

ОТЗЫВ

**официального оппонента
на диссертацию на соискание ученой степени
кандидата биологических наук Николаевой Ольги Владимировны
на тему: «Дивергентные метагеномы беспозвоночных животных»
по специальности 03.01.09 – «математическая биология, биоинформатика»**

Структура диссертации

Диссертация изложена на 151 страницах, состоит из Введения, трёх частей (Обзор литературы, Материалы и методы, Результаты и обсуждение), Заключение, Выводов и Списка литературы из 456 источников; содержит 32 рисунка и 23 таблицы.

В Введении сформулирована актуальность темы диссертационного исследования – определение положения организма на филогенетическом дереве, в особенности в случаях, когда простое сопоставление нуклеотидных и аминокислотных последовательностей по известным алгоритмам может оказаться недостаточным и когда приходится привлекать дополнительные филогенетические маркеры в виде элементов структуры геномов. Целью диссертации явилась оценка применимости различных типов анализа молекулярных данных для поиска связей между организмами по митохондриальной ДНК ряда таксонов беспозвоночных со значительной дивергенцией последовательностей. Соответственно сформулированной цели были поставлены задачи исследования: аннотировать митохондриальные геномы выбранных видов, сравнить порядок генов в мтДНК, провести филогенетический анализ по белок-кодирующим генам, в т.ч. с привлечением дополнительных филогенетических методов, определить взаимное положение и родство или независимость ортонектид (Orthonectida) и дициемид (Dicyemida).

Обзор литературы хорошо структурирован. Состоит из ясно изложенных кратких разделов, высвечивающих основные векторы диссертационной работы: структура митохондриального генома, подходы к

филогенетической реконструкции и необходимые для этого основные молекулярные маркеры, модели нуклеотидных замещений. Сложность филогенетического анализа с применением байесова подхода заключается в выборе единой или посайтной модели замещений с бóльшим или меньшим числом параметров, тем более что у беспозвоночных митохондриальный геном лабильней, чем у позвоночных — как длина кольцевой молекулы ДНК, так и порядок генов или блоков генов. Дано подробное описание объектов исследования: два вида киноринх — *Echinoderes svetlanae* и *Pycnophyes kielensis*, одного вида ортонектид *Intoshia linei*, одного вида дициемид *Dicyema* sp., и четырех видов волосатиков — *Gordionus alpestris*, *Gordionus wolterstorffii*, *Gordius* sp. и *Chordodes* sp. Описана история филогенетических исследований этих групп беспозвоночных.

В части «**Материалы и методы**» описаны поэтапно шаги по сбору и анализу материала: кем, где и когда собирались особи исследованных видов и как из них выделяли ДНК; подготовка и секвенирование геномных библиотек ДНК и транскриптомных библиотек РНК с последующей сборкой ридов и аннотированием митохондриальных последовательностей; выявление и анализ инвертированных повторов и шпилек в мтДНК волосатиков; определение нуклеотидного состава, нуклеотидной последовательности и порядка генов в мтДНК исследованных видов. Детально расписаны алгоритмы построения филогенетических деревьев по транслированным аминокислотным последовательностям митохондриальных белок-кодирующих генов с учётом различных эволюционных моделей. Для оценки монофилетичности таксономических групп приведены алгоритмы поиска синапоморфий и вероятностные оценки их состояний. Следует отметить большой объём, разнообразие и важность выполненной автором экспериментальной молекулярно-биологической работы и биоинформационного анализа полученных нуклеотидных и аминокислотных последовательностей.

Часть III «Результаты и обсуждение» состоит из двух разделов:

1. Структурная организация митохондриальных геномов.

2. Филогенетический анализ.

В первом разделе подробно исследована структура мт-геномов отдельно для представителей киноринх, волосатиков, ортонектид и дициемид. Особо отметим, что секвенировать и аннотировать митохондриальные геномы – сама по себе важная и большая работа, а с позиций основной цели диссертации она создаёт фундамент для детального филогенетического анализа молекулярной дивергенции исследованных организмов.

Киноринхи (Kinorhyncha)

Диссертант с коллегами впервые секвенировала и аннотировала митохондриальные геномы *Echinoderes svetlanae* и *Rhynophyes kielensis*, размеры которого типичны для многих животных – около 15 т.п.н. Оказалось также, что мтДНК обоих видов содержит типичный набор генов (13 белок-, 22 тРНК- и 2 рРНК-кодирующих генов), и дополнительно дублицированный ген метиониновой тРНК. Выявлен нуклеотидный состав в мт-геноме этих видов, смоделирована вторичная структура транспортных РНК – практически у всех по стандартному типу «клеверный лист». Определён порядок генов и показано частичное отсутствие характерных для билатеральных животных консервативных блоков; тем не менее, удалось смоделировать порядок мт-генов у общего предка киноринх. Совместная кластеризация двух пар генов метиониновой тРНК на филогенетическом дереве всех тРНК киноринх показывает, что дупликация произошла у общего предка киноринх, а дивергенция нуклеотидного состава в пределах каждого гена шла уже независимо в каждом из видов.

