EMR and Genomics at Mayo Clinic
From Discovery to Practice

Christopher G. Chute, MD, DrPH
Mayo Clinic

Genomics and health information technology systems: Exploring the issues
NHGRI  April 27, 2011
Outline

• Discovery and validation of genomic associations in clinical settings
• Translation of findings into practice
• Foreshadowing of routine clinical use
Enterprise Data Trust
Standards-based Clinical Data Repository

Enterprise Data Governance Program
1. Enterprise Data **Modeling** Activity
2. Enterprise **Metadata** Activity
3. Enterprise **Vocabulary** Environment

Clinical Systems & EMRs

Health Sci. Res.
- Comparative Effectiveness
- Genotype to Phenotype
- ...

Basic Science Genomics

© 2007 Mayo Clinic College of Medicine
eMERGE@Mayo
electronic MEdical Records and GEnomics

• NHGRI funded GWAS and EMR study
• Emphasis on high-throughput phenotyping
  • Disease and control cohort definitions
  • EMR data sources
  • Portable algorithms (14)
• Demonstrated Positive Predicative Value across five-member eMERGE consortium
## SHARP: Area 4: Secondary Use of EHR Data
### A $15M National Consortium

<table>
<thead>
<tr>
<th>Themes</th>
<th>Projects</th>
<th>Players</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Normalization</td>
<td>Clinical Data Normalization</td>
<td>IBM, Mayo, Utah, Agilex</td>
</tr>
<tr>
<td>Phenotype Recognition</td>
<td>Natural Language Processing (NLP)</td>
<td>Harvard, Group Health, IBM, Utah , Mayo, MIT, SUNY, i2b2, Pittsburgh, Colorado</td>
</tr>
<tr>
<td>Data Quality and Evaluation Framework</td>
<td><strong>High-Throughput Phenotyping</strong></td>
<td>CDISC, Centerphase, Mayo, Utah</td>
</tr>
<tr>
<td></td>
<td>UIMA and Scaling Capacity</td>
<td>IBM, Mayo</td>
</tr>
<tr>
<td></td>
<td>Data Quality</td>
<td>Mayo, Utah</td>
</tr>
<tr>
<td></td>
<td>Evaluation Framework</td>
<td>Agilex, MN HIE, Mayo, Utah</td>
</tr>
</tbody>
</table>
BORA: Biologically Oriented Repository Arch. Integration of Genomics and Clinical Phenotype

RLIMS - Core Labs

BioBank

Data Manager
- genomics raw data

BORA
- genomics data integration

EDT

multiple clinical data source & NLP

DDQB - ANALYTICS
Mayo Genome Consortia (MayoGC)

• A shared infrastructure for genotyped cohorts
  • “Pointers” to Biobank and patient identifiers
  • Well-defined eligibility criteria (consenting, QC, etc.)

• Drawn from research studies across Mayo Clinic
  • Enables study of novel EMR phenotypes

• Projects to date:
  • eMERGE Hypothyroidism Replication
  • GWAS liver enzymes, colon polyps, prostate volume, and Barrett’s esophagus

