

Getting Started

Searching protein domains on alternatively spliced regions of human gene TNNT1

According to RefSeq (NM_003283),

This gene encodes a protein that is a subunit of troponin, which is a regulatory complex located on the thin filament of the sarcomere. This complex regulates striated muscle contraction in response to fluctuations in intracellular calcium concentration.

Input

- Annotation of eight alternative transcripts from GENCODE Basic v24 ([Download](#))
- Chromosome 19 FASTA file from **GRCh38/hg38** ([Download](#))
- Reference file ([Download](#))
- HMM file ([Download](#))

Command line

Obtaining reference transcript sequence

```
$> astalavista -t astafunk --tref --gtf tnnt1.gtf --genome ~/example/genome/ > reference_tx.fasta
```

Creating reference file

```
$> hmmsearch --domtblout reference_file ~/Databases/Pfam/Pfam-A.hmm reference_tx.fasta
```

 **Obtaing a reduced HMM file**

```
$> grep -v "#" reference_file | awk '{print $5}' | sort | uniq > list-hmm-tnnt1
$> hmmfetch -f Pfam-A.hmm list-hmm-tnnt1 > database.hmm
```

or skip these commands and use directly the whole database Pfam-A.hmm as parameter for the option [--hmm]

Running AstaFunk

```
astalavista -t astafunk --genome ~/example/genome/ --gtf tnnt1.gtf --reference reference_file --hmm database.hmm
```

Output

AstaFunk identifies six complete alternative events between the eight alternative transcripts of the gene TNNT1 (**in the paper we present, just for an example, only two events**). See the standard output:

#	chr	gene_cluster_name	name_hmm	acc	description	bitscore	start_seq	end_seq	start_genomic	end_genomic	first_source	last_sii
EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	35.23096706 01	1	102	-55147129	-55134152	-55147168	-55147161

EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	98.64654794 89	85	205	-55141239	-55134200	-55147168	-55147161
EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	139.0575466 987	1	135	-55141281	-55134152	-55147168	-55147161
EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	139.0575466 987	1	135	-55141281	-55134152	-55147168	-55147161
EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	139.0575466 987	1	135	-55141281	-55134152	-55147168	-55147161
EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	158.3576407 894	69	205	-55141287	-55134152	-55147168	-55147161
EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	158.3576407 894	58	194	-55141287	-55134152	-55147168	-55147161
EHIT	chr19	ENST00000291901.12, ENST00000356783.9, ENST00000587465.6, ENST00000587758.5, ENST00000585321.6, ENST00000588981.5, ENST00000536926.5, ENST00000588426.5	Troponin	PF0099 2.17	Troponin	158.3576407 894	58	194	-55141287	-55134152	-55147168	-55147161