

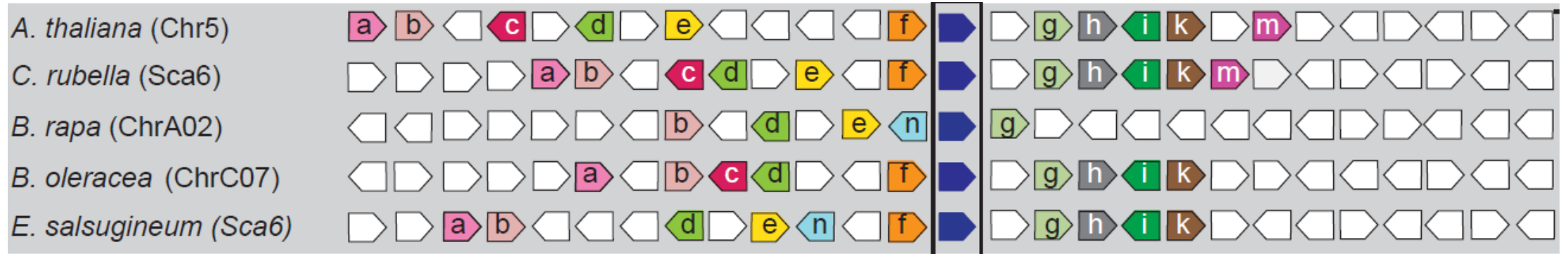
- JSMC Practical Course –  
Inferring Phylogeny Based on  
Sequence Information

Synteny and gene prediction

# Synteny

- Physical co-localization of genetic loci on the same chromosome
- Shared synteny: conservation of co-localization within two sets of chromosomes
- Shared synteny is one of the most reliable criteria for establishing the orthology of genomic regions in different species

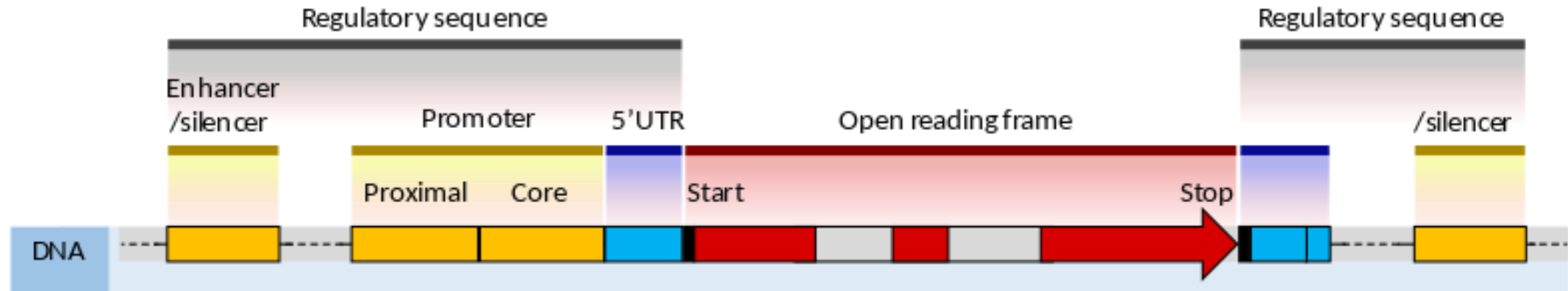
# Synteny



- Genes with the same color and letter have the same annotation
- Deletions and insertions of genes but order of some key genes is conserved

# Gene prediction

- Locating genes along a genomic DNA sequence
- Structure of a eukaryotic gene



# Gene prediction - *Ab-initio*

- Based only on genomic DNA sequence
- Searches for signals and content (specific sequences, codon usage, GC content) of protein coding regions and statistical properties of the given DNA sequence
- Example: GeneMarkS, Prodigal, Glimmer

# Gene prediction - *Homology Based*

- Given genome sequence is compared with a reference genome or protein to identify coding regions

