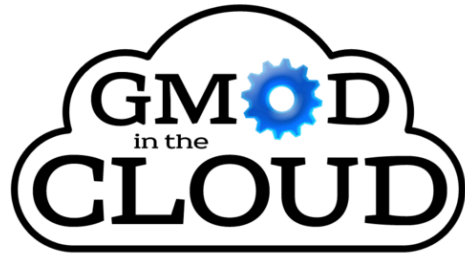


The genome annotation pipeline



MAKER
Annotate this!





GMOD in the Cloud toolset



GBrowse: Genome annotation viewer



Generic Model Organism Database project



Galaxy: Data analysis & integration



Chado: Biological database schema



JBrowse: Super-fast genome annotation viewer



BioMart: Data mining system



WebApollo: browser-based annotation editor



MAKER: Genome annotation pipeline



GBrowse_syn: Synteny viewer



Tripal: Chado web interface



InterMine: Data warehousing



CMap: Comparative map viewer



Pathway Tools: Metabolic, regulatory pathways



Canto: literature annotation tool

Eukaryotic genome annotation

MAKER: An easy-to-use annotation pipeline designed for ... - NCBI - NIH

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2134774/> ▼ Översätt den här sidan

av BL Cantarel - 2008 Citerat av 612 - Relaterade artiklar

We have developed a portable and easily configurable genome annotation pipeline called **MAKER**. Its purpose is to allow investigators to independently ...

[Abstract](#) · [Results](#) · [Discussion](#) · [Methods](#)

09/2018

MAKER2: an annotation pipeline and genome ... - BMC Bioinformatics

<https://bmcbioinformatics.biomedcentral.com/.../1471-2105-12-...> ▼ Översätt den här sidan

av C Holt - 2011 Citerat av 514 - Relaterade artiklar

22 dec. 2011 - We present **MAKER2**, a genome annotation and data management tool designed for second-generation genome projects. **MAKER2** is a ...

MAKER – developed as an easy-to-use alternative to other pipelines

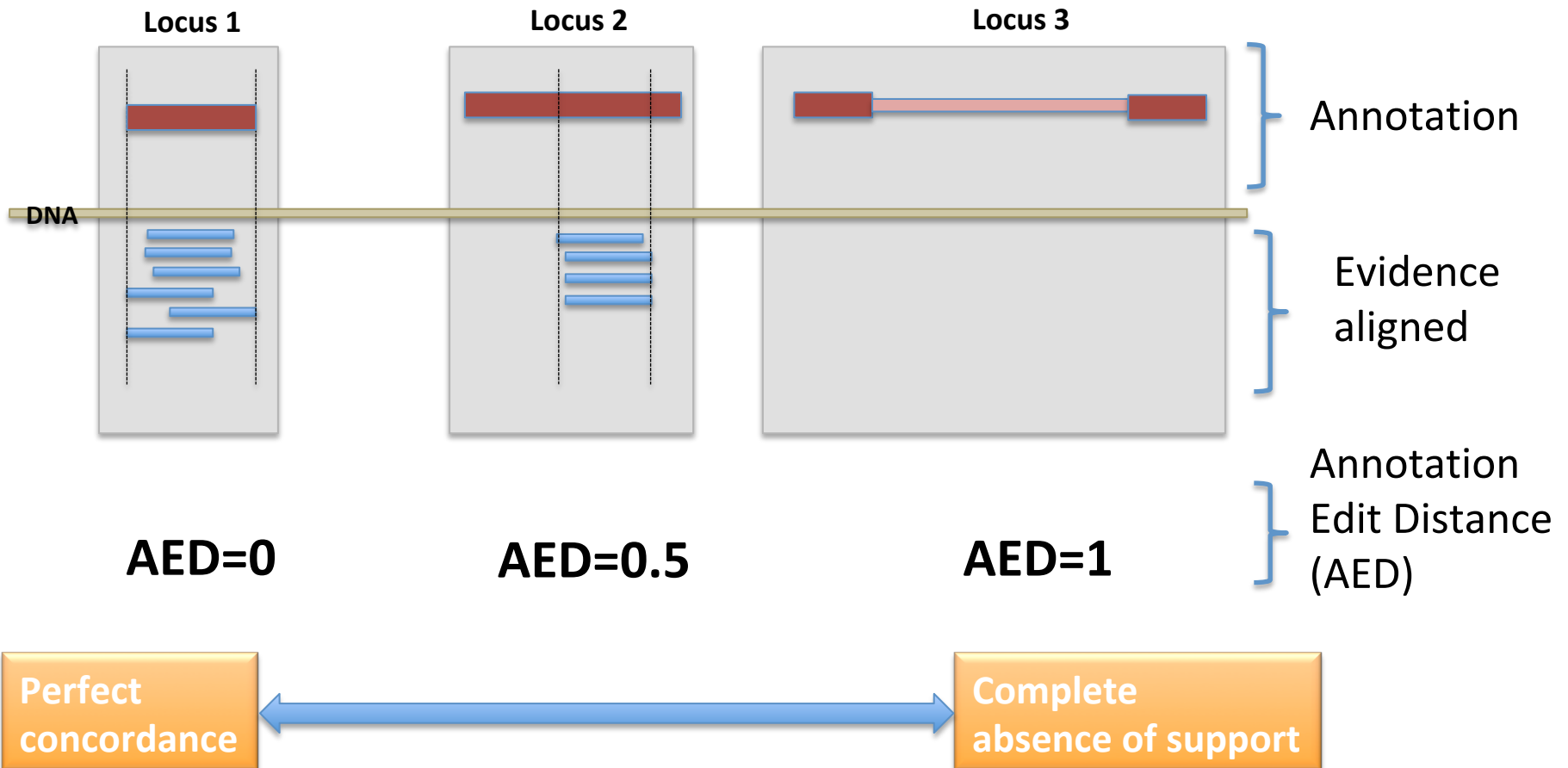
Why choose MAKER ?

