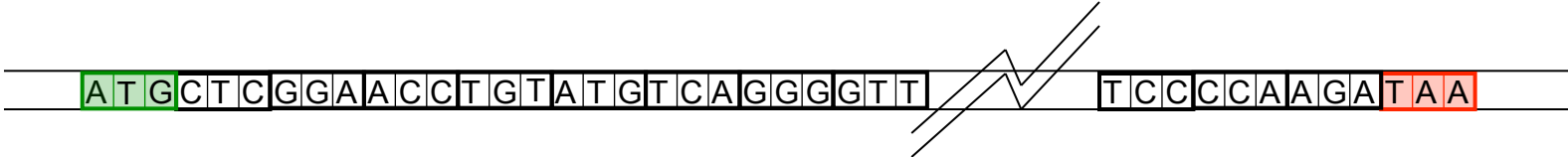


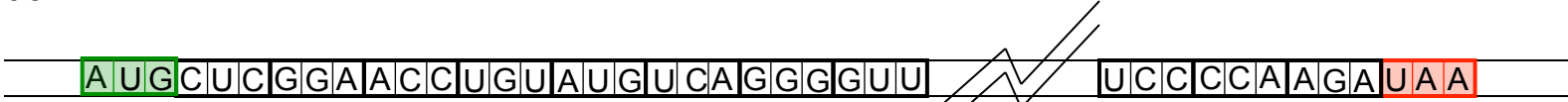
2. Gènes et protéines

- La séquence est-elle un bon modèle de l'ADN ?
- Les gènes, de Mendel à la biologie moléculaire
- Le code génétique
- **Un algorithme de traduction**
- Implémenter le code génétique
- Algorithmes + structures de données = programmes
- Les compromis de la conception d'algorithmes
- Les technologies de séquençage de l'ADN
- Le séquençage de génomes complets
- Comment trouver les gènes ?

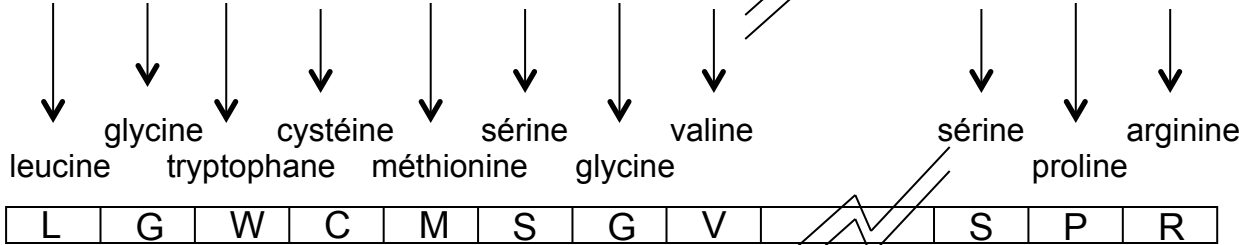
Transcription et traduction



Transcription



Traduction



Un algorithme de traduction

- Pour chaque triplet de la séquence en entrée
 1. le rechercher dans le tableau qui implémente le code génétique et retrouver l'acide aminé correspondant
 2. ajouter la lettre associée à l'acide aminé à l'extrémité de la séquence protéique

CTTATGGCATTGACTGATTCATGC#

CTT ATG GCA TTG ACT GAT TCA TGC



L M A L T D S C #

```
indexDNA, indexProtein: integer
DNAsequence, ProteinSequence: character string [1:*]
indexProtein, indexDNA ← 1
repeat
  ProteinSequence [indexProtein] ← lookupGeneticCode (DNAsequence
  [indexDNA], DNAsequence [indexDNA+1], DNAsequence [indexDNA+2])
  indexDNA ← indexDNA + 3
  indexProtein ← indexProtein + 1
until DNAsequence [indexDNA] = "#"
ProteinSequence [indexProtein] ← "#"
```

lookupGeneticCode est une fonction que nous écrivons séparément afin de fractionner la difficulté d'expression de l'algorithme

Et si la longueur de la séquence n'est pas un multiple de 3 ?

```
indexDNA, indexProtein: integer
```

```
DNasequence, ProteinSequence: character string [1:*]
```

```
indexProt, indexDNA ← 1
```

```
repeat
```

```
  ProteinSequence [indexProt] ← lookupGeneticCode (DNasequence  
  [indexDNA], DNasequence [indexDNA+1], DNasequence [indexDNA+2])
```

```
  indexDNA ← indexDNA+3
```

```
until DNasequence [indexDNA] = "#" or
```

```
  DNasequence [indexDNA+1] = "#" or
```

```
  DNasequence [indexDNA+2] = "#"
```

```
ProteinSequence [indexProt] ← "#"
```