

ПОСЛЕДОВАТЕЛЬНОСТЬ ГЕНА СУБЪЕДИНИЦЫ I ЦИТОХРОМОКСИДАЗЫ С В МОЛЕКУЛЯРНОЙ ТАКСОНОМИИ ЖИВОТНЫХ: ПРИНЦИПЫ, РЕЗУЛЬТАТЫ И ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ

Белорусский государственный университет, Минск, Республика Беларусь



Воронова Нина Владимировна, кандидат биологических наук, доцент кафедры зоологии биологического факультета Белорусского государственного университета

e-mail: voronoff@list.ru

Научные интересы связаны с изучением особенностей эволюционирования геномов в разных группах животных



Буга Сергей Владимирович, доктор биологических наук, профессор, заведующий кафедрой зоологии биологического факультета Белорусского государственного университета

e-mail: aphids@tut.by

Научные интересы связаны с вопросами фаунистики, биологии и экологии тлей



Курченко Владимир Петрович, кандидат биологических наук, доцент, заведующий лабораторией прикладных проблем биохимии биологического факультета Белорусского государственного университета

e-mail: kurchenko@tut.by

Научные интересы связаны с биохимическим и молекулярно-генетическим исследованием разнообразия животных и растений

Введение

В конце XX века систематика получила в свое распоряжение уникальный инструмент, открывший для исследователей новые возможности, – сравнительный анализ маркерных последовательностей ДНК. Сегодня геносистематика, в основе которой лежит представление о том, что последовательность ДНК может служить таким же видовым признаком, как и особенности морфологии, – одна из наиболее активно развивающихся областей биологии.

Значение развития геносистематики определяется тем, что во всех царствах живых существ количество видов, идентификация которых представляет проблему даже для специалистов, очень велико [1]. В то же время, точное определение вида в ряде случаев является чрезвычайно важным. Наиболее часто приводимый пример, иллюстрирующий важность точного таксономического определения, это комары рода *Anopheles* Meigen, 1818, переносящие возбудителя малярии. В настоящее время известно около 200 морфологически схожих видов, из которых более 60 переносят малярийного плазмодия и только около десятка видов переносят наиболее опасные его штаммы [2]. Морфологическая идентификация переносчиков трудоемка и под силу лишь небольшому числу специалистов, в то время как генетическая идентификация – это рутинная процедура.

Точное таксономическое определение позволяет решить целый ряд биологических и медицинских задач:

- выявление векторов опасных заболеваний человека в метапопуляциях морфологически сходных видов;
- выявление карантинных видов растений и животных;
- медицинская диагностика заболеваний человека и животных;
- поиск новых видов и видов-двойников среди вредителей сельского хозяйства и паразитов;
- контроль видового определения объектов при проведении биохимических, физиологических и других исследований растений и животных.

В 2003 году идея о существовании стабильных видоспецифических различий в последовательности ДНК привела к созданию глобального проекта по ДНК-штрихкодированию жизни [3]. Канадским Институтом биоразнообразия в Онтарио (Biodiversity Institute, Ontario, Canada) была создана база данных нуклеотидных последовательностей, в которой содержатся только короткие последовательности ДНК, позволяющие идентифицировать виды живых организмов всех царств (www.barcodinglife.com) [4].

В 2010 году был создан Всемирный консорциум CBOI, основная задача которого – пополнение существующей базы данных с тем, чтобы провести ДНК-паспортизацию всех известных видов живых существ. К сегодняшнему дню в базе данных по ДНК-штрихкодированию содержатся записи более чем по 160 тысячам видов живых организмов [5, 6].

