

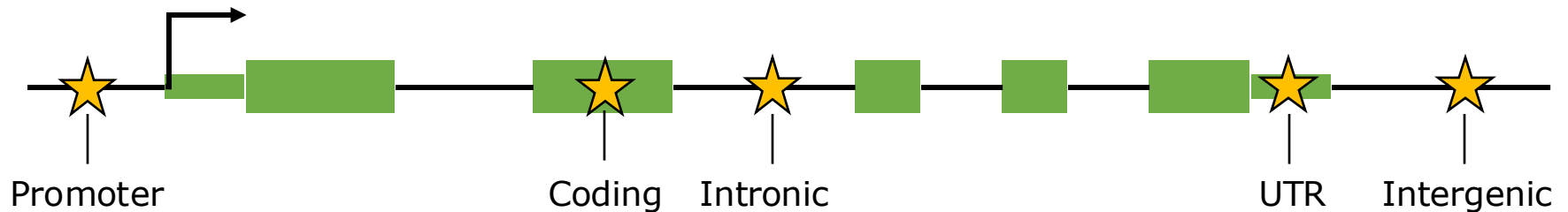
NGS - variant analysis

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

synonymous

missense

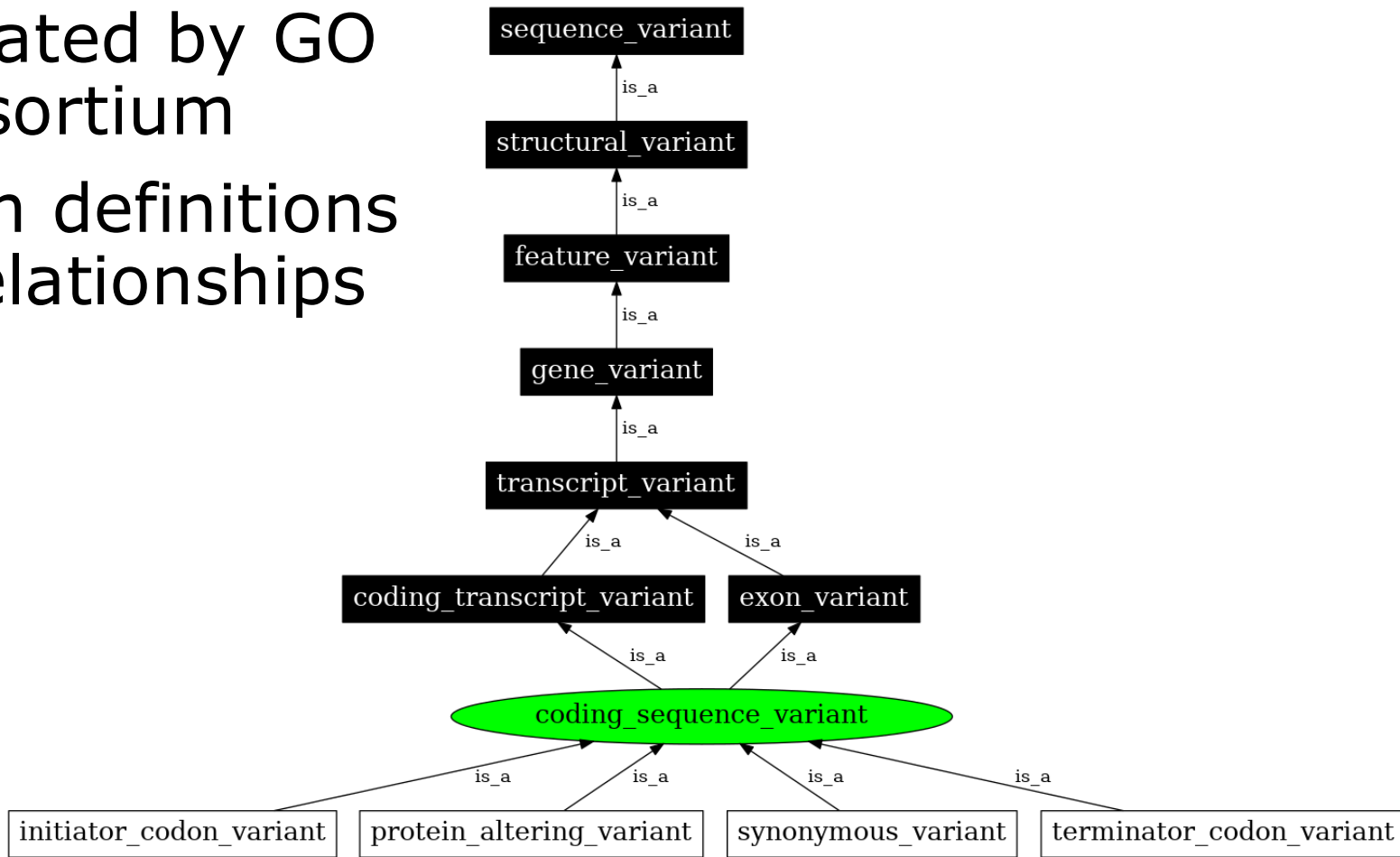
frameshift

stop gained

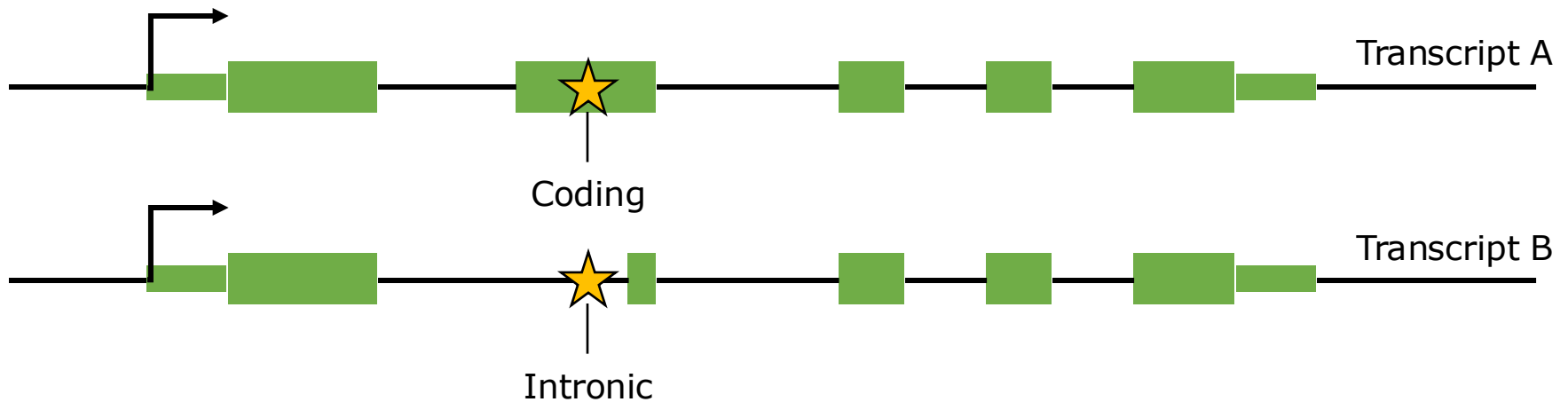
GCA AGG GAT AAA TGA TG-C AA
Ala Arg Asp Lys * Cys

Sequence ontology

- Initiated by GO consortium
- Term definitions + relationships




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

Web interface




- Point-and-click interface
- Suits smaller volumes of data

[Documentation](#)



Command line tool



- More options and flexibility
- For large volumes of data


[Documentation](#)

[Clone from GitHub](#)

[Download \(zip\)](#)

[Pull Docker image from DockerHub](#)

VEP



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](#)

