MCB 41

Henk Roelink

Office hours:

Fridays 1:00-2:30 (13:00-14:30)
21 Koshland

or

By appointment
roelink@berkeley.edu

Class reading: Chapter 11, page 148-151 (enhancers)
Sequencing Genomes

Why???

Curiosity
Evolution
Speciation
Model systems
Diseases
Relationships
Sequencing Genomes

How much DNA is there in a Genome?

- Prokaryote: $1 \times 10^6 - 10 \times 10^6$ bp
- Yeast: $12 \times 10^6$ bp
- Fish: $0.4 \times 10^9 - 3 \times 10^9$ bp
- Mammals: $1.5 \times 10^9 - 8 \times 10^9$ bp

- Bible: $3.5 \times 10^6$ letters
- Milky Way Galaxy: $10^{11}$ stars
How much DNA is there in a Genome?
How to Sequence a Genome

1/ DNA: get your organism and extract DNA
2/ Cut into small fragments and (sub)clone
3/ **Sequence** each subclone
4/ Assemble
5/ Compare with other sequences
Genes are very old
More “complexity” does not require more genes
Blue Ridge Mountains
Oldest Mountains in the USA, $10^9$ years old
These animals have largely the same genes

These animals have many genes in common

These animals have quite a few genes in common
If animals are made from the same gene products, what makes a mouse different from a frog or a bird?
Differences are Relative

The limb bones of different mammals are clearly equivalent. We can “morph” one into another just by changing the sizes and proportions. Genetic mutations altering the gene-control proteins, the diffusible molecules, and the patterns of expression of the limb-development genes, are sufficient to make these changes in living things. One mutation occurs at a time, and the internal control mechanisms of embryo development integrate the body parts to form a working whole.
There is more to a gene than the protein coding sequences.
How do you recognize these non-coding regulatory sequences (enhancers)?

http://ecrbrowser.dcode.org/
• Overall genomes look a lot alike
• Sequence information is much more valuable when compared with other species
• Sequences conserved over many many years are likely to be important
• The devil is in the details: relevant differences are harder to find than conserved sequences

Functional tests are cumbersome and costly, but really interesting!
Comparison of bat and mouse forelimb morphogenesis and $Prx1$ expression
Forelimb length of Prx1 mutants

How about sequencing your favorite organism?

1.5 year doubling time
Currently: **Knome** offers a $10^5$ personal sequence ($10^{-5}$/base, 4 month halving time)

A $1000$ ($10^{-7}$/base) genome 2 years away????
All of you will know your complete genome sequence if you so desire, and are willing to dish out <\$1000

• Ancestry
• Diseases / Life expectancy
• Partner matching
• Personality traits
• Personalized Medicine
Who might be interested in your DNA sequence?

• Your Health Insurance
• Your Life Insurance
• Your Mortgage Bank
• Your Employer
• The Police
• The CIA
• The Federal Government
• The US Olympic Committee
• Your Family (is your father really your father?)
Still want to sequence your genome?

Who will own the data?