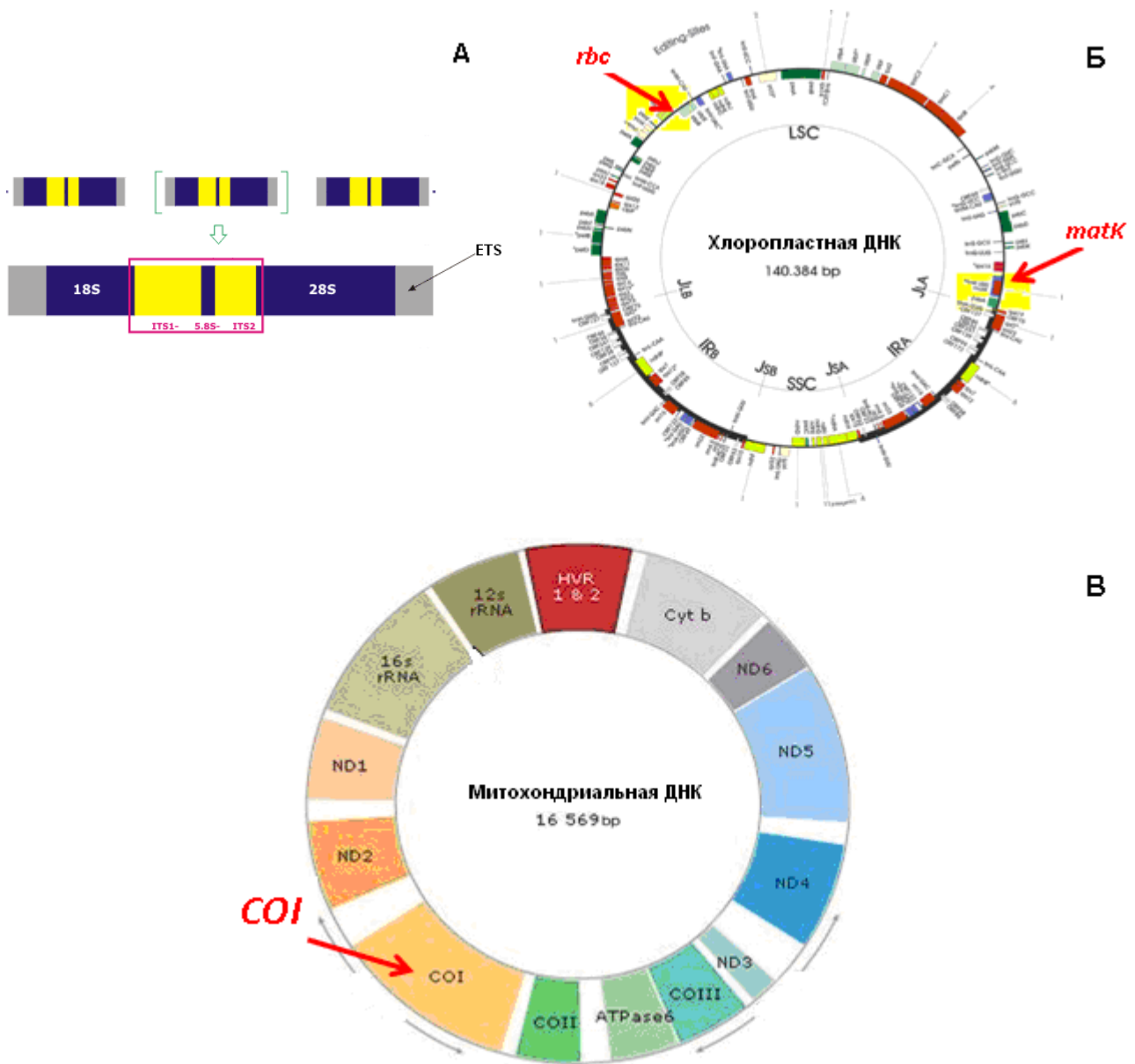
ДНК-штрихкодирование жизни

Основной сложностью в реализации проекта по ДНК-штрихкодированию являлся выбор такого участка генома, который бы отвечал предъявляемым требованиям. Он должен был быть:

- универсальным, то есть присутствовать во всех геномах;
- коротким, для того, чтобы он мог быть легко секвенирован;
- амплифицироваться под едиными праймерами, чтобы не было необходимости их написания для каждого нового организма;
- не иметь аллелей внутри одного генома, то есть всегда находиться в гомозиготном либо гемизиготном состоянии;
- и главное – частота возникновения гаплотипов этого участка генома в идеальном варианте, должна приближаться к золотому стандарту геносистематики – «один таксон – один гаплотип». То есть этот участок генома должен быть одновременно достаточно консервативным, чтобы оставаться неизменным у всех представителей одного вида и, в то же время, должен быть достаточно варибельным, чтобы разные виды, даже близкие, отличались друг от друга. В противном случае идея ДНК-штрихкодирования теряет смысл.

Уже к моменту создания, CBOI столкнулся с серьезной критикой, как со стороны генетиков, так и со стороны таксономистов [7, 8, 9]. Систематики высказывали сомнения в том, что последовательность ДНК длиной 500–600 нуклеотидов, ни скорость, ни характер эволюции которой не известны, будет отражать эволюционную историю вида в той же степени, что и морфология, анатомия или палеонтология. Генетики считали, что поиск участка генома, отвечающего заявляемым требованиям, сродни поиску святого Грааля.

Тем не менее, поскольку к двухтысячным годам в молекулярной таксономии уже был накоплен большой массив знаний, такие участки удалось выбрать. У растений информативными оказались хлоропластные гены белков фотосинтетического аппарата – рибулозо-бисфосфат карбоксилаза (*rbcL*) и матураза К (*matK*) [10]; у грибов – внутренние транскрибируемые спейсеры, находящиеся между генами рибосомальных РНК (ITS1, ITS2) [11]; у животных и простейших – митохондриальный ген субъединицы 1 цитохромоксидазы *c* (COI) (рисунок 1).



А – Внутренние транскрибируемые спейсеры между генами рРНК (ITS1, ITS2), Б – гены рубулозо-бисфосфат карбоксилазы (*rbcL*) и матуразы К (*matK*),
 В – ген субъединицы 1 цитохромоксидазы с (*COI*)

Рисунок 1 – Участки генома, используемые для ДНК-штрихкодирования видов

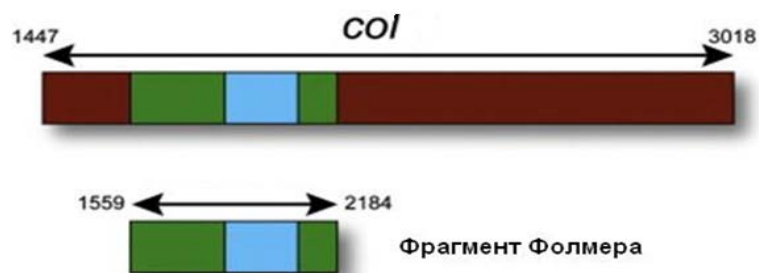


Рисунок 2 – Фрагмент Фолмера – участок гена субъединицы 1 цитохромоксидазы с (*COI*), используемый для ДНК-штрихкодирования

В качестве участка для ДНК-баркодинга животных была эмпирически выбрана 5'-терминальная область гена COI длиной 658 нуклеотидов или, так называемый, фрагмент Фолмера [12] (рисунок 2). Экспериментально было показано, что скорость нуклеотидных замен в этом участке митохондриального генома хорошо коррелирует со скоростью эволюции таксонов высокого ранга, то есть отрядов и классов. В связи с этим первоначально этот участок гена предлагался как маркер для построения филогений эволюционно удаленных групп. Однако позже было установлено, что нуклеотидные последовательности COI-5-региона обладают видоспецифичностью, что делает регион ценнейшим видовым маркером [3].

Ген COI как филогенетический маркер Animalia

Ген COI кодирует первую субъединицу мультимерного белка – цитохромоксидазы *c*, в состав которого у животных входит от 10 до 13 субъединиц. В дыхательной цепи эукариот этот белок образует IV, терминальный комплекс, катализирующий окисление кислорода и одновременный перенос протонов через мембрану (рисунок 3). Субъединица COI является ключевой для осуществления белком его функции, поскольку именно в этом протеине локализуются гем- и медь-содержащие домены – активные сайты, в которых происходит окисление O₂ (рисунок 4).

COI протеин у разных видов животных образован 510–530 аминокислотами. В нем выделяют 25 структурных регионов, в том числе 6 внешних и 5 внутренних петель, а также 12 участков, локализованных во внутренней мембране митохондрии (рисунок 5). COI разделяют на 5' и 3' секции. 5' секция состоит из первых 263 аминокислот и заканчивается M6 участком. В качестве ДНК-штрихкода используется фрагмент гена, кодирующий эту часть белка.

