What is next?

- *The "Big" Sequence Problem*
- *New Data Types & Homologous objects & Co-Modelling*
- *Species and Populations*
- *Questions*
- *Wrap-up*

The "Big" Sequence Problem

Genealogical Structure

Phylogeny
Ancestral Recombination Graph
Pedigree

Annotation (Hidden States)

Protein Genes
RNA
Signals
Evolving Annotation
Superimposed Annotations

Biological Knowledge

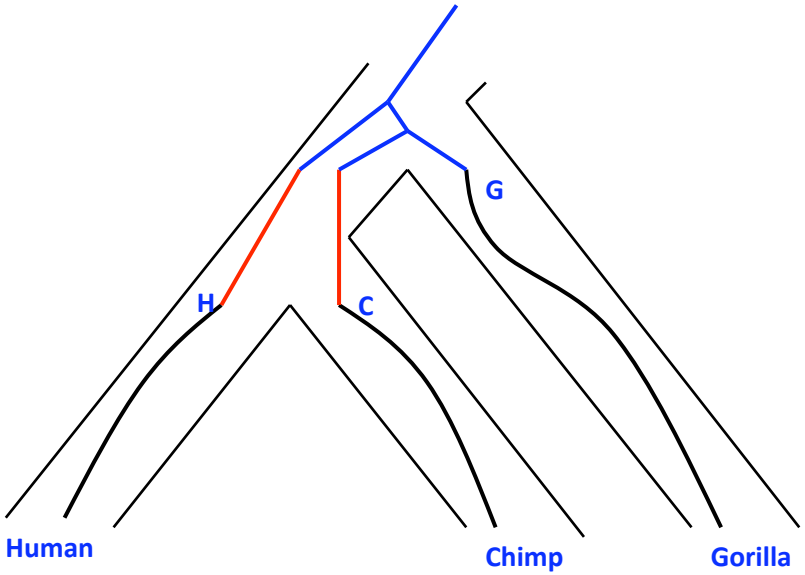
Priors on Annotations
Ontologies

Models of Evolution

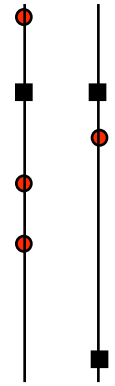
Substitutions
Insertion-Deletions
Genomic Events

- Combined or Sequential Solution?
- Realism in face of massive sequences
- Non-homologous and homologous analysis

Species, Populations, Individuals, Sequences and Mapping



Genetype -->Phenotype Function



Dominant/Recessive.

Penetrance

Spurious Occurrence

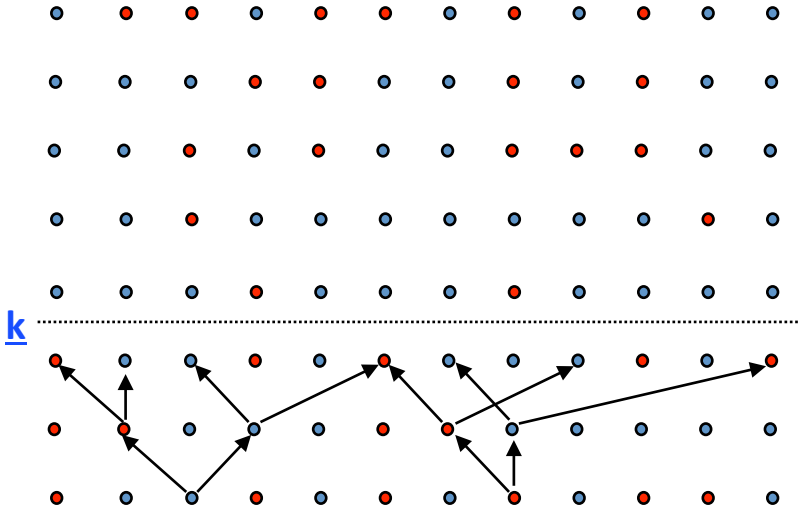
Heterogeneity

A set of characters.

Binary decision (0,1).

Quantitative Character.

genotype *Genotype → Phenotype* phenotype



- Mapping on the Tree of Life ??
- Knowledge based Genotype→Phenotype mapping ??
- The limits to pedigree inference

New Data Types & Homologous objects & Co-Modeling

High throughput

Expression Data
Protein Structure
Proteomics
…..

Homologous Objects

Genome Structure
Networks
- PIN
- Regulatory
- Transduction
- Metabolic (?)

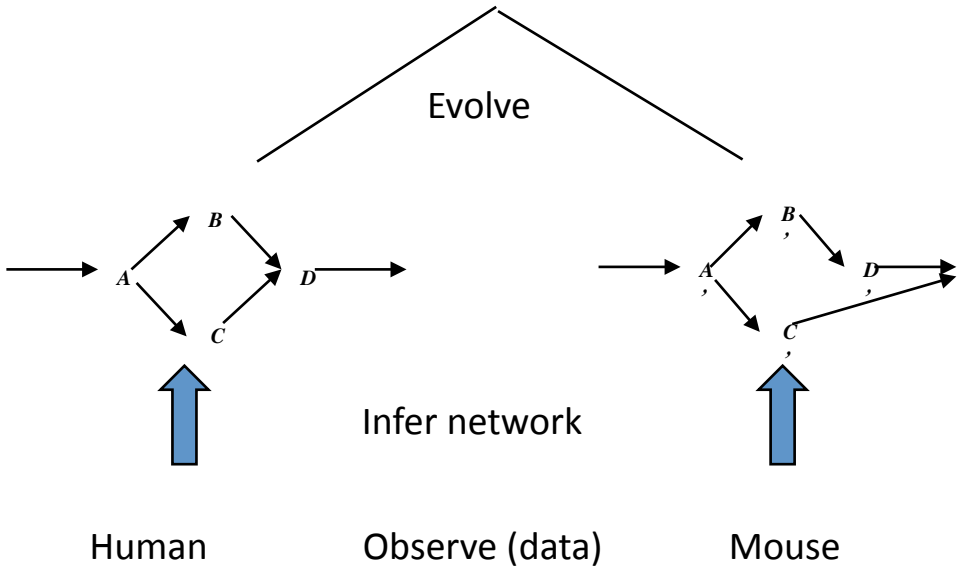
Protein Structure (?)
RNA Structure
Patterns (?)
Shape (?)
Dynamical Systems (?)
Molecular Motion
Phenotype (?)

Co-Modeling

Sequence-Gene Structure
Sequence-RNA Structure
Others???
Triple Modeling ??

The Necessity of Evolutionary Modeling

$$P(D_{mouse}, D_{human}) = P(D_{human} | N_{human}) P(D_{mouse} | N_{mouse}) P(N_{human}, N_{mouse})$$



Framework for Comparative Biology