|  |  |  |
| --- | --- | --- |
|  |  |  |
| BLAST  Database Searches |  | Week 4 - Tuesday  **Activity A (de novo)**  **Activity B (portal tool)** |
|  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Overview Become familiar with the principles and operation of heuristic methods for determining the best matches between a query sequence and sequences in large databases Learning Objectives  * Adjusting parameters to meet the goals of each search. * Using threshold settings to control output * Masking sequences to avoid corruption from simple repeats and common domains * Using E values to understand whether hits are meaningful  Activity & Description (A) Student groups will conduct a thought experiment to determine the likely impact of various search parameters and output thresholds.  Groups will randomly choose a slip of paper with a parameter or threshold on it and will be provided with web sites URLs that will provide insight. After conducting research they will formulate a hypothesis concerning its impact on a database search. Activity & Description (B) Student groups will conduct a BLAST analysis of various aa queries containing different types of repeated elements (repeats, poly Q, TM domains etc.) and investigate the impact of leaving them in or removing them from the query. Students will also test the effect of changing their assigned parameter using this query. In a follow-up exercise, output thresholds will be adjusted to visualize situations in which hits are limited by program settings rather than by the number of potentially significant hits in the database. Feedback & Discussion  1. Prepare a table of program parameters and thresholds and student predictions. Instruct students to evaluate predictions in the context of Activity B. 2. Project BLAST analysis using assigned queries. Discuss the various ways filtered sequences are represented. Discuss how the output is affected by looking at sequence hits and E values. Adjust threshold settings and discuss the impact of restricting output. | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Logistics  |  |  |  | | --- | --- | --- | |  | A | B | | Group Size | 2 | 2 | | Group Formulation | Any | Any | | Group Structure | -- | -- | | Seating/Computers | 2 | 2 | | Challenge/Room | 5 | 5 | | Presentation | WB | Proj. |  Student Assessment  |  | | --- | |  |  Workshop Assessment  |  | | --- | |  | |   Links: |