Волосатики (Nematophora)

Митохондриальные ДНК четырех исследованных видов волосатиков также имеют обычную кольцевую структуру и обычные для мтДНК животных размеры – порядка 15 т.п.н. со столь же типичным набором митохондриальных генов. Приведена общая характеристика нуклеотидного состава митохондриальных геномов волосатиков и порядка генов на мтДНК; восстановлен порядок генов у

общего предка этих видов. У всех этих видов выявлено общее свойство – наличие в белок-кодирующих генах множества палиндромов, состоящих из совершенных инвертированных повторов длиной от 15 до полутора сотен п.н. Интересно, что процесс формирования палиндромов в этом таксоне интенсивный и продолжается на внутривидовом уровне – большая часть палиндромов видоспецифична. Их наличие может приводить к появлению шпилек в нативном состоянии, что было подтверждено экспериментально, и влиять как на репликацию ДНК, так и на продукты ПЦР в экспериментах, а также ухудшать покрытие ридов и, в целом, осложнять сборку геномов при полногеномном секвенировании. Обсуждена функциональная роль палиндромов у волосатиков, в т.ч. их возможная регуляторная функция.

Ортонектида (*Intoshia linei*)

Её митохондриальный геном также обычного для многих животных размера – порядка 15 т.п.н. с тем же типичным набором генов, что и исследованные виды киноринх и волосатиков; выявлен порядок их расположения на мтДНК. Исследован нуклеотидный состав и смоделирована вторичная структура тРНК. В связи с неопределённостью имеющихся данных о первичной последовательности некодирующего региона мтДНК у ортонектид и наличием в этом участке длинных инвертированных повторов, превышающих длину ридов, были сконструирована пара праймеров и проведено прямое секвенирование по Сэнгеру. При этом пришлось преодолевать сложности из-за образования шпильки, для чего ДНК была предварительно обработана рестриктазой с сайтом узнавания в центре некодирующего региона. После чего каждый из исходных праймеров был использован в паре с дополнительным праймером из центра некодирующего региона. Для изучения у ортонектиды транскриптов генов мтДНК, их РНК-риды были собраны в пять (митохондриальных) контигов, каждый из которых содержал один или более генов; описана их структура. Секвенирование кДНК у паразитической и свободноживущей форм не выявило различий в экспрессии митохондриальных генов; однако наличие корреляции между %GC и числом ридов

говорит о том, что сходство паттернов экспрессии митохондриальных генов может являться артефактом секвенирования кДНК. Это ещё раз говорит о том, с какой тщательностью автор исследовал возникающие проблемы и неоднозначности, не ограничиваясь одним видом анализа, а применяя комплексный подход.

Дициемида (*Dicyema* sp.)

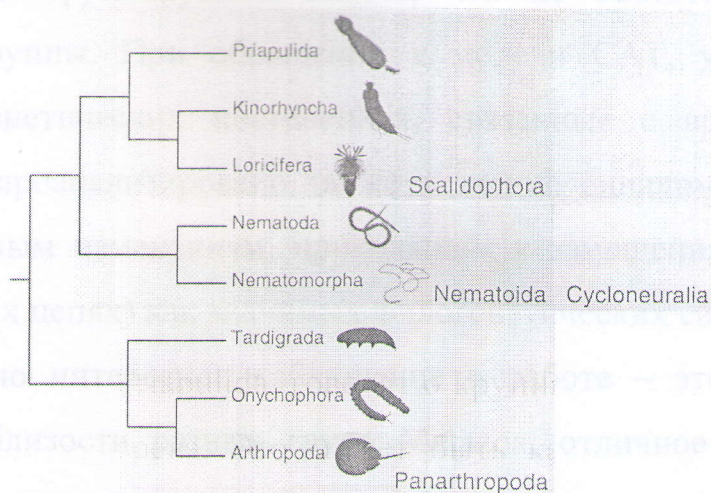
Исследован нуклеотидный состав и спрогнозирована вторичная структура митохондриальных тРНК. Необычной особенностью генома дициемиды является наличие множества кольцевых контигов. Это касается и митохондриального генома: в основном каждый контиг содержал один ген, хотя несколько содержали по два или три гена. Однако возникает вопрос: а все ли кольца истинные? Или часть из них – артефакт интерпретации? Действительно, одна и та же нуклеотидная последовательность может мыслиться и как линейная структура с двумя прямыми повторами, и как кольцо.

