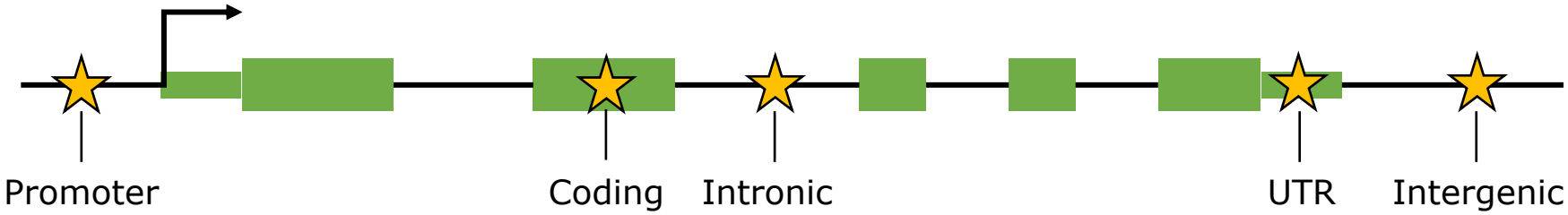
NGS – variant analysis

Variant annotation

Variant annotation

Functional annotation: relative to genomic features (often genes)

- Coding regions
- Promoter regions



Coding variants

GCA AGA GAT AAT TGT TGG CAA

Ala Arg Asp Asn Cys Trp Gln

GCA AGG GAT AAA TGA TG- C AA

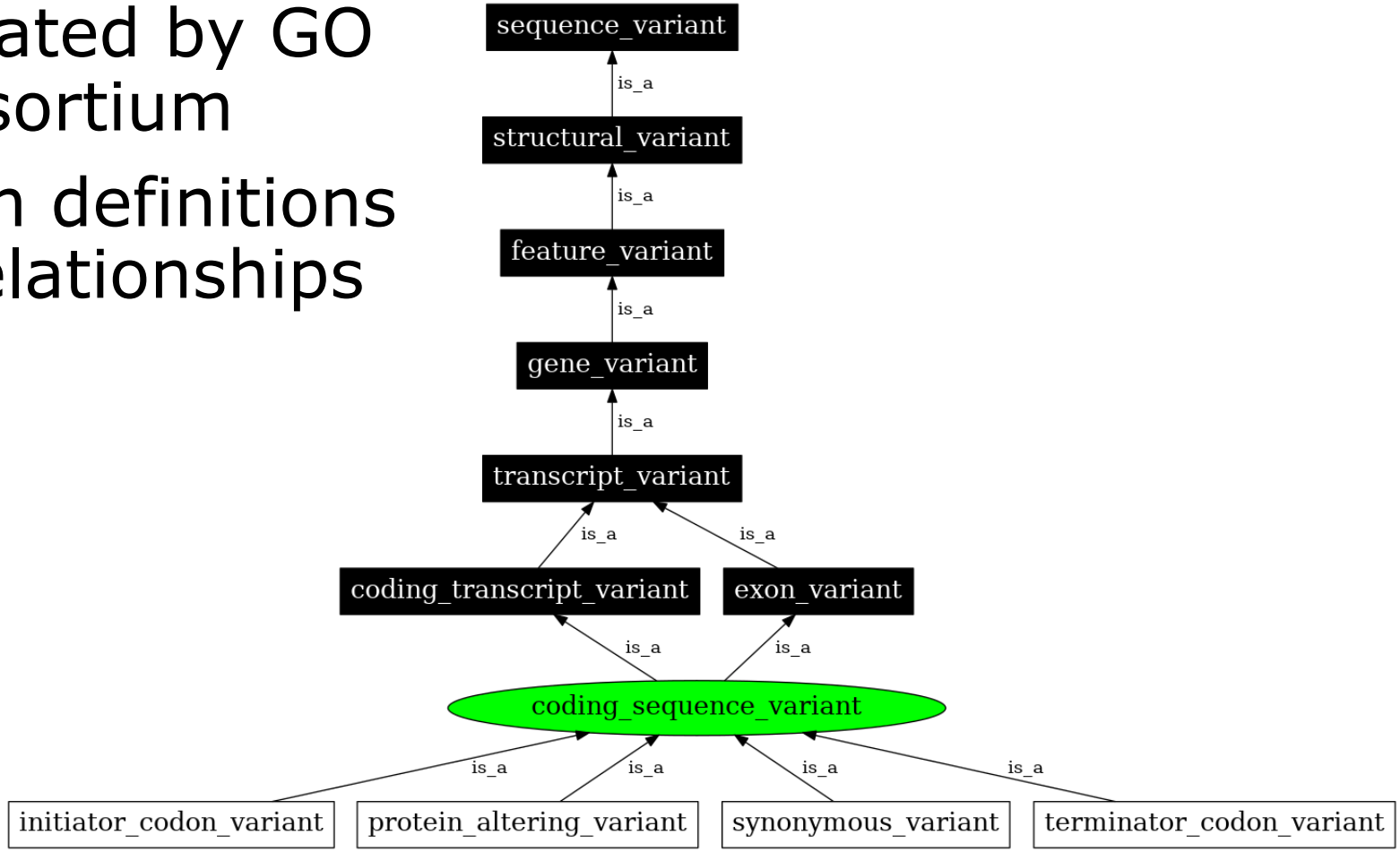
Ala Arg Asp Lys * Cys

http://pcingola.github.io/SnpEff/se_inputoutput/#effect-prediction-details

http://www.sequenceontology.org/browser/current_svn/term/SO:0001580

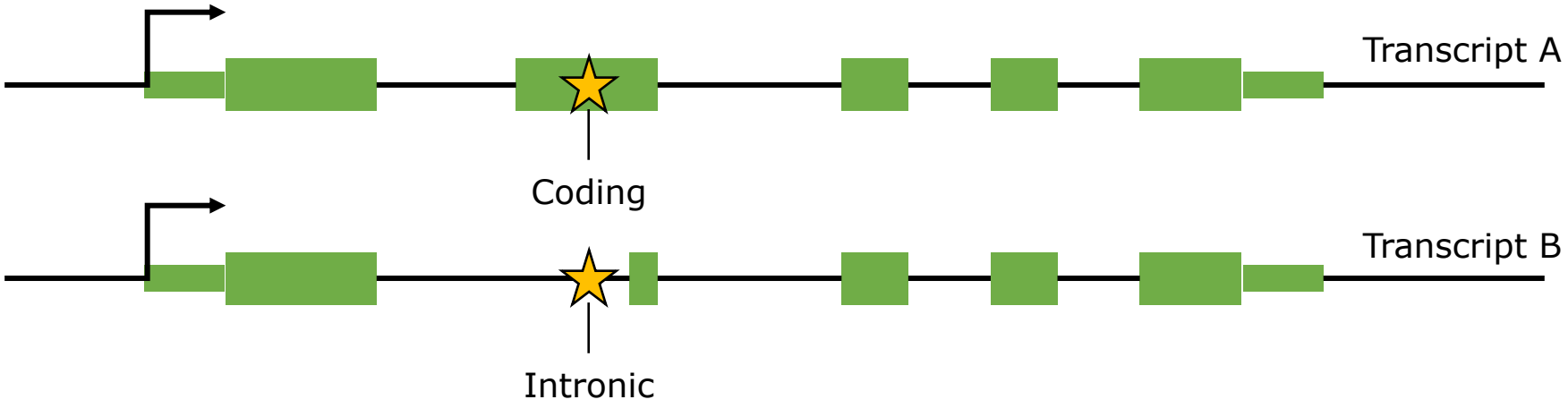
Sequence ontology

- Initiated by GO consortium
- Term definitions + relationships



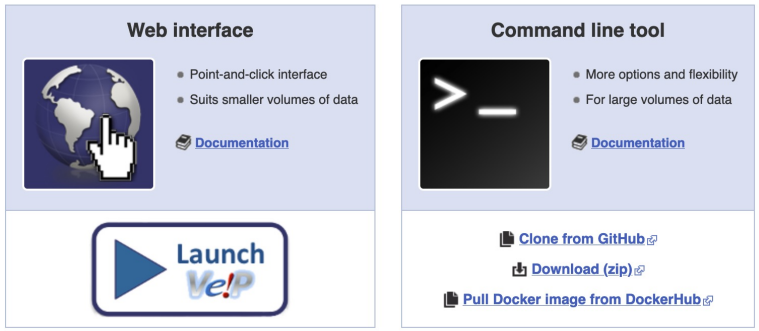
http://www.sequenceontology.org/browser/current_svn/term/SO:0001580

Isoforms



- Most genes have multiple isoforms
- Annotation for all isoforms?
- Single annotation/variant:
  - Effect with most impact
  - Most relevant transcript, e.g. based on ensembl transcript tags

Functional annotation tools



https://www.ensembl.org/info/docs/tools/vep/index.html

ANNOVAR

Command line:
https://annovar.openbioinformatics.org/
Web-based:
https://wannovar.wglab.org/

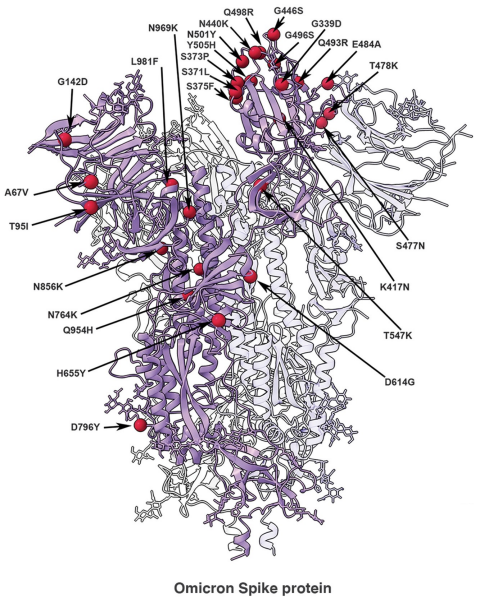
SnpEff

Genetic variant annotation and functional effect prediction toolbox. It annotates and predicts the effects of genetic variants on genes and proteins (such as amino acid changes).
Features:

- Supports over **38,000 genomes**.

Effect prediction scores

- One “high impact” mutation is not the other
- Prediction of effect based on e.g.:
  - Comparative genomics (conserved regions)
  - Protein structure/biochemistry
  - Experimental function (e.g. eQTL)
- For human genome many databases available, e.g.:
  - dbNSFP
  - MutationTaster



Omicron Spike protein