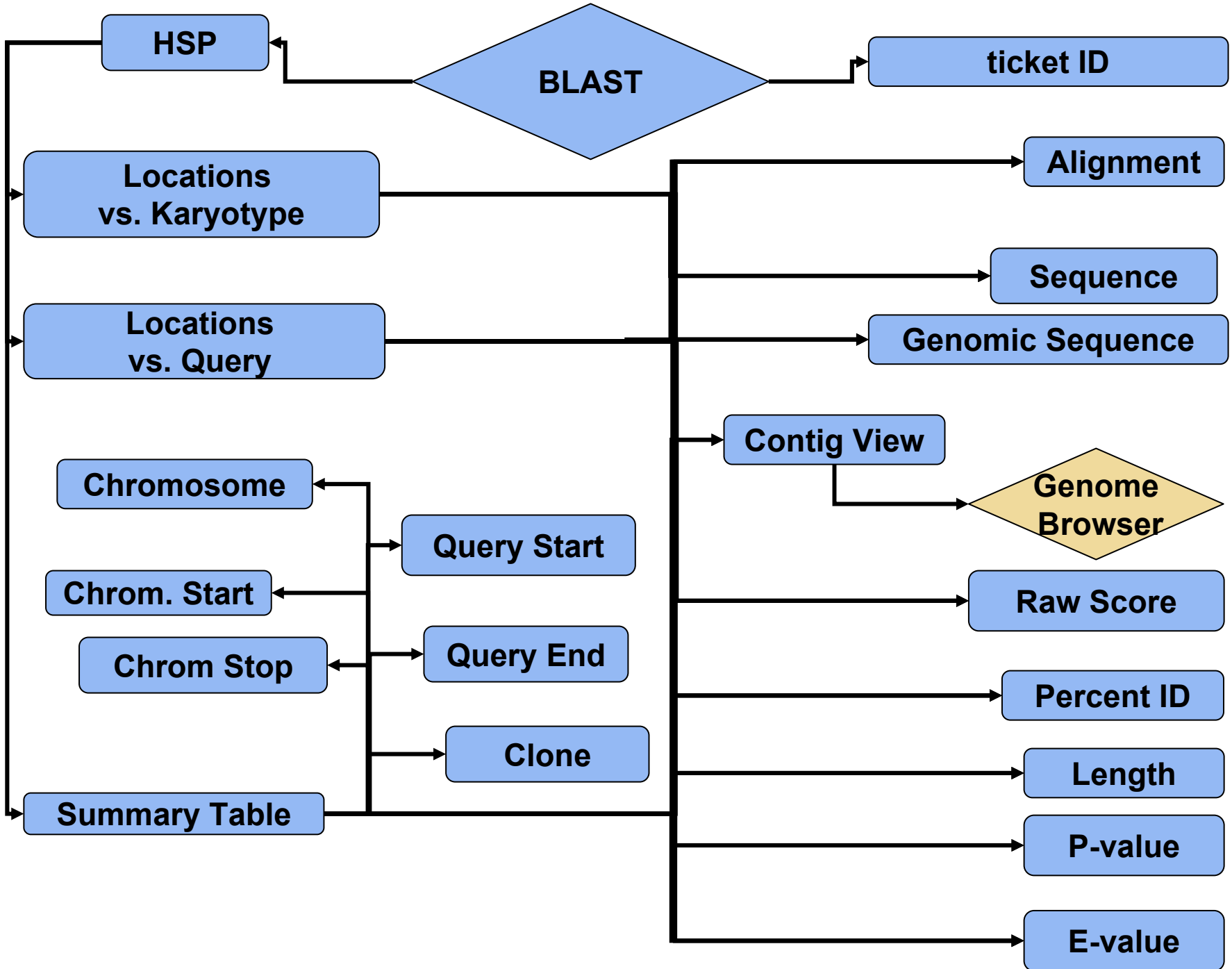
Welcome to the Gramene BLAST Tutorial

- This tutorial will show you how to conduct a BLAST search.
- With BLAST you may:
  - Search for sequence similarity matches in the Gramene database.
  - Select the best target database for your search.
  - Choose the best algorithm for your search.
  - Fine-tune search parameters.


HSP
BLAST
ticket ID
Locations vs. Karyotype
Alignment
Sequence
Locations vs. Query
Genomic Sequence
Contig View
Chromosome
Genome Browser
Query Start
Chrom. Start
Raw Score
Chrom Stop
Query End
Percent ID
Clone
Length
Summary Table
P-value
E-value

Tutorial Help

The hand icon indicates a link that allows you to go to the same page in your web browser.

If you are viewing this tutorial with Adobe Acrobat Reader, click the "bookmarks" on the left hand side of the Reader for easier navigation.

Action Options are noted in this font.

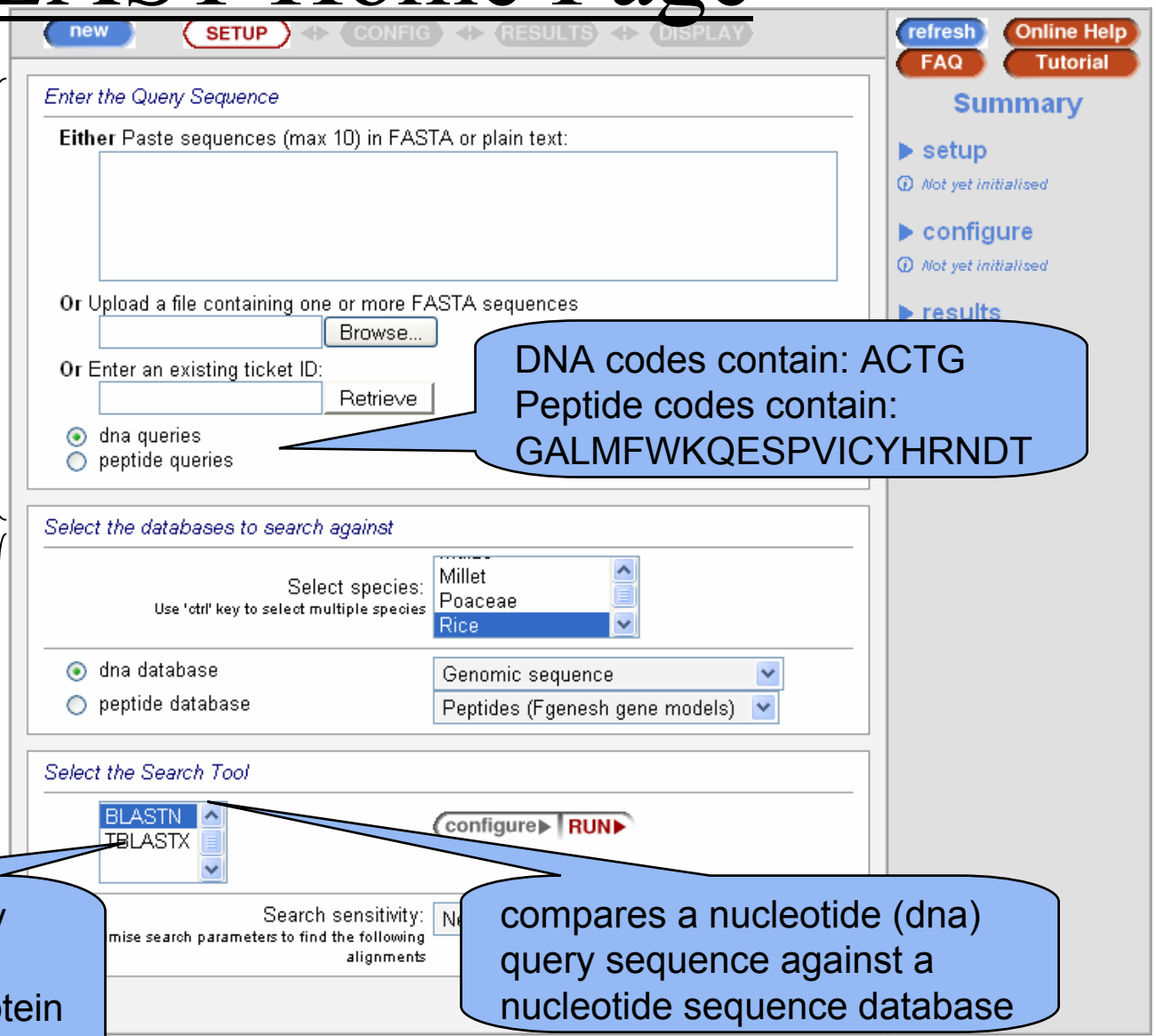
Notes or comments use this style font.