- Easy to use and to configure
- Scale to datasets of any size
- Multi-threaded and parallelized
- Everything is run through one command, no manual combining of data/outputs
- Metric for quality control (**A**nnotation **E**dit **D**istance)
- Distributed with accessory scripts (>30)
- All its capabilities...

MAKER

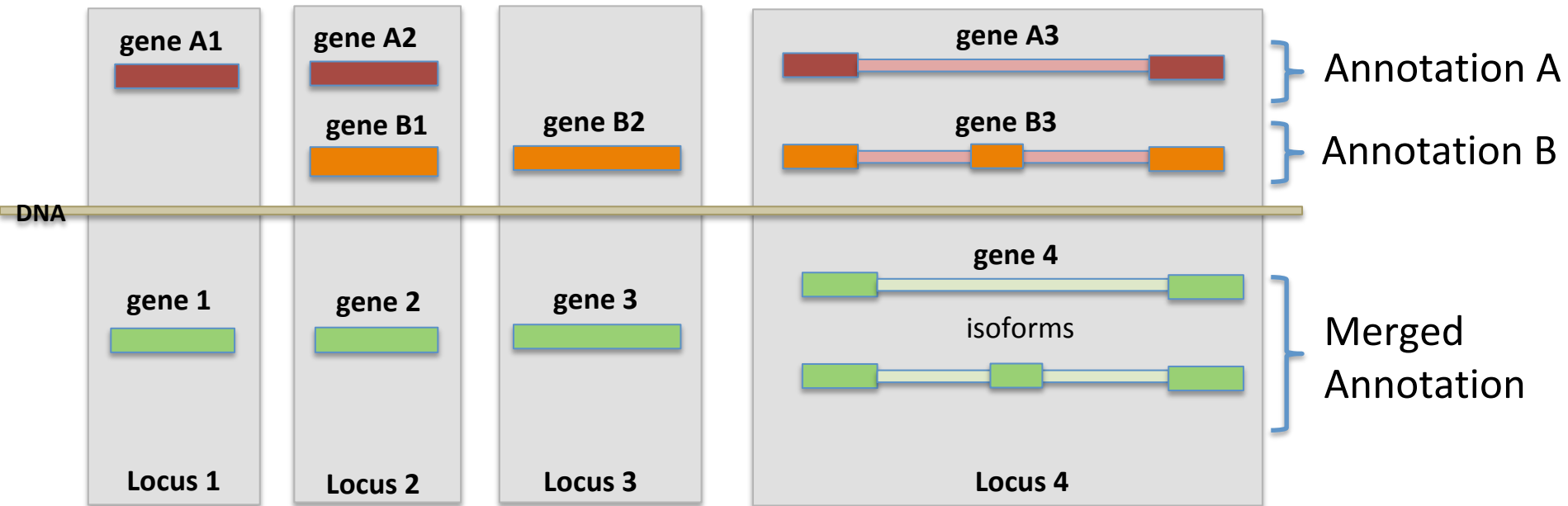
Capabilities:

- ☐ Add quality metrics to an annotation (AED)



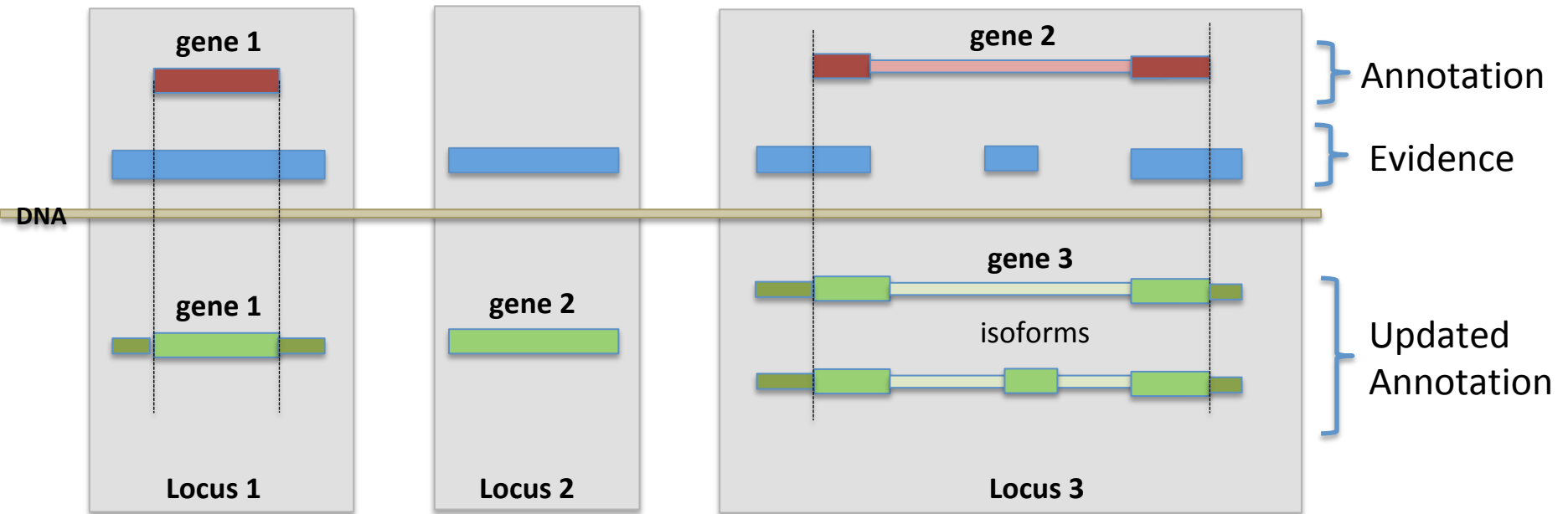
Capabilities:

- ☐ Merge multiple annotation sets



Capabilities:

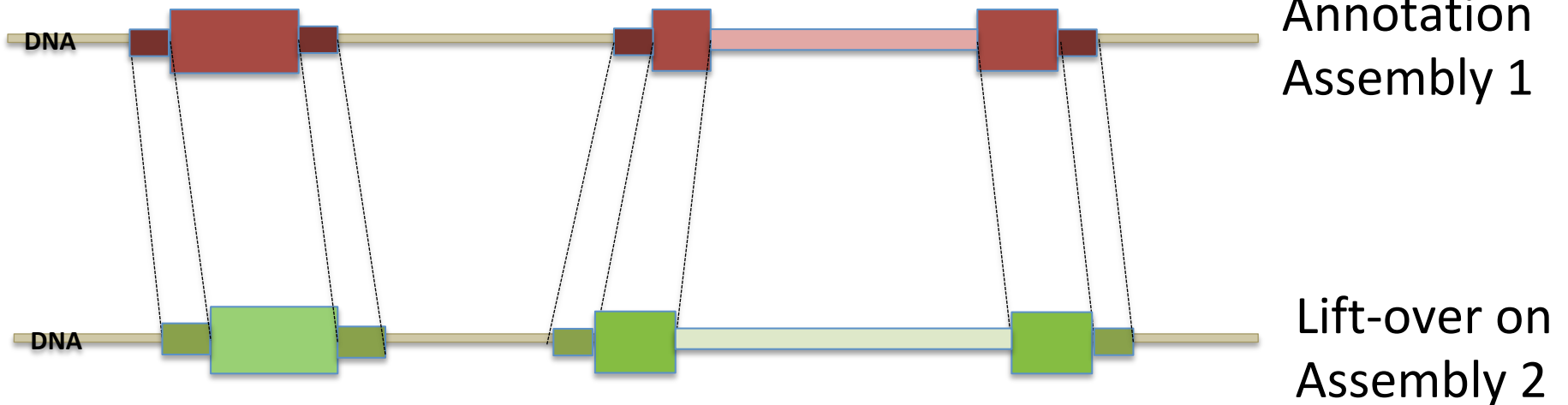
- ❑ Update an annotation in light of new evidence



MAKER

Capabilities:

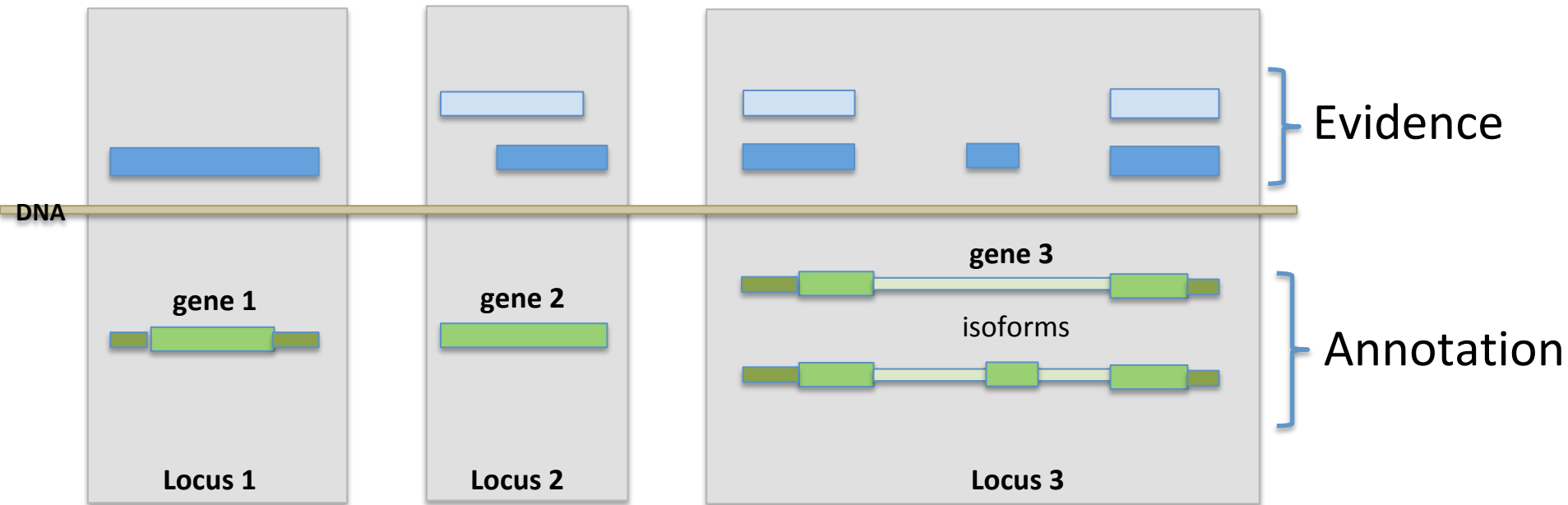
- ☐ Map annotation forwards to a new assembly



MAKER

Capabilities:

- ☐ Annotation pure evidence-based



- ⇒ Suitable for ab-initio training purpose (filtering needed)
- ⇒ Prediction not always complete ! (always_complete option)

MAKER



Capabilities:

- ☐ Annotation pure *ab-initio*



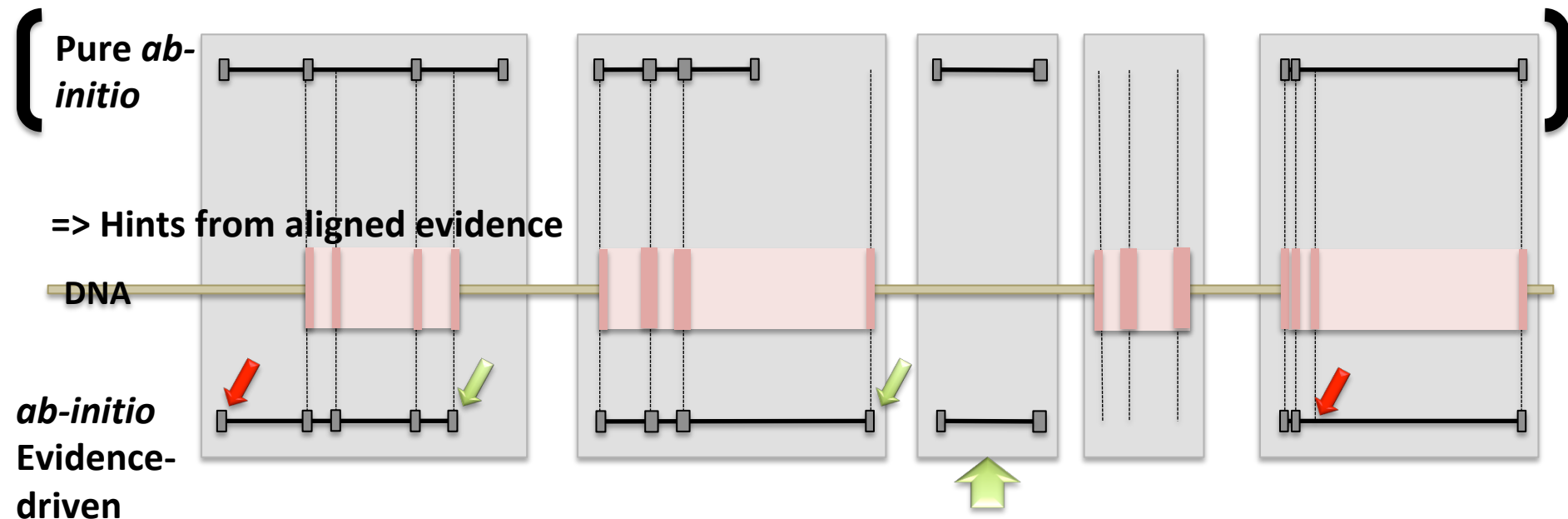
The same as standalone *ab-initio* ! So why MAKER ?

- ⇒ To take advantage of parallelization !
- ⇒ Can use several *ab-initio* tools (they can complement each other)

MAKER

Capabilities:

- ☐ Annotation *ab initio* evidence-driven



Ab-initio tools are better when hints are provided

Ab-initio predictions can fill gaps with no evidence

⇒ But may still be incomplete / partially wrong

But how does Maker work exactly?

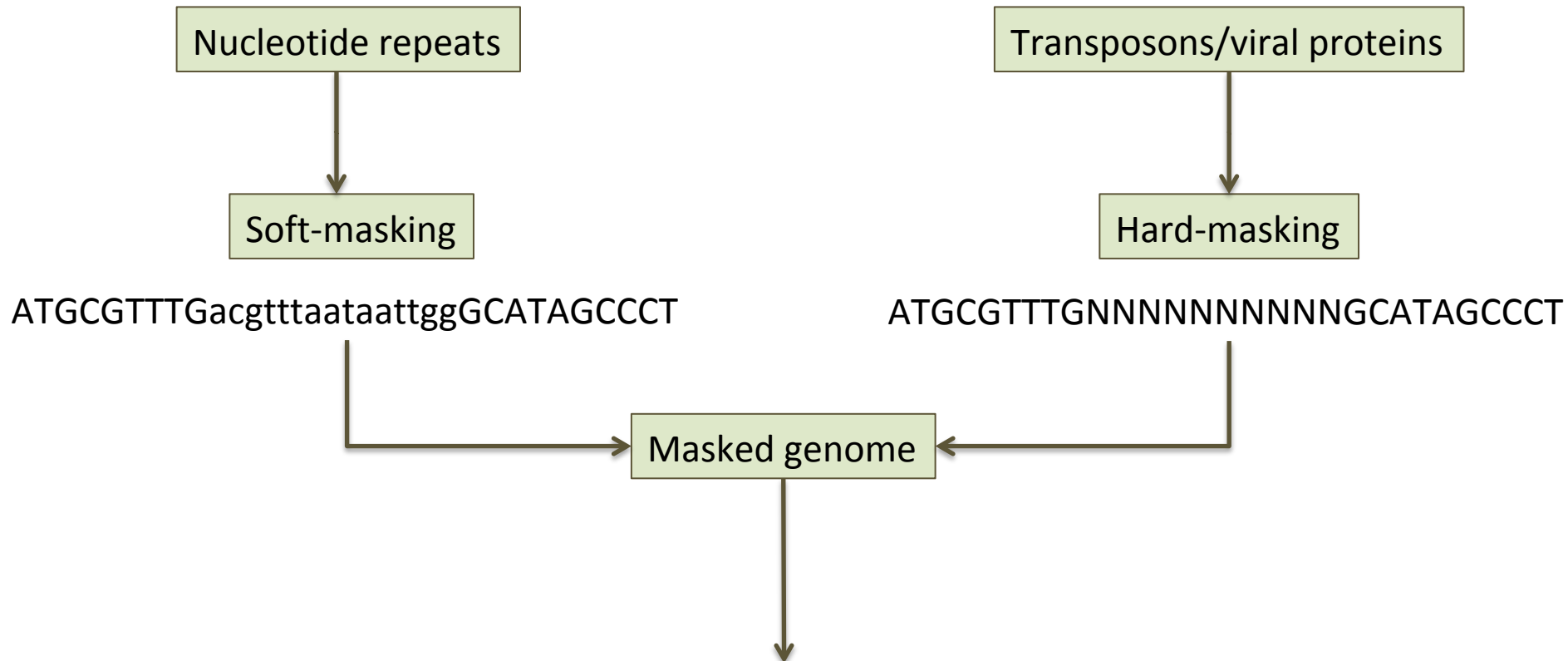
Use case

⇒ *Ab-initio* evidence-driven

Prerequisite:

- ☐ Evidence (proteins and/or transcripts)
- ☐ Hmm profile for *ab-initio* tool(s)
(Augustus comes with some pre-calculated profiles)

Step 1 - Repeat masking with repeatmasker

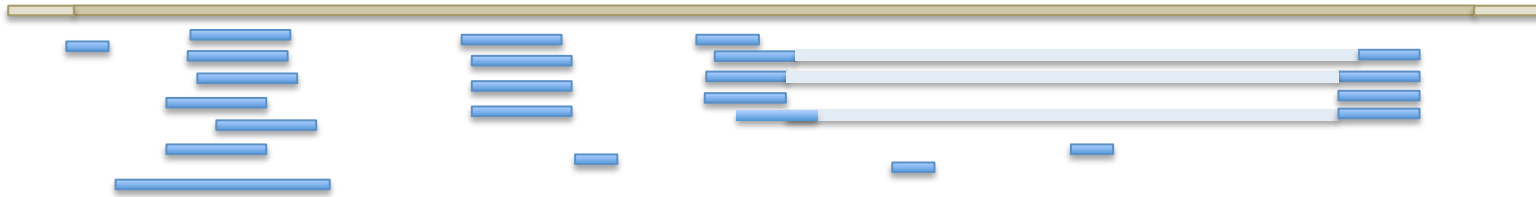


MAKER

Step 2 - Transcript and protein evidence alignment



Step 3 - Filtering and clustering alignments



Filtering is based on rules defined in the Maker configuration for a given project

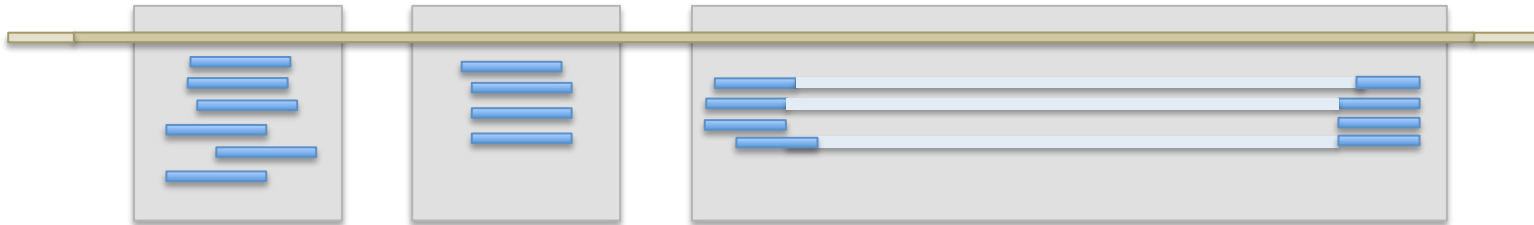
Example: EST alignment – 80% coverage and 85% identity

Default settings sensible for most projects, but can be changed!

MAKER

Step 3 - Filtering and clustering alignments

Clustering into 'loci'

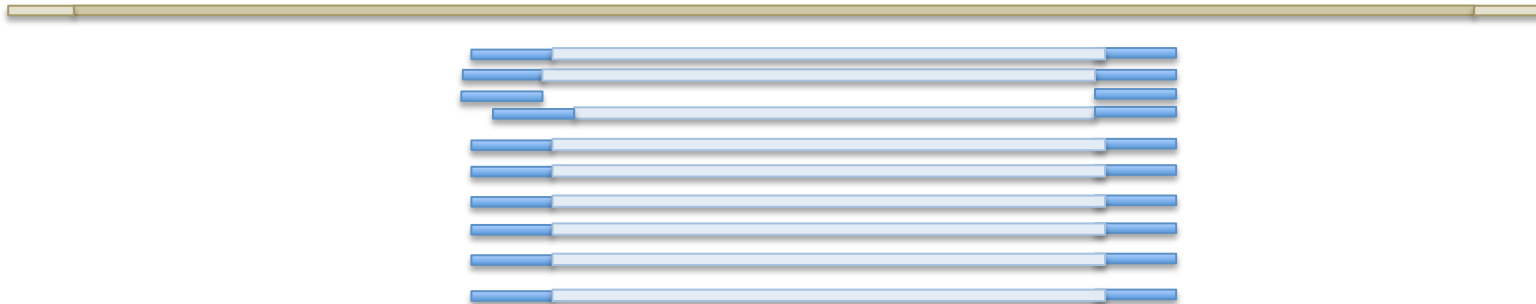


Importance of the quality of the data used:



=> Bad data can complicate clustering

Step 3 - Filtering and clustering alignments



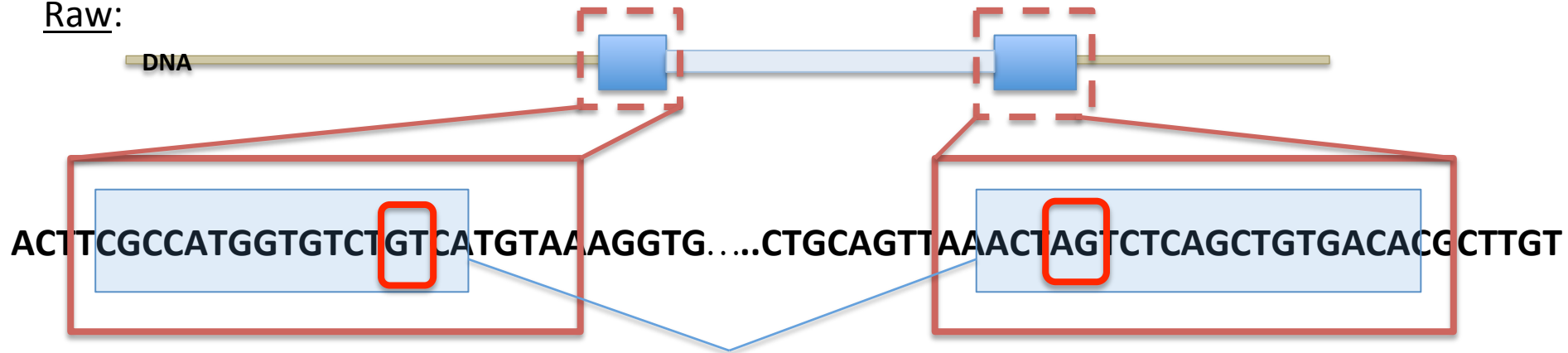
Amount of data in any given cluster is then collapsed to remove redundancy

