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?

François
Rechenmann

Genes: from Mendel to molecular biology

Genes: from Mendel…

- Between 1856 and 1863, Gregor Mendel studies the inheritance of traits, such as seed shape (round or wrinkled), between pea plants

- In 1866, Mendel's laws of inheritance

- No physical support was identified

- Work unknown by Darwin

- The laws were rediscovered 50 years later…
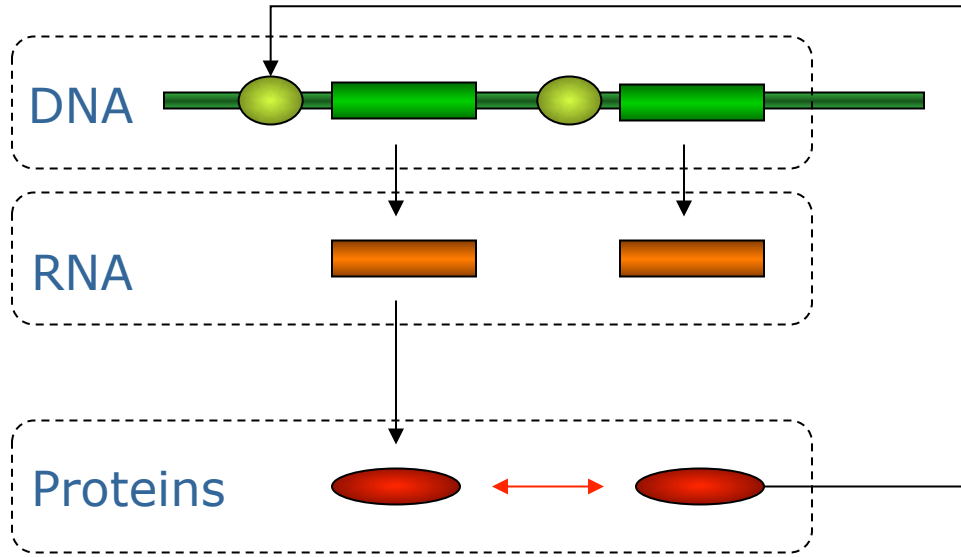
… to molecular biology

- Genes are regions of DNA which contain the information used by the cell to build proteins

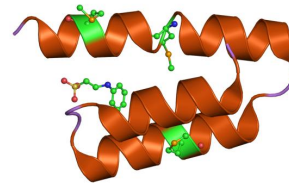- The "central dogma of molecular biology" (Francis Crick, 1956):

    one gene ➜ one protein

- At that time, quite a reasonable hypothesis

Translation

DNA

RNA

Proteins

Transcription

- From DNA to RNA
  - DNA: Deoxyribonucleic acid
  - RNA: Ribonucleic acid
- Single stranded
- Thymine ➔ Uracil
  - T ➔ U

ATTCGATCGGGTATTACG*

↓

AUUCGAUCGGGUAUUACG*

A transcription algorithm

```
i: integer
sequence: character string [1:*]
i ← 1
repeat
  if sequence[i] = "T"
    then sequence [i] ← "U"
  i ← i + 1
until sequence [i] = "*"
```

ATTCGATCGGGTATTACG*

↓

AUUCGAUCGGGUAUUACG*

Why transcription into RNA?

- DNA is the information archive to be preserved

- RNA is a copy to work on

- mRNA: messenger RNA
  - limited life time
  - subjected to degradation

Pictures & movies : material licensing

p. 3 : CC Bmdavll (talk) :
http://en.wikipedia.org/wiki/Image:Snow_pea_flowers.jpg

p. 3 : G. Mendel picture, Public domain

p. 5 : Public domain, *http://www.ebi.ac.uk/*