Evolutionary decomposition & structural characterization of functionally distinct protein isoforms

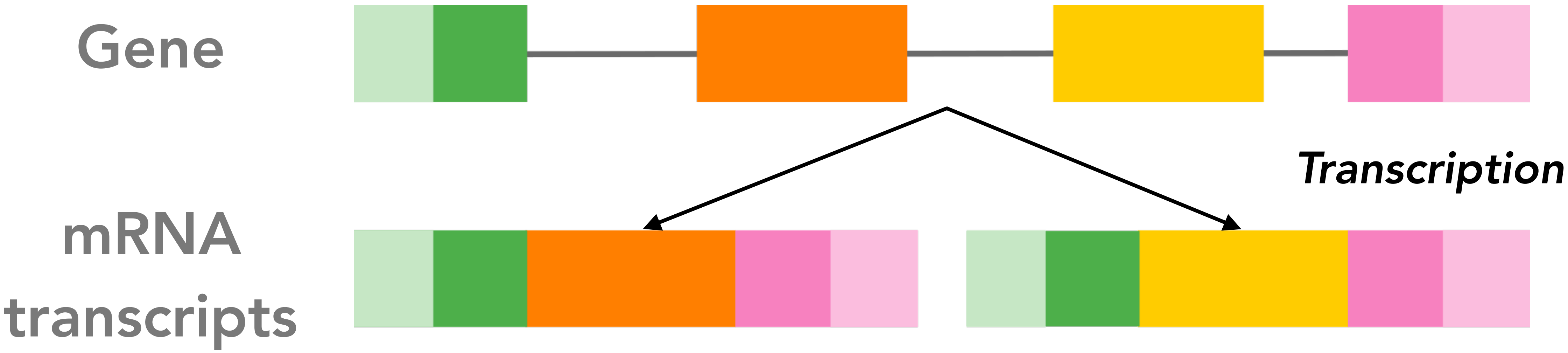
Elodie Laine, Laboratory of Computational and Quantitative Biology
Analytical Genomics team (dir. A. Carbone, www.lcqb.upmc.fr/AnalGenom/)

