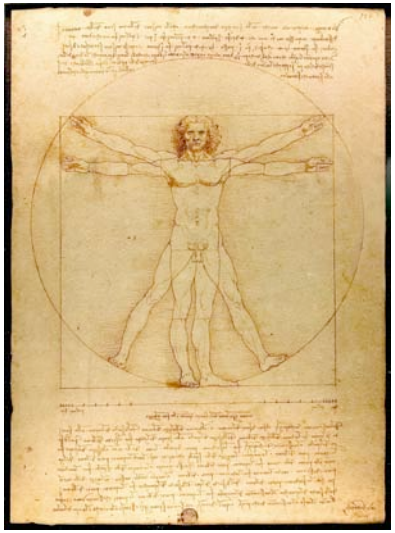
2. Genes and proteins

- The sequence as a model of DNA
- Genes: from Mendel to molecular biology
- The genetic code
- A translation algorithm
- Implementing the genetic code
- Algorithms + data structures = programs
- The algorithm design trade-off
- DNA sequencing
- **Whole genome sequencing**
- How to find genes?

Whole genome sequencing

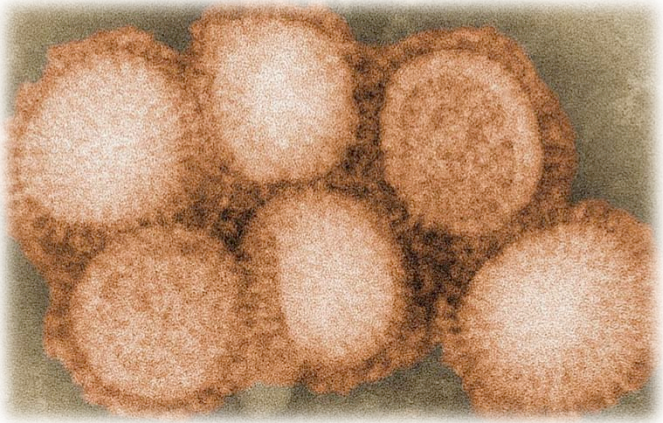
Whole genome sequencing

- *B. subtilis* project
  - 1989 – 1998
  - 35 laboratories
  - today some 100$, within a day
- Human genome project
  - 1990 – 2003
  - 2.7 billions 1991 dollars
  - tomorrow < 1000$?



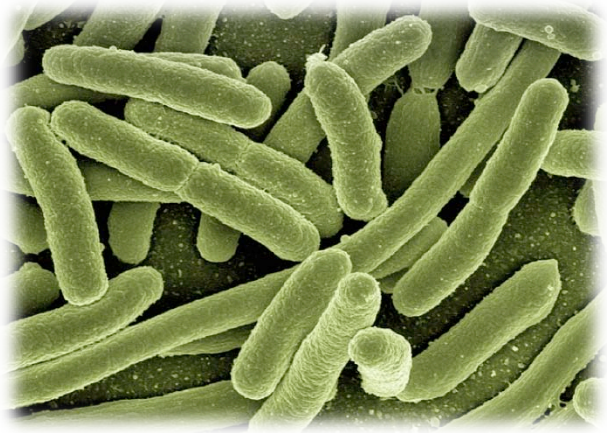
Some orders of magnitude

- Virus (influenza) $1.3\ 10^4$



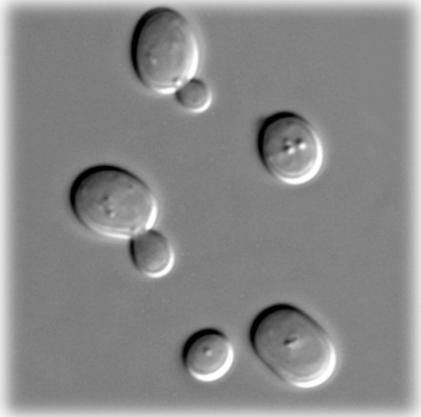
Some orders of magnitude

- Virus (influenza) $1.3\ 10^4$
- Bacteria (*E. coli*) $4.6\ 10^6$



Some orders of magnitude

- Virus (influenza) $1.3\ 10^4$
- Bacteria (*E. coli*) $4.6\ 10^6$
- Yeast $1.2\ 10^7$



Some orders of magnitude

- Virus (influenza) $1.3\ 10^4$
- Bacteria (*E. coli*) $4.6\ 10^6$
- Yeast $1.2\ 10^7$
- Drosophila $1.2\ 10^8$

Some orders of magnitude

- Virus (influenza) $1.3\ 10^4$
- Bacteria (*E. coli*) $4.6\ 10^6$
- Yeast $1.2\ 10^7$
- Drosophila $1.2\ 10^8$
- *H. sapiens* $3.4\ 10^9$

Some orders of magnitude

- Virus (influenza) $1.3\ 10^4$
- Bacteria (*E. coli*) $4.6\ 10^6$
- Yeast $1.2\ 10^7$
- Drosophila $1.2\ 10^8$
- *H. sapiens* $3.4\ 10^9$
- Maize $5.0\ 10^9$

Some orders of magnitude

- Virus (influenza) $1.3\ 10^{4}$
- Bacteria (*E. coli*) $4.6\ 10^{6}$
- Yeast $1.2\ 10^{7}$
- Drosophila $1.2\ 10^{8}$
- *H. sapiens* $3.4\ 10^{9}$
- Maize $5.0\ 10^{9}$
- Ameboa $6.7\ 10^{11}$

The sequence is only the starting point

- Sequence annotation, i.e. prediction of
  - Gene location
  - Gene / protein functions
  - Gene interactions
- Sequence comparison, between:
  - Species
  - Strains
  - Individuals

Pictures & movies : material licensing

p. 3 : Bonchien at en.wikibooks [Public domain], from Wikimedia Commons

p. 3 : Leonardo da Vinci [Public domain], via Wikimedia Commons

p. 4 : Cybercobra at en.wikipedia [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], from Wikimedia Commons

p. 5 : Public domain

p. 6 : Masur (Own work) [Public domain], via Wikimedia Commons

p. 7 : André Karwath aka Aka (Own work) [CC BY-SA 2.5 (http://creativecommons.org/licenses/by-sa/2.5)], via Wikimedia Commons

p. 8 : Rights reserved

p. 9 : liz west - CC BY 2.0

p. 10 : https://microbewiki.kenyon.edu/index.php/File:Amoeba_proteus_x_100.jpg#filelinks