Gramene Home Page

BLAST Home Page

Sequence Information: Enter a sequence that you have and are trying to locate, or are trying to find a similar sequence for

Search Information: where would you like to search for similar sequences, using which tool?

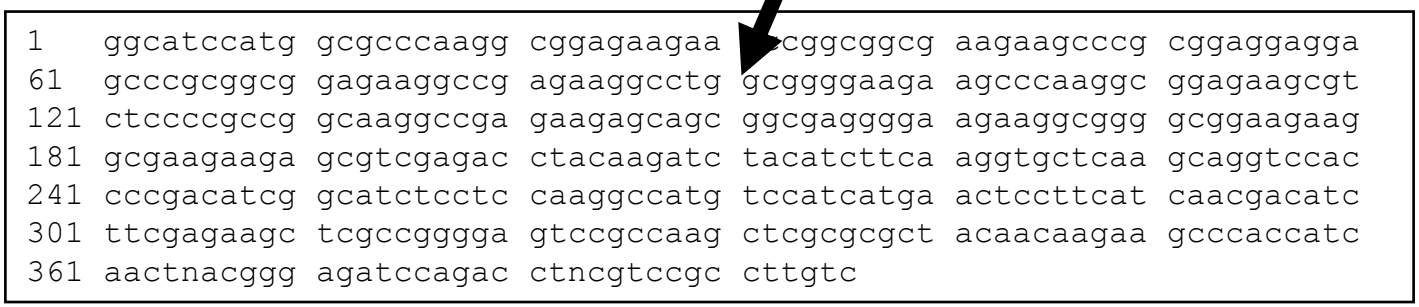


Example

BLAST - The Basic Local Alignment Search Tool – is used when searching for related (similar) sequences.

Problem: Finding a Nucleotide Match in the Rice Genome

You have an RFLP (restriction fragment length polymorphism) genetic marker with a **known hybridization sequence**, and want to know where it is located on the rice genome



```
1   ggcatccatg gcgcccaagg cggagaagaa gcggcggcg aagaagcccg cggaggagga
61  gcccgcggcg gagaaggccg agaaggcctg gcggggaaga agcccaaggc ggagaagcgt
121 ctccccgccg gcaaggccga gaagagcagc ggcgagggga agaaggcggg gcggaagaag
181 gcgaagaaga gcgtcgagac ctacaagatc tacatcttca aggtgctcaa gcaggtccac
241 cccgacatcg gcatctcctc caaggccatg tccatcatga actccttcat caacgacatc
301 ttcgagaagc tcgccgggga gtccgccaag ctcgcgcgct acaacaagaa gcccaccatc
361 aactnacggg agatccagac ctncgtccgc cttgtc
```

Step 1: Enter the Sequence

1. Paste in the **sequence**. You must **remove the sequence numbers** first!

Browse FAQ

Click for Help

You can search with up to 10 sequences at a time. Simply format them using FASTA format.

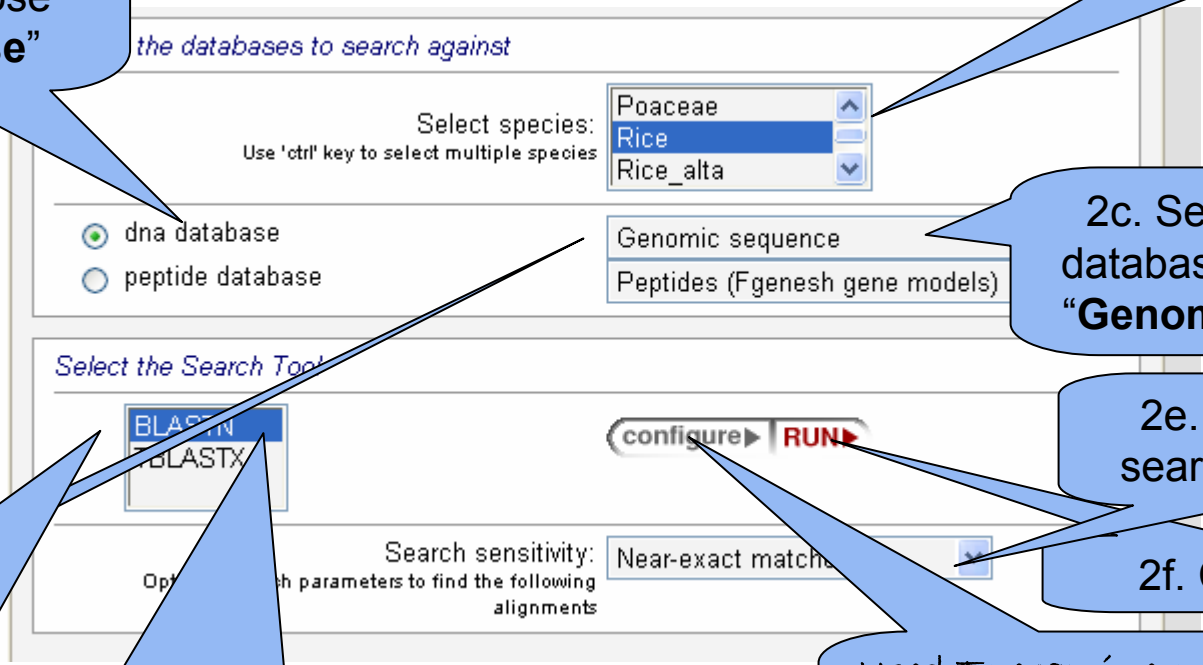
Alternatively, **save a sequence** to a file and use this box to upload it.