Alternative Splicing

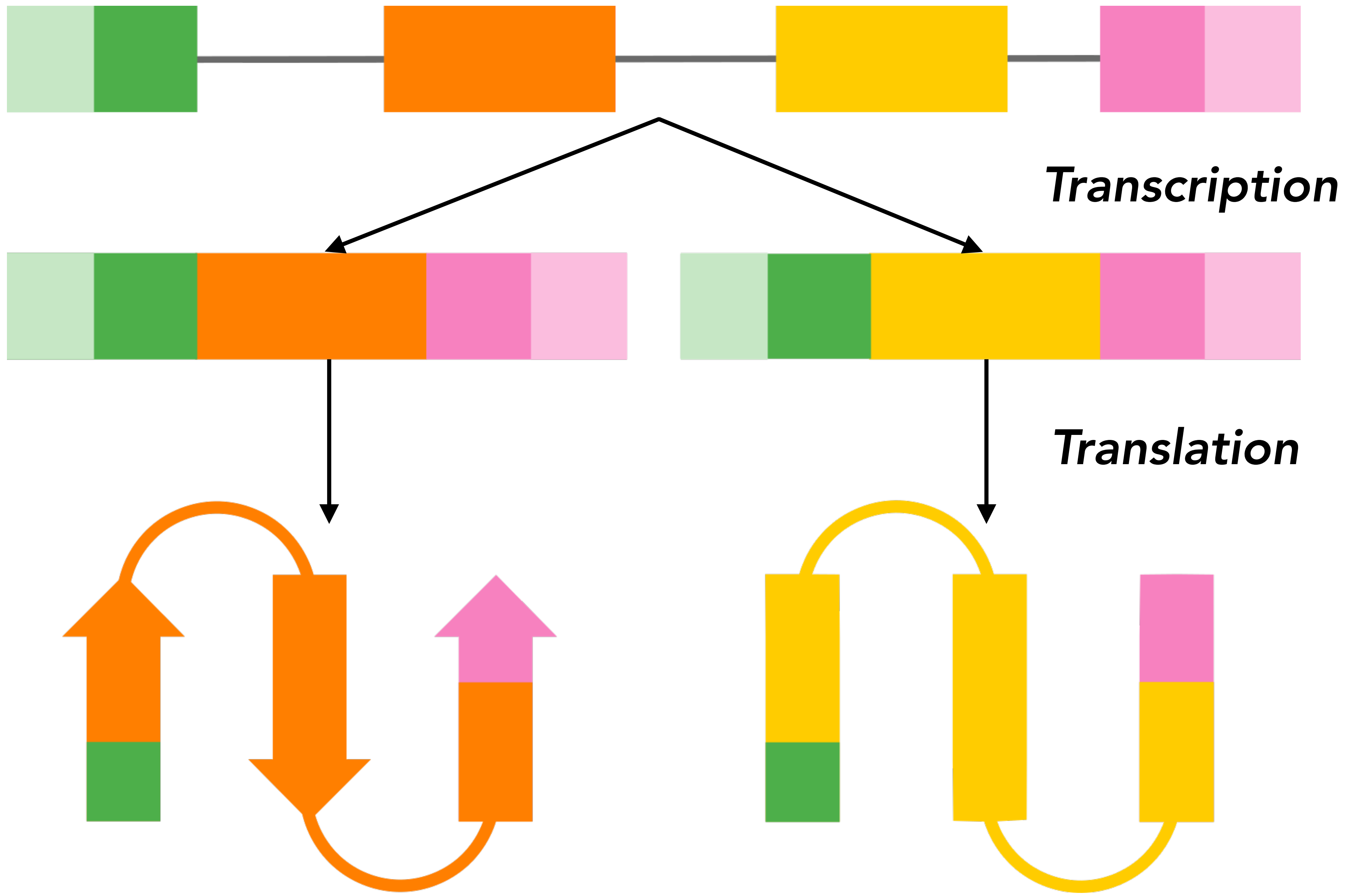
Gene

Alternative Splicing

Alternative Splicing



Alternative Splicing



Gene

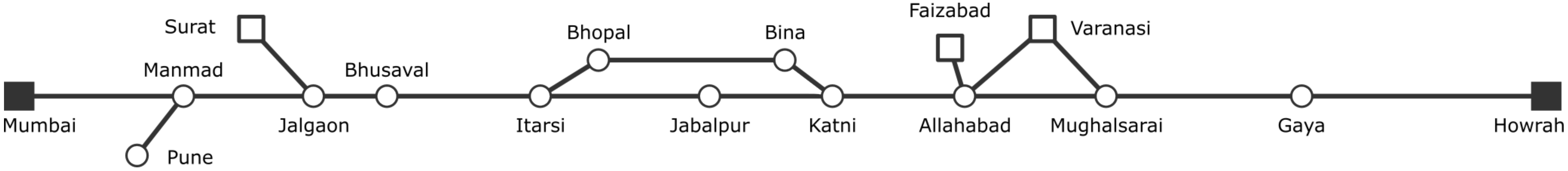
mRNA transcripts

Protein isoforms

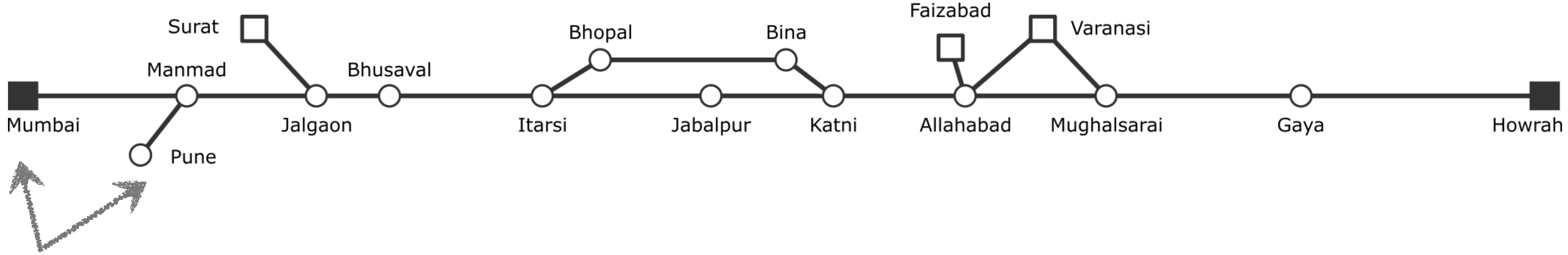
Transcription

Translation

Traveling in India

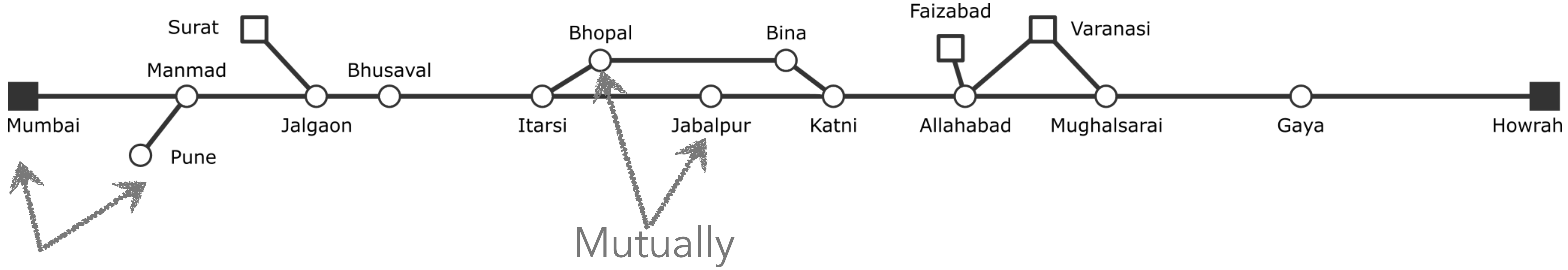


Traveling in India



Alternative starts

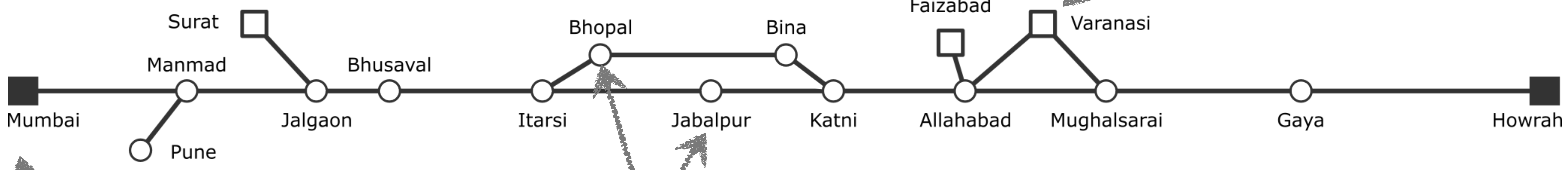
Traveling in India



Alternative starts

Mutually exclusive

Traveling in India

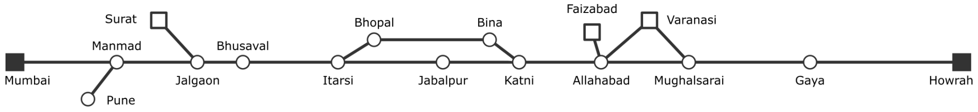


Insertion

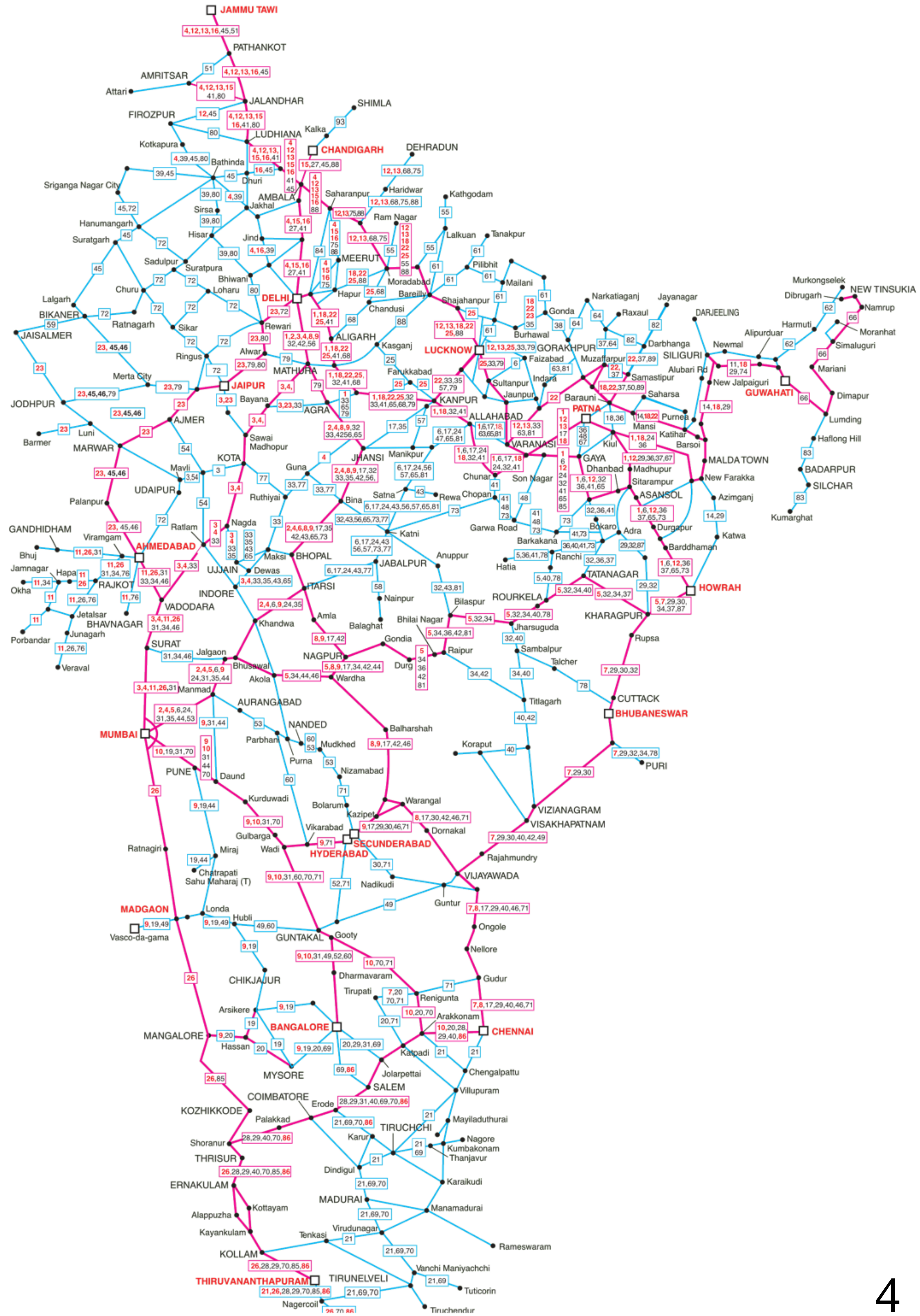
Alternative starts

Mutually exclusive

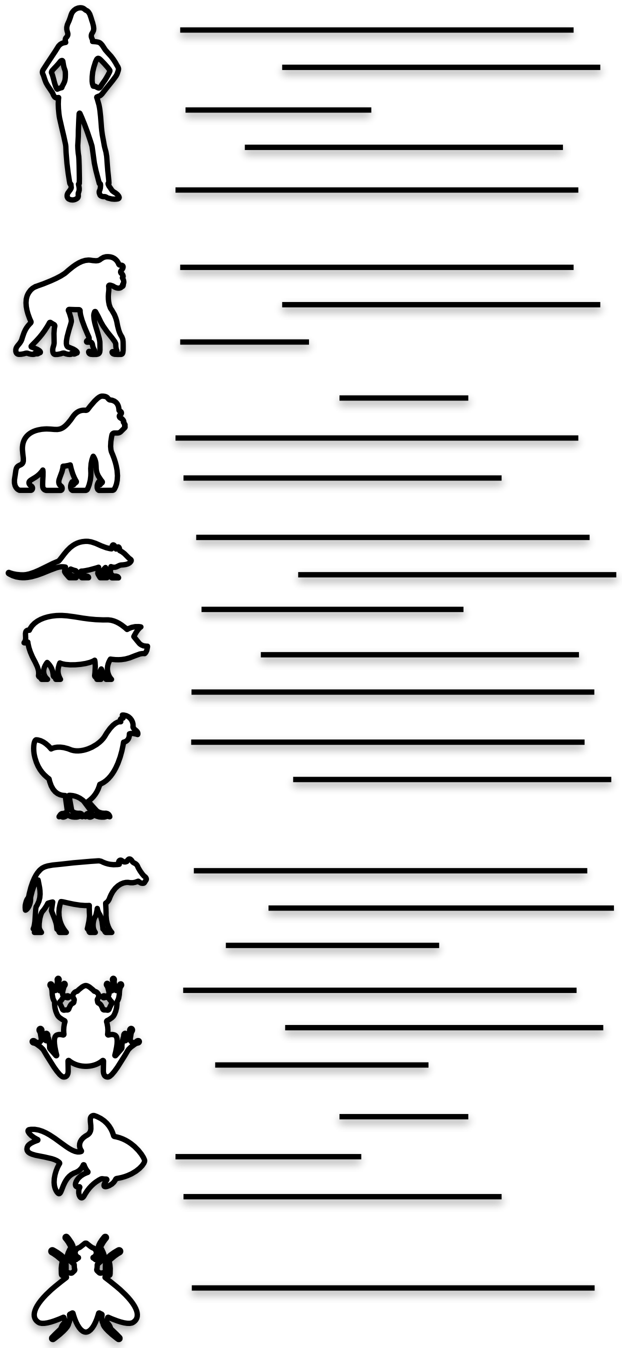
Traveling in India



With many lines, the network may become quite complex!



Main questions

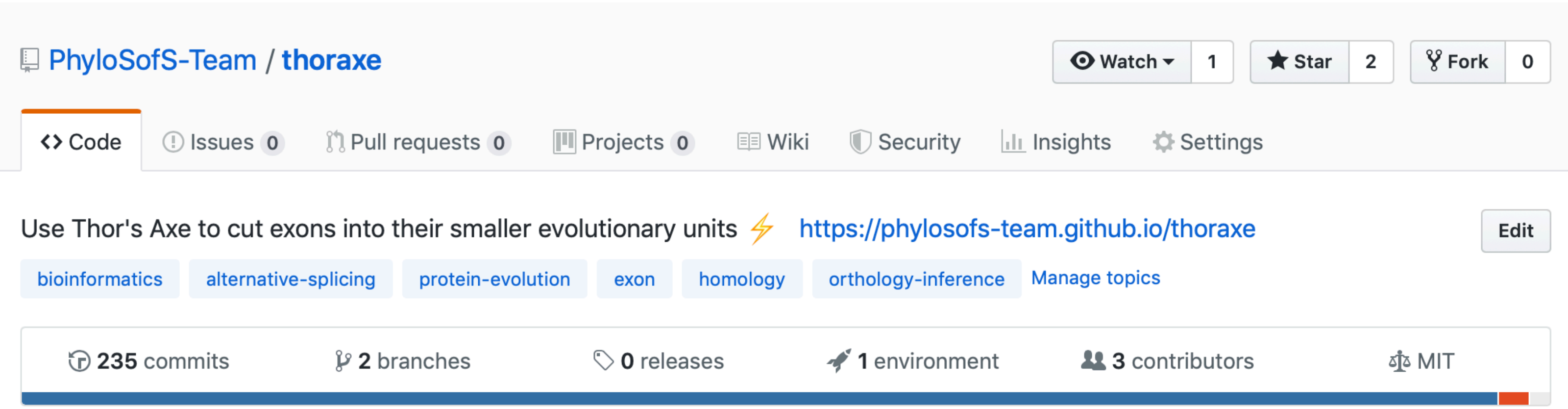


- How does alternative splicing impact protein **structure** and **interactions**?
- What are the **evolutionary ages** of observed alternative splicing events?
- How did alternative splicing induced new **functions** in evolution?

Methodology

ThorAxe

To find orthologous exons



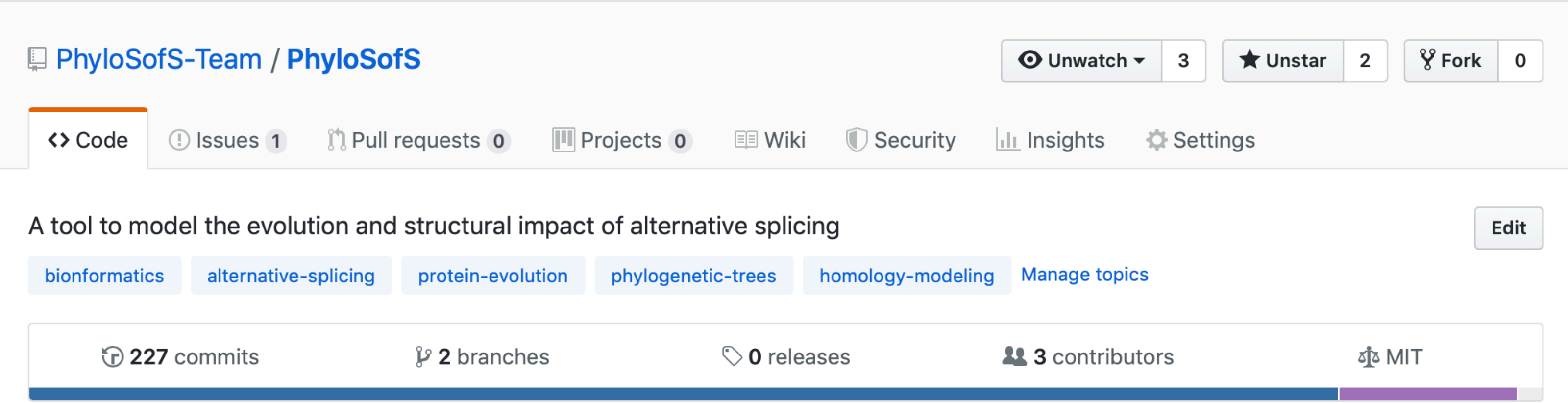
PhyloSofS-Team / **thoraxe**

Use Thor's Axe to cut exons into their smaller evolutionary units ⚡ https://phylosofs-team.github.io/thoraxe

235 commits · **2** branches · **0** releases · **1** environment · **3** contributors · MIT

