

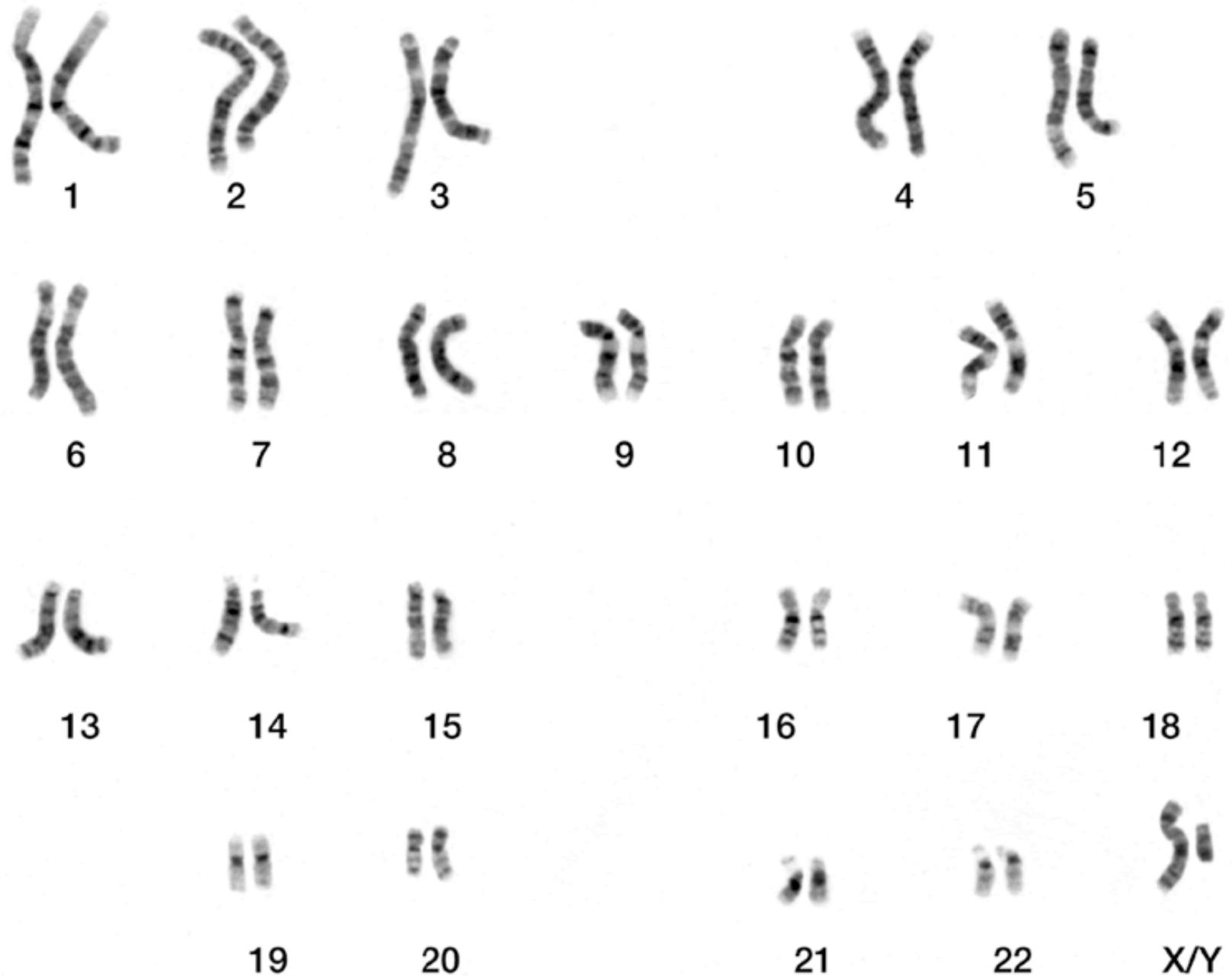
Genome Analysis and Comparison

- Chromosomes
- Genomic Maps
- Genome Comparison
- Genome Alignment

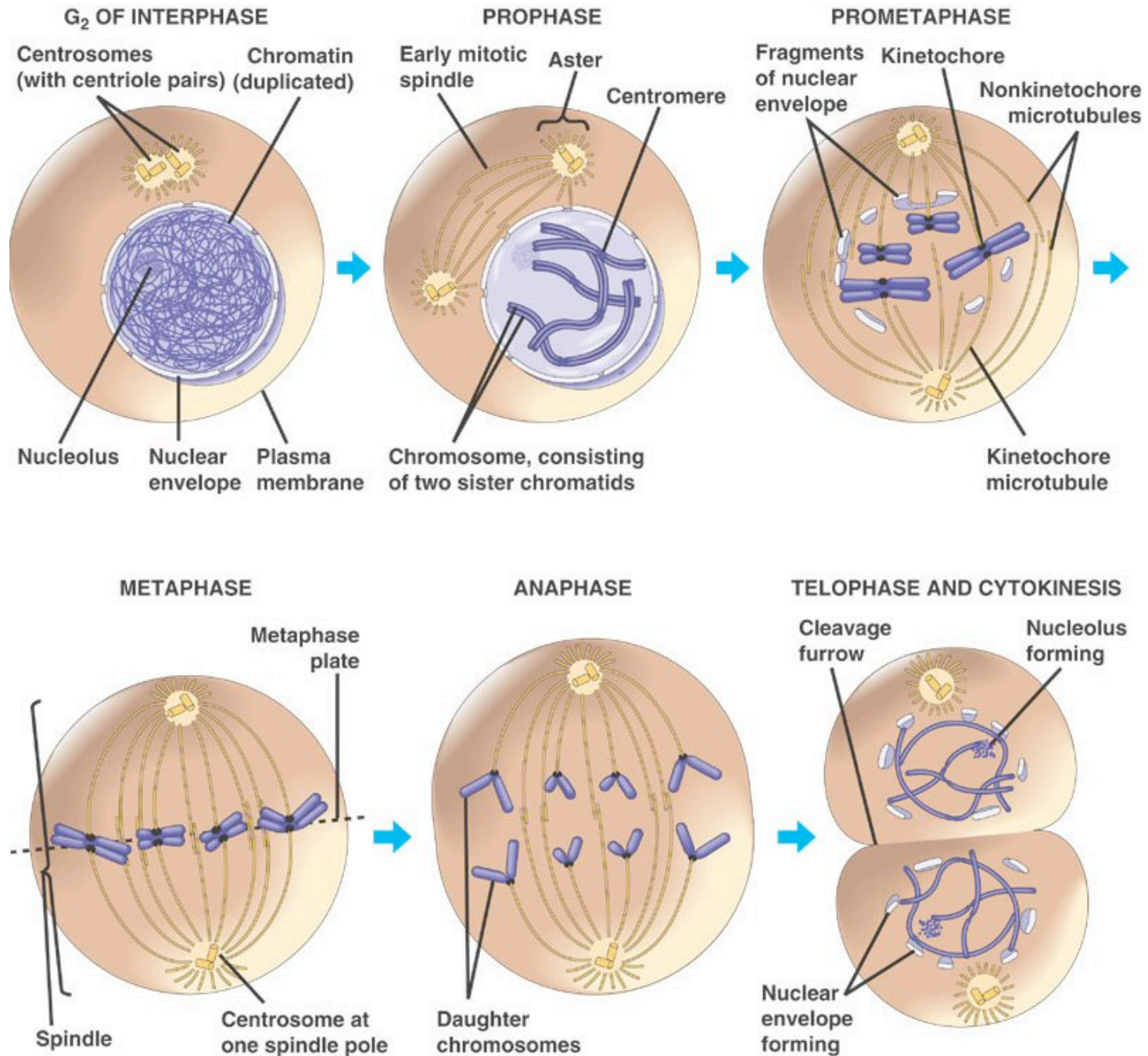
Comparative Genomics

Comparative genomics includes a comparison of gene number, gene content and gene location in both prokaryotic and eukaryotic groups of organisms. The availability of complete genome sequences makes possible a comparison of all of the proteins encoded by one genome, the proteome of that organism, with those of another.

