

## Тема работы:

Молекулярная эволюция белков, содержащих домен фосфолипазы A2, у плоских червей

## Состав коллектива:

1. Бочарникова Мария Евгеньевна, ФЕН НГУ (гр. 17410), почта: [m.bocharnikova@g.nsu.ru](mailto:m.bocharnikova@g.nsu.ru)
2. Афонников Дмитрий Аркадьевич, к.б.н., в.н.с. эволюционной биоинформатики и теоретической генетики, почта: [ada@bionet.nsc.ru](mailto:ada@bionet.nsc.ru)

## Научное содержание работы:

### 1. Постановка задачи:

Фосфолипазы — группа ферментов (класс гидролазы), которые катализируют процесс гидролиза фосфолипидов. В зависимости от положения гидролизуемой связи в фосфолипиде различают 4 основных класса фосфолипаз: А, В, С и D. Фосфолипазы A2 (КФ 3.1.1.4), отщепляют SN-2 ацильную цепь. Эти белки хорошо изучены у змей, а также широко представлены у млекопитающих и выполняют ряд важнейших базовых функций на уровне клеток и всего организма. Нарушение функций фосфолипаз A2 у человека часто ассоциировано с раком.

Целью данной работы является изучение особенностей структуры, функций и эволюции белков, содержащих фосфолипазный домен A2 у плоских червей. Для достижения цели поставлены следующие задачи:

- Идентификация последовательностей ФА2 у плоских червей и родственных организмов;
- Реконструкция филогении на основе последовательностей доменов ФА2;
- Анализ первичной структуры белков и предсказание их функции на основе биоинформатического анализа.

### 2. Современное состояние проблемы (на момент начала работы).

Паразитические плоские черви представляют угрозу здоровью населения в Сибири. В частности, в Западной Сибири зарегистрирована высокая заболеваемость описторхозом. Это заболевание может вызывать серьезные осложнения в организме хозяина (желчекаменную болезнь или рак). Механизмы патогенного воздействия паразитов на организм хозяина еще слабо изучены. Предполагают, что они могут быть связаны с индуцированием воспалительных процессов в тканях хозяина, механическим разрушением эпителиальных клеток хозяина и токсическим воздействием белков, секретлируемых паразитом. В этой связи изучение, фосфолипаз у паразитических червей представляет большой интерес. Необходимо отметить, что в отличие от змеиных ФА2 позвоночных, у плоских червей гомологи ФА2 изучены плохо. Неизвестны ни типы фосфолипаз, представленные в этих организмах, ни их структура и функции, а также их отношение к известным белкам этого большого семейства.

### 3. Подробное описание работы, включая используемые алгоритмы

Биоинформатический анализ данных включал следующие этапы (схема 1):