Threshold for the collapsing is also user-definable

MAKER

Step 4 - Polishing evidence alignments

Raw:



Blast-based alignments are only approximations, need to be refined

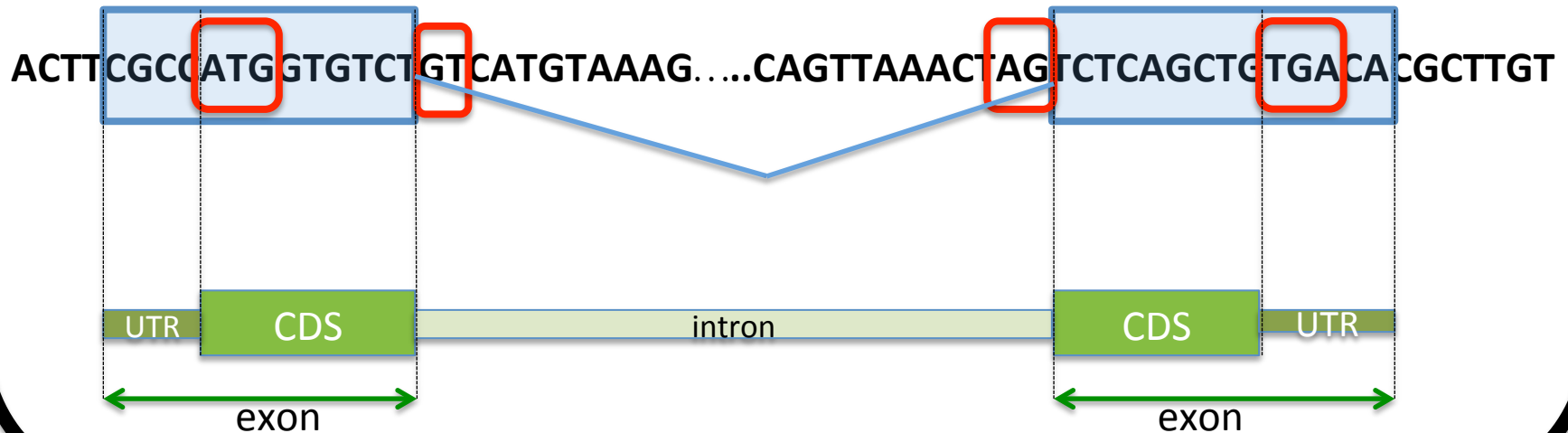
Exonerate is used to create splice-aware alignments

Polished:



Parenthesis !

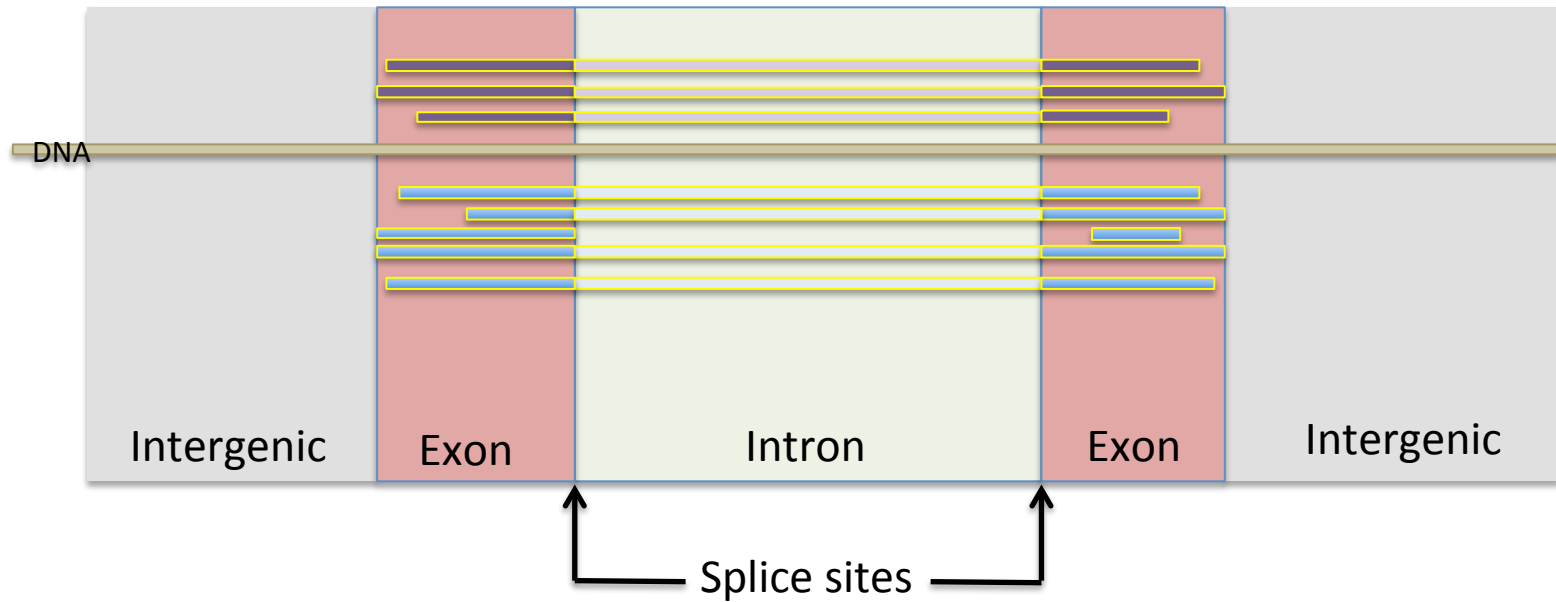
⇒ In a pure evidence based case, the last step will be the creation of gene model from polished alignments.