Mayo Genome Consortia (MayoGC): A Genotype-Phenotype Resource for Genome Wide Association Studies with an Application to the Analysis of Circulating Bilirubin: Bielinski, Chai, Pathak, Talwalkar, Limburg, Gullerud, Sicotte, Klee, Ross, Kocher, Kullo, Heit, Petersen, de Andrade, Chute (Accepted for publication at Mayo Clinic Proceedings)
<table>
<thead>
<tr>
<th>Study Name (NIH Grant Number)</th>
<th>Principal Investigator</th>
<th>Sample Size</th>
<th>Genotyping Platform</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase I (Completed)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electronic Medical Record Phenotypes and Community Engaged Genomic Associations (eMERGE) (NHGRI- U01 HG004599-01)</td>
<td>Dr. Christopher Chute Dr. Iftikhar Kullo</td>
<td>3197</td>
<td>Illumina Human 660W Quad-V1</td>
</tr>
<tr>
<td>Mayo Clinic Genome-wide Association Study of Venous Thromboembolism (NHGRI HG04735)</td>
<td>Dr. John Heit</td>
<td>2497</td>
<td>Illumina Human 660W Quad-V1</td>
</tr>
<tr>
<td><strong>Phase 2 (in process)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Molecular Epidemiology of NHL and CLL (NCI CA92153)</td>
<td>Dr. James Cerhan</td>
<td>1500</td>
<td>Illumina Omni Express 750</td>
</tr>
<tr>
<td>Haplotype-Based Genome Screen for Ovarian Cancer Loci (NCI R01 CA 114343)</td>
<td>Dr. Ellen Goode</td>
<td>535</td>
<td>Illumina 610 Quad</td>
</tr>
<tr>
<td>A1ATD Carriers and Lung Cancer Risk (NCI R01 CA 80127)</td>
<td>Dr. Ping Yang</td>
<td>1500</td>
<td>Illumina 610</td>
</tr>
<tr>
<td>BT SPORE - Project 4 in: Mayo SPORE in Brain Cancer (NCI P50 CA 108961)</td>
<td>Dr. Robert B. Jenkins</td>
<td>600</td>
<td>Illumina 610</td>
</tr>
<tr>
<td>A genome-wide association study for breast Ca in BRCA1 mutation carriers (NCI R01 CA 128978)</td>
<td>Dr. Fergus Couch</td>
<td>125</td>
<td>Illumina 610K</td>
</tr>
<tr>
<td>Triple Negative Breast Cancer Consortium (TNBCC)</td>
<td>Dr. Fergus Couch Dr. Celine Vachon</td>
<td>155</td>
<td>Illumina 660K</td>
</tr>
<tr>
<td>Collaborative Oncological Gene-Environment Study (COGS) – Breast Cancer</td>
<td>Dr. Stephen Thibodeau</td>
<td>500</td>
<td>Illumina custom iSelect (iCOGs 204K SNPs)</td>
</tr>
<tr>
<td>Collaborative Oncological Gene-Environment Study (COGS) – Ovarian Cancer</td>
<td>Dr. Ellen Goode</td>
<td>500</td>
<td>Illumina custom iSelect (iCOGs 204K SNPs)</td>
</tr>
<tr>
<td>Genomics of Primary Biliary Cirrhosis DK 80670</td>
<td>Dr. Konstantinos Lazaridis</td>
<td>1300</td>
<td>Illumina Immunochip (~200K SNPs)</td>
</tr>
<tr>
<td>PROGRESS (PSC Resource Of Genetic Risk, Environment and Synergy Studies) DK 84960</td>
<td>Dr. Konstantinos Lazaridis</td>
<td>1200</td>
<td>To be decided</td>
</tr>
</tbody>
</table>

**MayoGC**

n = 11,922 GWA data  
n = 2,800 iSelect data  
Total Sample Size = 14,722
Pharmacogenomics Research Network
Stimulus for Translation at Mayo

• Builds on work of Weinshilboum - TPMT
• Two clinical translational projects
• GWA study of the efficacy of aromatase inhibitors
  • NCIC-NCI MA.27 breast cancer adjuvant clinical trial
• GWA study of SSRI therapy of depression
  • SNRI therapy of patients who fail to respond
# CYP2D6 SNPs

## Psychiatry SNP Screening CHIP

<table>
<thead>
<tr>
<th>Allele</th>
<th>Enzyme Activity</th>
<th>5'UTR</th>
<th>Exon 1</th>
<th>Exon 2</th>
<th>Exon 3</th>
<th>Intron 3</th>
<th>Exon 4</th>
<th>Exon 5</th>
<th>Exon 6</th>
<th>Exon 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>*1</td>
<td>Normal</td>
<td>-1584</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>*2A</td>
<td>Increased</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>*2B</td>
<td>Decreased</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*2D</td>
<td>Decreased</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*3</td>
<td>None</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*4</td>
<td>None</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>*6</td>
<td>None</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*7</td>
<td>None</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*8</td>
<td>None</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*9</td>
<td>Decreased</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*10</td>
<td>Decreased</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*11</td>
<td>None</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*12</td>
<td>None</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*15</td>
<td>None</td>
<td>C</td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>*17</td>
<td>Decreased</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>*41</td>
<td>Decreased</td>
<td>G</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
</tr>
</tbody>
</table>

- **1707** DELETION of T in Exon 3 at 1707 Frame Shift Left
- **2549** DELETION of A in Exon 5 at 2549 Frame Shift to Left
- **138** INSERTION of T in Exon 1 at 138 Frame Shift to Right
- **2539** DELETION of T in Exon 3 at 1707 Frame Shift Left

**Frame Shift to Right:**
- 138
- 2539

**Frame Shift to Left:**
- 1707
- 2549

**Enzyme Activity:**
- Normal
- Increased
- Decreased
- None

**CYP2D6 SNPs:**
- Exon 1
- Exon 2
- Exon 3
- Intron 3
- Exon 4
- Exon 5
- Exon 6
- Exon 9

**Gene:**
- CYP2D6

**Alleles:**
- *1
- *2A
- *2B
- *2D
- *3
- *4
- *6
- *7
- *8
- *9
- *10
- *11
- *12
- *15
- *17
- *41
## Routine Psychiatry Practice - Screen Result

### Patient Genotype

<table>
<thead>
<tr>
<th>Gene</th>
<th>Genotype</th>
<th>Predicted Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2D6</td>
<td>*4/*41</td>
<td>Poor Metabolizer</td>
</tr>
<tr>
<td>CYP2C19</td>
<td>*1/*2</td>
<td>Intermediate Metabolizer</td>
</tr>
<tr>
<td>CYP1A2</td>
<td>See Table Below</td>
<td>Extensive Metabolizer</td>
</tr>
<tr>
<td>SLC6A4</td>
<td>L/L</td>
<td>High Activity</td>
</tr>
<tr>
<td>HTR2A</td>
<td>T/C</td>
<td>Intermediate Activity</td>
</tr>
</tbody>
</table>

### CYP1A2 Genotype

<table>
<thead>
<tr>
<th>Genotype</th>
<th>-3860G&gt;A - G/G</th>
<th>-2467T&gt;delT - T/T</th>
<th>-739T&gt;G - T/T</th>
<th>-729C&gt;T - C/C</th>
<th>-163C&gt;A - C/A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>125G&gt;C - C/C</td>
<td>558C&gt;A - C/C</td>
<td>2385G&gt;A - G/G</td>
<td>2473G&gt;A - G/G</td>
<td>2499A&gt;T - A/A</td>
</tr>
<tr>
<td></td>
<td>3497G&gt;A - G/G</td>
<td>3533G&gt;A - G/G</td>
<td>50590C&gt;T - C/C</td>
<td>5166G&gt;A - G/G</td>
<td>5347C&gt;T - T/C</td>
</tr>
</tbody>
</table>
Patient M

Antidepressants

**USE AS DIRECTED**
- desvenlafaxine (Pristiq®)
- fluvoxamine (Luvox®)
- selegiline (Emsam®)
- sertraline (Zoloft®)

**USE WITH CAUTION**
- citalopram (Celexa®)
- duloxetine (Cymbalta®)
- escitalopram (Lexapro®)
- mirtazapine (Remeron®)
- trazodone (Desyrel®)

**USE WITH CAUTION AND WITH MORE FREQUENT MONITORING**
- amitriptyline (Elavil®)
- bupropion (Wellbutrin®)
- clomipramine (Anafranil®)
- desipramine (Norpramin®)
- fluoxetine (Prozac®)
- imipramine (Tofranil®)
- nortriptyline (Pamelor®)
- paroxetine (Paxil®)
- venlafaxine (Effexor®)

Antipsychotics

**USE AS DIRECTED**
- quetiapine (Seroquel®)
- ziprasidone (Geodon®)

**USE WITH CAUTION**
- clozapine (Clozaril®)
- olanzapine (Zyprexa®)
- risperidone (Risperdal®)

**USE WITH CAUTION AND WITH MORE FREQUENT MONITORING**
- aripiprazole (Abilify®)
- haloperidol (Haldol®)
- perphenazine (Trilafon®)
1. APOB for familial hypercholesterolemia
2. BTK for X-linked agammaglobulinemia caused by mutations of the Bruton's tyrosine kinase gene
3. CXCR4 genotyping- determines whether a CCR5 antagonist may be an appropriate drug for a patient with HIV
4. CYP1A2-drug metabolism for drugs metabolized by this CYP enzyme (i.e. olanzapine)
5. CYP2C19-partial gene sequence based analysis for drugs metabolized by this CYP enzyme (several anti-seizure drugs, clopidogrel)
6. CYP2C9 – VKORC1 for warfarin response and resistance
7. CYP2D6-Luminex for drugs metabolized by this CYP enzyme
8. ENG and ACVRL1 sequencing-associated with hereditary hemorrhagic telangiectasia
9. FBN1 sequencing-Marfan Syndrome
10. HLA-B1502 for Identifying individuals of Asian ancestry who are at risk of developing Stevens-Johnson syndrome and toxic epidermal necrolysis when administered carbamazepine, phenytoin, or fosphenytoin therapy
11. HLA-B5701 for predicting likelihood of hypersensitivity reactions to abacavir in HIV-infected patients, based on the presence of the human leukocyte antigen (HLA)-B*5701 allele
12. LDLR for familial hypercholesterolemia.
13. PTP22 point mutation (1858C>T)-risk of erosive rheumatoid arthritis
14. TACI for common variable immunodeficiency.
15. TGFBR for diagnosis of Marfan syndrome.
16. TREC for determining immune reconstitution after BMT.
17. UGT1A1 for irinotecan sensitivity and diagnosis of Gilbert Syndrome and Crigler Najjar.
Summary

• Mayo has extensive discovery and evaluation efforts for genomic testing and practice
• Numerous experimental programs are being introduced into practice – e.g. warfarin
• Psychiatry has fully implemented genomic testing into standard workflow
• Pharmacogenomics is by far the major mode of genomic implementation at present