

1. Genomic texts

- The cell, atom of the living world
- At the heart of the cell: the DNA macromolecule
- DNA codes for genetic information
- What is an algorithm?
- Counting nucleotides
- **GC and AT contents of DNA sequence**
- DNA walk
- Compressing the DNA walk
- Predicting the origin of DNA replication?
- Overlapping sliding window

```
nbA,nbC,nbG,nbT, TotalNb, index: integer  
sequence: character string [1:*
```

```
nbA,nbC,nbG,nbT, TotalNb  $\leftarrow$  0  
index  $\leftarrow$  1
```

```
repeat
```

```
  case sequence [index] of
```

```
    "A": nbA  $\leftarrow$  nbA + 1
```

```
    "C": nbC  $\leftarrow$  nbC + 1
```

```
    "G": nbG  $\leftarrow$  nbG + 1
```

```
    "T": nbT  $\leftarrow$  nbT + 1
```

```
  endcase
```

```
  TotalNb  $\leftarrow$  TotalNb + 1
```

```
  index  $\leftarrow$  index + 1
```

```
until sequence [index] = "*"
```

```
display "Length of the sequence:" TotalNb
```

```
display "%A=" (nbA/TotalNb)*100, "    %C=", (nbC/TotalNb)*100, "    %G=", (nbG/  
TotalNb)*100, "    %nbT=", (nbT/TotalNb)*100
```

The input of our algorithm

AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAGAGTGTCTGATAGCAGC*

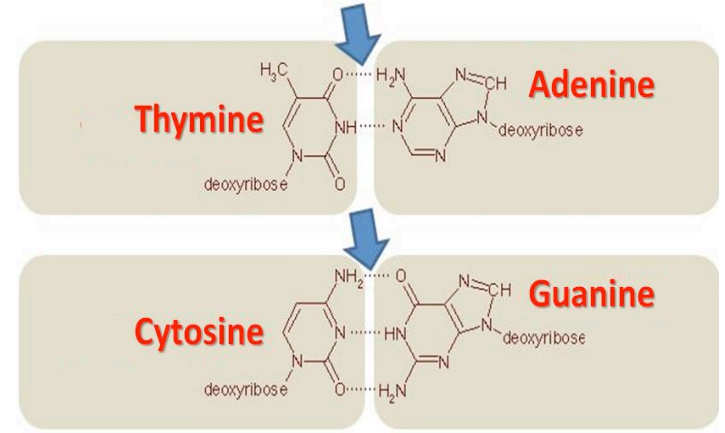
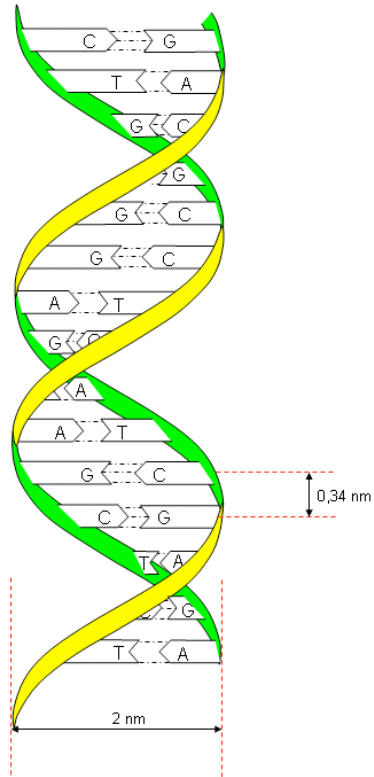
The output of our algorithm

AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAGAGTGTCTGATAGCAGC*

Length of the sequence: 70

%A = 28.57 %C = 17.14 %G = 24.28 %T = 30.00

A-T versus C-G



The output of our algorithm

AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAGAGTGTCTGATAGCAGC*

Length of the sequence: 70

%A = 28.57 %C = 17.14 %G = 24.28 %T = 30.00

Thus:

%GC = 41.43 (GC-content)

%TA = 58.57 (AT-content)

The output of our algorithm

AGCTTTTCATTCTGACTGCAACGGGCAATATGTCTCTGTGTGGATTAAAAAAGAGTGTCTGATAGCAGC*

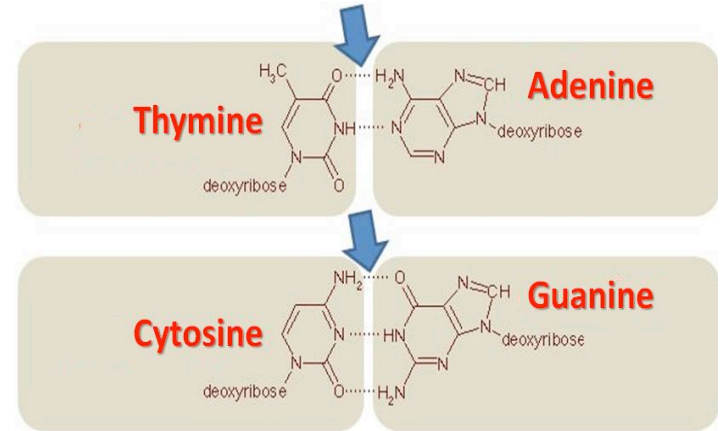
Length of the sequence: 70

%A = 28.57 %C = 17.14 %G = 24.28 %T = 30.00

Thus:

%GC = 41.43 (GC-content)

%TA = 58.57 (AT-content)



Pictures & movies : material licensing

p. 5 : Nature Education / Conditions of use for academic research : users may view, print, copy, download and text and data-mine the content

p. 5, p. 7 : Wikiality123 [GCC-BY-SA-3.0 (<http://creativecommons.org/licenses/by-sa/3.0/>)], from Wikimedia Commons