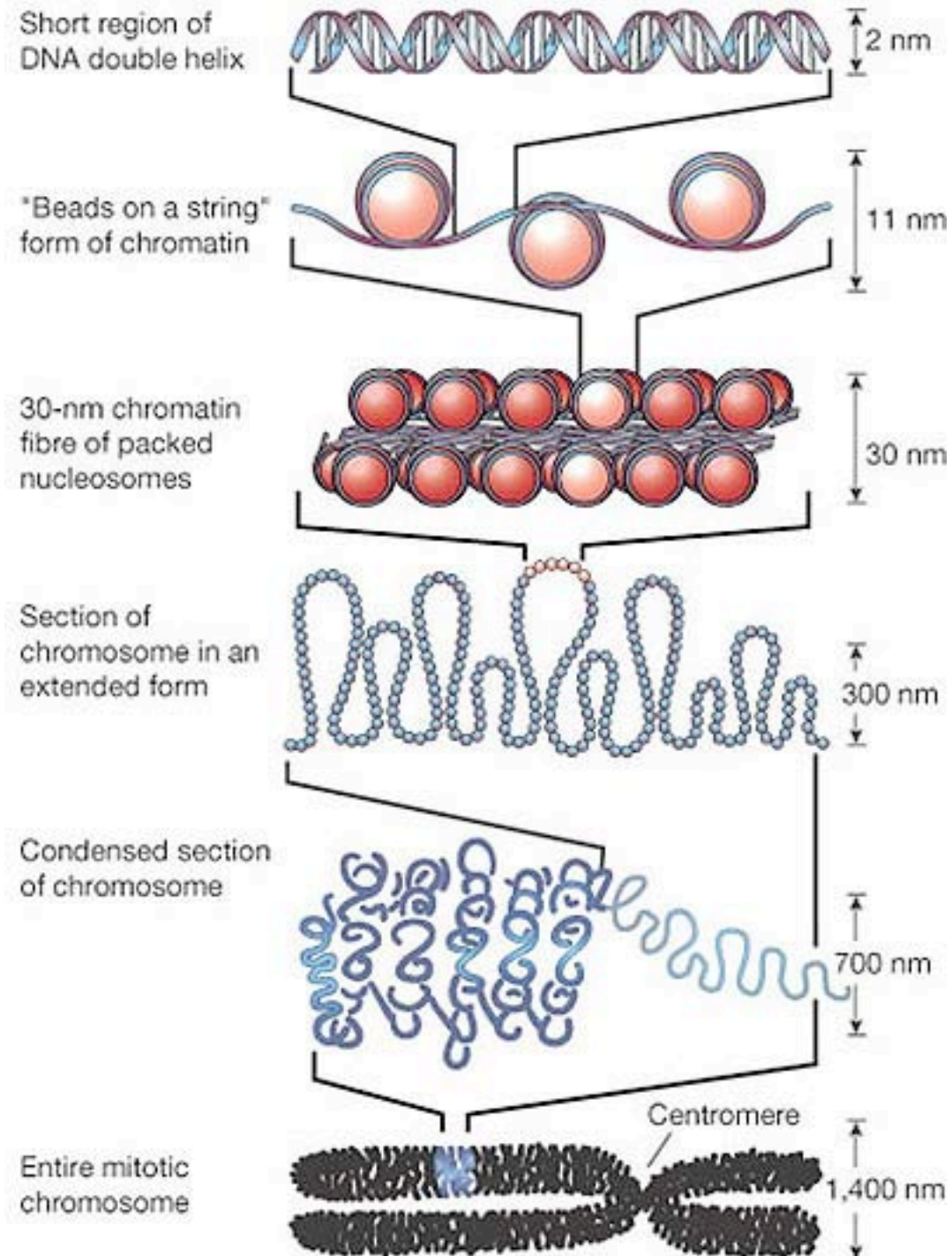
Human Chromosomes



Mitosis



The Chromosome Structure



Shotgun Sequencing

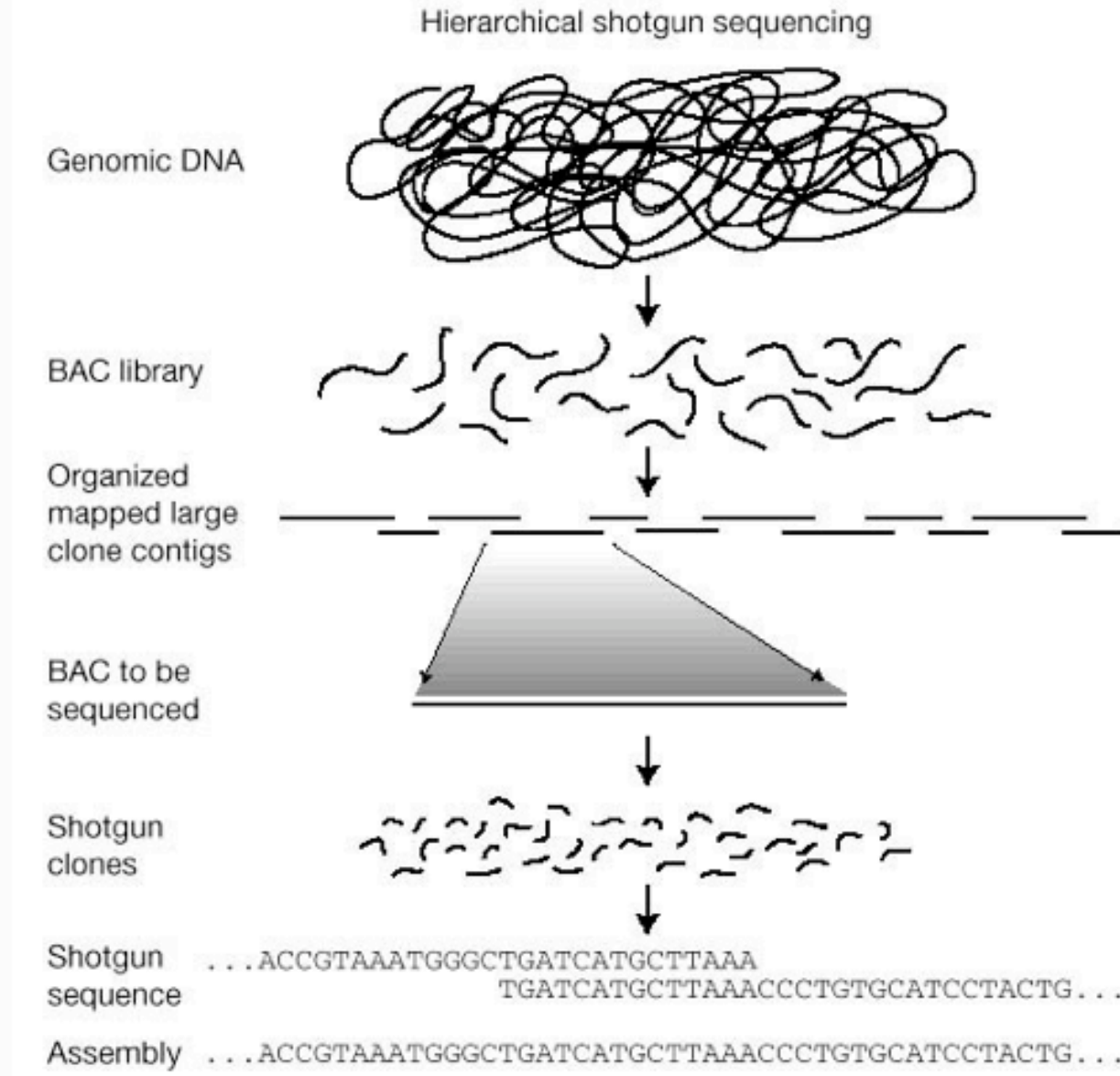
The classical 'chain termination' method of DNA sequencing is limited to 100-1000 basepairs.

To allow for sequence information of longer sequences, DNA can be broken randomly into shorter fragments.

The randomly terminated fragments can be re-assembled to the longer sequence.

Strand	Sequence
Original	AGCATGCTGCAGTCATGCTTAGGCTA
First shotgun sequence	AGCATGCTGCAGTCATGCT----- -----TAGGCTA
Second shotgun sequence	AGCATG----- -----CTGCAGTCATGCTTAGGCTA
Reconstruction	AGCATGCTGCAGTCATGCTTAGGCTA

Shotgun Sequencing



A library is constructed by fragmenting the target genome and cloning it into a large-fragment cloning vector; here, BAC vectors are shown. The genomic DNA fragments represented in the library are then organized into a physical map and individual BAC clones are selected and sequenced by the random shotgun strategy. Finally, the clone sequences are assembled to reconstruct the sequence of the genome.

Genomic Internet Resources

- **Entrez** <http://www.ncbi.nlm.nih.gov/Entrez>
- **LocusLink** <http://www.ncbi.nlm.nih.gov/LocusLink/>
- **GeneMap'99** <http://www.ncbi.nlm.nih.gov/genemap99/>
- **OMIM** <http://www.ncbi.nlm.nih.gov/Omim/>
- **Gene Ontology** <http://www.geneontology.org/>
- **HomoloGene** <http://www.ncbi.nlm.nih.gov/HomoloGene/>
- **BLAST** <http://www.ncbi.nlm.nih.gov>
- **ePCR** <http://www.ncbi.nlm.nih.gov/STS/>
- **GenBank** <http://www.ncbi.nlm.nih.gov/Genbank>
- **Genomic Biology page** <http://www.ncbi.nlm.nih.gov>
- **Mouse Genome Informatics (MGD/MGI)** <http://www.informatics.jax.org/>
- **Ensembl** <http://www.ensembl.org/index.html>

Useful Meta-Sites

ENTREZ

A cross-database search engine for biological and medical information.
PubMed, Blast, GeneViewer, GenBank, OMIM

OMIM - Online Mendelian Inheritance in Man

Compendium of human genes and genetic phenotypes.
Contains information on all known mendelian disorders and over 12,000 genes.
OMIM focuses on the relationship between phenotype and genotype.

GO - Gene Ontology

Links various levels of biological information.
Based on keyword association in databases.

Genomes of Model Organisms

Escherichia coli

One of the most frequently studied organisms.

Mycoplasma genitalium

The smallest known genome of all known cellular organisms.
Reduction of that genome shows that about 250-300 genes are required for *Mycoplasma*'s survival.

Bacillus subtilis

Like *E. coli* one of the most popular organisms in Molecular Biology.

Saccharomyces cerevisiae

Being a yeast, it is used to study the eukariotic cell.

Whole-Genome Alignment

BLASTN

Not designed for whole genome alignment.

MUMmer (Kurtz et al., Genome Biology 2004, 5:R12)

Residing at the core of the MUMmer package is the 'mummer' matching algorithm, which builds and searches a suffix tree data structure.

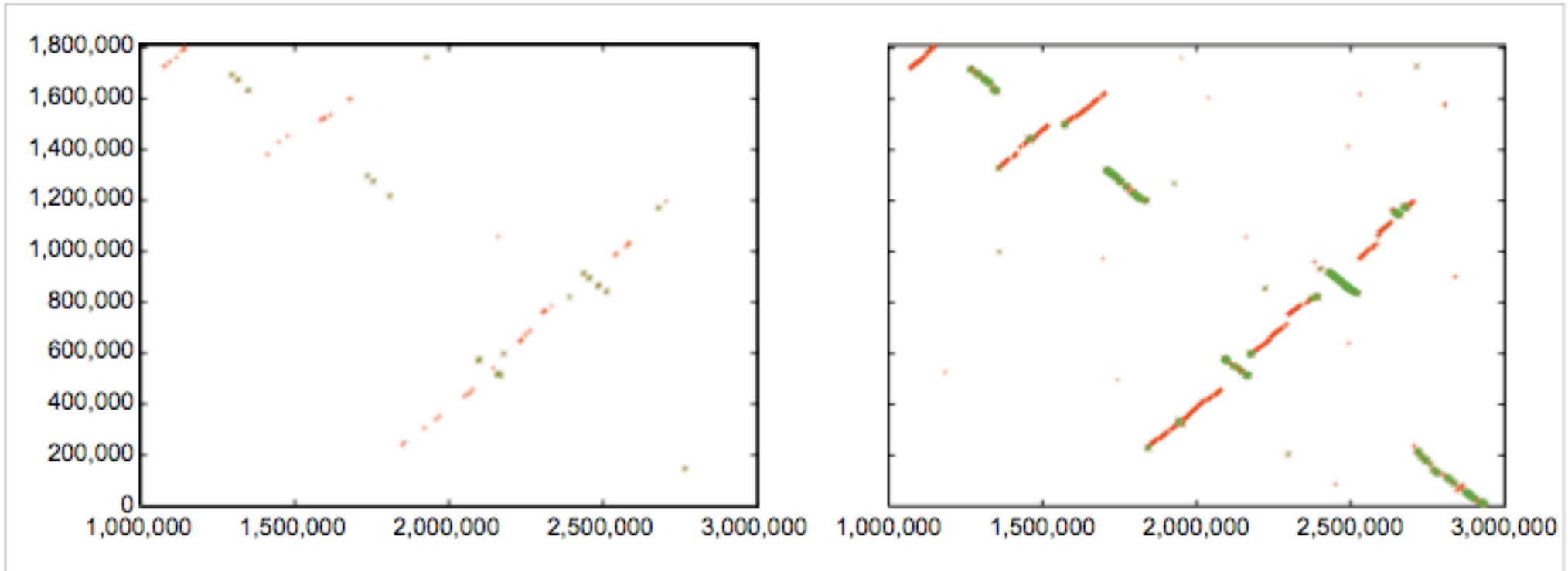
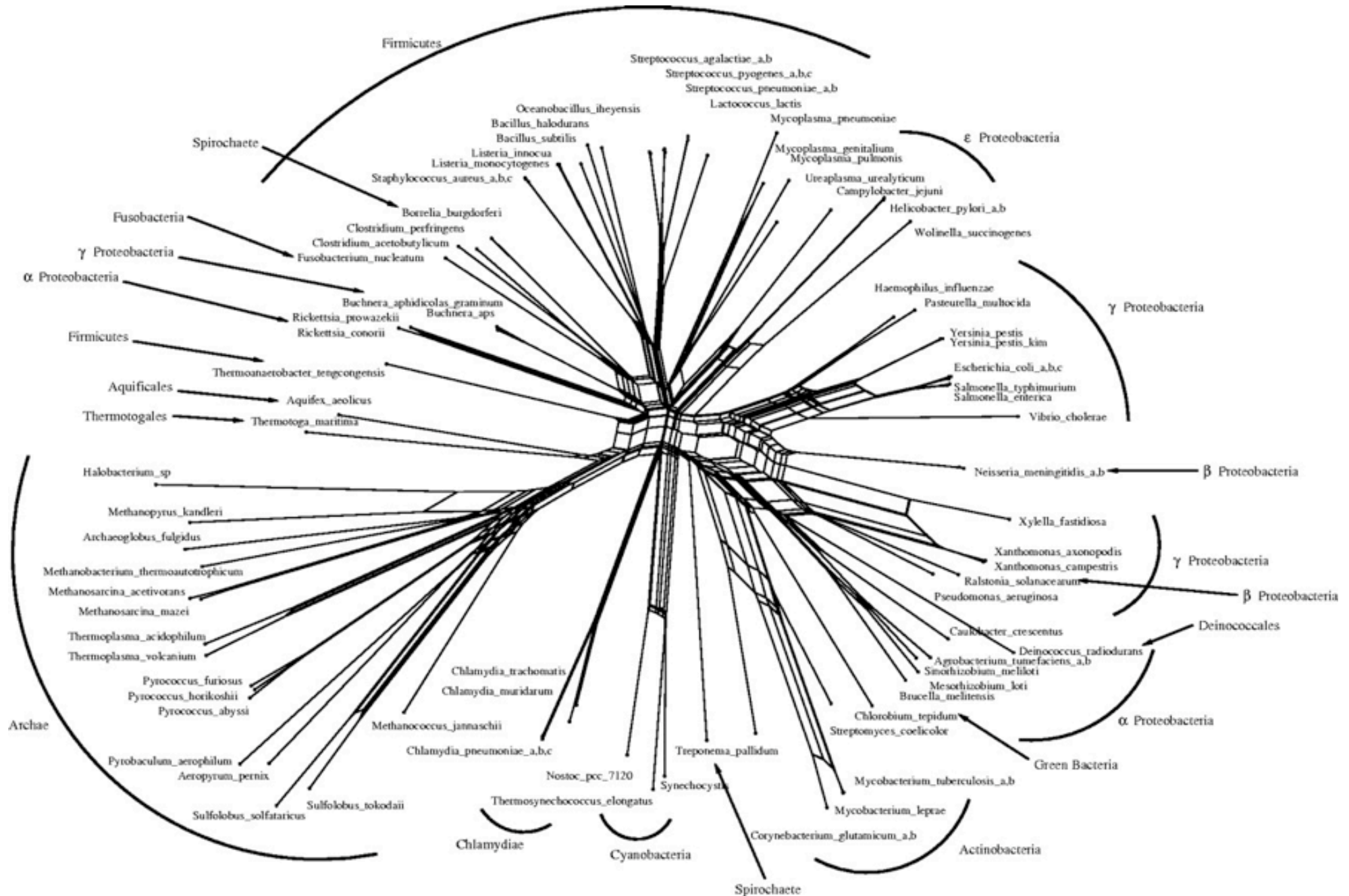


Figure 1

Dot-plot alignments of a 2.9 Mbp chromosome of *A. fumigatus* (x-axis) to a 2.1 Mbp scaffold of *A. nidulans* (y-axis). Left: nucleotide-based alignment with Nucmer. Right: amino-acid-based alignment with Promer. Aligned segments are represented as dots or lines, up to 3,000 bp long in the Nucmer alignment and up to 9,500 bp in the Promer alignment. These alignments were generated by the mummerplot script and the Unix program gnuplot.

Whole-Genome Phylogeny



Henz et al., Bioinformatics 2005 21(10):2329-2335

Learning Outcomes

- Chromosome Organisation and Structure
- Principles of Genomic Maps
- Shotgun Sequencing
- Genomic Reference Organisms and Tools
- Methods for Genome Comparison