При анализе вариабельности белка у разных видов насекомых оказалось, что карбоксильный конец высоко вариабелен и не пригоден для использования в геносистематике, участки M1, M3 и M4, все внешние петли и внутренняя петля I4 имеют стабильную среднюю вариабельность. Менее вариабельными оказались участки I1 и M6, и полностью инвариантна – петля E4, в области локализации второго протонного насоса. У млекопитающих также 3' секция в целом более консервативна [13].

На пространственной модели белка регион, используемый для видовой идентификации, выделен желтым цветом (рисунок 5). Эта часть белка образует наиболее важные в функциональном отношении домены: здесь располагаются биядерный центр, содержащий гем 3а и атом меди, гем а центр, первый из 2-х протонных каналов, а также участки, отвечающие за связывание с субъединицами III и VIa.

Локализация в этой части полипептидной цепи важнейших функциональных блоков имеет два существенных для геносистематики следствия. Во-первых, эволюция этого участка гена находится под действием, так называемого, функционального ограничения, то есть в этой части допустимы только такие аминокислотные замены, которые не нарушат функции белка, поскольку нарушение функции белка приведет к блокировке дыхательной цепи митохондрий. Такие мутанты будут немедленно элиминироваться, поскольку митохондрии, как известно, гаплоидны. Более того, учитывая важность эффективного функционирования дыхательной цепи и ее эволюционную древность, можно предположить, что отбор по этому признаку будет сугубо стабилизирующим. Следовательно, вероятность многочисленных или индивидуальных нуклеотидных замен будет минимальной, и таким образом, ген полностью отвечает принципу ограниченной вариабельности, что делает его высокоценным филогенетическим маркером. С другой стороны, общая высокая скорость нуклеотидных замещений в митохондриальных генах при необходимости сохранения аминокислотной последовательности приведет к тому, что наблюдаемые замены будут локализоваться в одних и тех же вырожденных сайтах. И поскольку появление нуклеотидных замен, это в чистом виде стохастический процесс, который происходит с определенной частотой при каждом акте репликации, с течением эволюционного времени

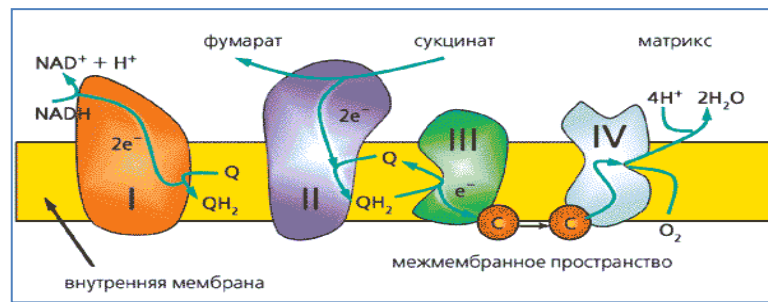


Рисунок 3 – Дыхательная цепь эукариот

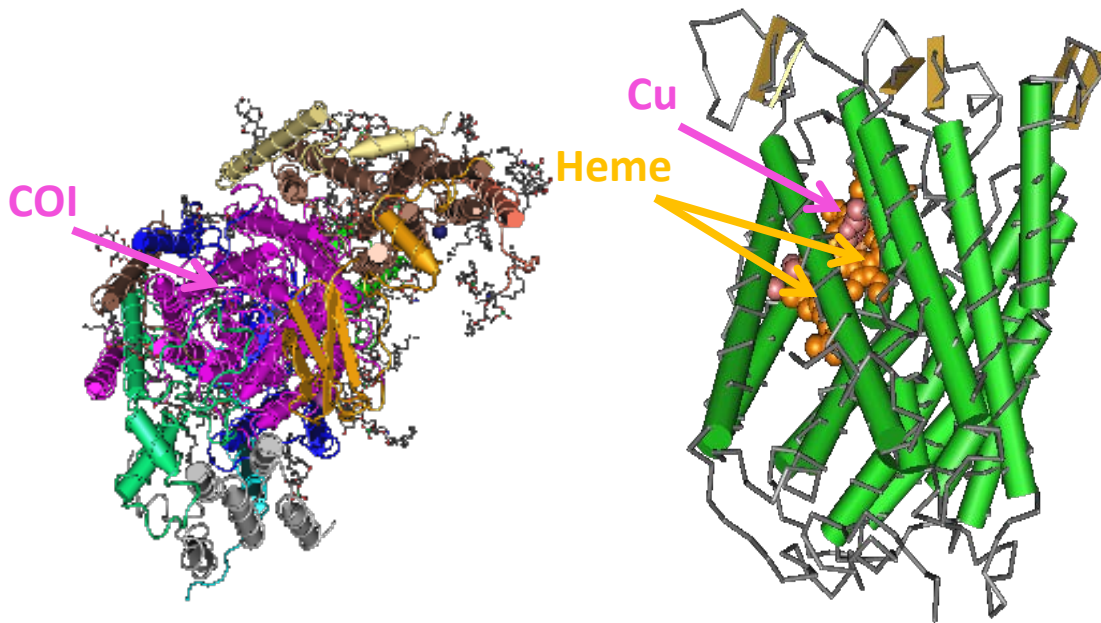


Рисунок 4 – Пространственная структура мультимерного белка цитохромоксидазы с (слева) и пространственная структура субъединицы 1 цитохромоксидазы с (справа)