1a. Select "**dna queries**" because this is a nucleotide sequence

Step 2: Choose Target Database

2a. Select your target organism(s), in this case "**Rice**"

2b. Select your database. For this example, choose "**dna database**"



2c. Select a specific database. In this case, "**Genomic Sequence**"

2e. Select your search sensitivity

2f. Click **RUN**

The available set of databases and search tools will change according to species and type of database, and change/grow with each release

2d. Select your search too, "**BLASTN**" for this example

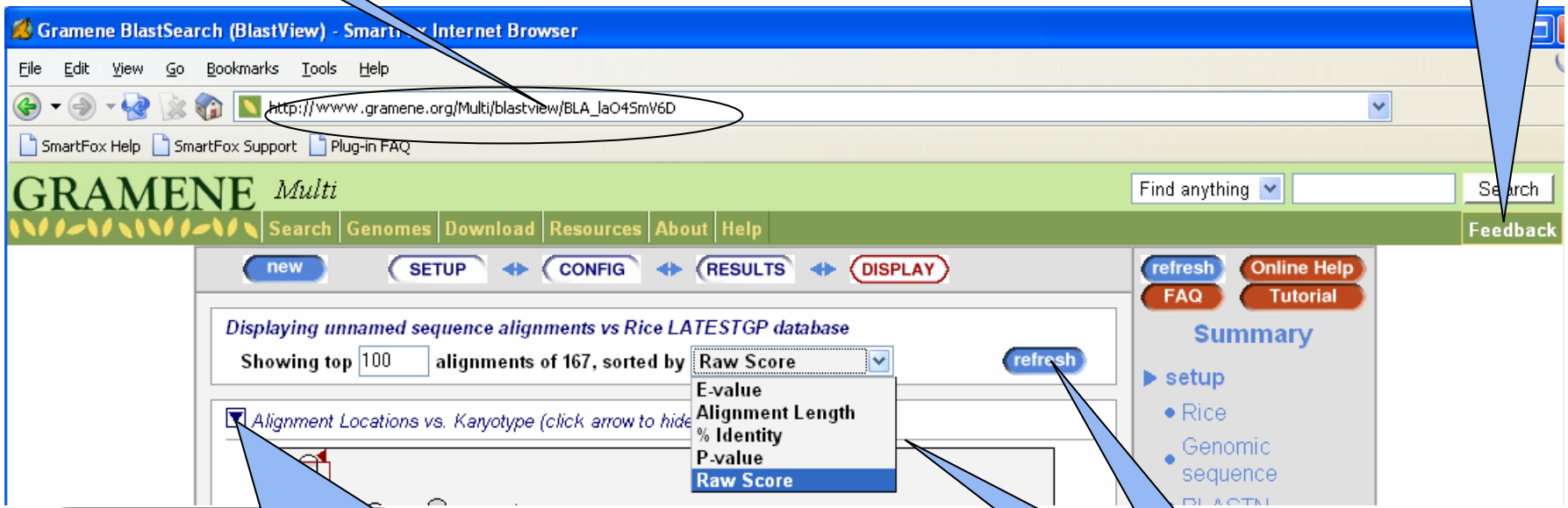
Used To examine or alter the optimized default parameters (see slide 16)

Use BLASTN for simple nucleotide against nucleotide searches
Use BLASTP for simple protein versus protein searches
Use TBLASTN for a protein query versus a DNA library search
Use BLASTX for DNA query versus protein library
Use TBLASTX for a translated DNA query against a translated DNA database

BLAST Queries

Ticket ID

Use Feedback to make enquiries to Gramene Staff about a search

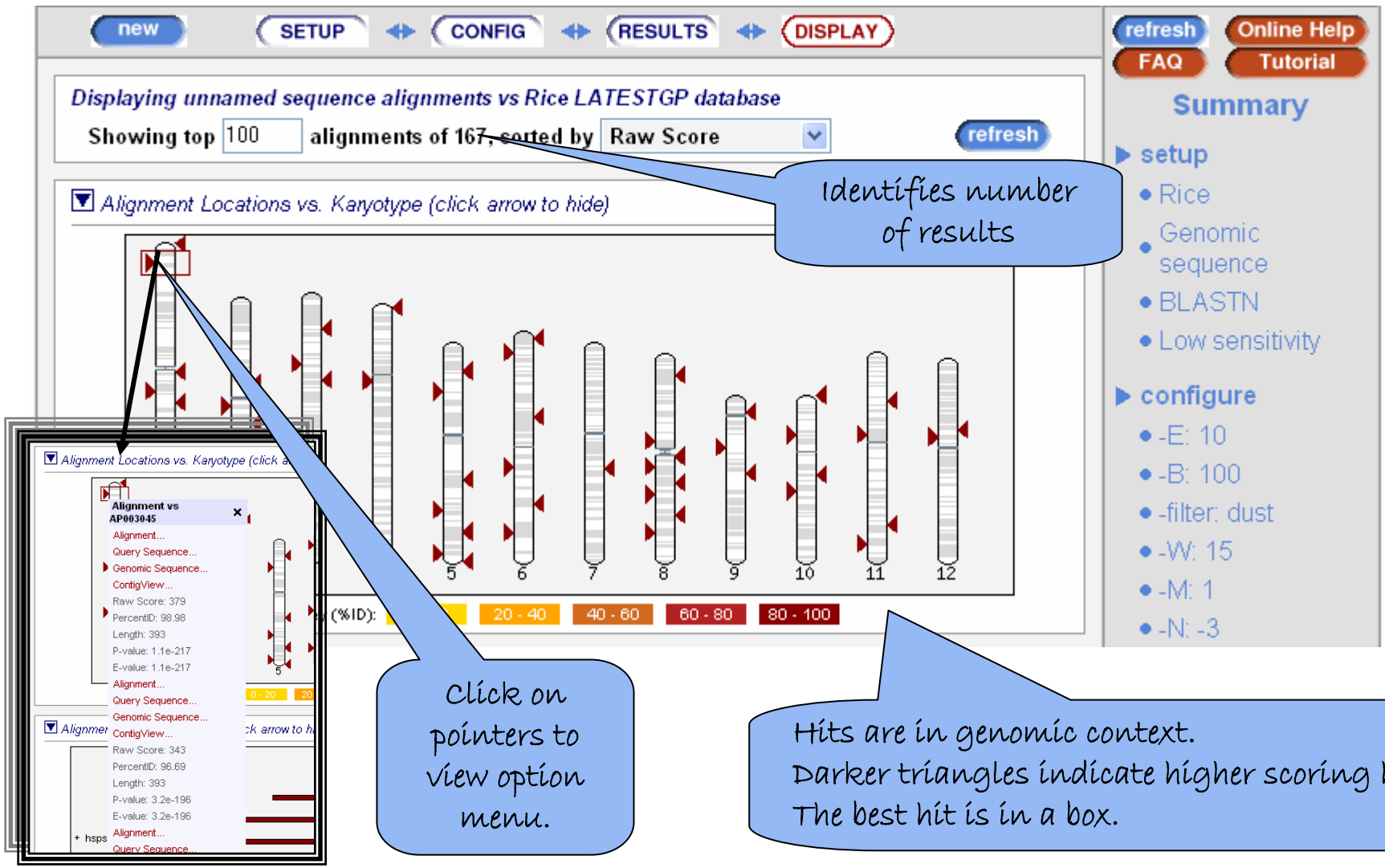


Click Arrows to hide or reveal results sections.