PhyloSofS

To reconstruct transcripts' phylogenies and predict isoforms' 3D structures

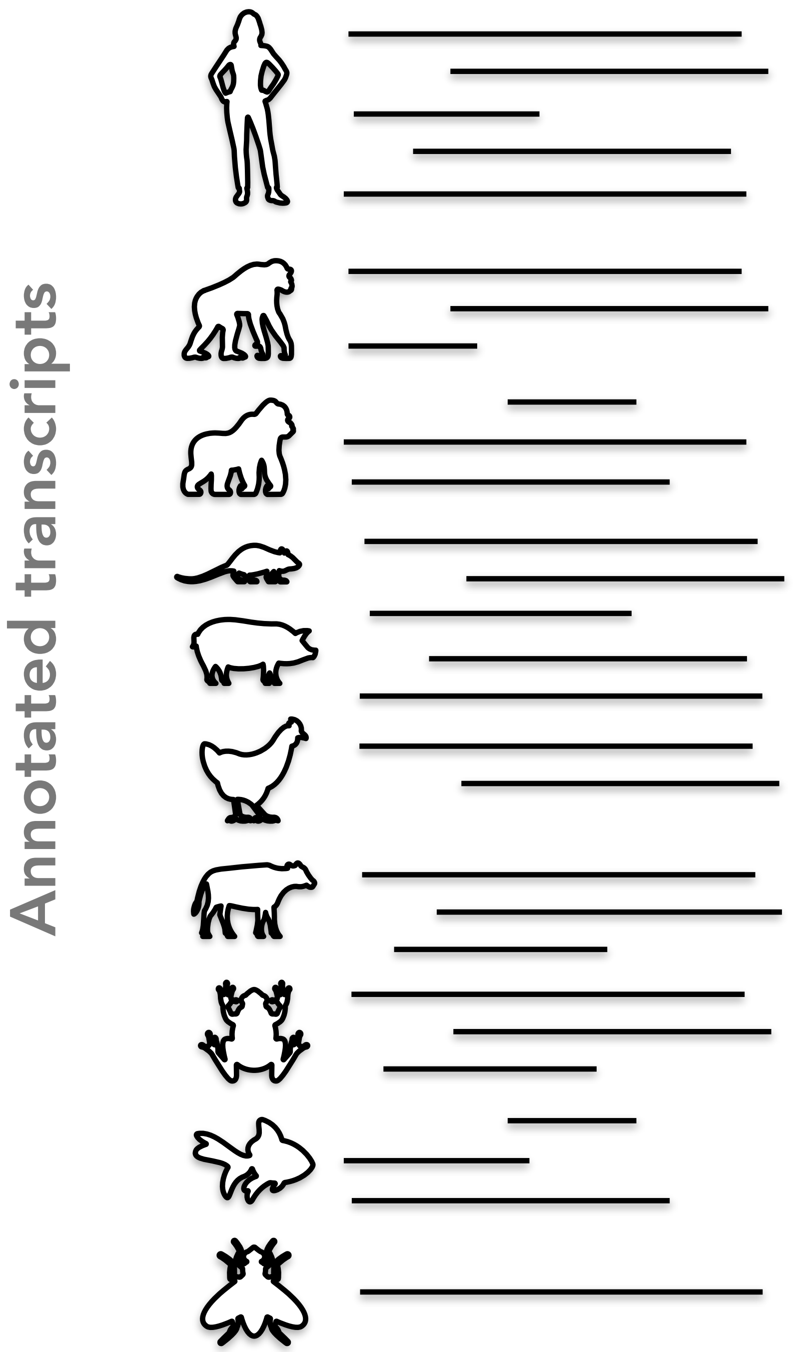


PhyloSofS-Team / **PhyloSofS**

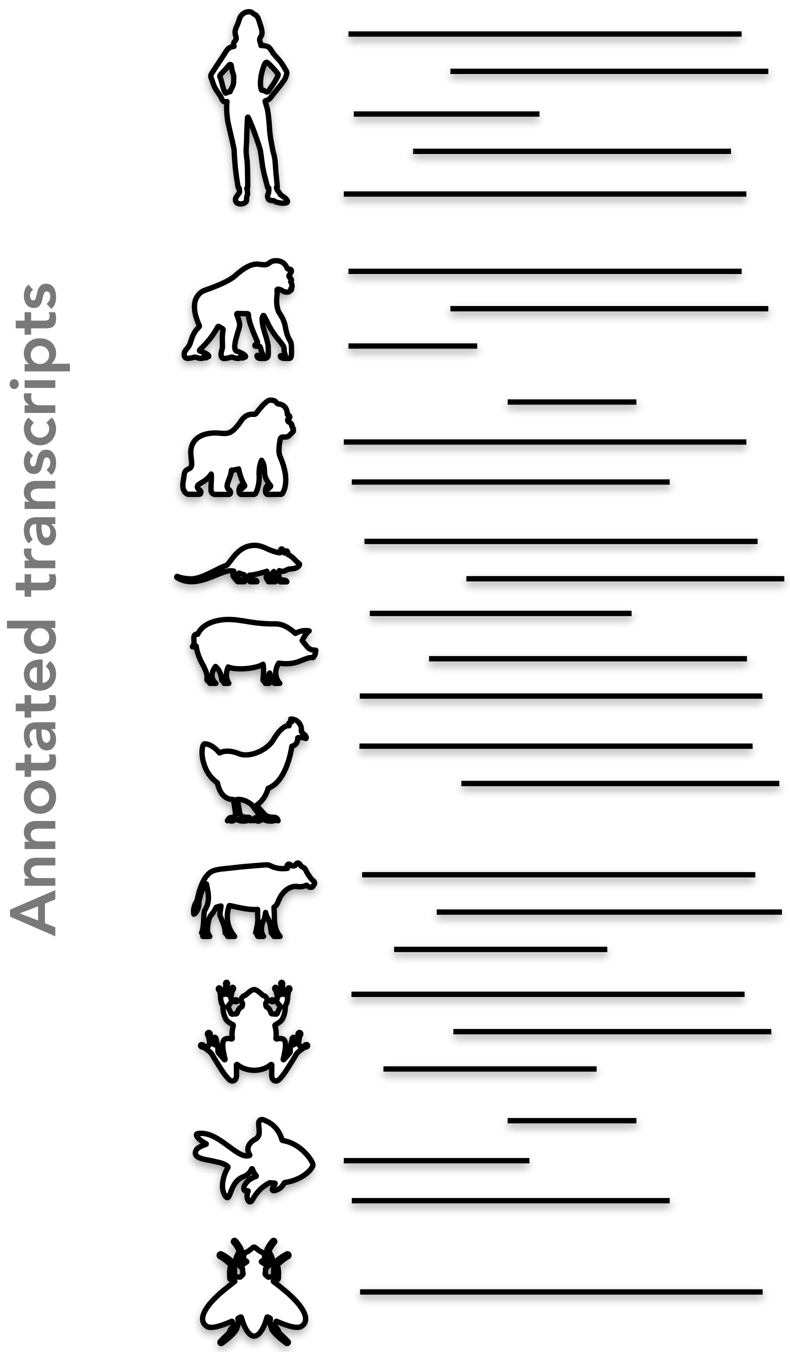
A tool to model the evolution and structural impact of alternative splicing