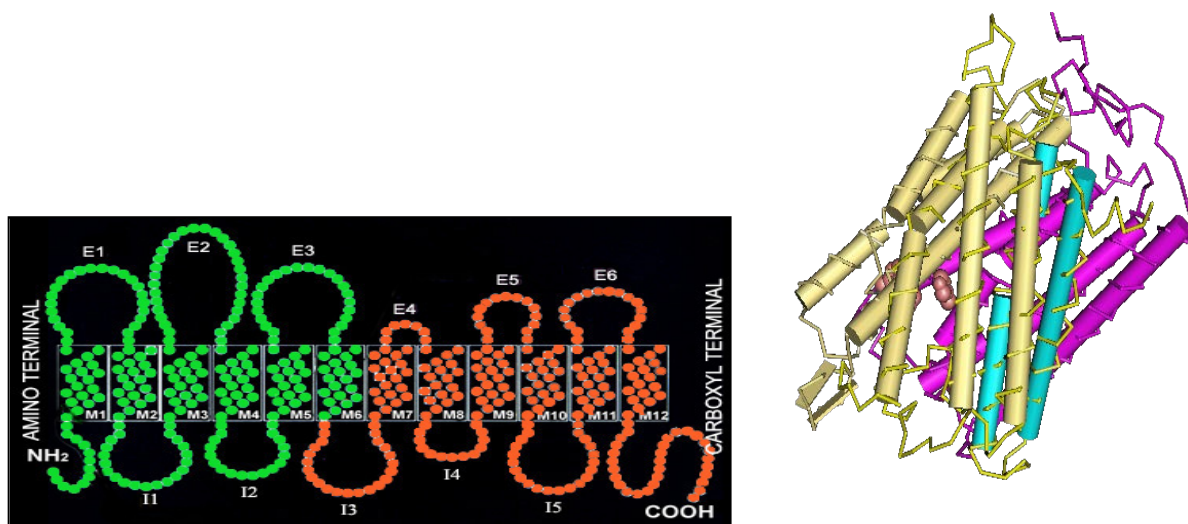


Рисунок 5 – Расположение молекулы субъединицы 1 цитохромоксидазы с в мембране митохондрии (слева) и регион, используемый для ДНК-штрихкодирования видов животных (справа, выделен желтым цветом)

нуклеотидные замены могут происходить «поверх» уже произошедших, маскируя их. Таким же образом, при неоднократных заменах в одном сайте, возникают нуклеотидные гомоплазии или, при обратных заменах, происходит полная потеря филогенетической информации.

И действительно, если в первое время идея генетической идентификации видов была встречена с большим энтузиазмом, то в последующие годы появились работы, которые ставили под сомнение применимость самого подхода [14–16]. Основной проблемой оказалось то, что последовательность COI далеко не всегда вела себя так, как это ожидалось, исходя из амбициозности проекта CBOL. В каких-то группах животных COI давала ожидаемые результаты, в других данные, получаемые при анализе этой последовательности, противоречили результатам, основанным на применении других подходов.

Информативность последовательности гена COI в разных таксонах животных

Чтобы понять причину возникающих противоречий и оценить применимость COI для конкретных объектов, мы проанализировали 253 работы разных авторов, основанных на применении COI в 42 группах животных (таблица 1).

Таблица 1 – Эффективность применения COI как геносистематического и филогенетического маркера в разных таксономических группах животных

Простейшие	
Амебы	Средняя внутривидовая вариабельность (до 3 внутривидовых гаплотипов)
Споровики	Внутривидовая вариабельность отсутствует
Инфузории	Средняя внутривидовая вариабельность (дивергенция между гаплотипами 6%)
Черви	
Трематоды	Высокая внутривидовая вариабельность (гаплотипизация более 85%)
Нематоды морские	Внутривидовая вариабельность низкая или отсутствует
Трихинеллы	Высокая внутривидовая вариабельность, большое число нуклеотидных замен
Полихеты	В разных группах: внутривидовая вариабельность отсутствует или высокая (дивергенция между гаплотипами 5%)
Олигохеты	Внутривидовая вариабельность низкая или отсутствует
Туникаты	
	В разных группах: внутривидовая вариабельность отсутствует или средняя (до 12 гаплотипов с дивергенцией 5,5%).
Губки	
	Внутривидовая вариабельность низкая или отсутствует
Кишечнополостные	
Шестилучевые кораллы	Отсутствие различий между сестринскими видами
Моллюски	
Брюхоногие моллюски	Высокая внутривидовая вариабельность (до 125 внутривидовых гаплотипов с высокой дивергенцией)
Двустворчатые моллюски	В разных группах: внутривидовая вариабельность отсутствует, средняя или высокая (до 10 гаплотипов с дивергенцией 50%)
Головоногие моллюски	Низкая внутривидовая вариабельность
Ракообразные	
Веслоногие раки	Высокая внутривидовая вариабельность (до 35 внутривидовых гаплотипов)
Десятиногие раки	В разных группах: внутривидовая вариабельность отсутствует или высокая (дивергенция между гаплотипами 4,6%)

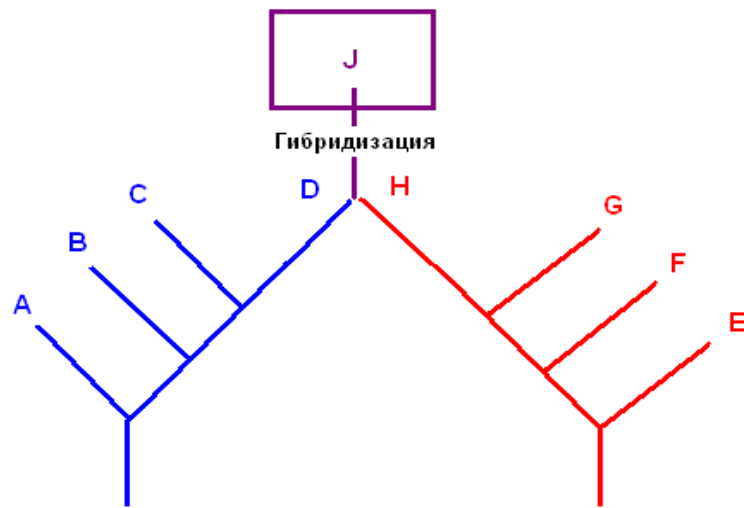


Рисунок 6 – Гибридизация, как причина парафилетического происхождения видов

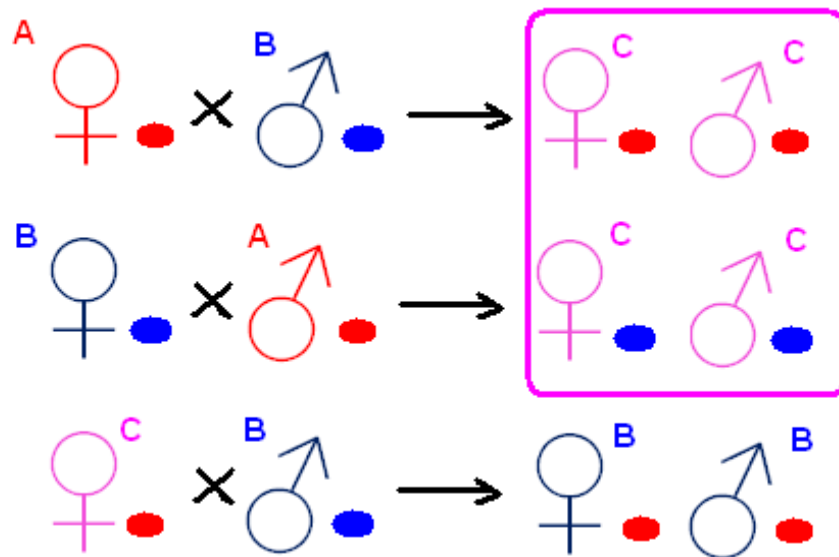


Рисунок 7 – «Перескакивание митохондрий» и образование внутривидовых гаплотипов митохондриальных генов в результате гибридизации (пояснения – в тексте)

