3. Gene prediction

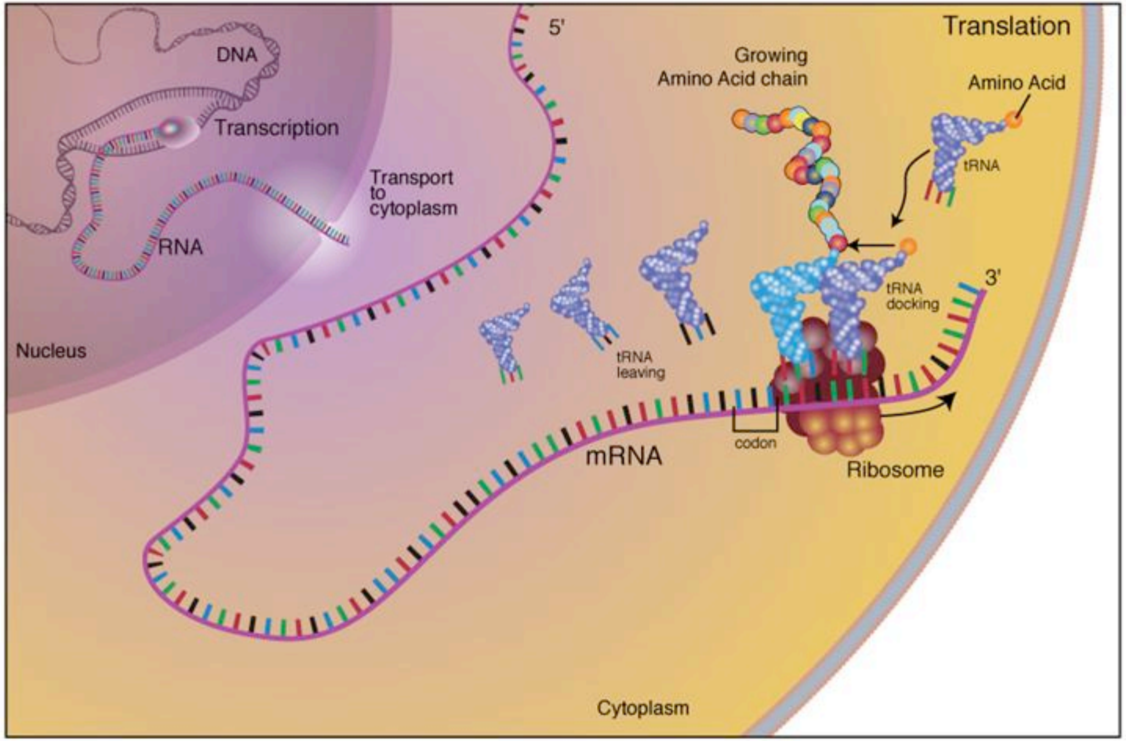
- All genes end on a stop codon
- A simple algorithm for gene prediction
- Searching for start and stop codons
- Predicting all the genes in a sequence
- **Making the predictions more reliable**
- Boyer-Moore algorithm
- Index and suffix trees
- Probabilistic methods
- Benchmarking the prediction methods
- Gene prediction in eukaryotic genomes