227 commits · **2** branches · **0** releases · **3** contributors · MIT

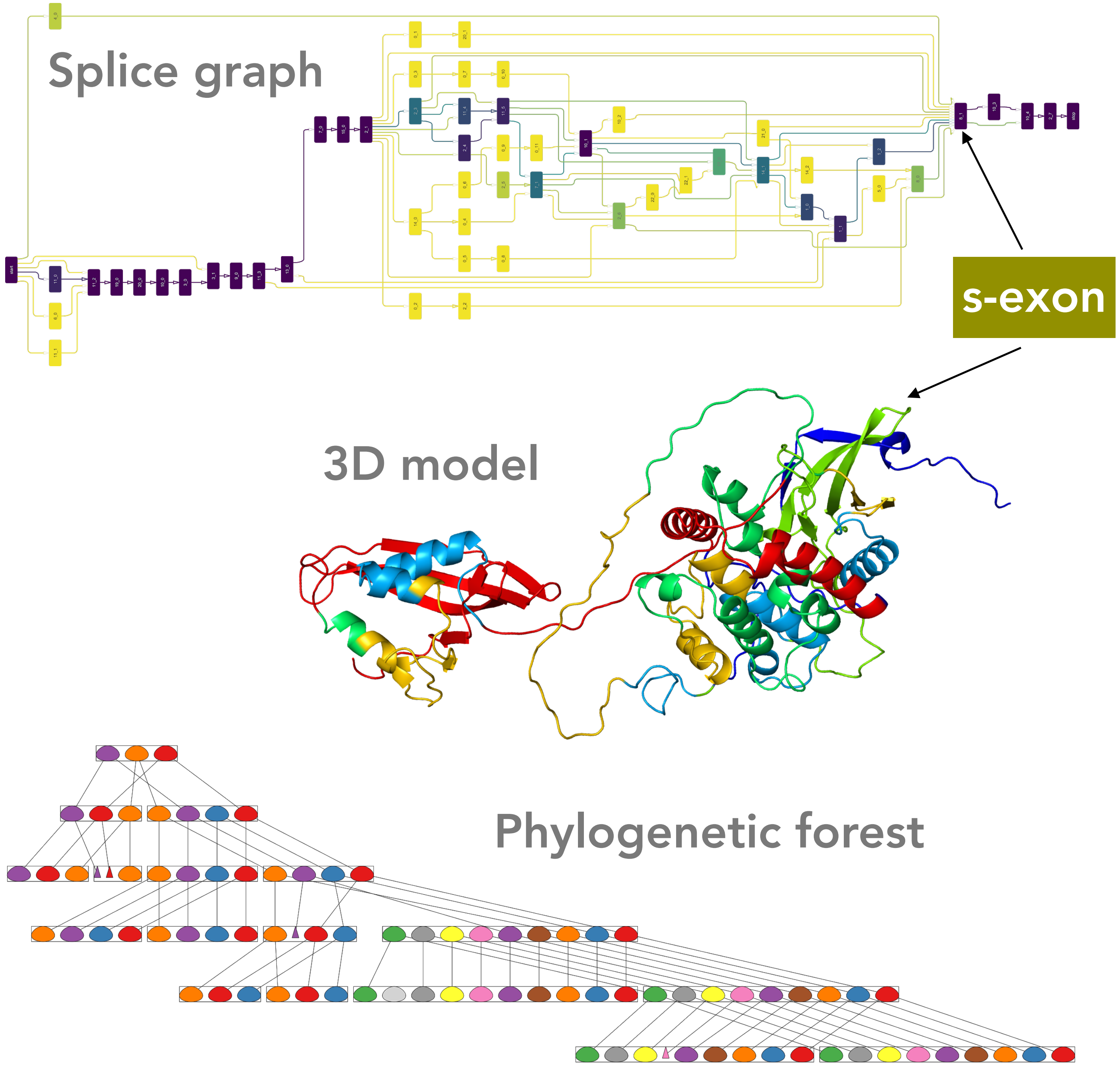
Inputs & Outputs



Inputs & Outputs



Annotated transcripts



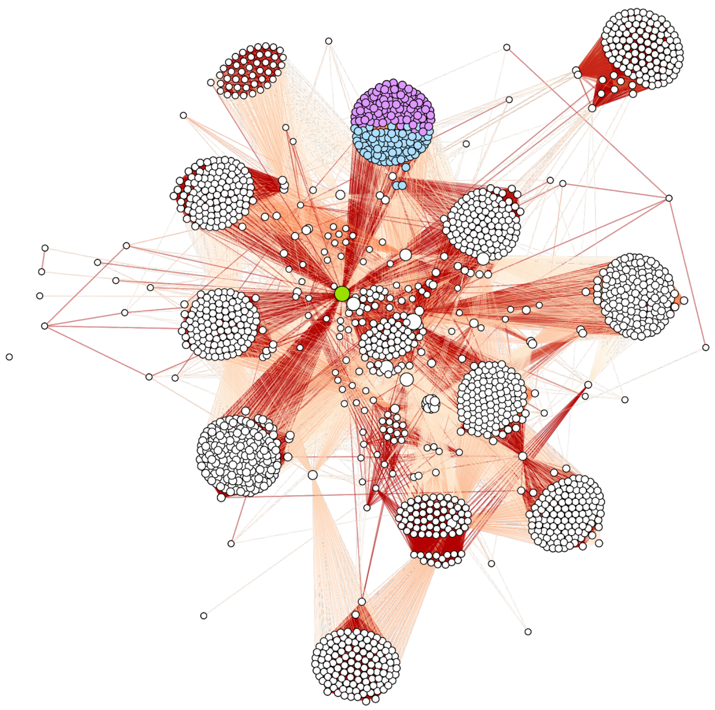
Splice graph

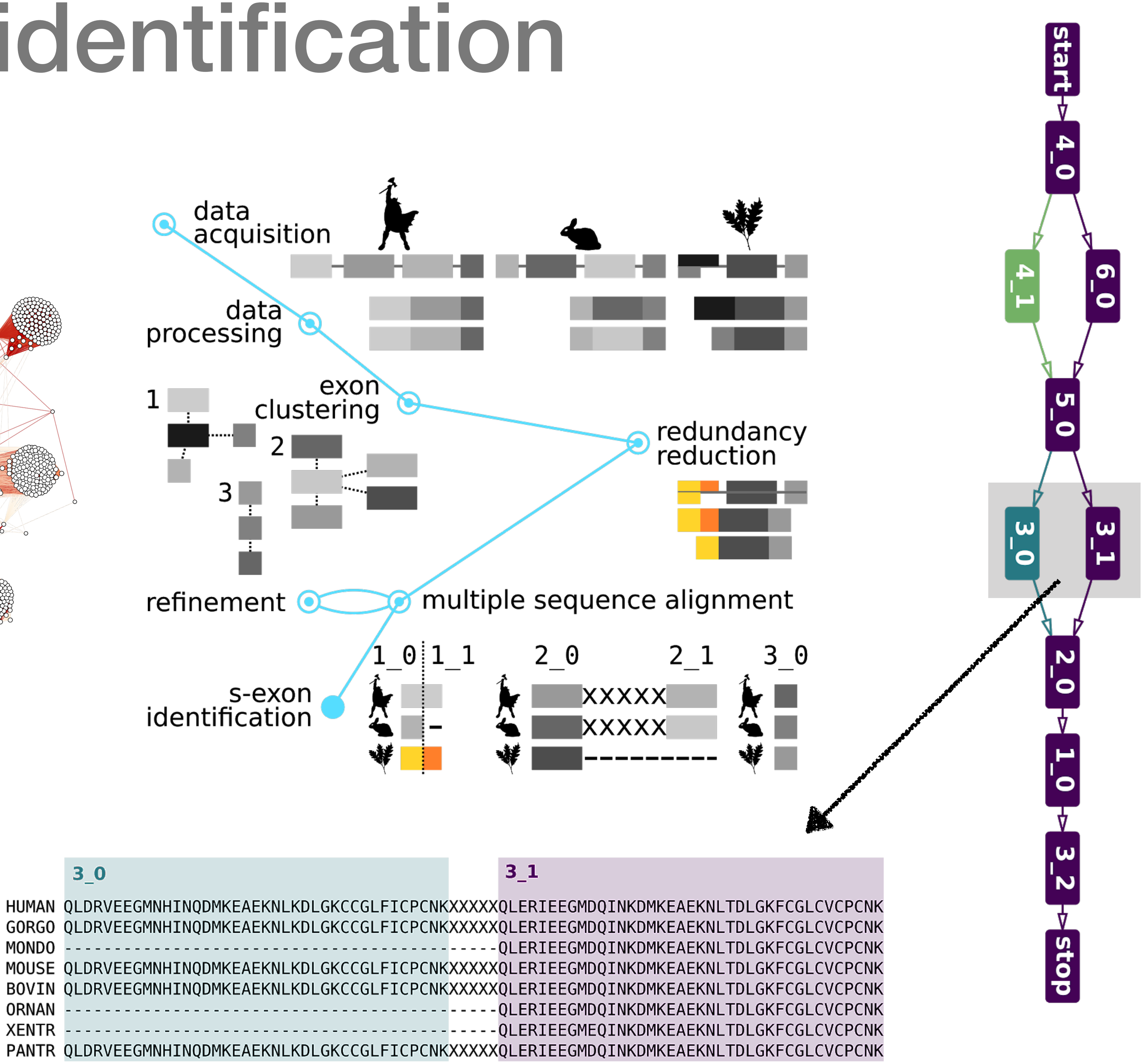
s-exon

3D model

Phylogenetic forest

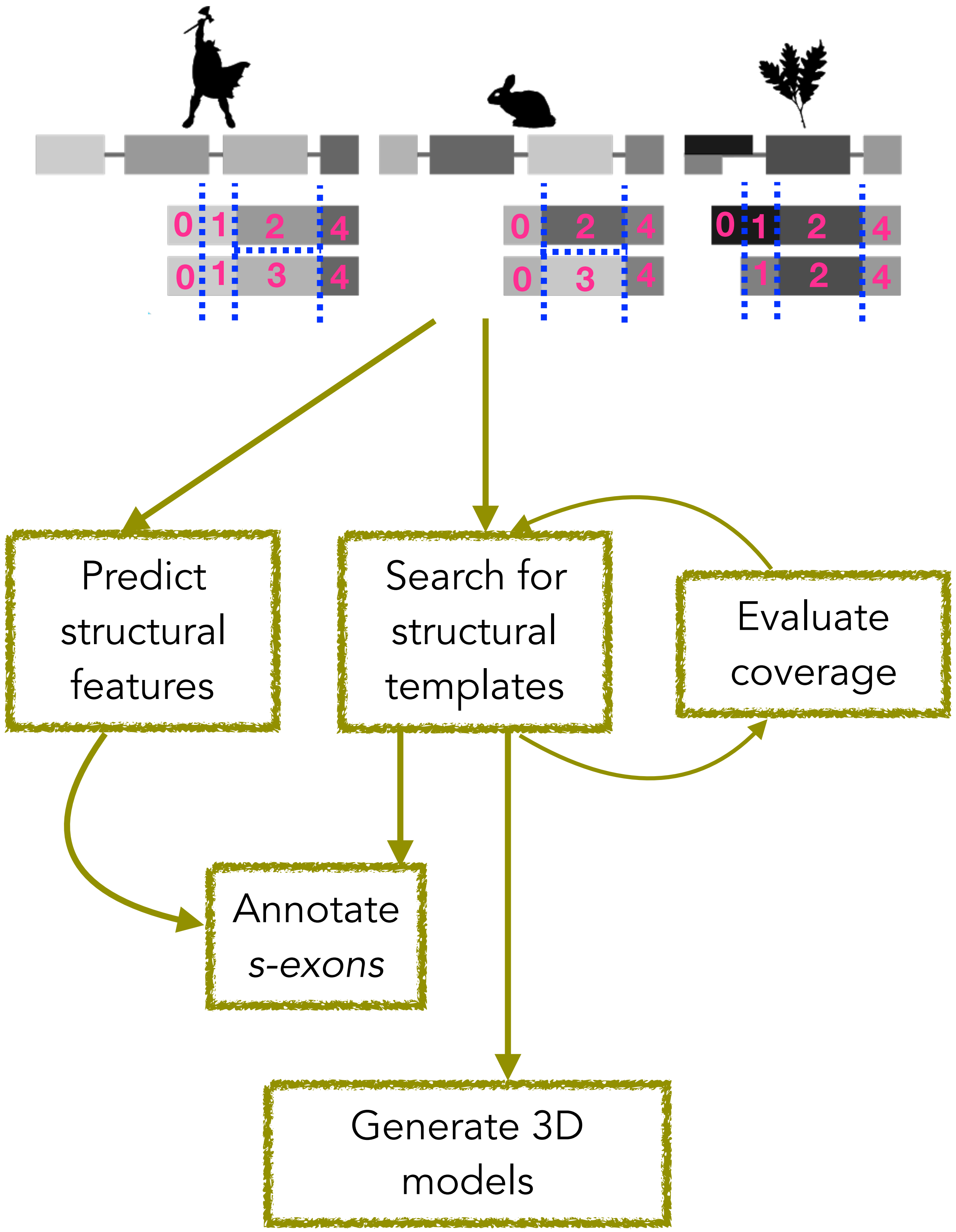
S-exon identification





```
HUMAN QLDRVEEGMNHINQDMKEAEKNLKDLGKCCGLFICPCNKXXXXXQLERIEEGMDQINKDMKEAEKNLTDLGKFCGLCVCPCNK
GORGO QLDRVEEGMNHINQDMKEAEKNLKDLGKCCGLFICPCNKXXXXXQLERIEEGMDQINKDMKEAEKNLTDLGKFCGLCVCPCNK
MONDO --------------------------------------------QLERIEEGMDQINKDMKEAEKNLTDLGKFCGLCVCPCNK
MOUSE QLDRVEEGMNHINQDMKEAEKNLKDLGKCCGLFICPCNKXXXXXQLERIEEGMDQINKDMKEAEKNLTDLGKFCGLCVCPCNK
BOVIN QLDRVEEGMNHINQDMKEAEKNLKDLGKCCGLFICPCNKXXXXXQLERIEEGMDQINKDMKEAEKNLTDLGKFCGLCVCPCNK
ORNAN --------------------------------------------QLERIEEGMDQINKDMKEAEKNLTDLGKFCGLCVCPCNK
XENTR --------------------------------------------QLERIEEGMEQINKDMKEAEKNLTDLGKFCGLCVCPCNK
PANTR QLDRVEEGMNHINQDMKEAEKNLKDLGKCCGLFICPCNKXXXXXQLERIEEGMDQINKDMKEAEKNLTDLGKFCGLCVCPCNK
```

Molecular modelling



Predict structural features

Search for structural templates

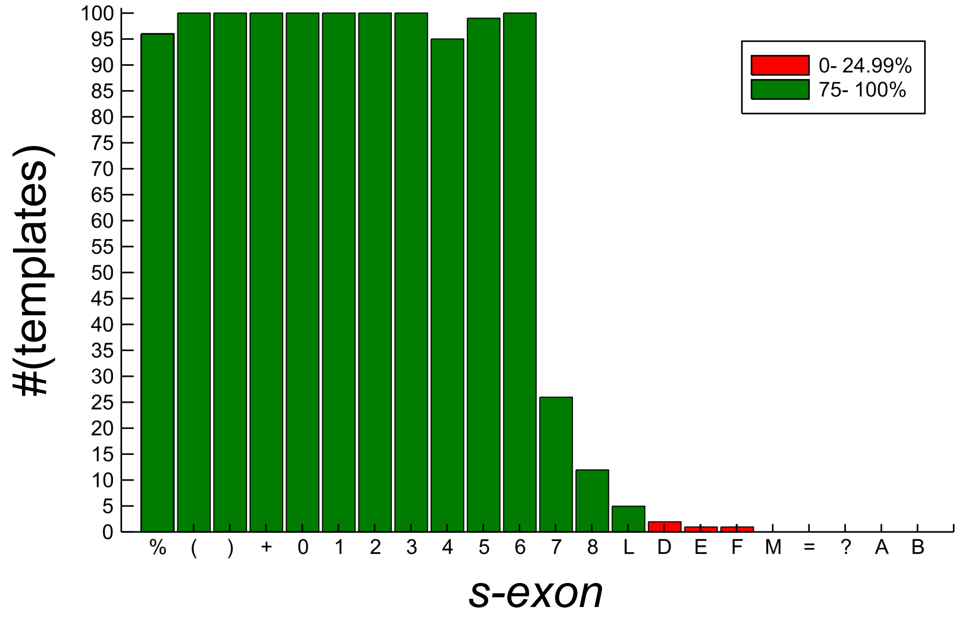
Evaluate coverage

Annotate *s-exons*

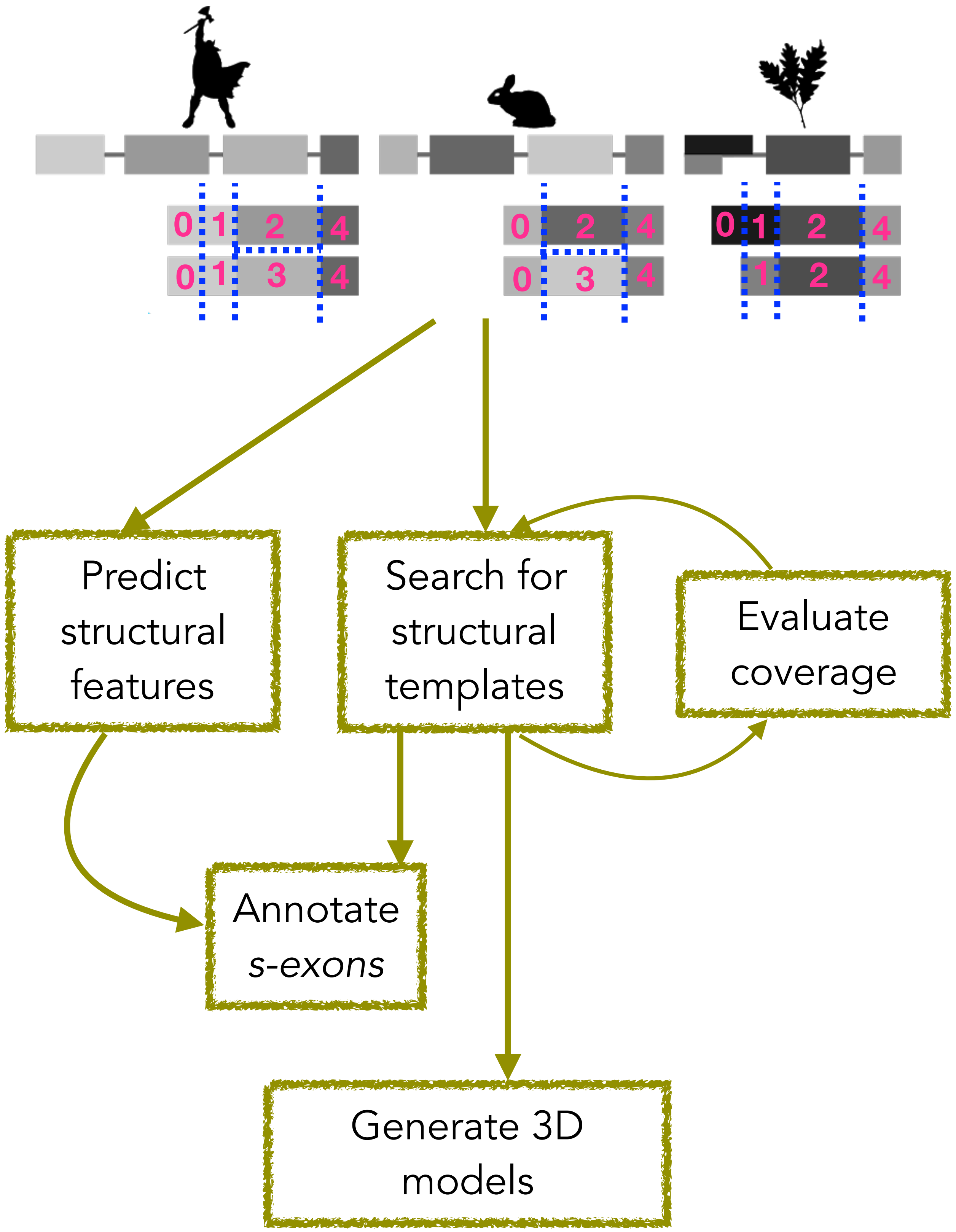
Generate 3D models

Structural template search is performed in an iterative way and centered on ***s-exons***.

