Дата последнего изменения: 27 декабря 2013 г.

**Отчет по первому блоку занятий**

**Раздел 1. Вирус**

В таблице 1 представлена информация обо всех гена вируса *Bovine astrovirus B170/HK*.

Таблица 1. **Гены вируса *Bovine astrovirus B170/HK* (идентификатор записи EMBL – HQ916314).** В таблице выделены цветом строчки, отвечающие самому длинному гену и самому короткому гену

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **№** | **Координаты начала** | **Координаты конца** | **Направление** | **Длина** |
| 1 | 23 | 3933 | Прямая цепь | 3911 |
| 2 | 23 | 2470 | Прямая цепь | 2448 |
| 3 | 3878 | 6235 | Прямая цепь | 2358 |

Белок строится из некоторой последовательности аминокислот, а каждая аминокислота кодируется в геноме тремя последовательными основаниями (триплетом). Так как каждый из данных участков является геном, кодирующим белок, то число оснований в нем кратно трем. Действительно, 2448, 2358 – длины генов делятся на три.

Для двух сопряженных участков цепи с координатами (23..2425) и (2425..3933) в файле приводится одно описание. Следовательно, эти два участка кодируют один белок, поэтому считаются одним геном. Также, в файле приводится отдельное описание гена, координаты которого пересекаются с координатами предыдущего гена ((23…2425)и(2425…3933) и (23…2470)). То есть, существуют некоторые варианты считывания с одного и того же участка. В первом случае, рибосома читает со старт-кодона 23-25 до кодона 2423-2425, потом откатывается на одну пару нуклеотидов назад, читает кодон 2425-2427 и далее – до стоп-кодона 3931-3933. Во втором случае, рибосома читает со старт-кодона 23-25 без отката до стоп-кодона 2468-2470. Скорее всего, что таким образом получаются различные модификации одного белка.

Длину генома можно посмотреть в поле ID. Число, после которого стоят буквы «BP», и есть искомая величина. Таким образом, длина генома данного вируса - 6317 BP (base pairs), то есть 6317 пар оснований.

Расположение генов этого вируса представлено на рисунке 1. Эти гены имеют общие участки, Крупнее участок перекрытия одной из пары генов изображен на рисунке 2.

Общая часть начинается с 3878 нуклеотида и заканчивается 3933 нуклеотидом. Эти гены имеют разные рамки считывания. Подробнее этот участок изображен на рисунке 2.

Рисунок 1. **Расположение генов *Bovine astrovirus B170/HK* по данным NCBI. Рисунок получен с помощью геномного браузера сайта NCBI:** [**http://www.ncbi.nlm.nih.gov/**](http://www.ncbi.nlm.nih.gov/)

**

Рисунок 2. **Участок генома *Bovine astrovirus B170/HK* с 2 перекрещенными генами. Рисунок получен с помощью геномного браузера сайта NCBI:** [**http://www.ncbi.nlm.nih.gov/**](http://www.ncbi.nlm.nih.gov/)

**Раздел 2. Прокариотический организм**

Выданный мне организм относится к роду *Thermanaerovibrio*, виду *acidaminovorans* и штамму *DSM 6589*. Идентификатор записи генома данного организма: NC\_013522. Этот прокариот относится к царству бактерий.

Геном этой бактерии содержит 1 848 474 пары нуклеотидов. Всего в нем 1 821 ген, и 1 730 из них кодируют белки.

АТФ-синтаза - комплекс из нескольких белковых субъединиц. Вместе они выполняют функцию синтеза АТФ из АДФ и неорганического фосфата. Упрощенно, это происходит за счет пропускания положительно заряженных ионов через мембрану по градиенту концентрации, еще используется разности потенциалов на мембране. Так как белки, входящие в АТФ-синтазу, необходимы клетке вместе, то целесообразно располагать их рядом в геноме. Поэтому, такие белки образуют оперон.

