Navigating the Platforms

Kristin M. Thompson
Group Leader, Cancer and Medical Project Management
Navigation Can be Complex

BSP
- IRB
- Sample intake/WC
- Sample plating
- QC reports

GAP
- Genotyping
- Expression
- Reports

GSP
- Sequencing
- Data reports
Navigation Can be Complex

Sponsored Research
- Budget tracking and reports

Firehose
- Project analysis

Project Team
- Project design and analysis

BSP
- IRB
- Sample intake/WC
- Sample plating
- QC reports

GAP
- Genotyping
- Expression
- Reports

GSP
- Sequencing
- Data reports
Project Managers are your GPS

Program Project Manager
- High level project oversight
- Budget tracking

Sponsored Research
- Budget tracking and reports

Firehose
- Project analysis

Project Team
- Project design and analysis

BSP
- IRB
- Sample intake/WC
- Sample plating
- QC reports

GAP
- Genotyping
- Expression
- Reports

GSP
- Sequencing
- Data reports

Coordination of Project Team
- Integration with Platforms
- Timelines
Formation of a PM Team

Organization
- Assembly of overarching challenges facing all projects
- Overall picture of projects to aid in prioritization
- Balance of Project Manager allocation

Consistency
- Unified voice between Project Teams and Platforms
- Consistent information transfer to project teams, platforms and PIs

Supervision
- Training
- Career Development
<table>
<thead>
<tr>
<th>Project Manager</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noel Burtt</td>
<td>MPG Program Manager, Metabolic disease</td>
</tr>
<tr>
<td>Christine Stevens</td>
<td>Neurological, Metabolic, Inflammation</td>
</tr>
<tr>
<td>Deb Farlow</td>
<td>Metabolic, Cardiovascular</td>
</tr>
<tr>
<td>Namrata Gupta</td>
<td>Infectious, Cardiovascular, Inflammation</td>
</tr>
<tr>
<td>Elizabeth Nickerson</td>
<td>NHGRI Medical</td>
</tr>
<tr>
<td>Maria Cortes</td>
<td>SIGMA</td>
</tr>
<tr>
<td>Carrie Sougnez</td>
<td>Cancer</td>
</tr>
<tr>
<td>Erica Shefler</td>
<td>Cancer</td>
</tr>
</tbody>
</table>
Whole Exome Project Highlights

- Projects range from 2 samples (Congenital diaphragmatic hernia) to 3300 samples (ESP_NHLBI_Gabriel)
- Over 9000 exomes sequenced or in process for Medical projects
  - Over 2500 additional samples for Cancer Projects
- Common Challenges adding to timeline delays across projects and action plans to address them
  - Intermittent library construction failures
    - New low input protocol allows for faster turnaround of new library because additional material is at GSP
  - Topoffs for sufficient coverage
    - Early communication between platform and project team regarding turnaround vs. cost aids in better planning
# Whole Genome Projects

<table>
<thead>
<tr>
<th>Initiative</th>
<th>Narrative</th>
<th>Samples</th>
<th>Status</th>
<th>Collaborators/Leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4X Coverage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2D_Altshuler _NIDDK</td>
<td>Low pass whole genome sequencing to catalogue genetic variation associated with type 2 diabetes</td>
<td>2,600</td>
<td>136 complete – additional samples on hold due to contamination issues in the sequencing process. Increased capacity of automated WGS processing will get timeline back on schedule. Pilot due in early October to evaluate contamination resolution</td>
<td>David Altshuler</td>
</tr>
<tr>
<td>1000 Genomes</td>
<td>Sequence 2500 samples to 4x genome-wide and deep whole exome - Create catalog of 1% variants in reference genomes</td>
<td>205</td>
<td>212 complete for full project, 13 underway (in LC or sequencing) – waiting on contamination control changes to LC</td>
<td>1000 Genomes Consortium, David Altshuler</td>
</tr>
<tr>
<td>Autism_Walsh_NIMH</td>
<td>Variant discovery and homozygosity mapping in Autism patients</td>
<td>5</td>
<td>5 Complete Delayed distribution of data checking for contamination</td>
<td>Chris Walsh</td>
</tr>
<tr>
<td><strong>20-30X Coverage (Standard Protocol for Cancer Projects)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCKD1 Kidney</td>
<td>Find causal gene in MCKD1</td>
<td>6 + 8 new</td>
<td>6 complete - more coverage and additional affected samples queued</td>
<td>Mark Daly</td>
</tr>
</tbody>
</table>
## Custom Hybrid Selection