First iteration



Molecular modelling



Predict structural features

Search for structural templates

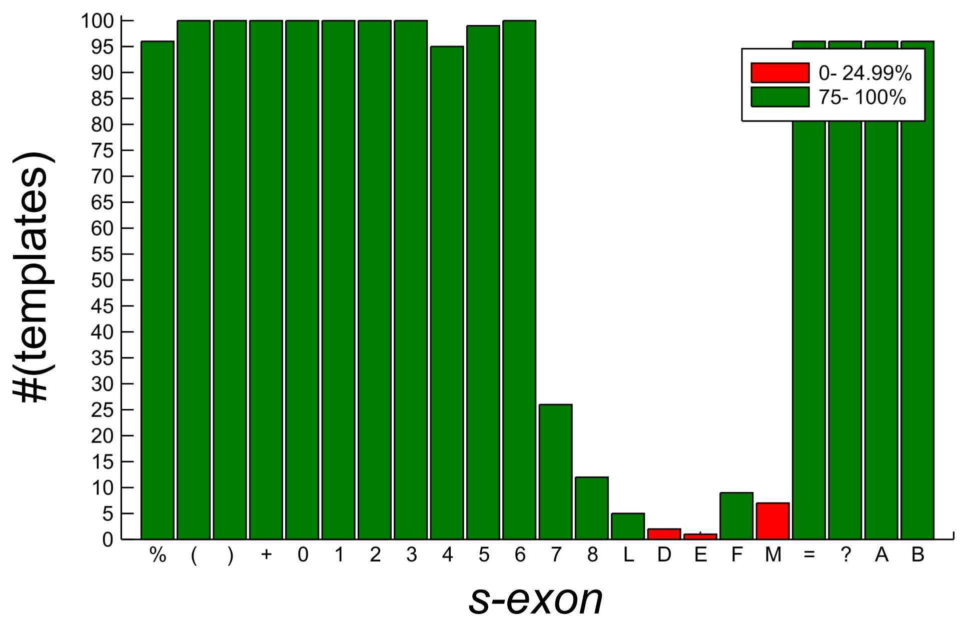
Evaluate coverage

Annotate *s-exons*

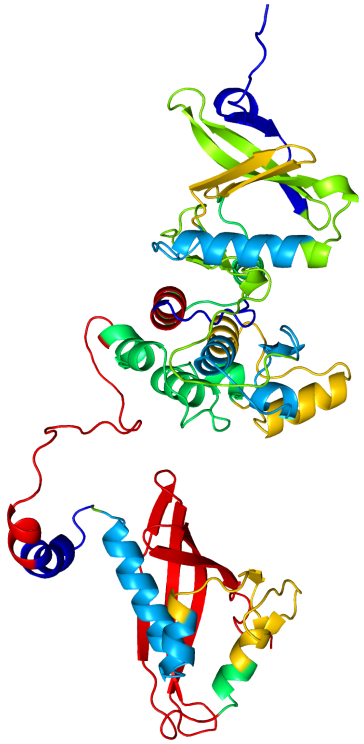
Generate 3D models

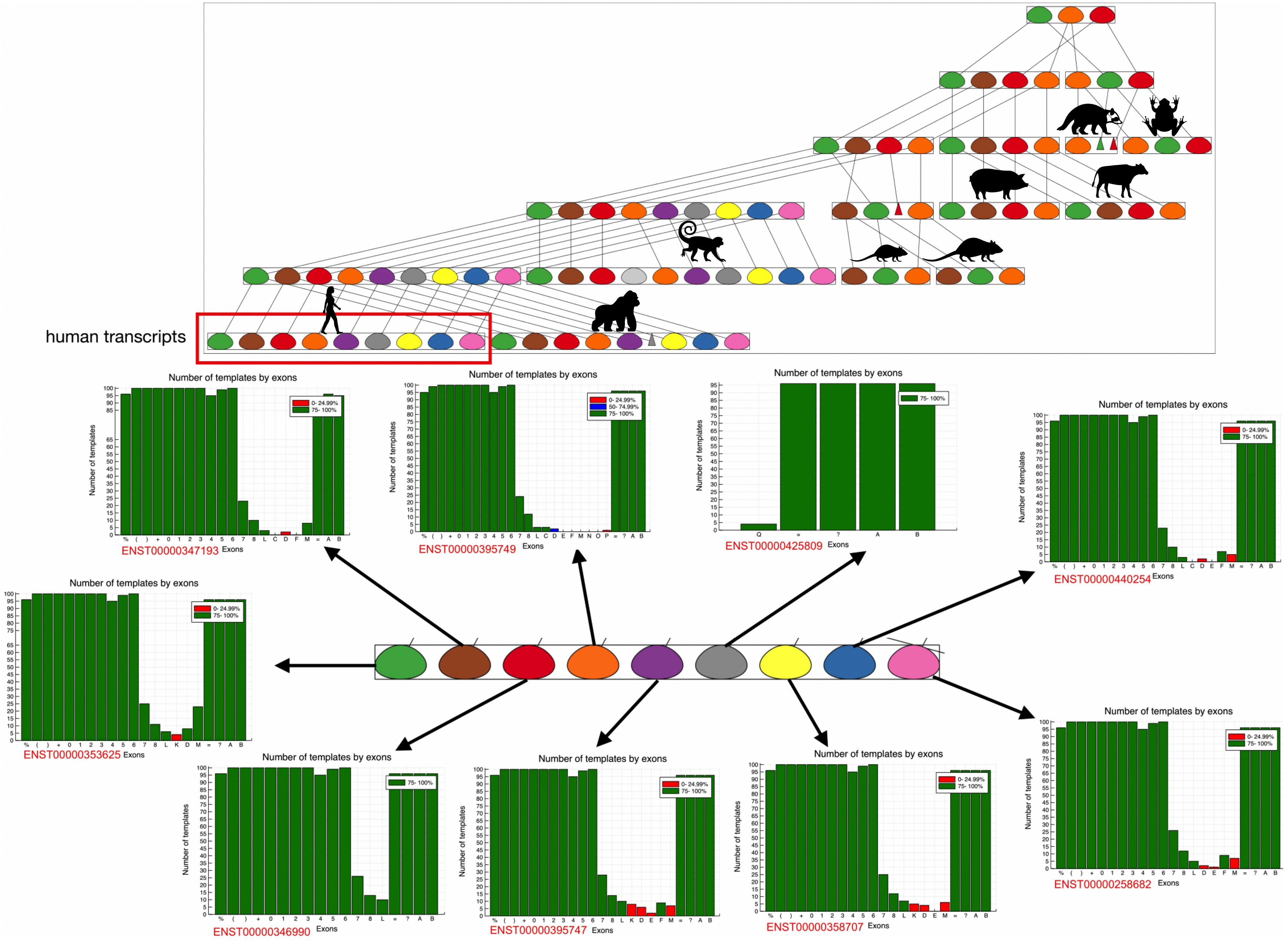
Structural template search is performed in an iterative way and centered on ***s-exons***.

