Crowdsourced analysis with Quilt
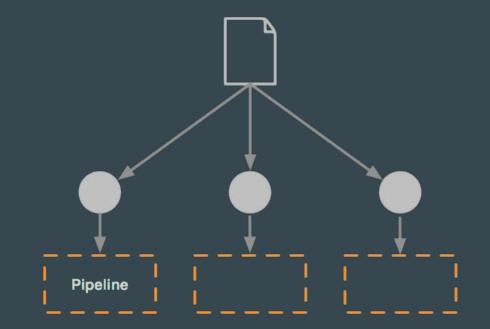
• • •

Aneesh Karve, CTO
Quilt Data, Inc.

Today: divisive computing

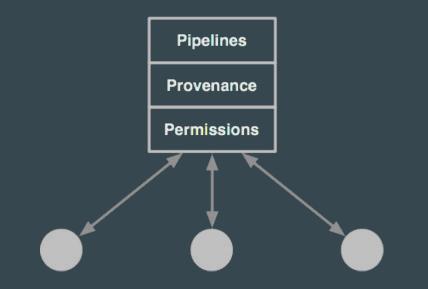# Today: divisive computing

Task	Who's doing it?	Value
ETL	Many	Low
Analyze	Few	High
Share	Few	High
Reuse	Few	High

Tomorrow: crowd computing

Task	Who's doing it?	Value
ETL	Few	Low
Analyze	Many	High
Share	Many	High
Reuse	Many	High

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?
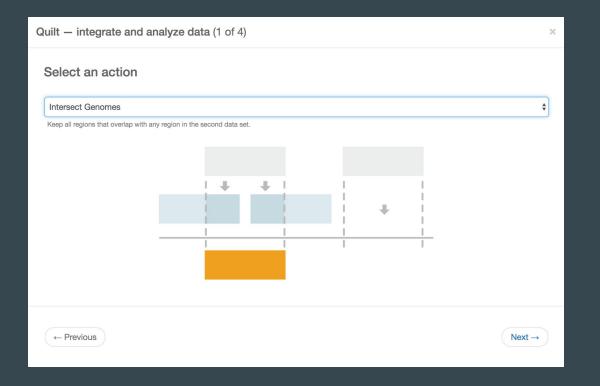
1: Genome-wide CRISPR library

- 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 = (me1 + ac) - me3
- Find NANOG enhancers in ESCs
  - (H3K4me1 + H3K27ac) - H3K4me3

Select an action

Intersect Genomes

Keep all regions that overlap with any region in the second data set.

← Previous

Next →

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