У *Thermanaerovibrio* *acidaminovorans DSM* в такой оперон входит 9 генов. Они сходны по функциям (кроме одного, продукт которого не представлен в базе данных NCBI), расположены довольно близко на одной (комплементарной цепи). Я не могу сказать точно, к какому типу относится АТФ-синтаза этой бактерии, так как в продуктах упоминает V-тип, а также F0 – синтаза. Но АТФ-синтаза типа VOV1 встречается преимущественно у архей и в вакуолях растений. Поэтому, скорее всего она относится к типу FOF1. Расположение генов в этом опероне представлено на рисунке 3. Краткая характеристика генов этого оперона – в таблице 2.



Рисунок 3. **Оперон генов, кодирующих АТФ-синтазу.** Рисунок получен при помощи базы данных NCBI

Таблица 2. **Характеристика оперона АТФ-синтазы.** Гены представлены в порядке 5’ - 3’, при этом координаты уменьшаются, так как оперон расположен на обратной цепи

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № | Координаты гена | Идентификатор белка | Продукт гена по данным NCBI | Перевод названия продукта гена |
| 1 | 552 528..553 895 | YP\_003317034.1 | H(+)-transporting two-sector ATPase |  |
| 2 | 553 902..555 686 | YP\_003317035.1 | H(+)-transporting two-sector ATPase |  |
| 3 | 555 699..556 310 | YP\_003317036.1 | hypothetical protein |  |
| 4 | 556 307..556 606 | YP\_003317037.1 | vacuolar H+transporting two-sector ATPase F subunit |  |
| 5 | 556 623..557 042 | YP\_003317038.1 | H+transporting two-sector ATPase subunit C |  |
| 6 | 557 078..559 009 | YP\_003317039.1 | V-type ATPase 116 kDa subunit |  |
| 7 | 559 000..560 091 | YP\_003317040.1 | H+transporting two-sector ATPase subunit C |  |
| 8 | 560 088..560 432 | YP\_003317041.1 | ATP synthase F0 subunit B |  |
| 9 | 560 429..561 055 | YP\_003317042.1 | V-type ATPase subunit D |  |

**Раздел 3. Белок**

Другой интересный ген этой бактерии - ген родового белка второй стадии споруляции (stage II sporulation E family protein). Краткая информация о нем представлена в таблице 3.

Таблица 3. **Основные свойства гена, кодирующего белок второй стадии споруляции *(stage II sporulation E family protein)* бактерии *Thermanaerovibrio* *acidaminovorans DSM*** **(индентификатор YP\_003318195.1)**

|  |  |  |
| --- | --- | --- |
| Свойство | Квалификатор в записи генома | Значение |
| Локус гена в геноме | /locus\_tag | Taci\_1699 |
| Имя гена | /gene | - |
| Идентификатор в базе Gene | /db\_xref | GeneID:[8631572](http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene&cmd=Retrieve&dopt=full_report&list_uids=8631572) |
| Начало в геноме | - | 1780257 |
| Конец в геноме | - | 1780883 |
| Цепь | - | Обратная |
| Длина гена (в парах нуклеотидов) | - | 627 |
| Идентификатор белка в базе данных RefSeq | /protein\_id | YP\_003318195.1 |
| Длина белка (количество аминокислот) | - | 208 |

Близлежащее окружение гена, кодирующего белок *stage II sporulation E family protein* из организма *Thermanaerovibrio acidaminovorans DSM* приведено на рисунке 4. Все гены, кроме правого на комплементарной цепи, правый - на прямой. Считая с лева на право, первый и второй гены пересекаются (длина общего участка 4 нуклеотида), как и второй с третьим (общий участок такой же длины). Следующие не перекрываются и расстояния между ними 56 127 нуклеотидов. Их рамки считывания относительно гена, кодирующего белок *stage II sporulation E family protein* соответственно: +2, +3, +1 (исходный ген), +2, -1. Функции белков, кодируемых соседними генами довольно, сильно отличаются, поэтому они не могут образовывать оперон.

Рисунок 4. **Близлежащее окружение гена, кодирующего белок второй стадии споруляции *(stage II sporulation E family protein)* бактерии *Thermanaerovibrio* *acidaminovorans DSM***. Исходный ген в рамочке. Рисунок получен с помощью геномного браузера сайта NCBI: <http://www.ncbi.nlm.nih.gov/>