Primer on Reading Frames and Phase

Wilson Leung 08/2016

A genomic sequence has 6 reading frames

- 3 reading frames per strand (+ / -)
  - 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

A codon could be derived from nucleotides in adjacent exons

Spliced mRNA

<table>
<thead>
<tr>
<th>CTG</th>
<th>AGA</th>
<th>GAT</th>
<th>TTT</th>
<th>CCG</th>
</tr>
</thead>
</table>

Phase 0

<table>
<thead>
<tr>
<th>CTG</th>
<th>AGA</th>
<th>GT</th>
<th>...</th>
<th>AG</th>
<th>GAT</th>
<th>TTT</th>
<th>CCG</th>
</tr>
</thead>
</table>

Phase 1

<table>
<thead>
<tr>
<th>CTG</th>
<th>AGA</th>
<th>G</th>
<th>GT</th>
<th>...</th>
<th>AG</th>
<th>AT</th>
<th>TTT</th>
<th>CCG</th>
</tr>
</thead>
</table>

Phase 2

<table>
<thead>
<tr>
<th>CTG</th>
<th>AGA</th>
<th>GA</th>
<th>GT</th>
<th>...</th>
<th>AG</th>
<th>T</th>
<th>TTT</th>
<th>CCG</th>
</tr>
</thead>
</table>

Phase depends on the reading frame

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

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

![Diagram showing incompatible donor and acceptor phases with a frame shift in translation.]

Phase 0 donor site is incompatible with phase 2 acceptor site.

Calculating the exon acceptor phase:

\[ P_i = (3 - ((L_{i-1} - P_{i-1}) \mod 3)) \mod 3 \]

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
\( (L_{i-1} - P_{i-1}) \mod 3 \) = Donor phase of previous exon
\( 3 - ((L_{i-1} - P_{i-1}) \mod 3) \) = Acceptor phase of current exon
\( (3 - ((L_{i-1} - P_{i-1}) \mod 3)) \mod 3 \) = Change to phase 0 if acceptor phase is 3

Questions?

![Image of a question mark]

[Link to image: http://www.flickr.com/photos/wingedwolf/5471047557/sizes/m/in/photostream/]