Семестр 2, блок 2, практическая работа 1

**Базы последовательностей белков**

В работе изучается способ отыскания информации о белке YP\_004339259.1, синтезируемый бактерией Hippea maritima, штамм DSM10411, в различных базах данных.

**1.** С помощью системы uniprot, расположенной по адресу uniprot.com, найдены следующие идентификаторы записей о белке в разичных базах:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| База данных | Uniprot | UniRef50 | RefSeq | PDB |
| Идентификатор | F2LXL4\_HIPMA | UniRef50\_P33038? | WP\_013681244.1 NC\_015318.1YP\_004339259.1 | P330381NAW |

Идентификатор для UniRef50 не найден: опробованы способы поиска по сопоставлению с идентификаторами в других базах или с разными формами названия организма (“UniRef50\_P33038@ OR “F2LXL4\_HIPMA” и подобное) – результат либо семь с половиной тысяч позиций, либо 0.

**2.** С помощью сервиса “history – compare” проведено сравнение состояния записи о белке на моменты 2015-03-11 и 2011-05-31. Сервис выделил несколько изменений, среди которых:

Замена идентификатора белка:



Приведены альтернативные именования белка в различных номенклатурах:



Добавлена информация о других базах данных, содержащих записи с этим белком:



**3.** Для сравнения были выбраны S7THW2 и Y1I1W6, особых различий между ними замечено не было, разве что молекулярная масса белков разная и у обоих оппонентов моего белка составляет 44,824 против 45,027 (см. Excel-таблицу). Остальные поля из категорий “function” и “sequences” не содержали либо различий, либо записей.

**4.** В записях по своему белку требуемые примеры найдены не были. Поэтому пользуюсь примером из раздела «помощь» для сульфидных связей белка-ингибитора альфа-амилазы WDAI-3, кодируемого геном IHA-B1-2 организма Triticum aestivum (пшеница мягкая). Связи в записи представлены следующим образом:

