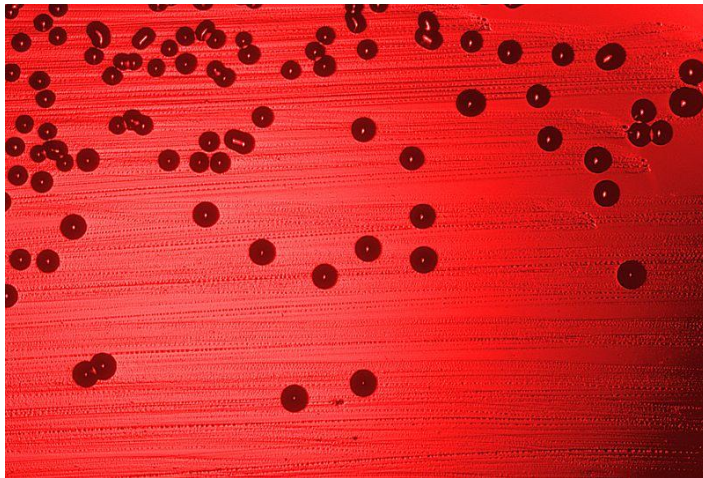


Genome projects

Jereesh A S

Whole genome sequencing

- *Haemophilus influenzae* (1995)
 - commensal bacterium which resides in the human respiratory tract was the first organism to have its entire genome sequenced



Haemophilus influenzae

<http://science.sciencemag.org/content/269/5223/496>

1,830,140 base pairs of DNA

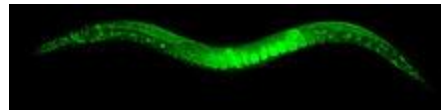
Whole genome sequencing

- yeast *Saccharomyces cerevisiae* (1996)
 - The first eukaryotic genome was sequenced.
- Has a genome of only around 12 million nucleotide pairs
- <https://dx.doi.org/10.1126%2Fscience.274.5287.546>



Whole genome sequencing

- The worm *Caenorhabditis elegans* (1998)
 - The first animal to have its whole genome sequenced.
- Has a genome of only around 100,258,171 nucleotide pairs.



- <https://dx.doi.org/10.1126%2Fscience.282.5396.2012>



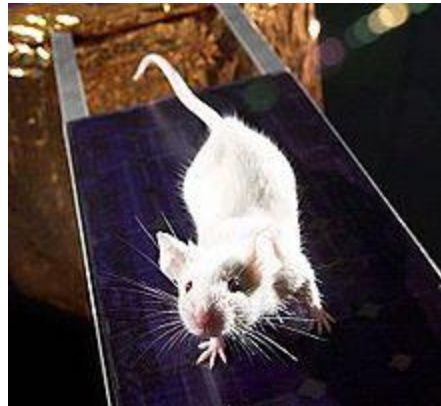
Whole genome sequencing



- In 1999, the entire DNA sequence of human [chromosome 22](#).
- By the year 2000, the second animal and second invertebrate (yet first insect) genome was sequenced - that of the fruit fly [*Drosophila melanogaster*](#)
- The first [plant](#) genome - that of the model organism [*Arabidopsis thaliana*](#) - was also fully sequenced by 2000

Whole genome sequencing

- The genome of the laboratory mouse *Mus musculus* was completed in 2002



Whole genome sequencing

In 2004, the [Human Genome Project](#) published the human genome.

articles

Finishing the euchromatic sequence of the human genome

International Human Genome Sequencing Consortium*

**A list of authors and their affiliations appears in the Supplementary Information*

The sequence of the human genome encodes the genetic instructions for human physiology, as well as rich information about human evolution. In 2001, the International Human Genome Sequencing Consortium reported a draft sequence of the euchromatic portion of the human genome. Since then, the international collaboration has worked to convert this draft into a genome sequence with high accuracy and nearly complete coverage. Here, we report the result of this finishing process. The current genome sequence (Build 35) contains 2.85 billion nucleotides interrupted by only 341 gaps. It covers ~99% of the euchromatic genome and is accurate to an error rate of ~1 event per 100,000 bases. Many of the remaining euchromatic gaps are associated with segmental duplications and will require focused work with new methods. The near-complete sequence, the first for a vertebrate, greatly improves the precision of biological analyses of the human genome including studies of gene number, birth and death. Notably, the human genome seems to encode only 20,000–25,000 protein-coding genes. The genome sequence reported here should serve as a firm foundation for biomedical research in the decades ahead.

Thank you