1. Выбор последовательностей фосфолипаз A2 позвоночных для поиска их ортологов у плоских червей. Для этого мы использовали работу Хуанга и соавторов [Huang, 2015]. В этой работе представлены фосфолипазы нескольких классов: G7, G6, G5, G4, G3, G16, G15, G12, G10 для 11 организмов. Для нашего анализа мы отобрали белки человека (*Homo sapiens*), коровы (*Bos taurus*), мыши (*Mus musculus*), крысы (*Rattus norvegicus*), лягушки (*Xenopus laevis*), курицы (*Gallus gallus*), рыбы (*Danio rerio*).
2. Выбор геномов плоских червей и моллюсков. При решении этой задачи мы воспользовались базами данных Wormbase Parasite (<https://parasite.wormbase.org/index.html>), которая содержит полногеномные последовательности, аминокислотные последовательности белков и их аннотации.
3. В качестве внешней группы для филогенетического анализа были взяты белковые последовательности трех моллюсков из базы данных MolluskDB (<https://ensembl.molluscdb.org/index.html>). Это тихоокеанская устрица (*Crassostrea gigas*), морское блюдечко (*Lottia gigantea*) и филиппинская мидия (*Modiolus philippinarum*).
4. Для выделения ортологических групп была использована программа Orthofinder. Она использует в качестве входных данных последовательности белков различных организмов, а на выходе их классификацию по ортогруппам и, дополнительно, реконструированное филогенетическое дерево видов (рисунок 1). В дополнение к аминокислотным последовательностям геномов плоских червей и моллюсков мы подавали на вход программе, дополнительно, последовательности фосфолипаз из работы Хуанга и соавт. [Huang, 2015]. Это было сделано для того, чтобы после разбиения белков по ортогруппам идентифицировать семейства фосфолипаз червей и моллюсков по присутствию в этих ортогруппах классифицированных белков позвоночных.
5. Для фосфолипаз группы 3 мы провели дополнительный филогенетический анализ. Для этого при помощи программы hmmscan пакета HMMER идентифицировали в них домены Phospholipase A2 (идентификаторы PF00068.20, PF05826.13), Group XII secretory phospholipase A2 precursor (PF06951.12). Последовательности доменов включая 3 терминальных остатка были выделены из белков и выравнены программой Mafft. Филогенетическое дерево по полученному выравниванию было реконструировано программой IQ-TREE. Визуализация филогенетического дерева осуществлялась при помощи программы arhaeopterix (<https://sites.google.com/site/cmzmasek/home/software/archaeopterix>).

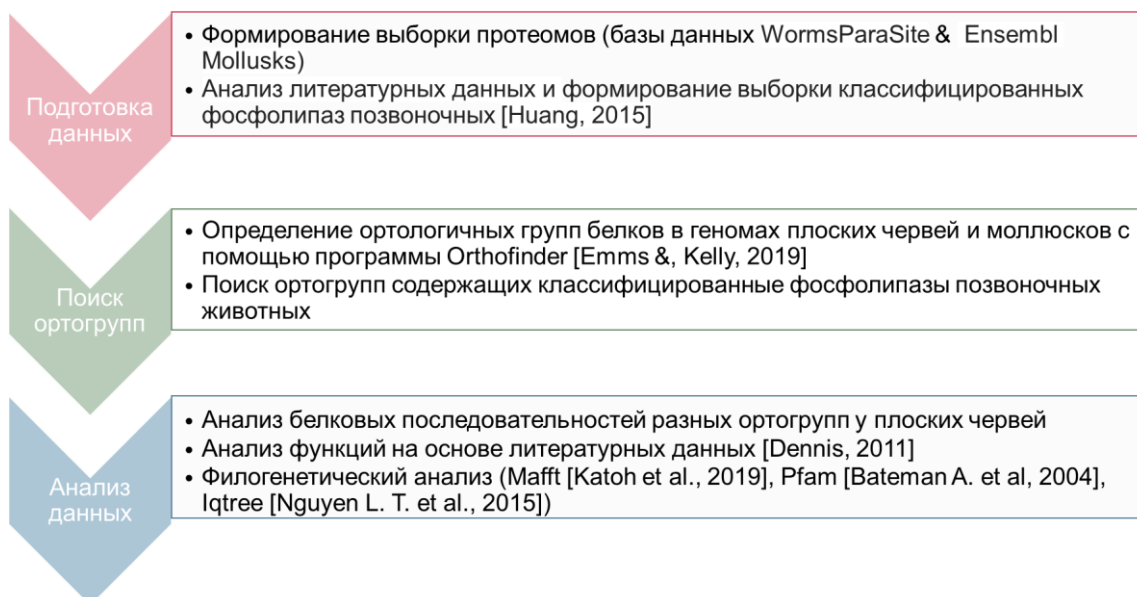


Схема 1 – Этапы методики для поиска гомологов ФА2 у плоских червей

#### 4. Полученные результаты.

Анализ геномов моллюсков и плоских червей при помощи программы Orthofinder позволил выделить 43221 ортогрупп, из которых 8 содержат ФА2 позвоночных: 3 группы представлены у всех организмов, 4 только у моллюсков и свободноживущих червей, 1 группа широко представлена у трематод (особенно интересна, так как данная ФА2 млекопитающих связана с канцерогенезом, поэтому на основе выровненных последовательностей и литературных данных мы построили схему доменов ФА2 плоских червей – рисунок 3). Одна из групп фосфолипаз позвоночных (G2C) не была обнаружена ни у моллюсков, ни у плоских червей.