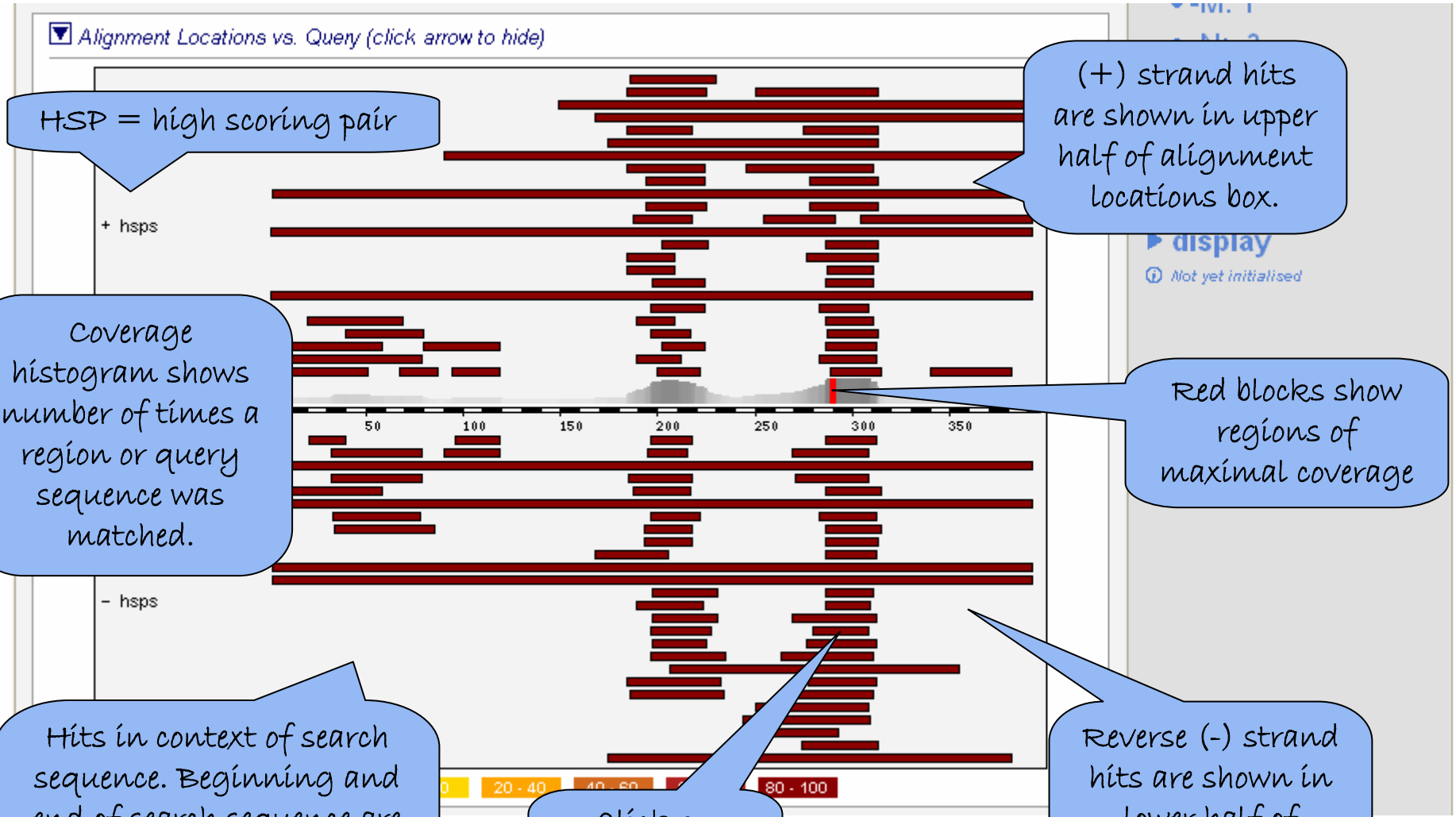
To sort results select an option and refresh

Results are stored on Gramene's server for one week, so that they can be accessed later with the **Ticket ID** or a bookmark to the results page.

Step 3: Results - Genomic Context



Step 3 Results - Search Sequence



HSP = high scoring pair

(+) strand hits are shown in upper half of alignment locations box.

Coverage histogram shows number of times a region or query sequence was matched.

Red blocks show regions of maximal coverage

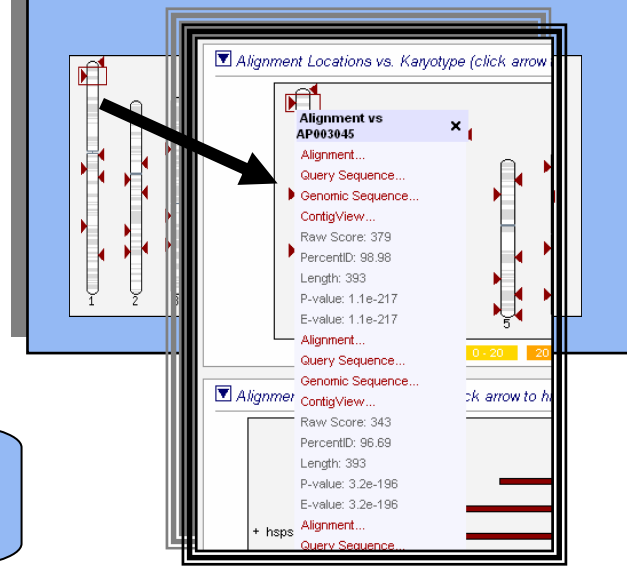
Hits in context of search sequence. Beginning and end of search sequence are at left and right ends. Darker colors are higher-scoring hits.

Click on blocks to view option menu.