Филогения Ecdysozoa

Автор построила филогенетическое дерево (рис. 25 диссертации) для клады Линяющих (Ecdysozoa) по восстановленным (по ДНК) аминокислотным последовательностям 12-ти белков, кодируемых митохондриальными генами, с использованием модели GTR+CAT+Г. Автор внесла свой вклад в дискутируемую проблему филогении и систематики этой группы животных. Подтверждена монофилия ряда таксономических групп (рис. 25). Однако, сопоставляя дерево рис. 25 с, данными в работе Giribet & Edgecombe (2017), можно видеть, что это не распространяется на всю топологию дерева. Например, Nematoda и Nematomorpha, оставаясь согласно мт-дереву рис. 25 монофилетичными, не образуют монофилетичную кладу Nematoida из-за группы Loricifera+Kinorhyncha. Однако это ещё можно объяснить, если обратить внимание на подкладу с 75%-ой поддержкой. В то же время стопроцентная кластеризация на дереве рис. 25 морских пауков Rysnagonida с вышеуказанной кладой плюс Tardigrada и полное отделение от других членистоногих (Arthropoda), а также полное отделение

Loricifera+Kinorhyncha от Priapulida, говорят о крупных систематических факторах смещения в топологии дерева. Это может быть как отбор по мт-генам, так и случайностями (генетическом дрейфе) по мтДНК как однолокусного маркера в ходе длительной эволюции: последнее можно сгладить с помощью мультилокусных (широкогеномных) маркеров.

G. Giribet and G. D. Edgecombe



Ещё более интригующим фактом в данной части исследования стало «сосуществование» альтернативных филогенетических сценариев, не отвергаемых статистическими тестами.

Филогения Lophotrochozoa

Автор исследовала филогению этого таксона также по 12-ти митохондриальным белкам с использованием модели GTR+CAT+Г. Показано, что основные таксоны Lophotrochozoa монофилетичны (за исключением Annelida). Также было проверено положение Mesozoa, формально объединяющей две групп животных – Orthonectida и Dicyemida. Подтверждена полифилия Mesozoa: ортонектиды и дициемиды образуют далёкие друг от друга ветви в пределах Lophotrochozoa, указывая на независимое положение этих двух групп Mesozoa. В то же время, следует отметить, что ряд таксонов Lophotrochozoa (не только Annelida) достаточно рыхлые, что видно по низким уровням поддержки: например, группы с длинными ветвями, моллюски и др. Поэтому не исключена вероятность появления Orthonectida и Dicyemida вместе при моделировании филогенетического дерева Lophotrochozoa.

В связи с этим следовало бы раскрыть фразу: «Несмотря на то что реконструкция филогении Lophotrochozoa по митохондриальным данным указывает на полифилию Mesozoa, статистические тесты не выявляют значимых различий между гипотезами о монофилии и полифилии этой группы». Интересно, что без использования модели сайт-специфичных замен (CAT) все представители с длинными ветвями группируются вместе, что не соответствует данным об эволюции этой группы. При обращении к модели CAT, устраняются грубые ошибки в филогенетических построениях, связанные с артефактом длинных ветвей. В работе проанализировано также значение синапоморфий (по порядку генов и по геномным изменениям, приводящим к замещениям аминокислотных остатков в белковых цепях) как значимых филогенетических сигналов.

Ещё одно интересное наблюдение в работе – это совершенно иное представление о близости разных групп Metazoa, отличное от эволюционного моделирования филогений, полученное методами количественной таксономии по частотам встречаемости различных аминокислот в белках, кодируемых митохондриальными генами. Это говорит о том, что подход к анализу филогенетических отношений между удалёнными таксонами должен основываться на самых разных моделях дивергенции нуклеотидных и аминокислотных последовательностей и методах их анализа.

Завершая анализ диссертационной работы, следует подчеркнуть широкий фронт детальных исследований, выполненных автором: от выявления структурной организации геномов рассмотренных организмов, в т.ч. вторичной структуры тРНК, – до изучения филогении Metazoa и положения исследованных таксонов на филогенетических деревьях, полученных с использованием различных моделей дивергенции нуклеотидных и аминокислотных последовательностей. Работа – несомненно уровня, достойного для защиты кандидатской диссертации.

Теперь перейду к возникшим у меня вопросам и замечаниям – помимо тех, что были подняты выше по ходу обсуждения результатов работы.

