

Primer on Reading Frames and Phase

Wilson Leung 12/2023

1

A genomic sequence has 6 reading frames

- 3 reading frames per strand (+ / -)

Score = 57.0 bits (136), Expect = 1e-18
 Identity = 25/25 (100%), Positives = 25/25 (100%), Gaps = 0/25 (0%)
 Frame = 12

Query 2 VEKVASIVFHFDIFPITVICNSNAY 76
 VEKVASIVFHFDIFPITVICNSNAY
 Sbjct 1 VEKVASIVFHFDIFPITVICNSNAY 25

- Frame: base to start translation relative to the first base of the sequence
 - 1: Start translation at **first** base of sequence
 - 2: Start translation at **second** base of sequence
 - 3: Start translation at **third** base of sequence

2

A codon could be derived from nucleotides in adjacent exons

Spliced mRNA: CTG AGA GAT TTT CCG

Phase 0: CTG AGA | GT ... AG | GAT TTT CCG

Phase 1: CTG AGA | G | GT ... AG | AT TTT CCG

Phase 2: CTG AGA | GA | GT ... AG | T TTT CCG

Donor: GT, Intron: ..., Acceptor: AG

3

Splice donor and acceptor phases

- Phase: Number of bases between the complete codon and the splice site
 - Donor phase: Number of bases between the **end of the last complete codon** and the splice donor site (GT/GC)
 - Acceptor phase: Number of bases between the splice acceptor site (AG) and the **start of first complete codon**
- Phase **depends on the reading frame** of the CDS

4

Phase depends on the reading frame

Genomic sequence: GTTTATTTTCAGCGSTTACGCCAAT

Splice acceptor: AG

- Acceptor phase:
 - Phase 2 relative to **frame 1**
 - Phase 0 relative to **frame 2**
 - Phase 1 relative to **frame 3**

5

Compatibility of donor and acceptor site

- Donor and acceptor phases of adjacent exons must be **compatible** with each other:
 Donor phase + acceptor phase = 0 or 3

CCA AAT G | GT ... AG | TT CTC GAT

Translation: P N V L D

6

Incompatible donor and acceptor phases introduce a frame shift in the translation

Translation: P N G F S

Phase 0 donor site is incompatible with phase 2 acceptor site

7

Calculating the exon acceptor phase

$$P_i = (3 - ((L_{i-1} - P_{i-1}) \% 3)) \% 3$$

where: $i > 1$

- P_i = Acceptor phase of exon i
- L_{i-1} = Length of previous exon
- P_{i-1} = Acceptor phase of previous exon

- $L_{i-1} - P_{i-1}$ = Previous exon length starting from the first complete codon
- $(L_{i-1} - P_{i-1}) \% 3$ = Donor phase of previous exon
- $3 - ((L_{i-1} - P_{i-1}) \% 3)$ = Acceptor phase of current exon
- $(3 - ((L_{i-1} - P_{i-1}) \% 3)) \% 3$ = Change to phase 0 if acceptor phase is 3

8

Questions?

<http://www.flickr.com/photos/wingedwolf/5471047557/sizes/m/in/photostream/>

9