Let's get back on track !

MAKER

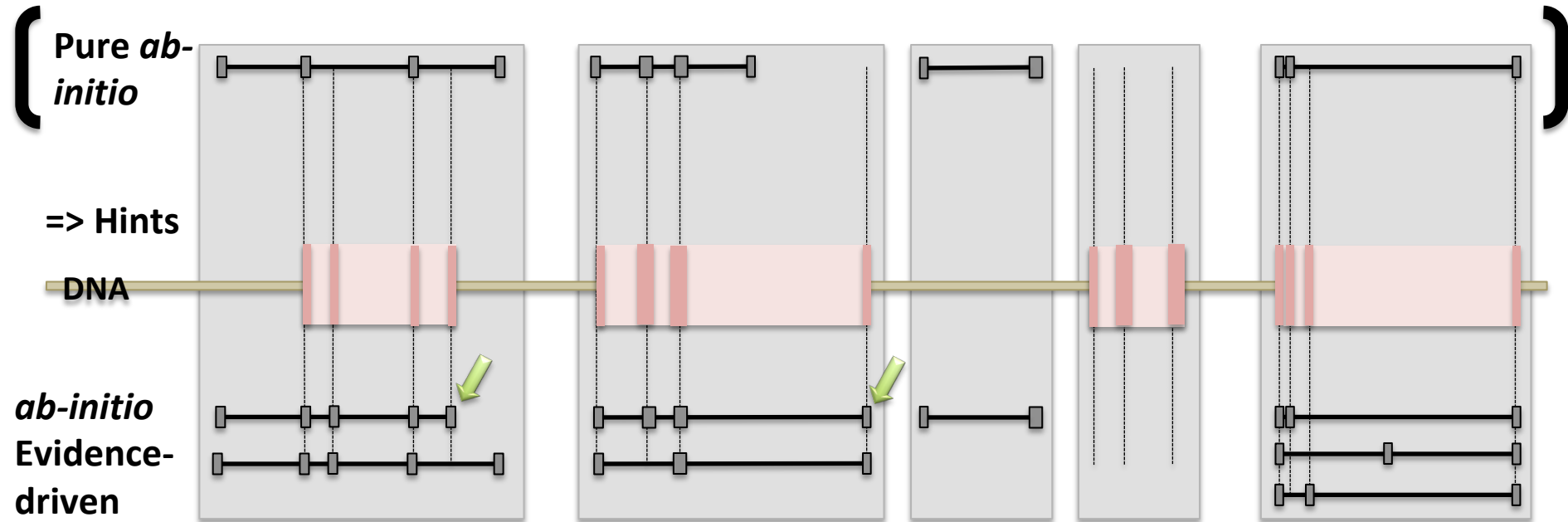
Step 5 – Generating hints



Hints are passed to *ab-initio* tools that accept them

MAKER

Step 6 - *Ab-initio* gene prediction (evidence-driven)