- Филогении, построенные по кодирующим участкам нуклеотидных последовательностей, будучи не селективно нейтральными, могут не быть равномерно тикающими молекулярными часами. Автору следовало обсудить это возможное ограничение на использование белок-кодирующих генов в филогенетических целях.
- В связи с последним замечанием ещё одно: почему бы не сравнить филогении по исходным нуклеотидным последовательностям в целях сравнения с полученными филогенетическими деревьями по транслированным аминокислотным последовательностям митохондриальных белок-кодирующих генов?
- Следовало бы обсудить не приводит ли использование метода GTR с самой общей матрицей параметров к ложному поиску решений в пространстве параметров огромной размерности.
- И к этому ещё один вопрос: почему было не сравнить полученные филогении с «примитивным» бутстреп-анализом?
- Хотелось бы видеть привязки полученных филогенетических деревьев к временной и палеонтологической шкалам.
- Филогенетическая сеть рис. 7 диссертации наводит на вопрос, с какими свойствами тРНК или аминокислот связаны два филогенетических кластера?
- Выводы диссертации местами сформулированы нечётко. Например, в выводе 1 хотелось бы видеть, каким образом «вероятность ... может быть ... применена в филогенетическом анализе». В выводе 3 было бы лучше написать «должны применяться» вместо «могут применяться». В выводе 4 вторую часть было бы лучше сформулировать так: «в отличие от дициемид, что подтверждает гипотезу полифилетичности Mesozoa».
- Кроме того, в выводах не упомянуто о большой проделанной работе по структурной организации митохондриальных геномов у представителей исследованных четырёх таксонов, тем более что первой задачей диссертационной работы было: «Аннотировать митохондриальные геномы» (стр. 6 диссертации).
- Следует отметить большое количество опечаток в работе, а ряд из них присутствует и в автореферате. Например, список литературы 456, а не 472 (стр. 7

автореферата); митохонемы на стр. 6 диссертации, бел^оков в подписи к рис. 25, 27 диссертации; и другие.

- Первая фраза Заключения выглядит таинственно (стр. 112 диссертации): «Дивергентность в митохондриальных геномах в первую очередь проявляется в высоких уровнях отличия нуклеотидных последовательностей по сравнению с другими организмами». Как это понять?

- Что касается термина «дивергентный геном». На стр. 6 диссертации дано его определение: «т. н. диверегные митогеномы – те, эволюция которых проходила быстрее, чем митогеномов «в среднем». Во-первых, здесь опечатка в определяемом термине (вероятно «дивергентные» вместо «диверегные»); во-вторых, «т. н.» означает, что этот термин был где-то ранее определён, но ссылка на источник отсутствует; в-третьих, что значит «быстрее»?

Указанные замечания не умаляют значимости диссертационного исследования. Опечатки и неточности, конечно, следует потом исправить. А что касается предложений по сопоставлению полученных филогений с другими, полученными иными способами – это пожелание к дальнейшему развитию исследования.

Проделанная Николаевой О.В. работа имеет большое теоретическое и практическое значение для филогенетики. Отмеченные вопросы относятся лишь к уточнению и прояснению текста диссертации и не влияют на высокую оценку полученного материала и его анализа. Диссертация отвечает требованиям, установленным Московским государственным университетом имени М.В. Ломоносова к работам подобного рода, оформлена согласно приложениям № 5, 6 Положения о диссертационном совете Московского государственного университета имени М.В. Ломоносова. Содержание диссертации соответствует паспорту специальности 03.01.09 – «математическая биология, биоинформатика» (по биологическим наукам), а также критериям, определенным пп. 2.1-2.5 Положения о присуждении ученых степеней в Московском государственном университете имени М.В. Ломоносова, Личный вклад автора в работу большой и детально описан на стр. 8-9 диссертации. Публикаций по материалам

диссертационного исследования достаточно. Автореферат отвечает содержанию диссертации. Защищаемые положения доказаны.

Таким образом, соискатель **Николаева Ольга Владимировна** заслуживает присуждения ей ученой степени кандидата биологических наук по специальности 03.01.09 – «математическая биология, биоинформатика» (по биологическим наукам).

Официальный оппонент:

доктор биологических наук

зав. лабораторией ФГБУН Институт общей генетики им. Н.И. Вавилова РАН

(ИОГен РАН), Москва 119991, улица Губкина, д.3

Животовский Лев Анатольевич

 28 мая 2022г.

Контактные данные:

тел.: +7(916)0739468

Специальность, по которой официальным оппонентом защищена диссертация: «генетика» 03.00.15 (по спецификации года защиты оппонента).

Адрес места работы:

119991, г. Москва, ул. Губкина, д. 3,

Лаборатория генетических проблем идентификации

Тел.: +7(916)0739468, e-mail: levazh@gmail.com

Подпись зав. лабораторией ИОГен РАН профессора Л.А. Животовского удостоверяю:


**Зам. директора
по научной работе**

БРУСКИН С.М.