After 4 iterations

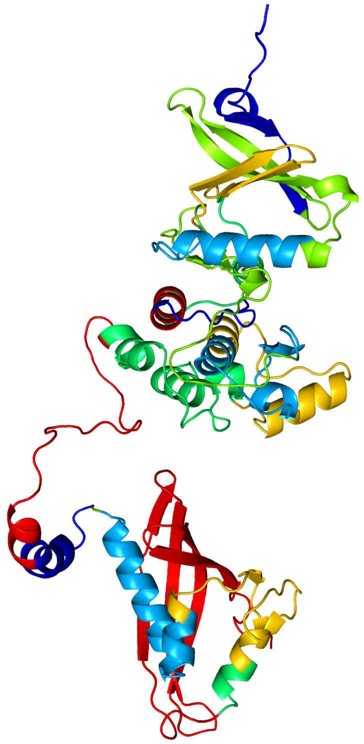


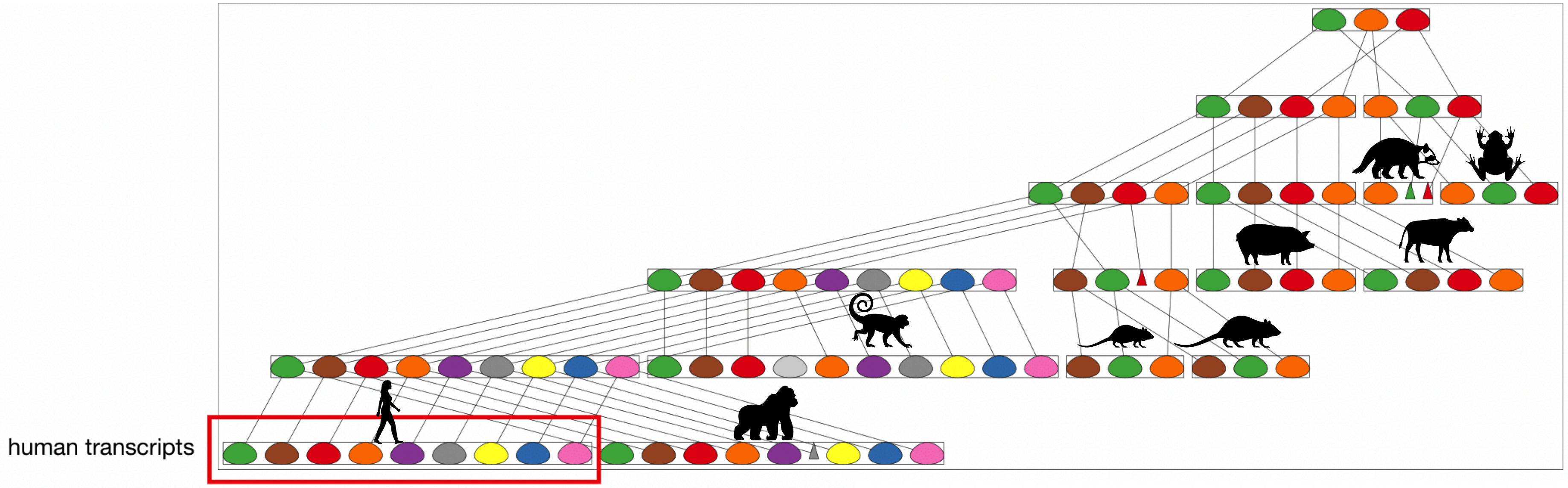
Case study - CAMK2B

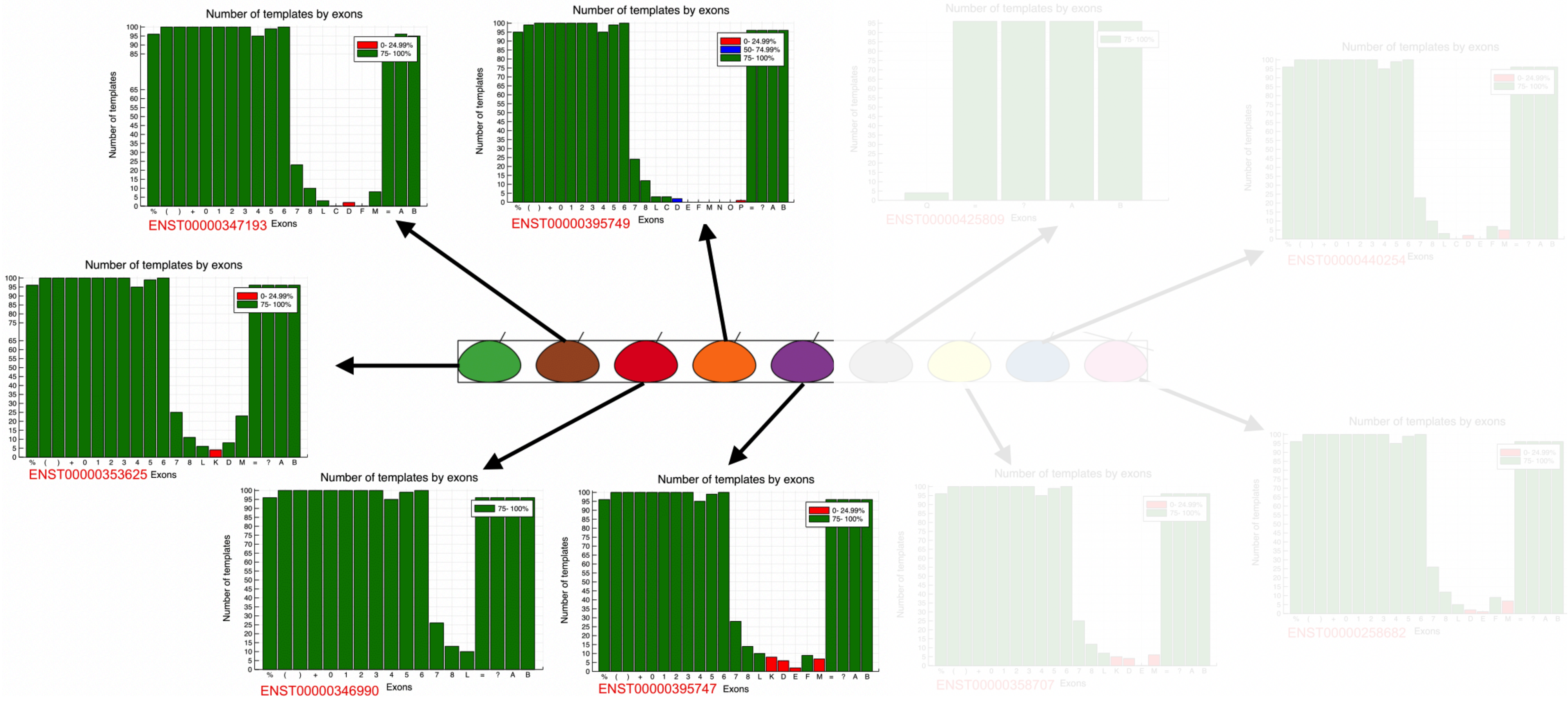




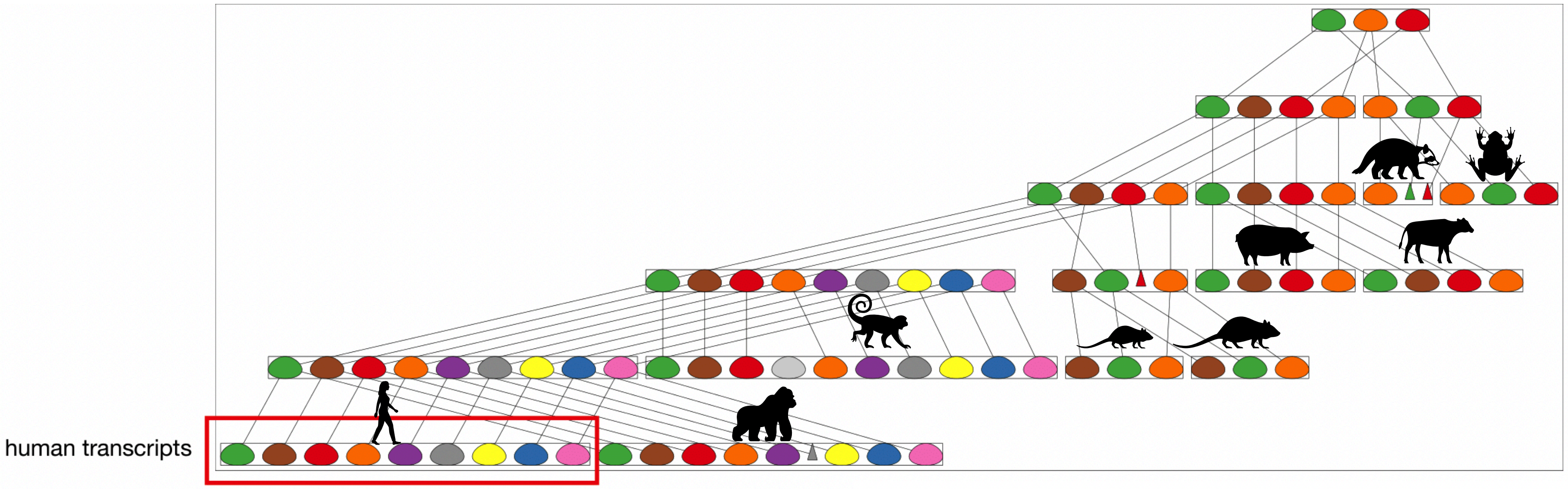
Case study - CAMK2B

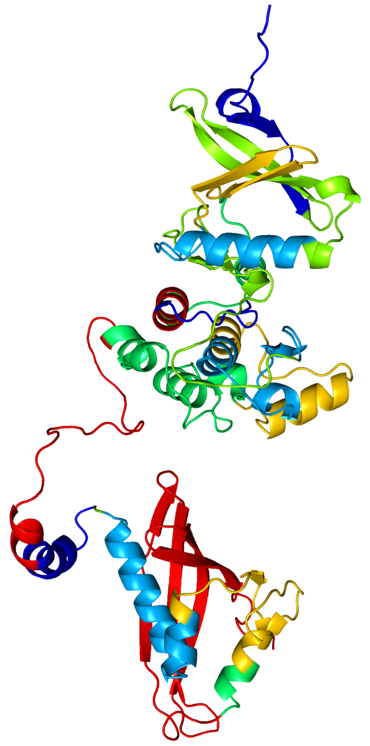


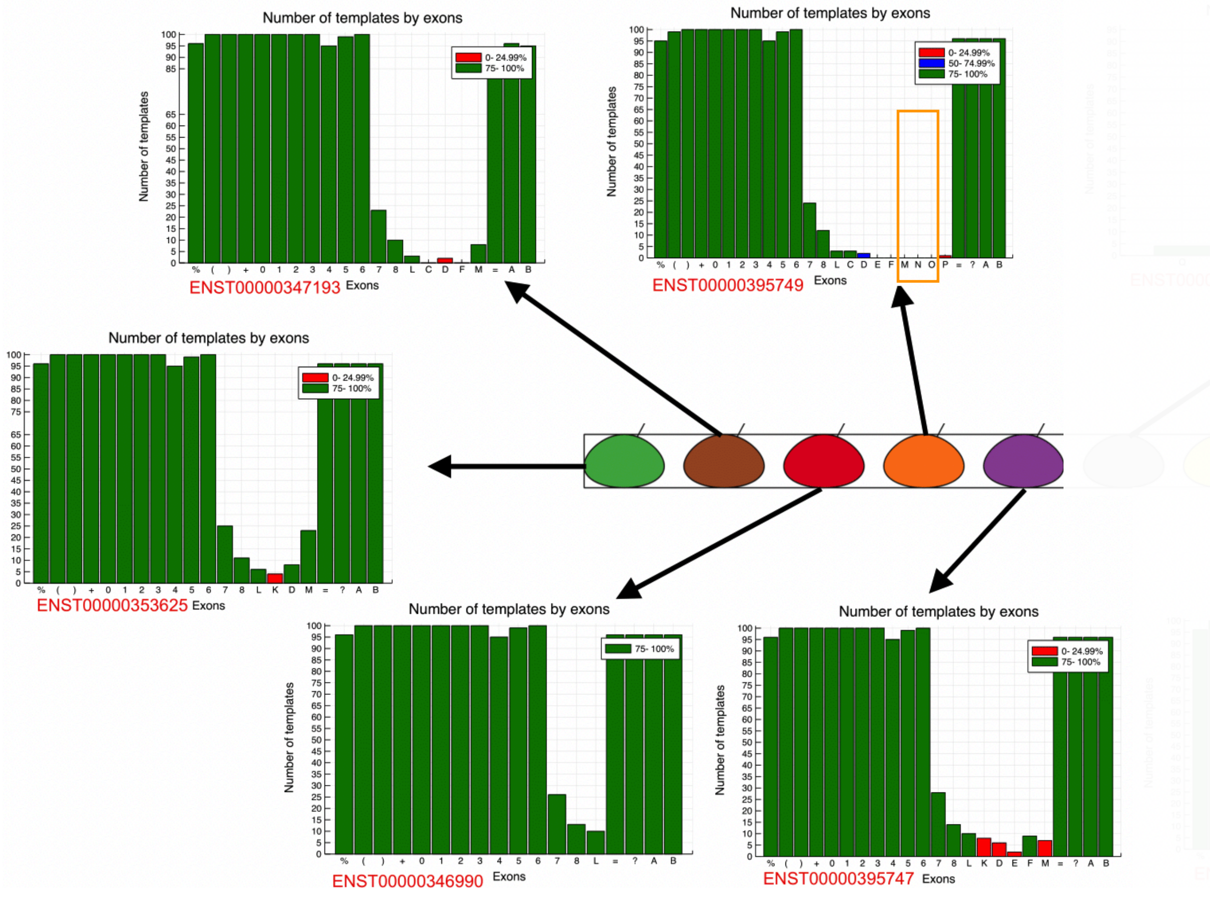


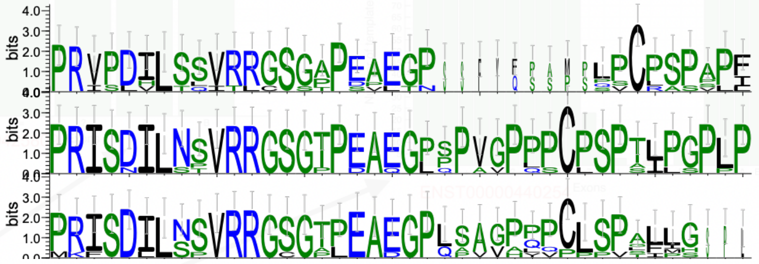


Case study - CAMK2B



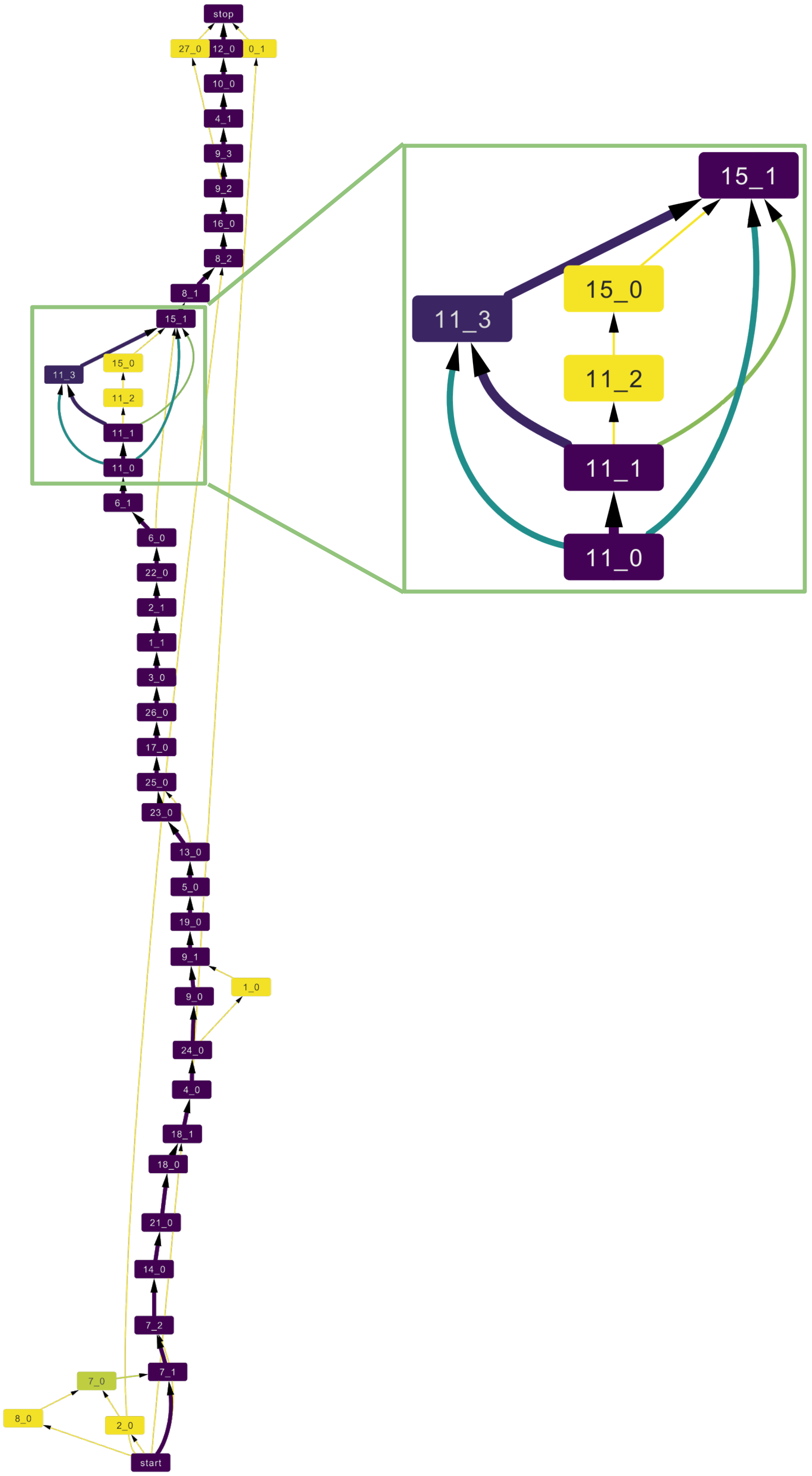




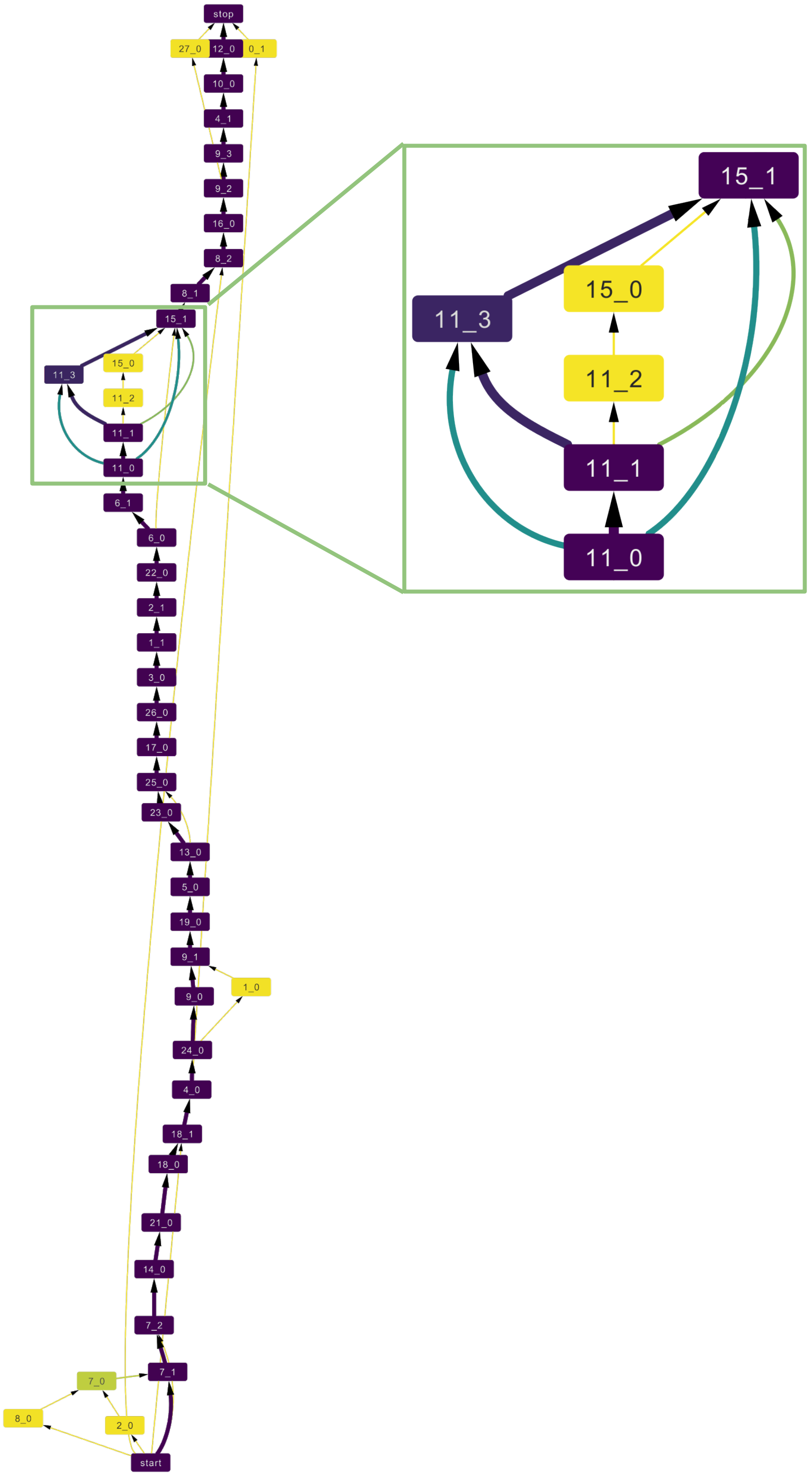


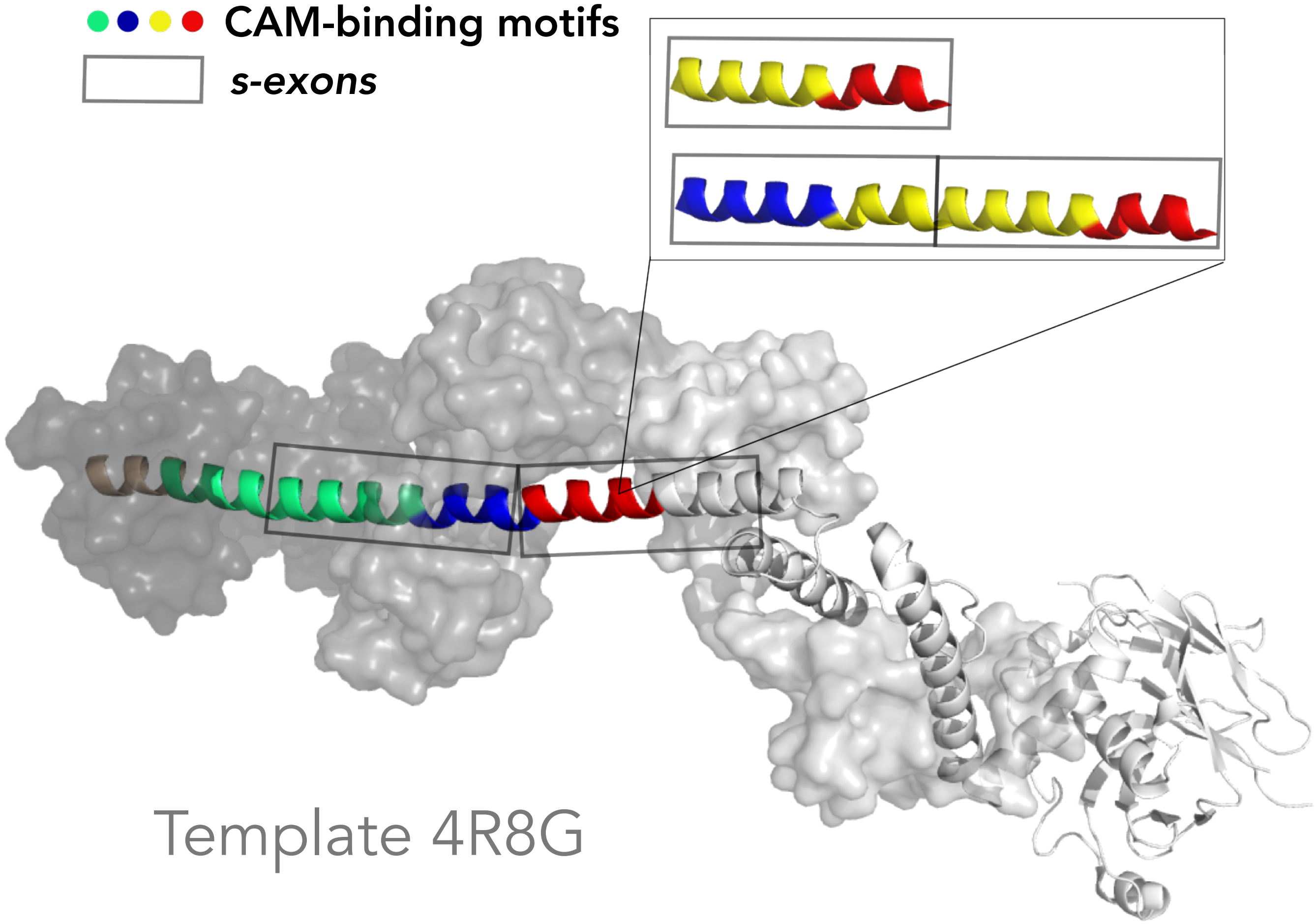
Proline-rich regions, likely to be involved in interactions

Case study - MYO1B



Case study - MYO1B

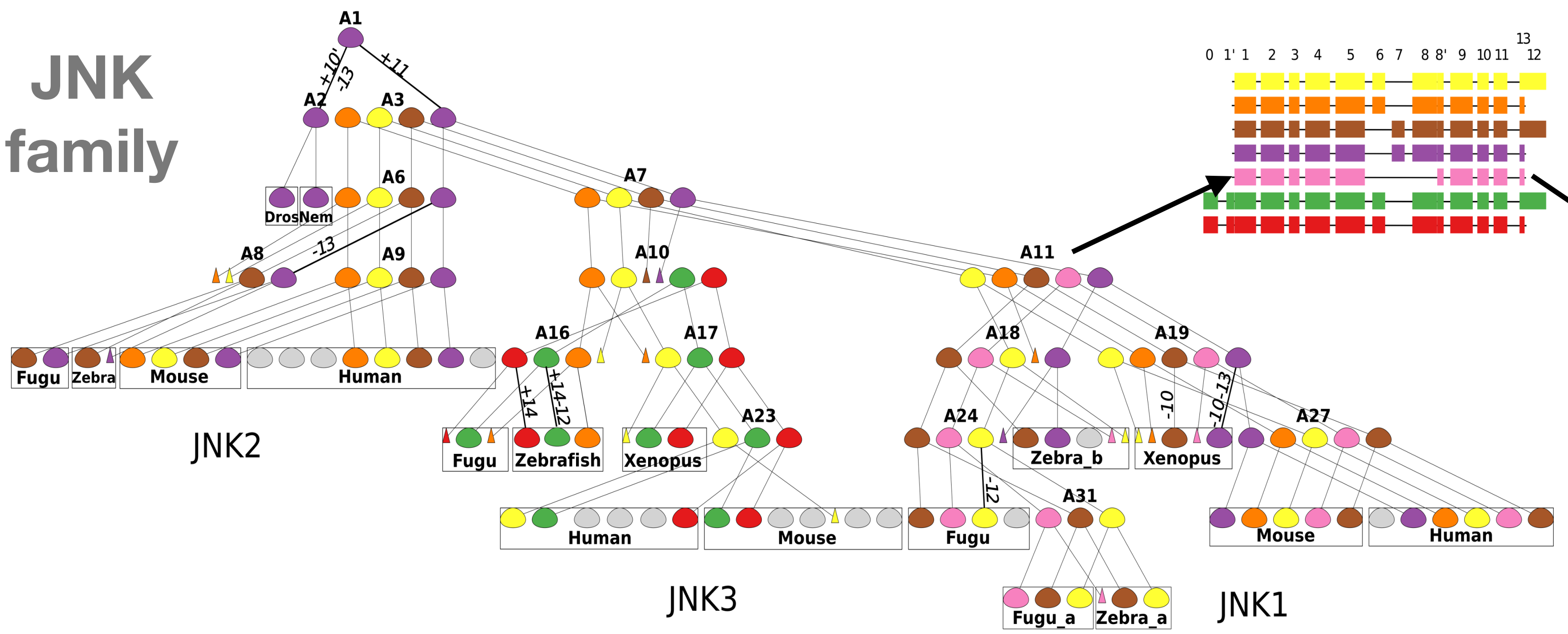