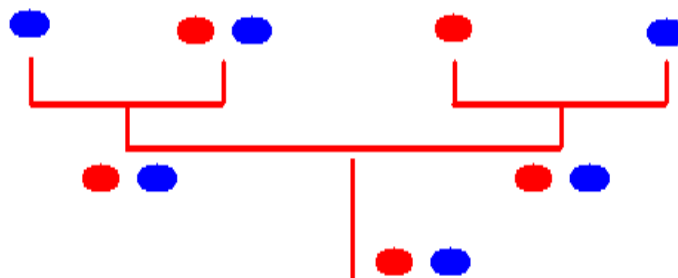


Рисунок 8 – Гетероплазмия. Нарушение принципа дихотомического появления гаплотипов (пояснения – в тексте)

Таблица 1 (продолжение) – Эффективность применения COI как геносистематического и филогенетического маркера в разных таксономических группах животных

Паукообразные	
Скорпионы	Различия между подвидами. Высокая межвидовая вариабельность, трудно интерпретируемые результаты
Гамазовые клещи	Высокая внутривидовая вариабельность (дивергенция между гаплотипами до 18%)
Иксодовые клещи	Внутривидовая вариабельность низкая или отсутствует
Паразитиформные клещи (другие)	Высокая внутривидовая вариабельность (до 32 внутривидовых гаплотипов)
Пауки	Внутривидовая вариабельность низкая (дивергенция между гаплотипами 1,5%)
Насекомые	
Ногохвостки	Внутривидовая вариабельность низкая (дивергенция между гаплотипами 1%)
Стрекозы	Внутривидовая вариабельность низкая или отсутствует
Сетчатокрылые	Высокая внутривидовая вариабельность (до 24 внутривидовых гаплотипов с дивергенцией 6%)
Чешуекрылые	Высокая внутривидовая вариабельность (до 59 внутривидовых гаплотипов с 1,5% дивергенцией).
Наездники	Межпуляционная вариабельность, высокая насыщенность последовательности
Пчелы	Средняя внутривидовая вариабельность. Гетероплазмия. Высокая изменчивость аминокислотного состава белка
Муравьи	Средняя внутривидовая вариабельность
Пилильщики	Внутривидовая вариабельность низкая или отсутствует
Осы	Средняя внутривидовая вариабельность
Псиллиды	Высокая внутривидовая вариабельность (до 30 внутривидовых гаплотипов с низкой дивергенцией)
Тли	Внутривидовая вариабельность низкая или отсутствует
Клопы	Внутривидовая вариабельность низкая или отсутствует
Москиты	Средняя внутривидовая вариабельность (до 12 внутривидовых гаплотипов с дивергенцией 2%)
Мухи	Внутривидовая вариабельность отсутствует. Редко – отсутствие различий между сестринскими видами
Жуки	Низкая внутривидовая вариабельность, иногда – межпопуляционная вариабельность. Дивергенция между видами до 30%
Рыбы	
Костистые рыбы	Внутривидовая вариабельность низкая или отсутствует. Межвидовая дивергенция 2–150 нуклеотидов
Рептилии	
Черепахи	Внутривидовая вариабельность низкая или отсутствует
Птицы	
	Внутривидовая вариабельность отсутствует
Млекопитающие	
Парнокопытные	Низкая внутривидовая вариабельность
Приматы	Средняя внутривидовая вариабельность (дивергенция между гаплотипами 2%)
Человек	Средняя внутривидовая вариабельность (дивергенция между гаплотипами 2,5%)
Собаки	Средняя внутривидовая. Гетероплазмия

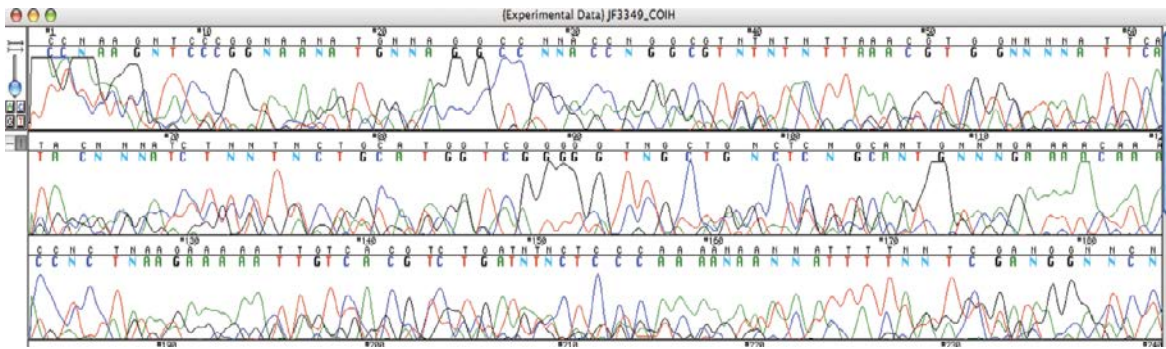


Рисунок 9 – Визуализация Numt (Nuclear mitochondrial DNA) (По Buhay, 2009)

