Determine CDS Coordinates

E L G I N
GAACCTTTGGCATAAAC

**tBLASTn**: query- CDS for *D. melanogaster* exon; subject- nt sequence of interest (turn off low complexity filter and use no compositional adjustment).

For highly conserved alignments: look for splice sites nearest the exon boundaries. For less conserved alignments: attempt to extend the exon boundaries to include additional sequence from the ORF, while still maintaining correct splice phases; conserve exon size

**Yes**: check phase compatibility with previous splice donor site

- check for predicted acceptor splice site markers in gander

**No**: acceptor site must be AG; look for another possibility

- go to [http://flybase.org](http://flybase.org) to check the intron sequence of interest: is the non-canonical splice site present for this gene in *D. melanogaster*?

**is the acceptor site “AG”?**

- GT: note the phase and make sure it aligns with next acceptor site.

**what is the splice donor site?**

- GC: look for more evidence to support this non-canonical, alternative splice site

- check for predicted donor splice site markers in gander

- check for nearest GT that will fit the donor phase

- no GT possible because it violates minimum intron size rule (~40 nt) when looking downstream or leads to loss of conserved sequence when looking upstream

**check for predicted acceptor splice site markers in gander**

**continue looking for other evidence to justify GC**

- other lines of evidence possible to use to verify one GT is better than another

---

1. EST evidence to prove canonical or alternative splice site exists (BLAST)
2. Gene prediction tracks in gander to support alignment to selected site or demonstrate alternative solution
3. BLASTx track to verify level and length of conservation
4. Check RNA-Seq data if available
5. Check closely-related species for choice of splice site (BLAT analysis): sequencing error?
6. Perform CLUSTALw to test level of conservation between species

---

Jeannette Wong. 08.05.2010