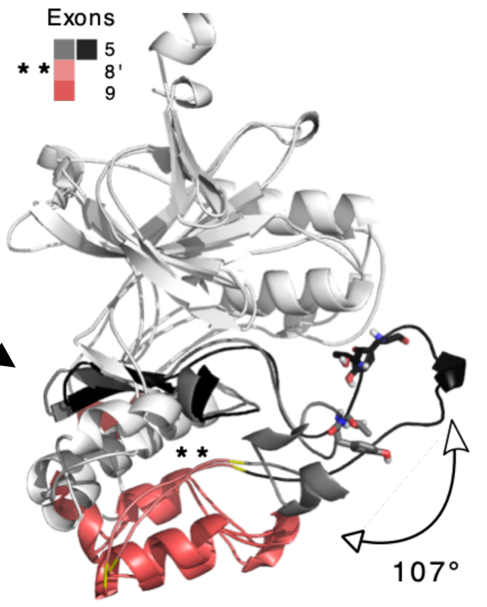


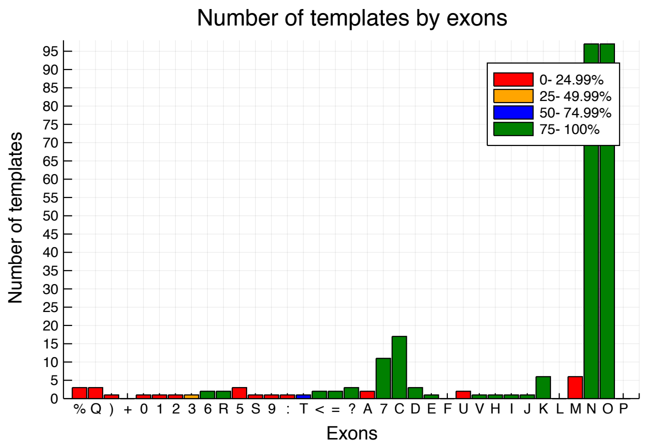
Template 4R8G

A few more examples

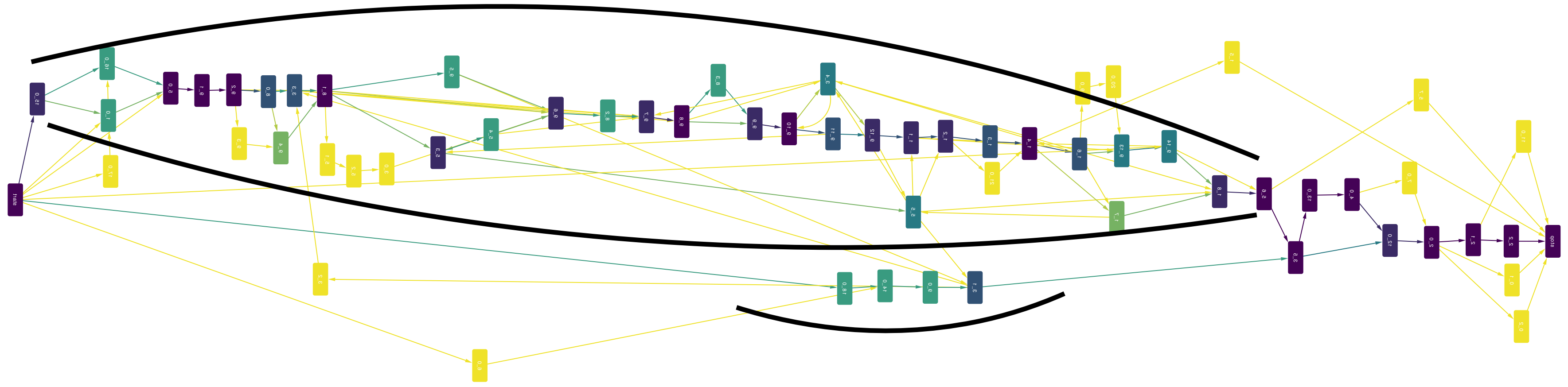
JNK family

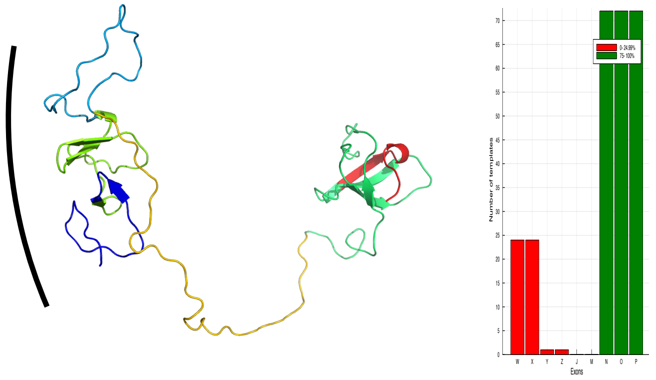






NEBL





Summary of the results

50 human genes…

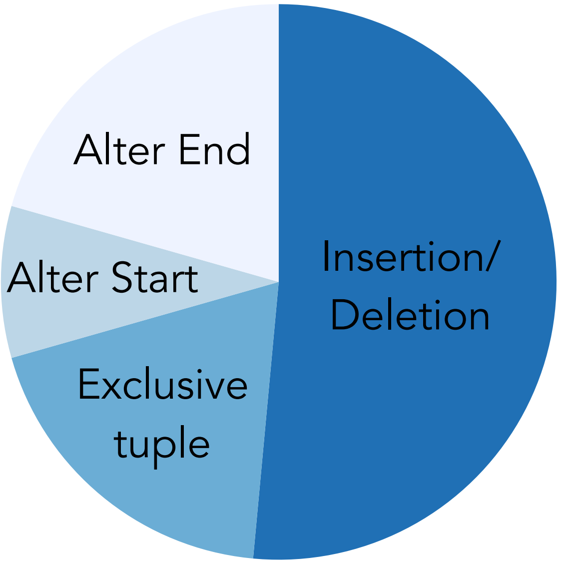
- belonging to 16 families
- across 12 species
- ~900 transcripts (Ensembl)
- [8-91] *s-exons* (ThorAxe)

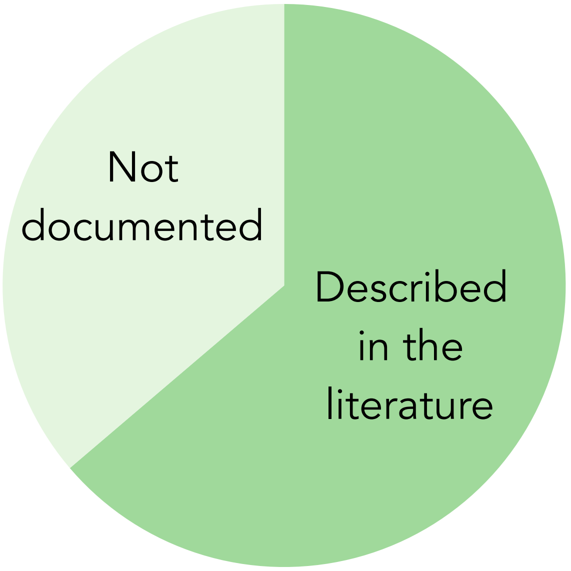
Summary of the results

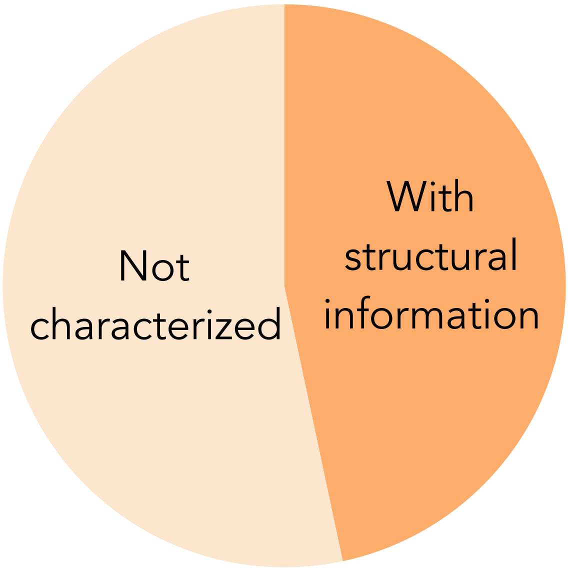
50 human genes…

- ◉ belonging to 16 families
- ◉ across 12 species
- ◉ ~900 transcripts (Ensembl)
- ◉ [8-91] *s-exons* (ThorAxe)

