

Практикум 10. PFAM.

В рамках практикума для изучения я выбрал семейство белков **PF00231** (ID: ATP-synt, name: ATP synthase) – АТФ-синтазы.

F-АТФазы (также известные как АТФ-синтазы, F1F0-АТФазы или H(+)-транспортирующие двухсекторные АТФазы) (7.1.2.2) состоят из двух связанных комплексов: F1, который является каталитическим ядром и F0, который представляет собой встроенный в мембрану протонный канал. Оба комплекса (F1 и F0) являются роторными двигателями. Два ротора вращаются в противоположных направлениях, но ротор F0 обычно сильнее: используя энергию протонного градиента, он заставляет ротор F1 вращаться в обратном направлении, что запускает синтез АТФ. У бактерий эти АТФазы могут работать и в обратную сторону, гидролизуя (расщепляя) АТФ для создания протонного градиента.

В Pfam seed выравнивание содержит 657 последовательностей, full – 13,978.

В семействе АТФ-синтаз было выделено (по доменным архитектурам) подсемейство со следующей доменной архитектурой: **PF12349 - PF00231** (рис. 1):

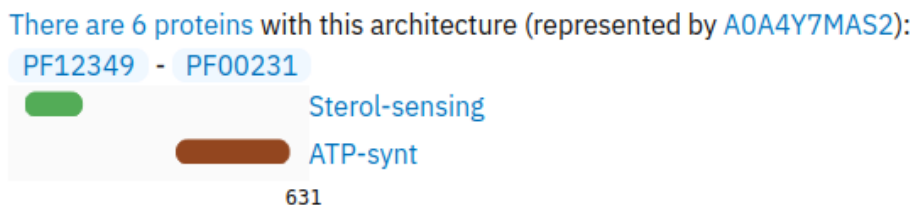


Рис.1. Выбранная доменная архитектура подсемейства белков, выделенного из семейства **PF00231**.

Последовательности 6 белков, входящих в данное подсемейство, были скачаны и выровнены при помощи программы MAFFT с параметрами по умолчанию.

В полученном выравнивании был выделен домен Sterol-sensing (позиции 1-352) (Рис.2):

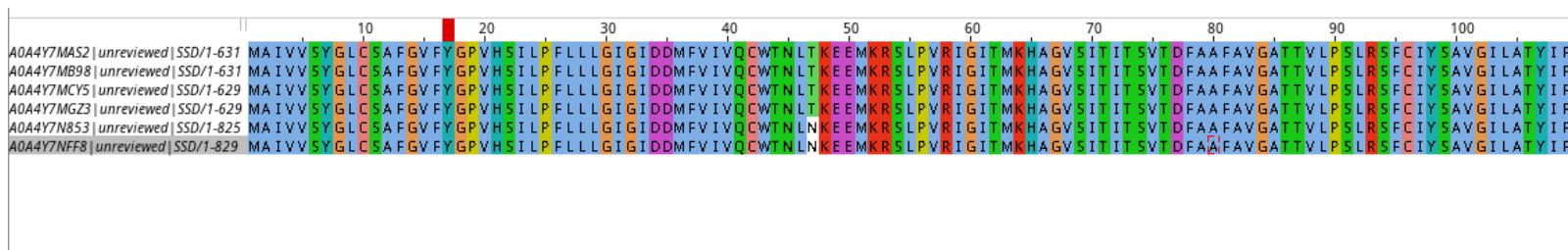


Рис.2. Часть выравнивания домена Sterol-sensing.

Белки с этим доменом – встроенные в мембрану транскрипционные факторы, которые увеличивают синтез липидов в животных клетках.

НММ-профиль

С использованием программы hmmbuild по выравниванию был построен НММ-профиль домена Sterol-sensing:

```
hmmbuild --amino sterol.out domal.fa
```

В семействе в PFAM 24к белков, поэтому ограничимся только последовательностями членистоногих (3222 штук). Все белки из выделенного ранее подсемейства – дафний.

При помощи программы hmmsearch был осуществлен поиск ранее построенным профилем по этим белкам:

```
hmmsearch -o search.out sterol.out sterol.fasta
```

Было найдено 2690 белков (из 3222 у членистоногих).

Из подсемейства, выделенного ранее, было найдены все белки (6 из 6).

Вес находок для них варьировался от 832.2 до 823.7.

Количество находок, которые имеют вес $\geq 823.7 - 6$. Порогом для выделения подсемейства я возьму 812: разрыв между min.score для подсемейства и max.score не в подсемействе: 823.7 vs 800.3.

Таблица 1. Результаты выделения подсемейства профилем по выбранному порогу

	Positive	Negative
True	6	2684
False	0	0