<table>
<thead>
<tr>
<th>Initiative</th>
<th>Bait size and/or # of genes</th>
<th>Pooled or barcoded</th>
<th>Total Sample</th>
<th>Status</th>
<th>Collaborators/Leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV</td>
<td>255 genes (target is 0.44 Mb)</td>
<td>Pooled in BSP (12-13 samples)</td>
<td>1000</td>
<td>LC failure</td>
<td>Paul De bakker</td>
</tr>
<tr>
<td>Height/BMI</td>
<td>2Mb</td>
<td>Pooled in BSP (12 samples)</td>
<td>1152</td>
<td>pilot initiating</td>
<td>Joel Hirschhorn</td>
</tr>
<tr>
<td>SCZ and Bipolar</td>
<td>510 genes</td>
<td>Pooled in BSP (most 20 samples)</td>
<td>~1900</td>
<td>in progress</td>
<td>Pamela Sklar, Stanley Center</td>
</tr>
<tr>
<td>CVD Allelic Spectrum : FHS/JHS</td>
<td>216 genes</td>
<td>individual barcodes</td>
<td>1600/2004</td>
<td>FHS: 876 complete; 144 finishing topoffs; 617 in redos JHS: 2004 samples requested from Jackson Heart Study</td>
<td>David Altshuler, Christine Seidman</td>
</tr>
<tr>
<td>Pfizer_T2D_MI</td>
<td>.05Mb/ 311 genes</td>
<td>individual barcodes</td>
<td>1800</td>
<td>1300 complete</td>
<td>David Altshuler</td>
</tr>
<tr>
<td>Challenge_Lipids</td>
<td>2.5 Mb bait, target is 1.8 Mb</td>
<td>individual barcodes</td>
<td>475</td>
<td>complete</td>
<td>Sekar Kathiresan</td>
</tr>
<tr>
<td>Challenge_Lipids</td>
<td>2.5 Mb bait, target is 1.8 Mb</td>
<td>individual barcodes</td>
<td>285</td>
<td>in progress</td>
<td>Sekar Kathiresan/Dan Rader</td>
</tr>
<tr>
<td>Challenge_Lipids</td>
<td>2.5 Mb bait, target is 1.8 Mb</td>
<td>individual barcodes</td>
<td>95</td>
<td>in progress</td>
<td>Sekar Kathiresan/Dan Rader</td>
</tr>
<tr>
<td>T2D_ME_NIDDK</td>
<td>2.05Mb bait size with 1.4MB of target for 571 genes/regions</td>
<td>individual barcodes</td>
<td>10,000</td>
<td>Array is in house, samples arriving this Fall</td>
<td>David Altshuler</td>
</tr>
<tr>
<td>Prostate_Reich_NIH</td>
<td>0.8MB</td>
<td>individual barcodes</td>
<td>100</td>
<td>Project Planning Underway</td>
<td>David Reich</td>
</tr>
<tr>
<td>Inflammatory Bowel Disease - GWAS hits</td>
<td>4MB</td>
<td>individual barcodes</td>
<td>5000</td>
<td>Project planning underway</td>
<td>Mark Daly + IBD consortium</td>
</tr>
<tr>
<td>Breast Cancer risk – GWAS hits</td>
<td>5MB</td>
<td>individual barcodes</td>
<td>5200</td>
<td>Project planning underway</td>
<td>Peter Kraft</td>
</tr>
<tr>
<td>MCKD1 Kidney</td>
<td>N/A (Nimblegen)</td>
<td>15</td>
<td></td>
<td>in progress</td>
<td>Mark Daly</td>
</tr>
</tbody>
</table>
Large vs. Small Projects

From a project management perspective many parts of a project are agnostic to sample number

IRB
- Generally there is a single IRB document per project; can have multiple cohorts

Sample collection
- Intensity varies with the number of cohorts not the number of samples

Sample QC
- Assays are formatted for high-throughput (48-384)

Sequencing
- Capacity per week is 100s of exomes and 10s of genomes and is expected to increase to 100s
Sample tracking is currently tracked in excel spreadsheets and does scale by sample number.

Efforts are underway to generate a database that will aid in tracking of projects.
How To Initiate a Project

Standard offerings – Genotyping, Exomes, Genomes and custom hybrid selection

Email- MPGNewProject@Broadinstitute.org

If you are interested in a non-standard projects contact the platform involved directly