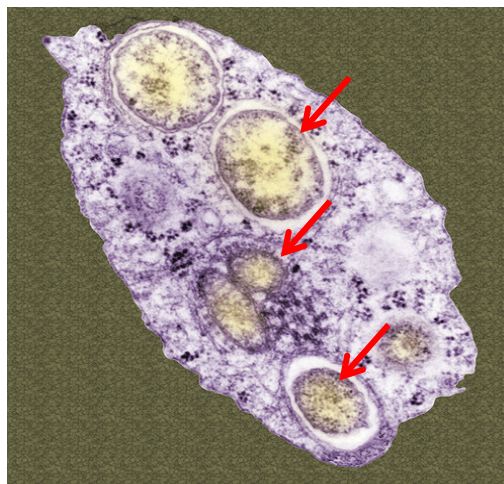


Рисунок 10 – *Wolbachia* внутри клетки насекомого

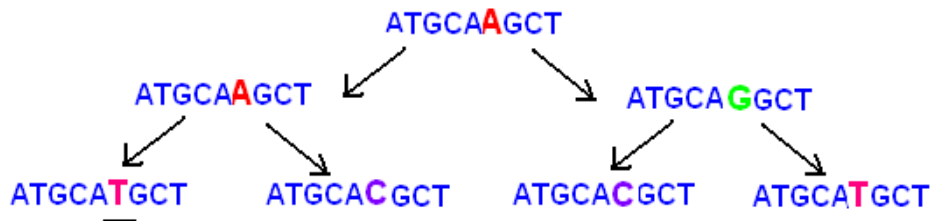


Рисунок 11 – Гипотетическая схема насыщения последовательности ДНК

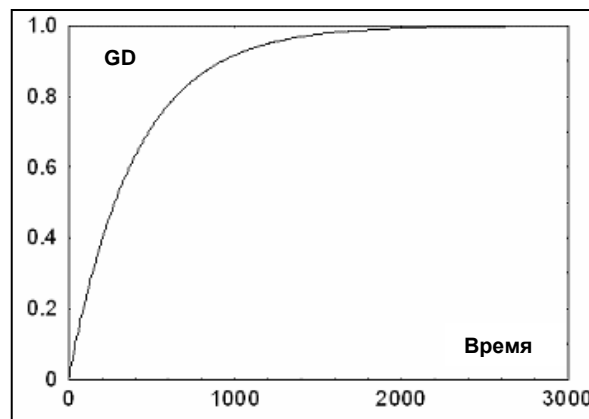


Рисунок 12 – График зависимости величины генетической дистанции от времени дивергенции таксонов

Нам удалось обнаружить не так много работ, посвященных простейшим, однако все эти работы были успешны [17, 18, 19]. И, хотя у амёб и инфузорий обнаруживались внутривидовые гаплотипы COI, количество их было не велико, причем уровень дивергенции между внутривидовыми гаплотипами был всегда ниже, чем различия между разными видами. То есть на основании этих исследований можно заключить, что COI показал себя как эффективный филогенетический маркер при работе с данными таксонами.

Черви – большая парафилетическая группа, включающая, по меньшей мере, 10 типов животных. Представители трех из них – плоские, круглые и кольчатые – достаточно хорошо изучались при помощи COI. Как и ожидалось, существуют особенности в поведении COI в этих таксонах. В частности, у свободноживущих морских нематод и полихет почти отсутствует внутривидовая вариабельность, а там, где она наблюдается, это может быть связано с существованием не описанных подвидов, о чем говорят авторы [20]. У паразитических – другая картина [21, 22]. У этих червей наблюдается очень высокая частота нуклеотидных замен в последовательности гена COI, а у трематод, конкретно у птичьих шистосом, частота обнаружения гаплотипов достигает 85%, то есть до 85% особей в популяциях одного вида могут иметь уникальную последовательность COI, что крайне нетипично для этого гена.

У моллюсков и ракообразных [23] также существуют различия в поведении COI в разных таксонах: от крайне высокого уровня внутривидовых различий у двустворчатых, когда отдельные гаплотипы одного вида могут различаться на 50% [24], до почти полного отсутствия внутривидовой вариабельности у кальмаров [25]. В то же время, в другом типе, у кораллов, не только нет внутривидовой вариабельности по последовательности COI, но у видов одного рода эта последовательность часто оказывается идентичной, что лишает COI ценности как маркера видовой идентификации [26].

У паукообразных COI оказался наиболее информативным у иксовых клещей [27], в то время как у других паразитиформных [28], также как у скорпионов [29], обнаружена очень высокая внутривидовая вариабельность, что крайне затрудняет работу с ними.

Насекомые, это, как известно, очень крупный инфракласс беспозвоночных, к которому относится до 75% всех известных видов в царстве животных, причем это древняя группа, возраст которой определяется в районе 400 млн. лет. Учитывая, что насекомые представляют собой подавляющее большинство животного мира планеты, изучение их эволюции, в том числе с использованием геносистематических маркеров, привлекает внимание большого числа исследователей.

Как можно было ожидать исходя из эволюционной удаленности отдельных отрядов насекомых, поведение последовательности COI сильно различается. Например, у стрекоз последовательность COI оказалась очень эффективна и для видовой идентификации, и для построения филогенетических систем [30]. У чешуекрылых последовательность COI также применяется очень широко, однако количество внутривидовых гаплотипов всегда очень большое, до полусотни, причем различия между гаплотипами внутри одного вида может быть таким же, как различия между видами в других отрядах насекомых [31].

В разных семействах перепончатокрылых различия по поведению COI также довольно значительны [32, 33]. Причем, что интересно, у пчел при средней внутривидовой вариабельности наблюдается высокая изменчивость аминокислотного состава белка, что редкость и, кроме того, у пчел обнаружена гетероплазмия, то есть у одной особи могут быть митохондрии с различающимся геномом [34]. А у наездников наблюдается географическая изменчивость последовательности внутри одного вида [35].