Search for RBS pattern

- RBS: Ribosome Binding Site

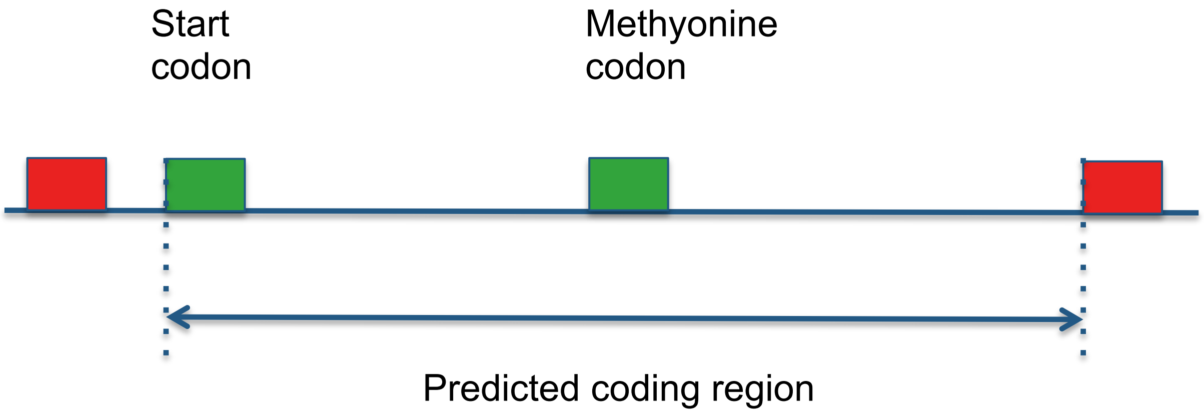
Search for RBS pattern

- RBS: Ribosome Binding Site



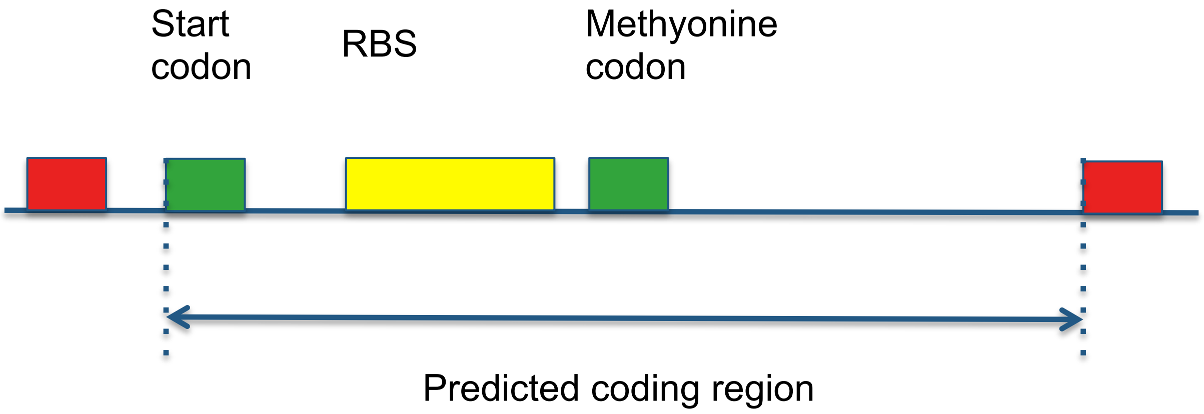
Search for RBS pattern

- RBS: Ribosome Binding Site
- The position of the start codon of a predicted coding region may be revised after RBS has been located



Search for RBS pattern

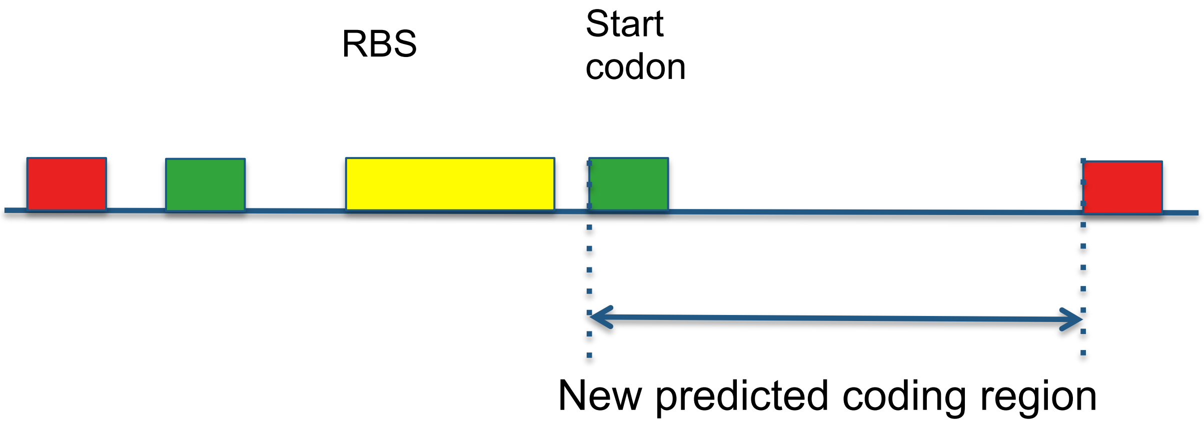
- RBS: Ribosome Binding Site
- The position of the start codon of a predicted coding region way be revised after RBS has been located



Search for RBS pattern

- RBS: Ribosome Binding Site

- The position of the start codon of a predicted coding region way be revised after RBS has been located



RBS patterns are highly variables

- The RBS pattern is organism-dependent
- Within the same genome, several variants of the pattern may occur
  - Adapted and efficient search pattern algorithm are required

...AUAAAGGAGGUAAAUAAUG

...AUAAAGGAAAUAAAUAAUG

...AUAACAGAGGUAAAUAAUG

...AUAACAGGAGUAAAUAAUG

- Some genes are “RBS-less”

Pictures & movies : material licensing

p. 4 : National Human Genome Research Institute's Talking Glossary (http://www.genome.gov/glossary/).