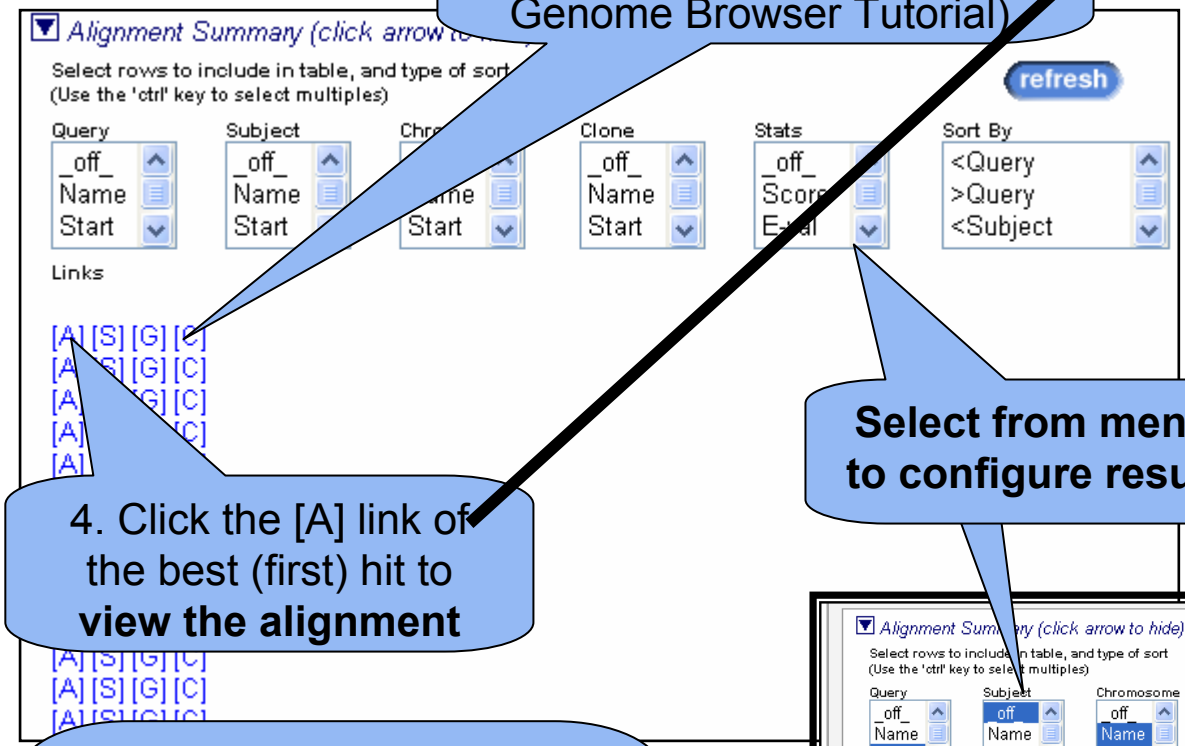
Reverse (-) strand hits are shown in lower half of alignment locations box.

Step 3: Results – Alignment Summary

Shortcut: You can also click on the best hit in the genome graphic (from Slide 10) to open a menu for that alignment.

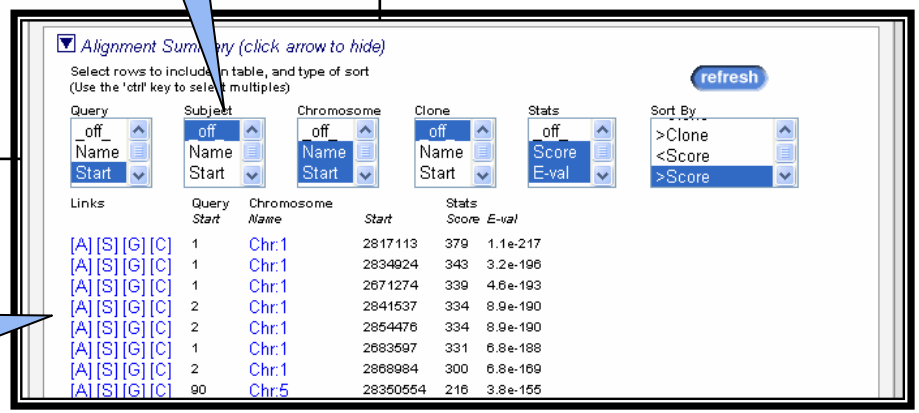


5. Click the [C] for "**contig" view** of the genome (see Genome Browser Tutorial)



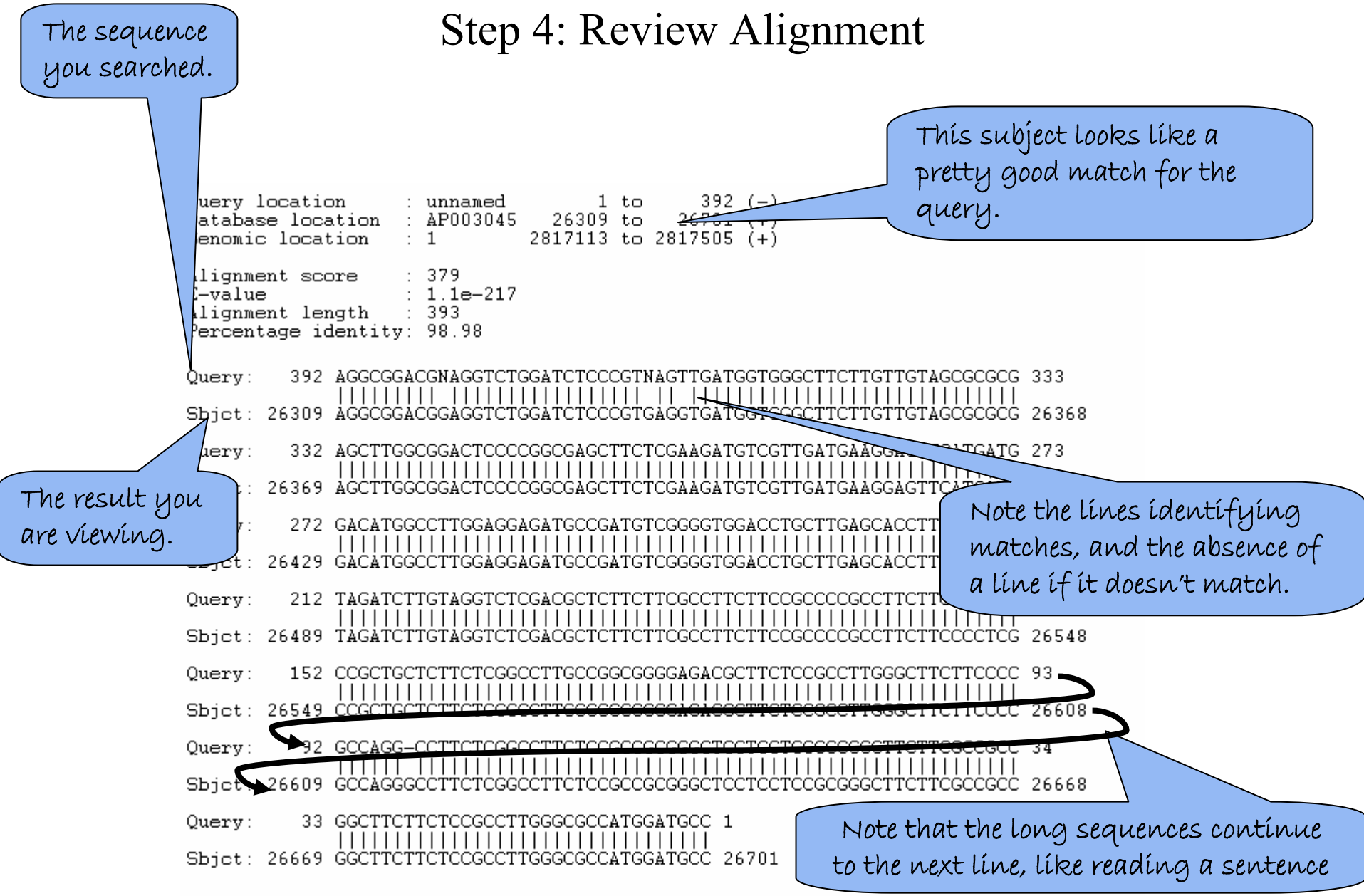
4. Click the [A] link of the best (first) hit to **view the alignment**

