

BIOINFORMATICS INTERNSHIP

Session twenty three

RECAP

- BLAST
- Data give and output
- What is important

The screenshot shows a BLAST search result page. At the top, there are tabs for Descriptions, Graphic Summary, Alignments (selected), and Taxonomy. Below the tabs, the alignment view is set to 'Pairwise'. A dropdown menu for alignment view is open, showing options: Pairwise with dots for identities, Query-anchored with dots for identities, Query-anchored with letters for identities, Flat query-anchored with dots for identities, and Flat query-anchored with letters for identities. The search results show 34 sequences selected. The first result is 'PREDICTED: creatine kinase B-type isoform X1 [Erinaceus europaeus]' with sequence ID XP_016041590.1 and length 406. The alignment is displayed with a score of 758 bits (1956), an expect value of 0.0, and 100% identity. The alignment is shown in a table with columns for Query, Sbjct, and Score. The alignment is annotated with letters (1), a plus sign (2), a white space (3), and dashes (4). A 'Related Information' link is present. The second result is 'Homo sapiens chromosome 14, GRCh38.p12 Primary Assembly' with sequence ID NC_000014.9 and length 107043718. The alignment is displayed with a score of 383 bits (207), an expect value of 2e-103, and 100% identity. The alignment is shown in a table with columns for Query, Sbjct, and Score. The alignment is annotated with letters (1), a plus sign (2), a white space (3), and dashes (4). A 'Related Information' link is present.

A Alignment view dropdown menu

B Download button

C Matched database sequence

D Alignment display

E Related Information link

F Sort by dropdown menu

1 Letter match

2 Plus sign match

3 White space match

4 Dashed line match

Score 758 bits(1956) Expect 0.0 Method Compositional matrix adjust. Identities 369/406(91%) Positives 376/406(92%) Gaps 25/406(6%)

Query 1 MPFSNSHNALKLRFAEDEFDPDLSAHNNHMAKVLTPELYAE LRAKSTPSGFTLDDVIQTG 60

Sbjct 1 MPFSNSHNA KLRFAEDEFDPDLS HNNHMAKVLTPELYAE LRAKSTPSGFTLDDVIQTG 60

Query 61 VDNPGHPYINTVGC VAGDEESYEVFKDLFDPIIEDRHGGYKPSDEHKTDLPDNLQ-- 116

Sbjct 61 VDNPGHP+INTVG VAGDEESYEVFKDLFDPIIEDRHGGYKPSDEHKTDLPDNLQVRGE 120

Score 383 bits(207) Expect 2e-103 Identities 207/207(100%) Gaps 0/207(0%)

Query 116 VDNPGHPYINTVGC VAGDEESYEVFKDLFDPIIEDRHGGYKPSDEHKTDLPDNLQ-- 116

Sbjct 116 VDNPGHP+INTVG VAGDEESYEVFKDLFDPIIEDRHGGYKPSDEHKTDLPDNLQVRGE 120

Features: [creatine kinase B-type](#)
[creatine kinase B-type isoform X1](#)

Related Information
[Gene](#) - associated gene details
[Genome Data Viewer](#) - aligned

MOTIF FINDING – ITS BACK!

- Lets chat about our project goals

LET'S UNDERSTAND CODE

- Having shared my code:
 - please unscramble the comments so that they make sense
 - Remove comments that don't belong
 - Run code and find the single line feed (the bug in the code)
 - Remove/ solve