У хоботных насекомых, также как у двукрылых, в целом отмечается хорошая стабильность последовательности COI [36, 37]. Даже при наличии внутривидовых гаплотипов, как у moskitov, уровень различий между ними всегда ниже, чем между отдельными видами [38, 39]. В той же степени эффективным COI оказался у жесткокрылых. У многих видов жуков обнаружена географическая изменчивость последовательности COI,

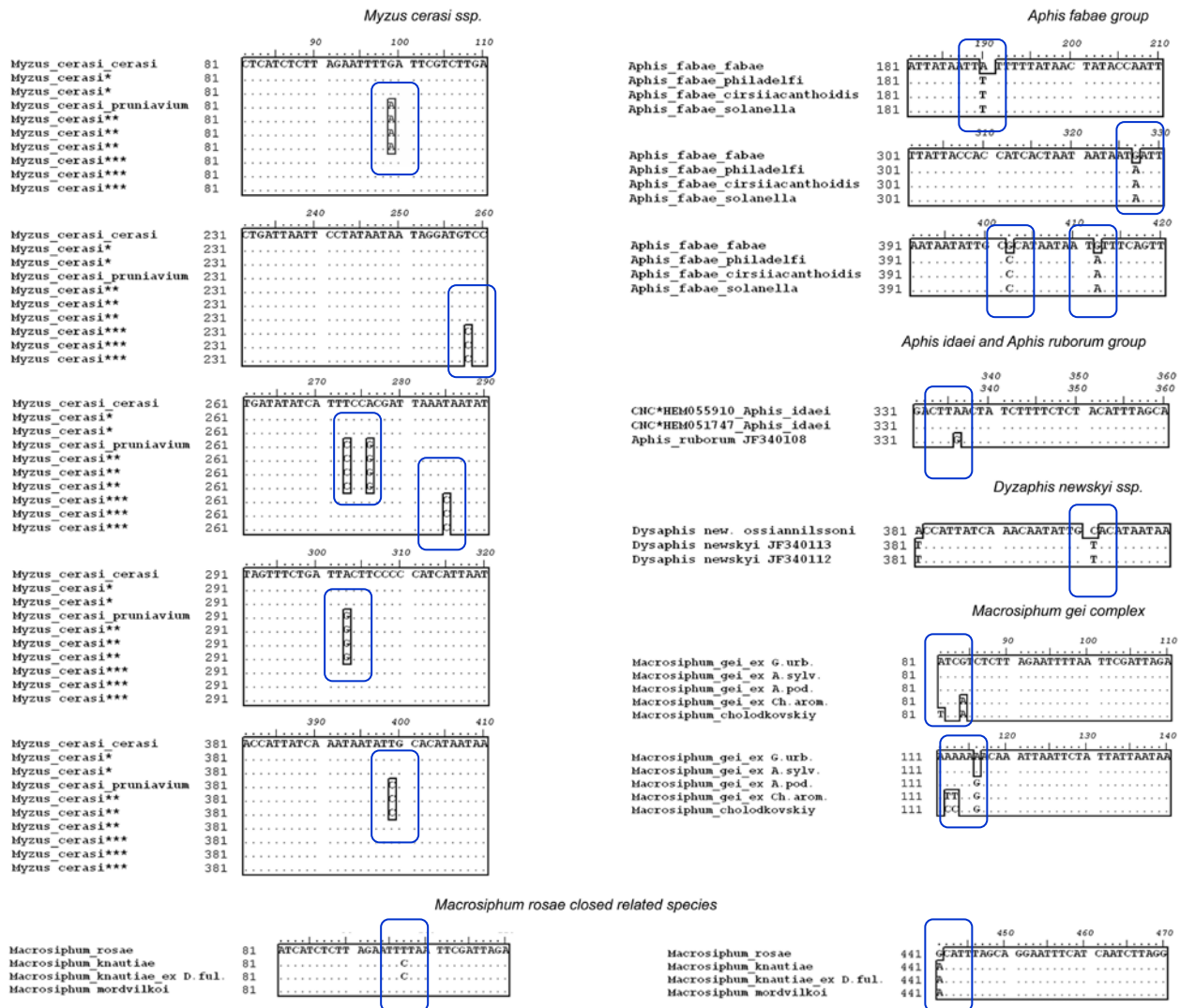


Рисунок 13 – Нуклеотидные последовательности COI видов и подвидов тлей фауны Беларуси

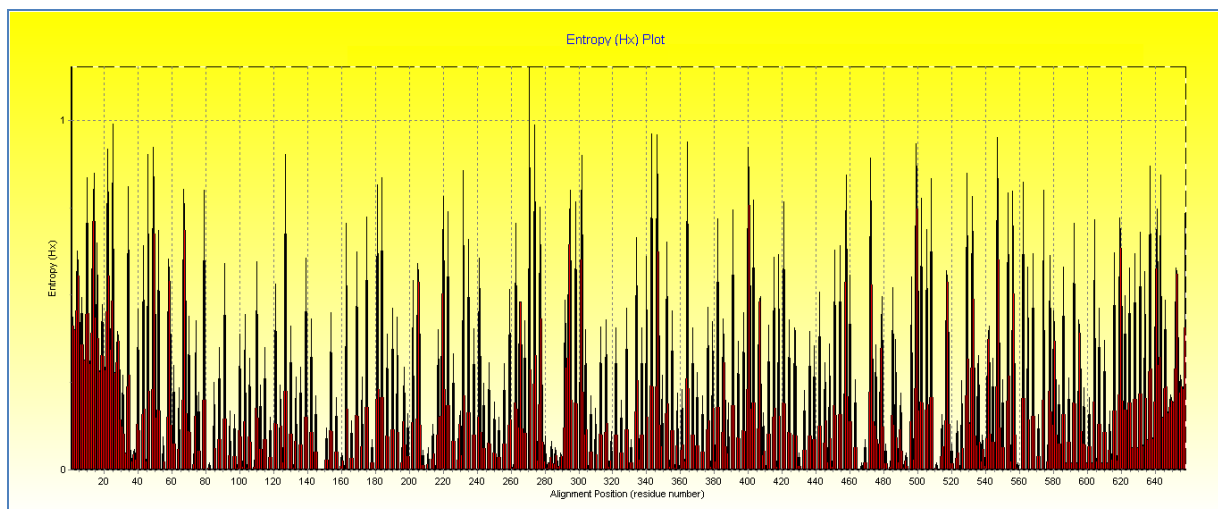


Рисунок 14 – Энтропия нуклеотидного состава последовательности COI-5 секции тлей семейства Aphididae

хотя многие авторы считают, что в этих случаях необходимо описывать подвиды, соответствующие отдельным гаплотипам [40].

Что касается позвоночных животных, то у них в целом последовательность COI оказывается высокоэффективной, что связано с общим низким уровнем внутривидовой вариабельности [41–43]. Интересно отметить, что у человека существует внутривидовой полиморфизм по COI и даже отмечена связь определенных гаплотипов COI с некоторыми видами рака [44]. У некоторых хищных, в частности у собак, отмечена гетероплазмия [45].