69 conserved AS events involving 105 *s-exons*



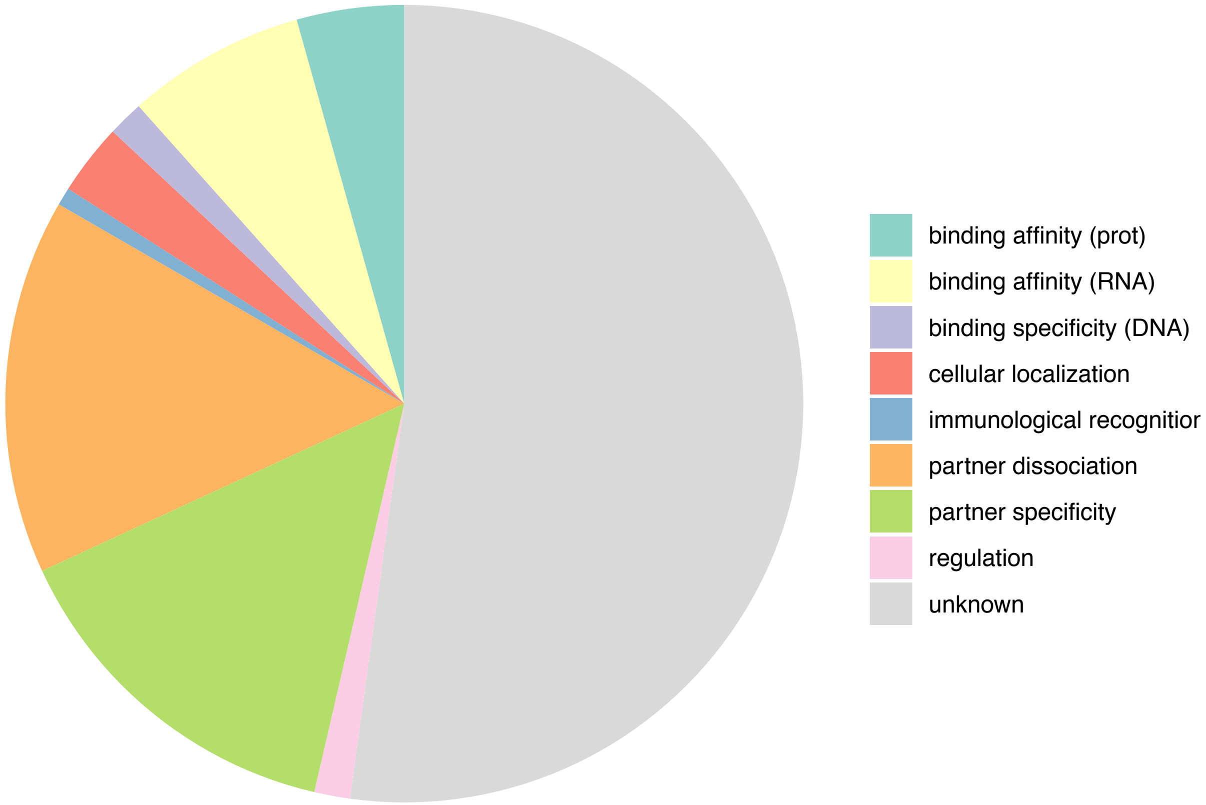




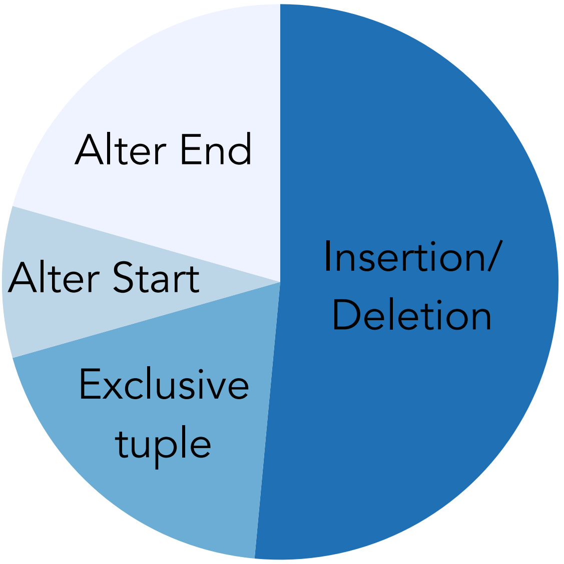
Summary of the results

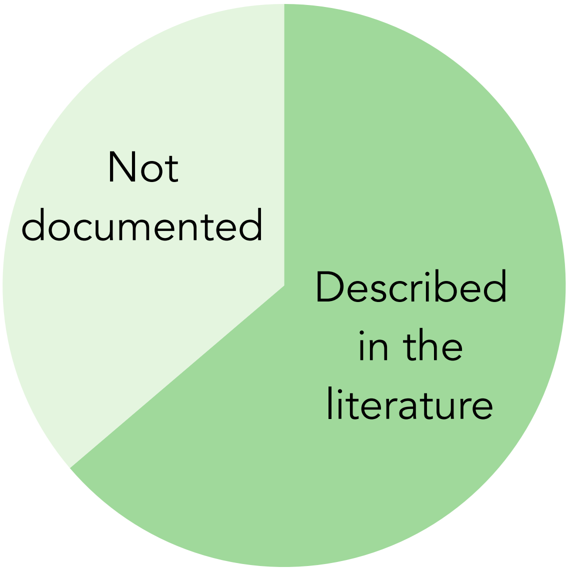
50 human genes…

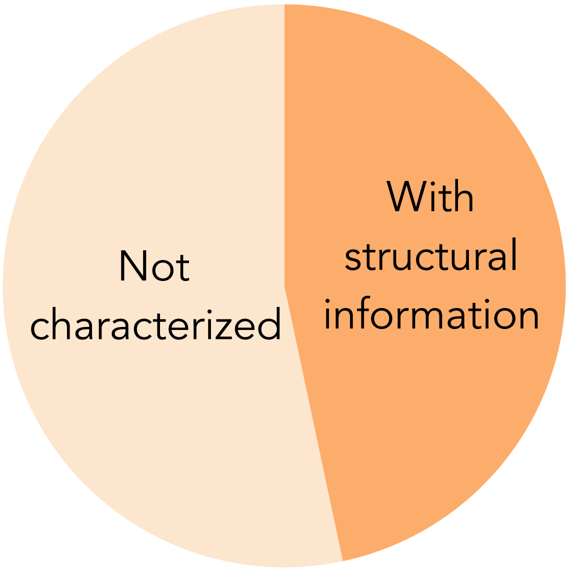
- belonging to 16 families
- across 12 species
- ~900 transcripts (Ensembl)
- [8-91] *s-exons* (ThorAxe)



69 conserved AS events involving 105 *s-exons*

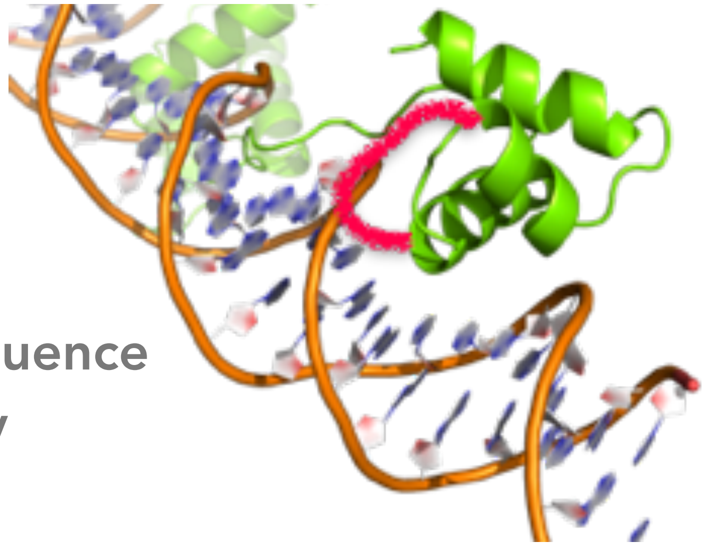






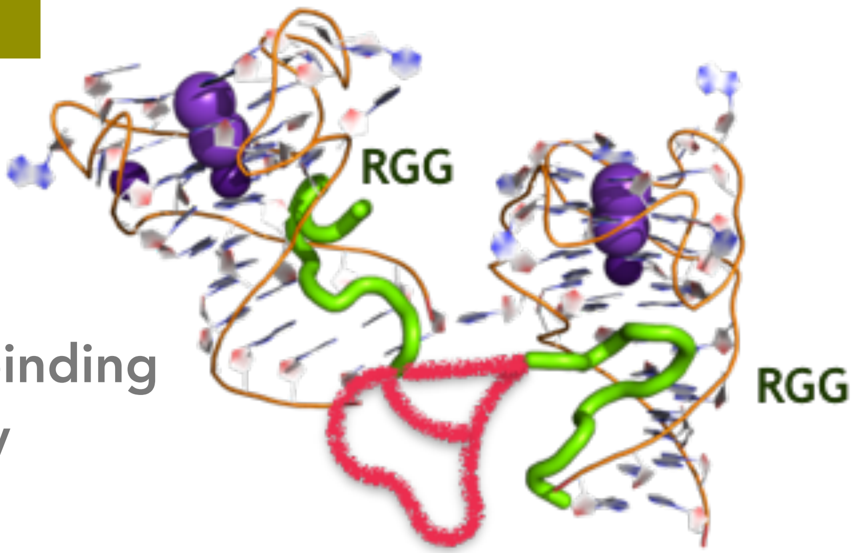
Protein interactions

PAX6



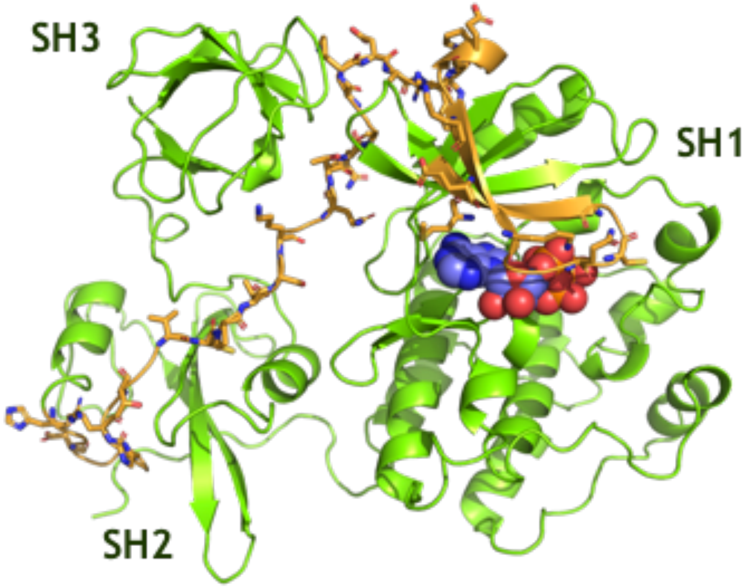
DNA sequence specificity

FMR1



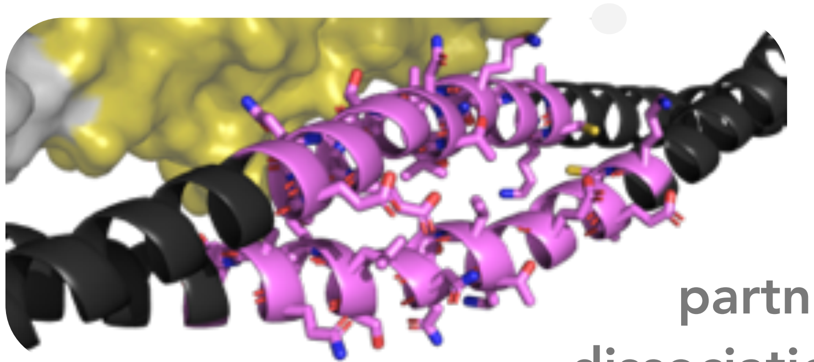
RNA binding affinity

FYN



substrate binding

TPM1



partner dissociation

Perspectives

- Transfer protein interaction annotations
- Characterize disordered regions
- Compare conformational preferences

- Include RNA-seq data to better estimate conservation levels
- Automatically detect events of interest in the splice graph

ThorAxe & PhyloSofS
https://github.com/PhyloSofS-Team

Analytical Genomics @LCQB

Robert Koch Institut

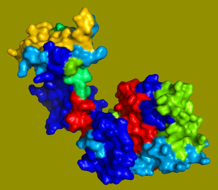
THANKS!


Hugues Richard


Diego Javier Zea


Antoine Labeeuw


Sofya Laskina


Alexander Korzec


Adel Ait-Hamlat


Lélia Polit