Как дополнительный результат, программа Orthofinder реконструирует филогенетическое дерево видов, которое представлено для плоских червей и моллюсков на рисунке 2. Структура дерева хорошо согласуется с современными представлениями об эволюции плоских червей. Моллюски отделяются в отдельную кладу, которая является внешней, по отношению к плоским червям.

#### 5. Иллюстрации, визуализация результатов.

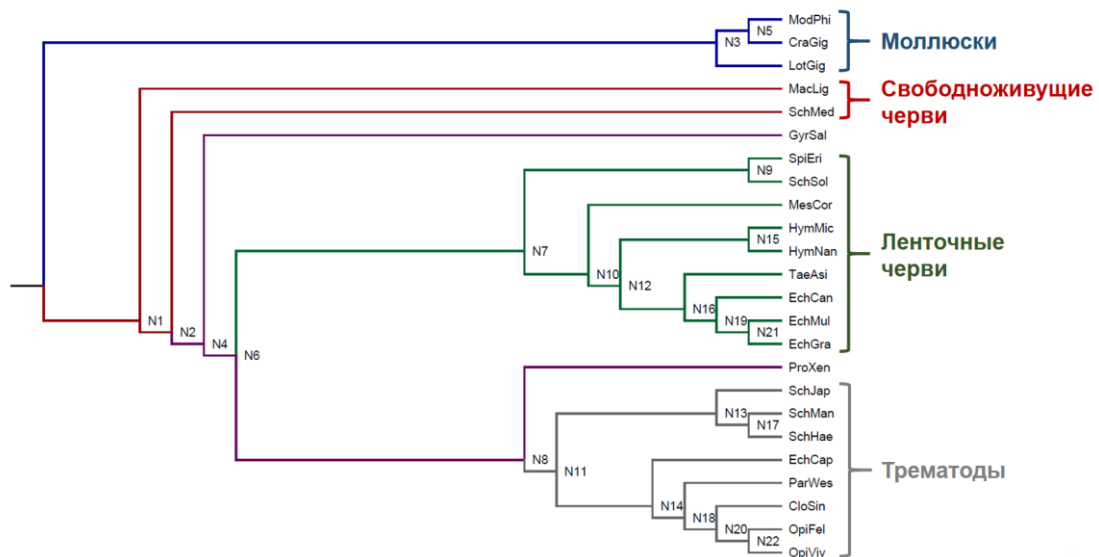


Рис. 1 – Филогенетическое дерево видов, реконструированное программой Orthofinder

