Crowdsourced analysis with Quilt

Aneesh Karve, CTO
Quilt Data, Inc.
Today: divisive computing
Today: divisive computing

<table>
<thead>
<tr>
<th>Task</th>
<th>Who’s doing it?</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETL</td>
<td>Many</td>
<td>Low</td>
</tr>
<tr>
<td>Analyze</td>
<td>Few</td>
<td>High</td>
</tr>
<tr>
<td>Share</td>
<td>Few</td>
<td>High</td>
</tr>
<tr>
<td>Reuse</td>
<td>Few</td>
<td>High</td>
</tr>
</tbody>
</table>
Tomorrow: crowd computing

<table>
<thead>
<tr>
<th>Task</th>
<th>Who’s doing it?</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETL</td>
<td>Few</td>
<td>Low</td>
</tr>
<tr>
<td>Analyze</td>
<td>Many</td>
<td>High</td>
</tr>
<tr>
<td>Share</td>
<td>Many</td>
<td>High</td>
</tr>
<tr>
<td>Reuse</td>
<td>Many</td>
<td>High</td>
</tr>
</tbody>
</table>
Tomorrow: crowd computing
Barriers to crowd computing

- Disincentives to share (competition)
- Paywalls
- High fragmentation, low findability
- High technical barriers
Lowering the barriers

- Social layer (incentive, impact)
- Data management layer (provenance)
- Unified storage & schemas
- Lower technical barriers (no coding)
Experiments in crowd computing

- Can bench scientists create targeted CRISPR screens?
- Future: Will they share gRNA effectiveness?
1: Genome-wide CRISPR library

- Find SpCas-9 PAMs in hg19
- Intersect with DNase Hypersensitive Sites from ENCODE
- Filter for off-target effects
- https://quiltdata.com/app/table/1592/
2: Foundations for CRISPR screens

- Load, tag ChIP-Seq peaks from ENCODE into one database
- Intersect and subtract intervals without coding
3: Users identify targets

- Enhancers = (mel + ac) - me3
- Find NANOG enhancers in ESCs
  - (H3K4mel + H3K27ac) - H3K4me3
Select an action

Intersect Genomes

Keep all regions that overlap with any region in the second data set.
4: Users design CRISPR screen

- Enhancers + Genome-wide CRISPR library
  - gRNAs that target NANOG enhancers
- Next: sharing experimental results
Partner with Us

- Free storage, compute, visualization
- CRISPR, variant interpretation, ChIP-Seq, etc.
- QuiltData.com
- aneesh@quiltdata.io