

8/5/07

Tim Hubbard -

Sanger -

- Founded in 1992
- Strong Open Data release policy + open infrastructure resources
- Expanded in 2000/1 - to more model organisms
- funded by Wellcome Trust
- Campus includes EBI, and conference centers

Mission - use large scale sequencing

Large scale resources

- mouse KOs
 - Zebrafish KOs
 - make lists of genes for all to work on -
2006/7 are chromatin related
-) Can request KOs online

Still sequencing

- Zebrafish, Gorilla,
- human variation
 - Reference genome
 - Seq. variation
 - Cancer genome
 - Hapmaps
- copy number variation

Sequencing Revolution

- Big drop in cost
- Need informatics resources

Sequencing 200 genomes

- Use ancestral recomb graph w/ low error rates - tested on yeast

Cancer resequencing

Leena Peltonen joining Sanger

Biological Data

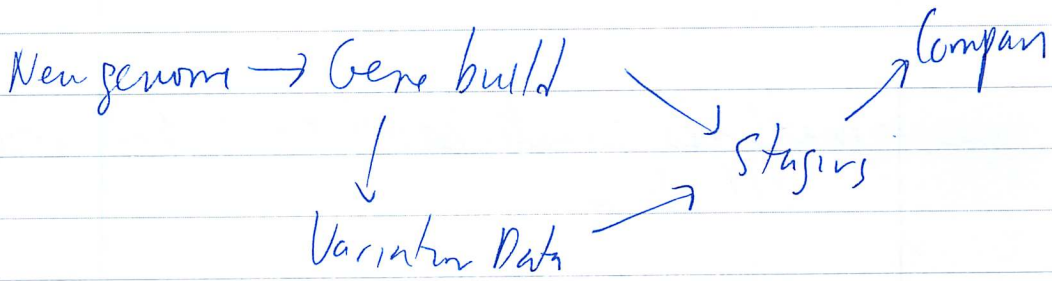
- Annotation + gene ID
- Gene classification

Ever closer union between Sanger/EBI data resources

Ensembl

- Genome annotation
- Comp genomics
- Infrastructure

Data flow



Access to Data

Download

BioMart: structured data search engine

DAS

Pedigrees? not done

What about rest of genome

Predictions

structure

Genes

pragmatic
in between
pure

comp. modelling

evidence

molecular assembly; fragments

Comp.
ab-initio

Why hasn't ab. initio worked?

- Methods aren't based on what organism, cell does
 - motif based prediction is more like what cell does.
- 