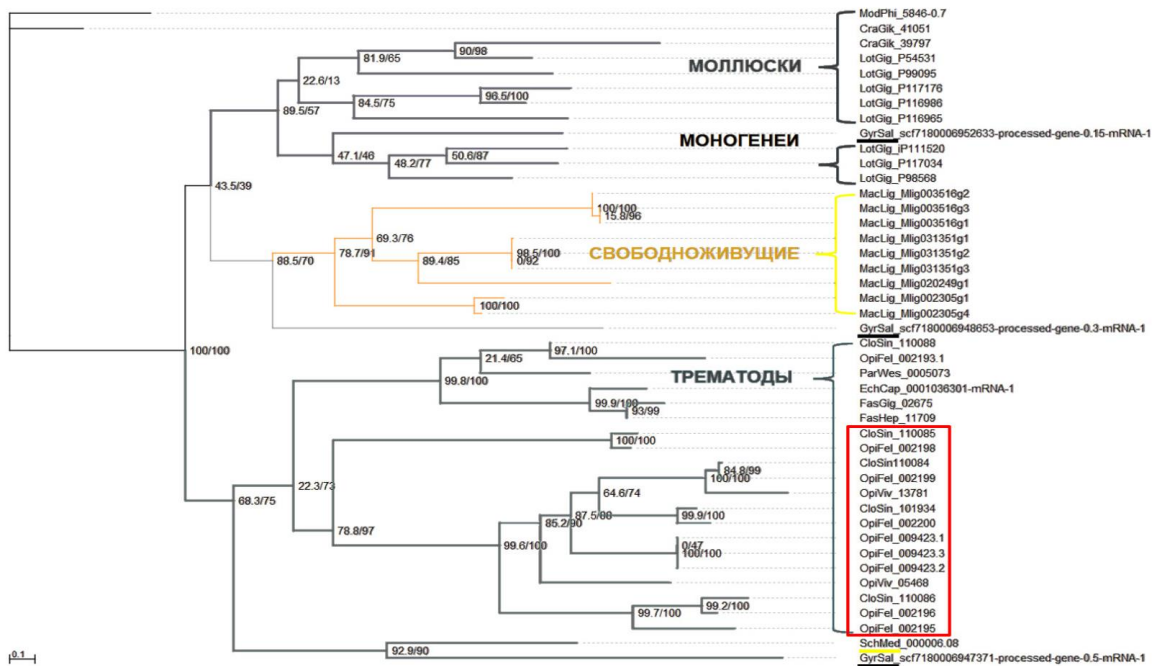


Рис. 2 - Филогенетическое дерево у ФА2 группы 3. Многократно дублицированные гены фосфолипаз трематод обведены красным прямоугольником.

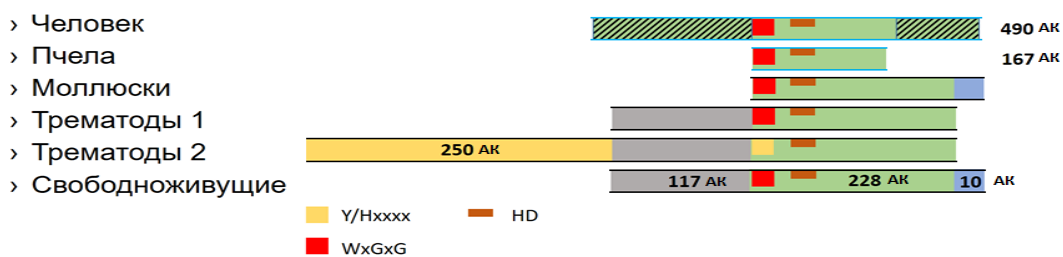


Рис.3 - Структурная организация фосфолипаз группы Г3

Эффект от использования кластера в достижении целей работы.

Почти все вышеперечисленные программы запускались на кластере НГУ. Ключевой программой нашего исследования является – Orthofinder, которую мы запускали исключительно на кластере НГУ. Соответственно, работа на кластере НГУ имела крайне положительный эффект.

Перечень публикаций, содержащих результаты работы (если есть). Если имеется, указать импакт-фактор журнала (Thomson Reuters, РИНЦ,...).

1. Бочарникова М.Е., Биология, Биоинформатика. МНСК-2021: Биология – сборник тезисов. – стр. 6

Опционально: ваши впечатления от работы вычислительной системы и деятельности ИВЦ НГУ, а также предложения по их совершенствованию.

Работать очень удобно. Во время работы не возникало проблем, разве что, периодически кластер не работал по техническим причинам, но об это оповещали и, в целом, никаких заминок данные перерывы в нашем исследовании не сделали. Предложений по совершенствованию нет, на мой взгляд система отлажена отлично.