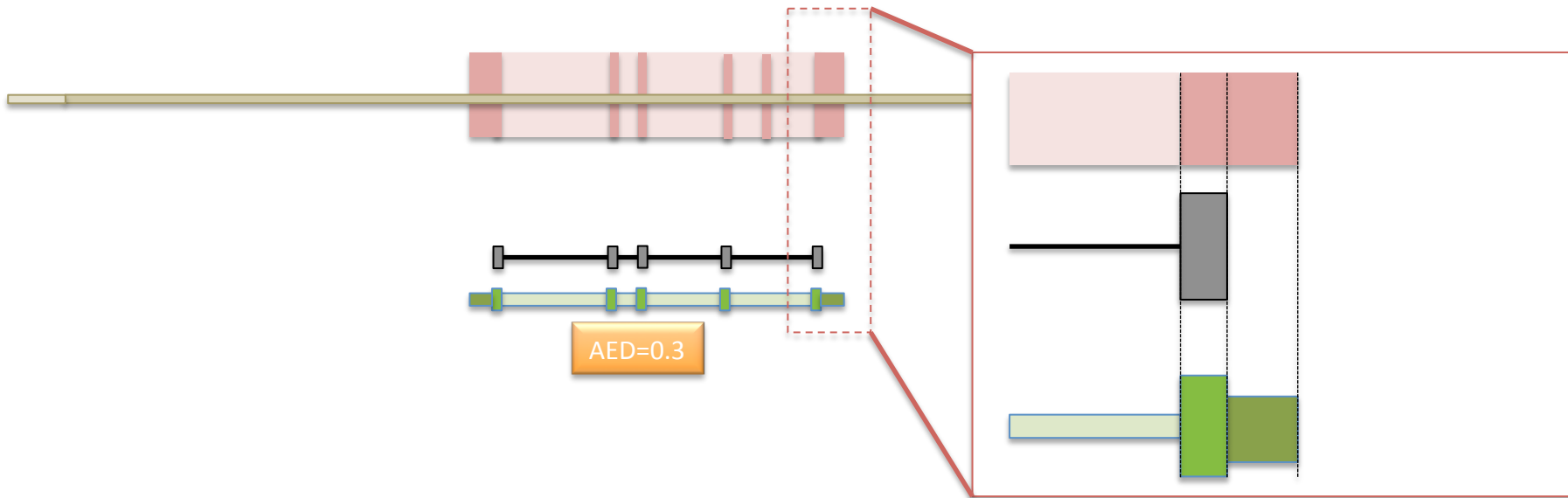


Ab-initio predictions are refined when hints are provided

Isoforms accepted if parameter activated

MAKER

Step 7 - Annotation



- Add UTRs
- Add quality control metrics

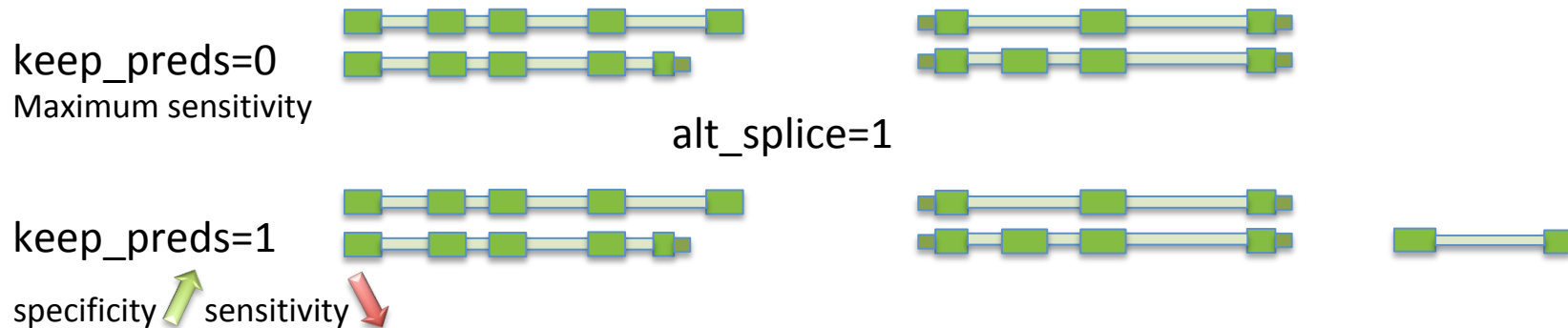
MAKER

Step 8 – Selecting gene models

- ⇒ selected in agreement with the available evidence
- ⇒ The minimum agreement threshold can be chosen

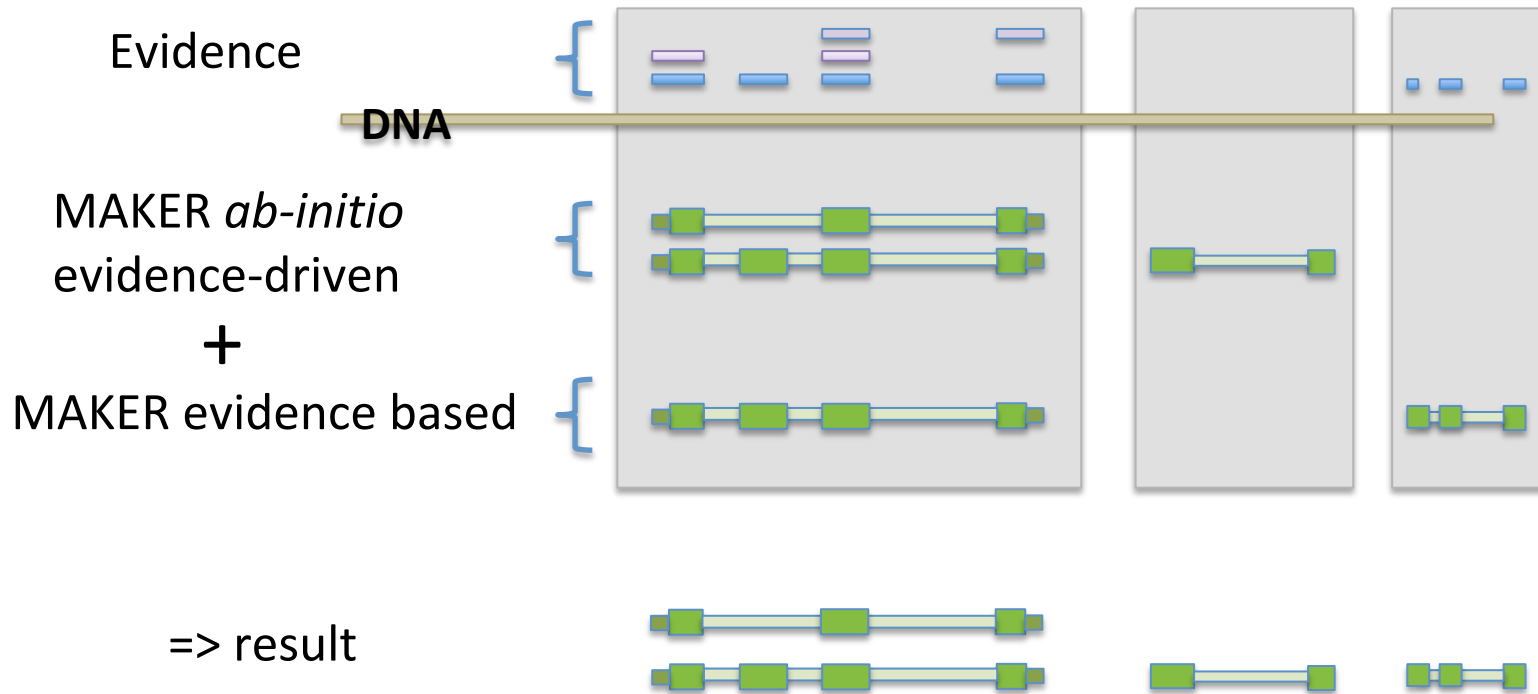


Final MAKER Annotation:



MAKER

Tip - complete the annotation



MAKER

What next ?



MAKER
Annotate this!

Output = Annotation in **gff3** format



Genome browser

CHADO

Biological
database schema



Tripal: Chado web interface

Web  **pollo**

Browser-based annotation editor

Submission

Downstream analysis

...

THE END