Select from menus to configure results



Click these links to view:
[A] alignment
[S] query sequence
[G] target (genome) sequence
[C] on genome ("ContigView")

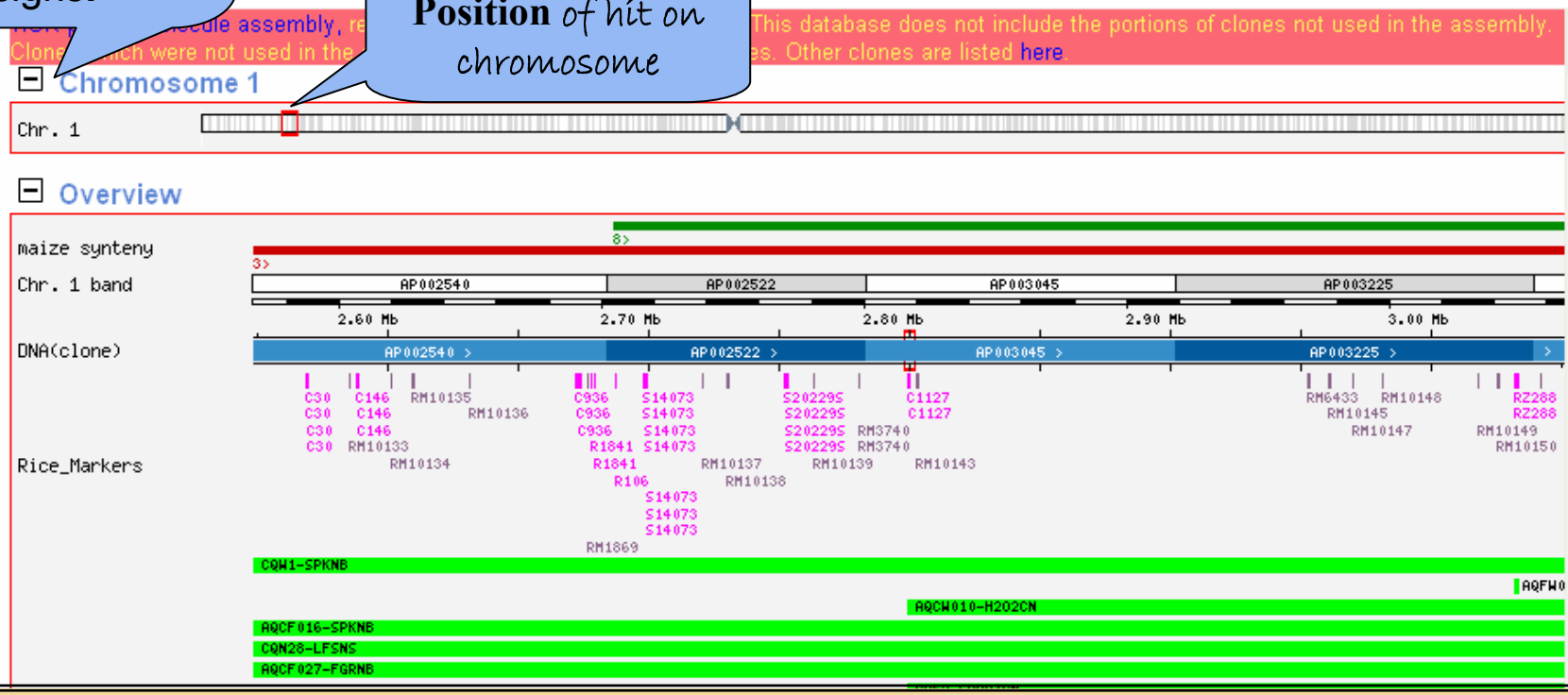
Step 4: Review Alignment



Step 5: Review Genome with Genome Browser

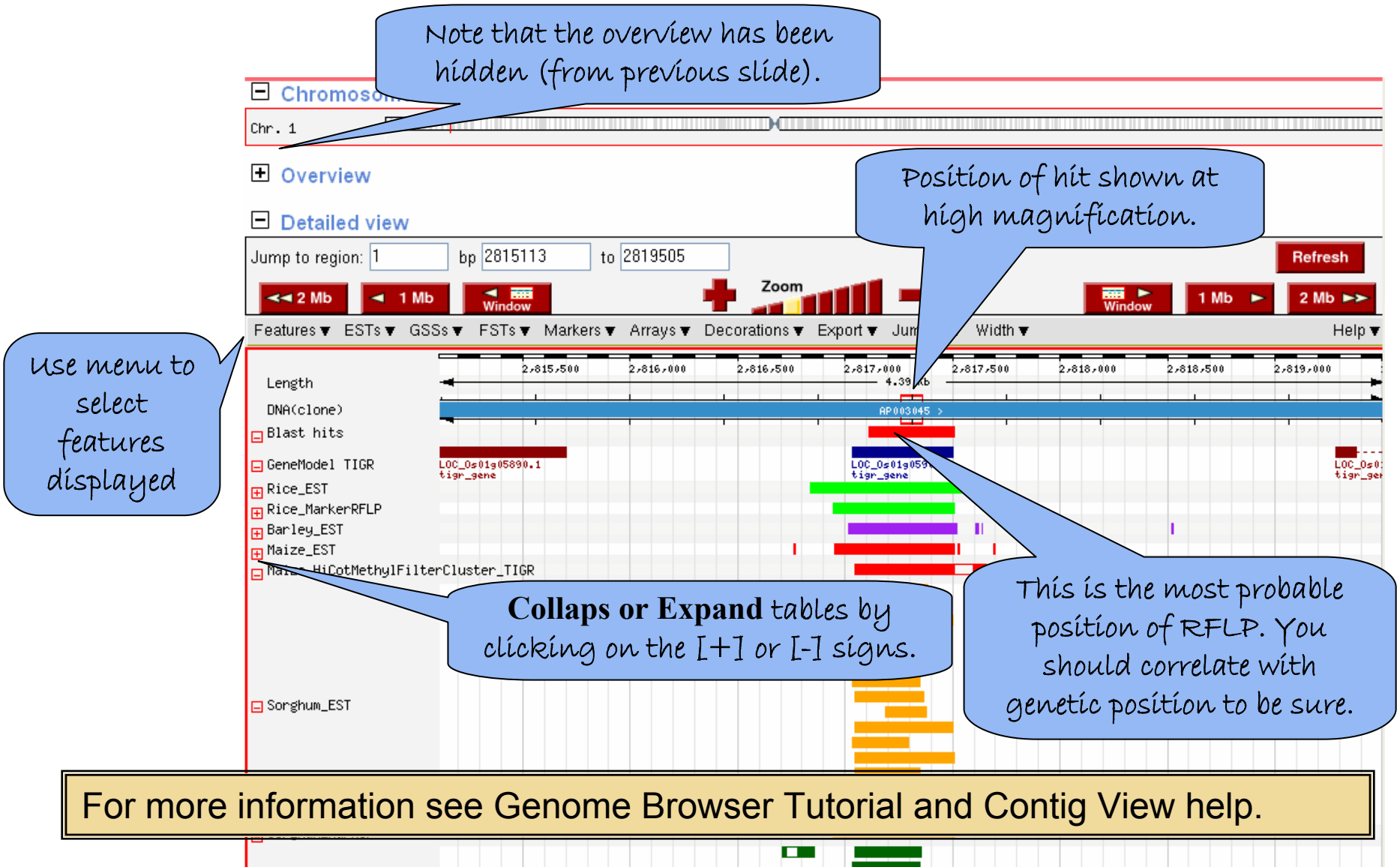
Show or Hid tables by clicking on the [+] or [-] signs.

Position of hit on chromosome

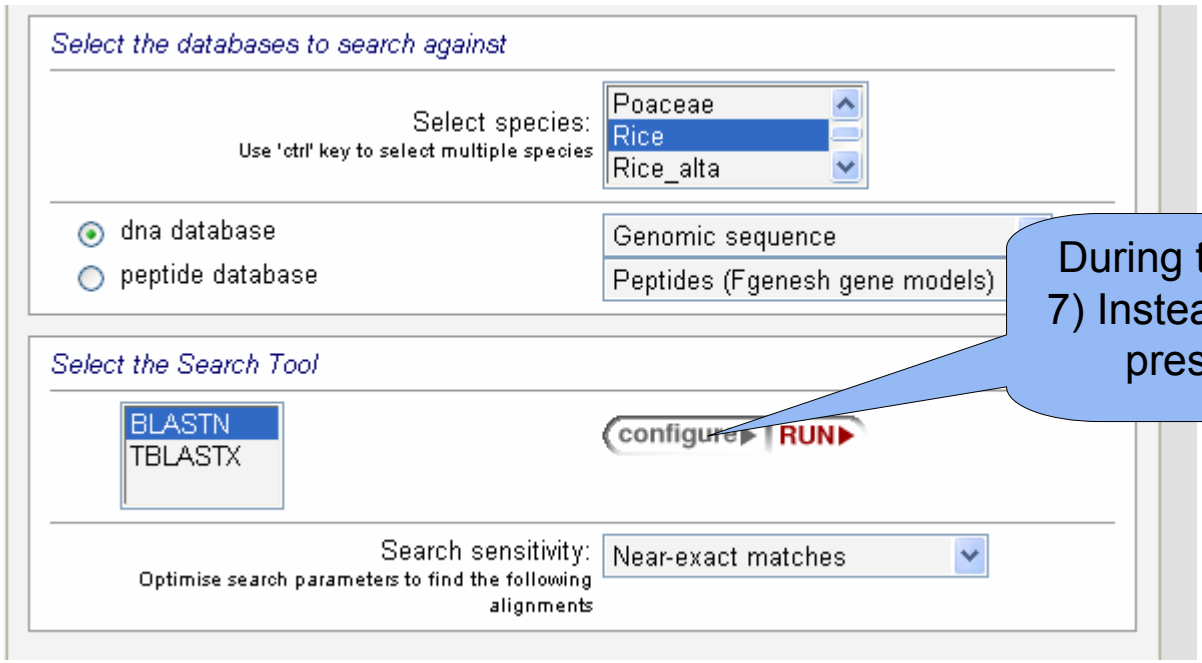


For more information see Genome Browser Tutorial and Contig View help.

