-8} \)

Only 7 of 152 published genetic associations are established.
The power for various sample sizes was calculated using the CaTS Power Calculator assuming prevalence = 0.01, risk allele frequency = 0.2, $\alpha = 10^{-7}$ and expected Hardy-Weinberg proportions in cases and controls. Under these conditions, a nonparametric allele frequency difference with an odds ratio of 1.2 reproduces the power in the multiplicative model using Power for Association With Errors (PAWE).

Association → Biology

Associated DNA \( (p<5 \times 10^{-8}) \)

- genes
- iRNA
- orf

Fine mapping

Indel, inversion
SNP, CNV

Functional markers

Gene constitution
Gene expression

BIOLOGY!
Summary

Genetics
- Sex; Race
- >30 Genes

Environment
- Epstein-Barr virus;
- Immune history

Unified understanding of disease etiology

Better diagnosis, prognosis, therapy & prevention
LLAS2

- Lupus Susceptibility genes in multiple ancestries
- >32,216 SNPs in 16,901 samples.

<table>
<thead>
<tr>
<th>Population</th>
<th>Cases</th>
<th>Controls</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>EA</td>
<td>4220</td>
<td>3803</td>
<td>8023</td>
</tr>
<tr>
<td>Asian</td>
<td>1311</td>
<td>1342</td>
<td>2653</td>
</tr>
<tr>
<td>AA</td>
<td>1566</td>
<td>1891</td>
<td>3457</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1511</td>
<td>791</td>
<td>3203</td>
</tr>
<tr>
<td>Gullah</td>
<td>156</td>
<td>131</td>
<td>287</td>
</tr>
<tr>
<td>Totals</td>
<td>8608</td>
<td>7827</td>
<td>16,435*</td>
</tr>
</tbody>
</table>

* 466 samples of `Other’ or missing population information.
Ancestral Identification for LFRR
SLE Affecteds. September 2010

- 30% European
- 8% Asian
- 8% Amerindian
- 6% Gullah
- 3% Hispanic
Center for Autoimmune Genomics & Etiology

- **Goal** - Identify & exploit initiating events in inflammatory disorders

- **Expertise & Capacity**
  - Human Genetic Variation *(service)*
  - Chromatin Epigenetics
  - Epidemiology, Compliance, **Biobank** *(service)*
  - Biology
  - Informatics & Literature

- **Logistics**
  - People – 11 New Faculty
  - Space - 20,000 square feet for labs and offices
  - Infrastructure – informatics, data management & analysis,
    compliance…
  - Financing external and internal (70% to 30%)
  - Time (yesterday)
ImmunoChip

• 184 p<5x10^{-8} associations

• Crohn’s (63)
• Type 1 Diabetes (40)
• SLE (36 with 12 shared)
• Ulcerative Colitis (28)
• Celiac Disease (30)
• Multiple Sclerosis (26)
• Rheumatoid Arthritis (25)
• Psoriasis (19)
• Ankylosing Spondylitis (8)
• AITD (2)
• Primary biliary cirrhosis (1)
• IgA deficiency (1)

196,000 SNPs
1000 genomes data
5000 AIMs
6400 in HLA
~740 SNPs / assn.
SNPs ~85% saturation
150,000 sold
$39 each

Ordered Feb 2010
Available Apr 2010
Particularly Recruiting Minority SLE Patients & Family Members

With patient’s permission, send names and phone numbers to:

**TOLL FREE:** 1-888-OK-LUPUS
(1-888-655-8787)

**FAX:** 1-405-271-3045 *(secure)*

**Email:** Lupus-Recruiters@lupus.omrf.org

**Web:** [http://lupus.omrf.org](http://lupus.omrf.org)
Physician/Scientists

Thank you for visiting the site of the Lupus Multiplex Registry and Repository (LMRR)

The LMRR is the only resource of clinical data, serum, plasma and DNA of multiplex lupus families that is available to scientists (with approval) for their research.

To learn more about the LMRR, click here:

- List of Publications
- What is the LMRR?
- Become an Approved User
- Flyers to Advertise Study
- Physician Referral Fax
- Email Inquiries

http://lupus.omrf.org
The Lupus Genetic Studies