Step 5: Review Best Hit on Genome

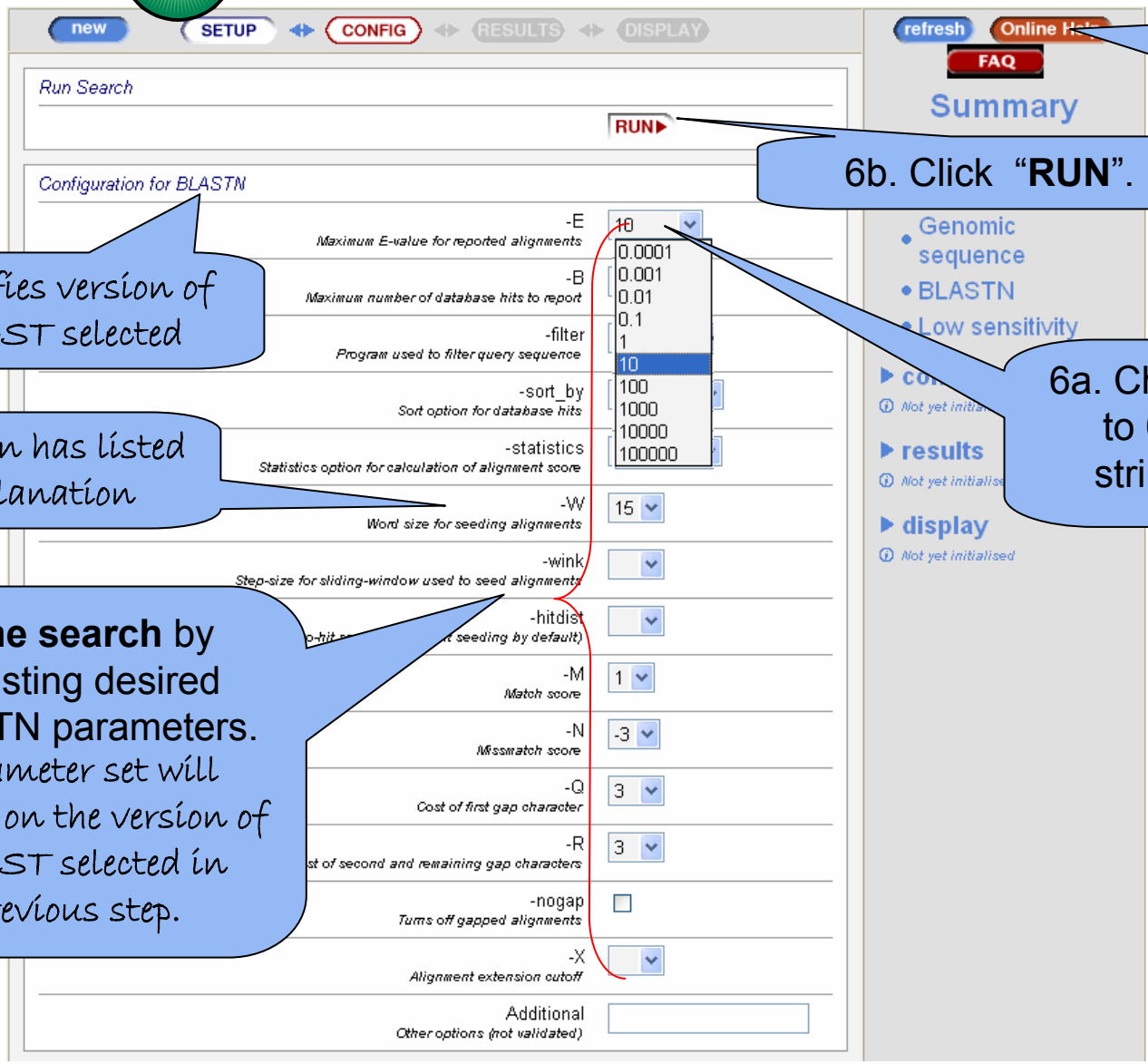


Step 6: Adjusting Settings



During the search (Slide 7) Instead of hitting RUN, press **configure**.

Step 6: Adjust Settings - Configure



See **help** for "config" assistance

6b. Click "**RUN**".

Identifies version of BLAST selected

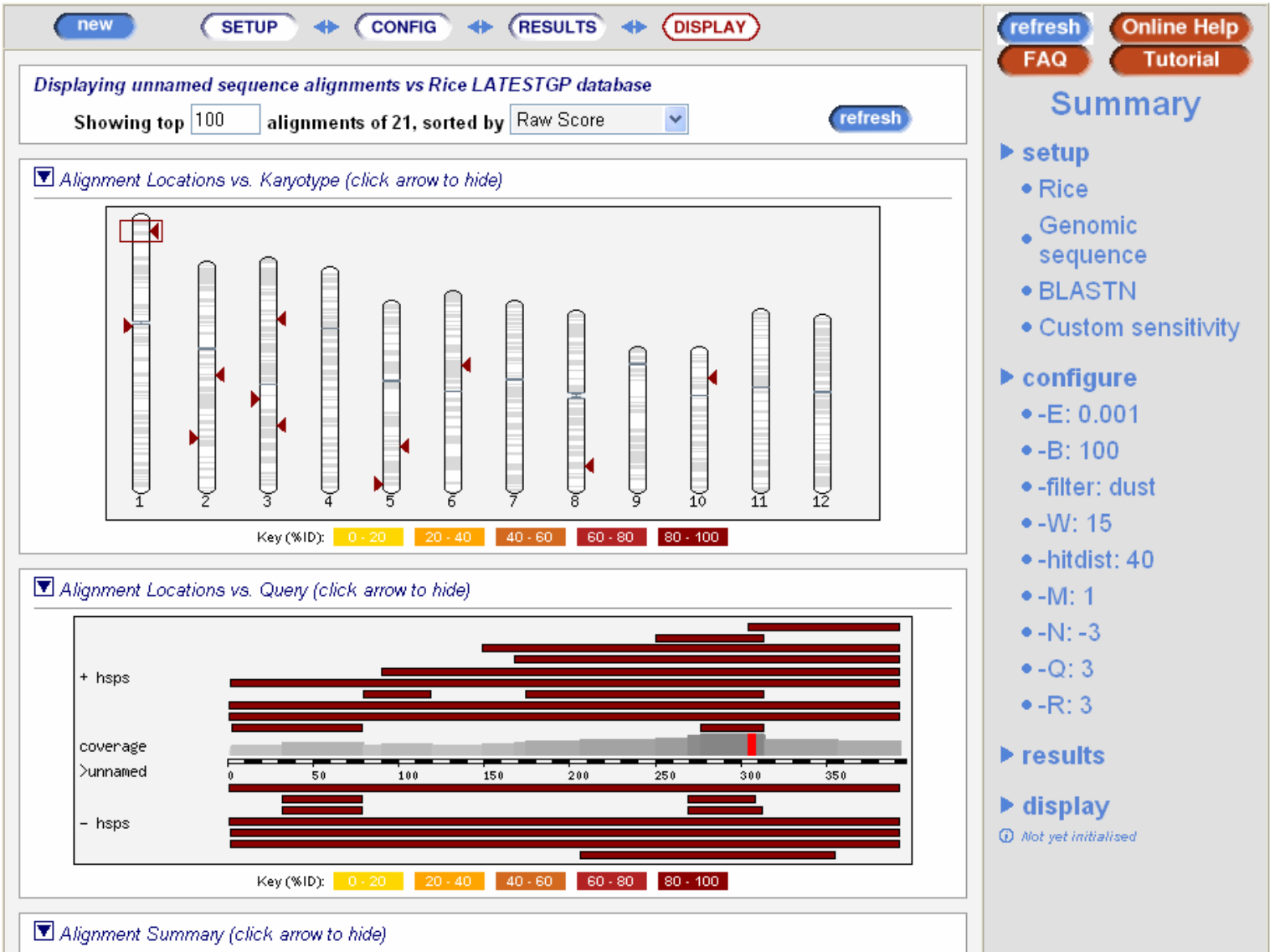
6a. Change E-value cutoff to 0.001 to increase stringency of search.

Each option has listed it's explanation

Tune search by adjusting desired BLASTN parameters. Parameter set will depend on the version of BLAST selected in previous step.

Click "CONFIG" if you want to change parameters & try again (repeat slide 17).

Step 7: Refined Results



urrent
rameters
n in the
panel.

More BLAST Information

Web sites that provide more BLAST information and advice on setting parameters include:

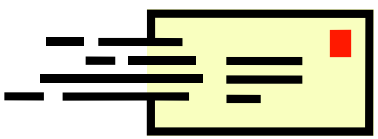
- the WU-BLAST web site
- EBI 2can introductions for protein or nucleotide BLAST, and
- NCBI-BLAST (*a good tutorial although focused on a slightly different implementation.*)

A book on the BLAST family of sequence similarity search algorithms (Joseph Bedell, Ian Korf and Mark Yandell, **BLAST,** O'Reilly & Associates, 2003, ISBN: 0-596-00299-8) provides profound theoretical background, as well as a protocol section covering common practical search problems. The pre-defined optimized parameter sets are based on recommendations in this book

Contact Gramene

Use the feedback button, located at the top of every page, to provide **feedback** or to **ask questions** about Gramene or your search needs.

or



Email Gramene at gramene@gramene.org