Как видно, эффективность применения COI у животных оказалась различной. В каких-то таксонах COI показывает себя как удобный, корректный, информативный маркер, в других – таких как кораллы, двустворчатые моллюски, шистосомы, скорпионы – попытки использовать COI оказались почти полностью безуспешны. Мы попытались выяснить, какие особенности наследования митохондрий могут вызывать такие эффекты.

Причины снижения эффективности использования последовательности COI в отдельных таксонах животных

Существует много причин, которые могут приводить к низкой эффективности COI в конкретных группах животных. Вероятно, самая распространенная, это парафилитическое происхождение видов, то есть формирование вида более, чем из одного корня. Например, в результате гибридизации (рисунок 6).

Во-первых, в гибридизационном видообразовании могут принимать участие как самцы, так и самки обоих видов, и в результате у гибрида сформируются гаплотипы митохондриальных генов, поскольку их митохондрии будут иметь различное происхождение. При множественной гибридизации количество митохондриальных гаплотипов внутри формирующегося вида будет возрастать. Во-вторых, возможны обратные скрещивания, приводящие к, так называемому, перескакиванию митохондрий. То есть при скрещивании гибрида с одной из родительских форм будут появляться потомки, которые морфологически имитируют родительский негибридный вид, свободно с ним скрещиваются, однако их хромосомы соответствуют другому негибридному виду (рисунок 7). Подобное часто встречается у чешуекрылых и представляет собой большую проблему для их геносистематики [46].

Частоту гибридизационного видообразования у животных часто недооценивают, считая, что подобный тип видообразования больше свойственен растениям. Однако, по мнению многих эволюционистов, процент парафилитических видов у животных в действительности очень высок (таблица 2) [47].

Таблица 2 – Частота встречаемости парафилитических видов в разных классах животных (по Funk, Omland, 2003)

Таксон	Число:			Процент парафилитических видов
	Публ.	Родов	Видов	
Млекопитающие	139	103	469	17,0
Птицы	74	87	331	16,7
Рептилии	56	45	147	22,4
Амфибии	35	26	137	21,3
Рыбы	100	99	371	24,3
Членистоногие	143	126	702	26,5
Другие беспозвоночные	37	41	162	38,6
Среднее				23,1

Другой проблемой, оказывающей негативное влияние на возможности геносистематики, является гетероплазмия и патернальная передача митохондрий.

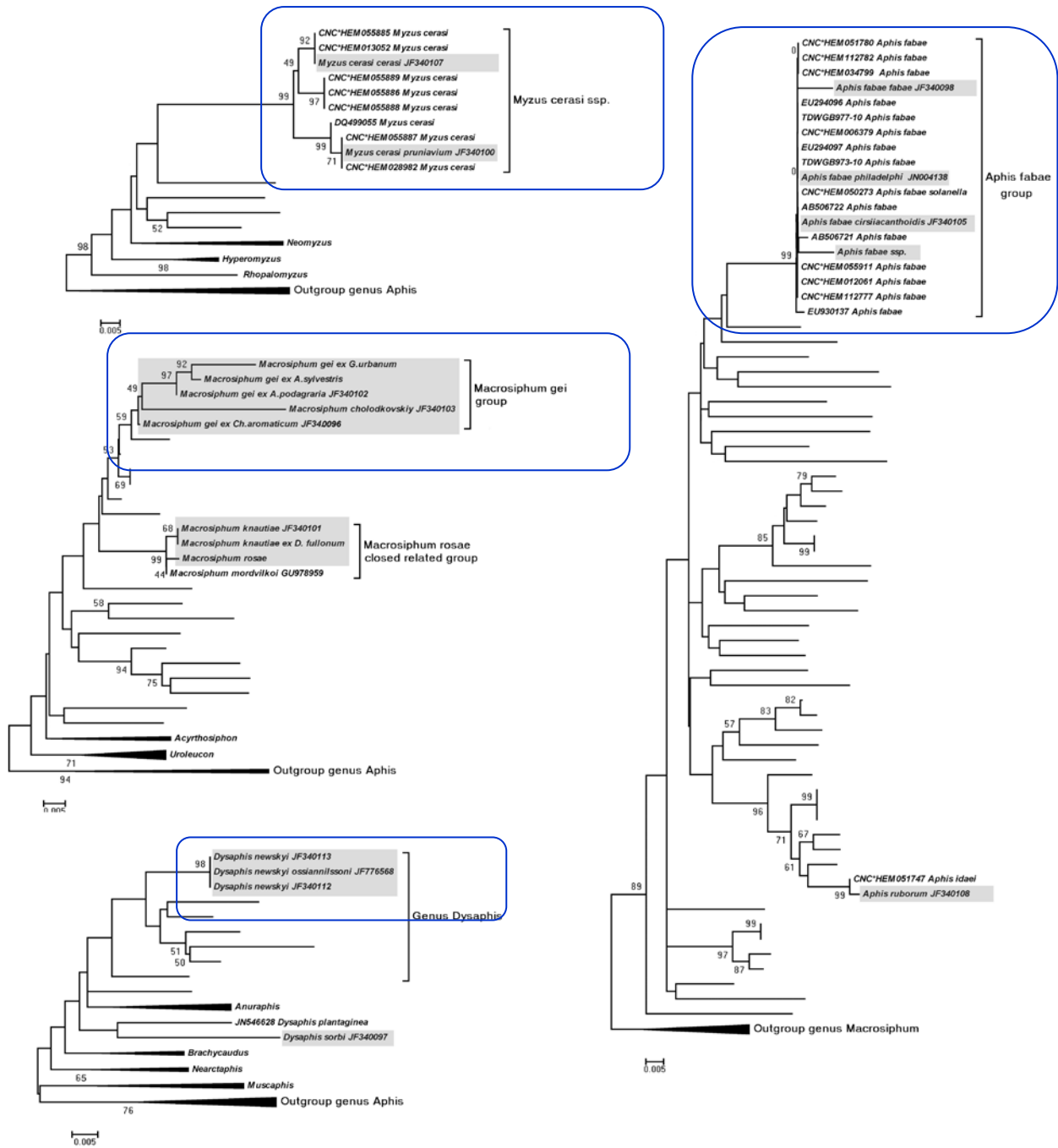


Рисунок 15 – Идентификация подвидов у тлей на основе последовательности COI

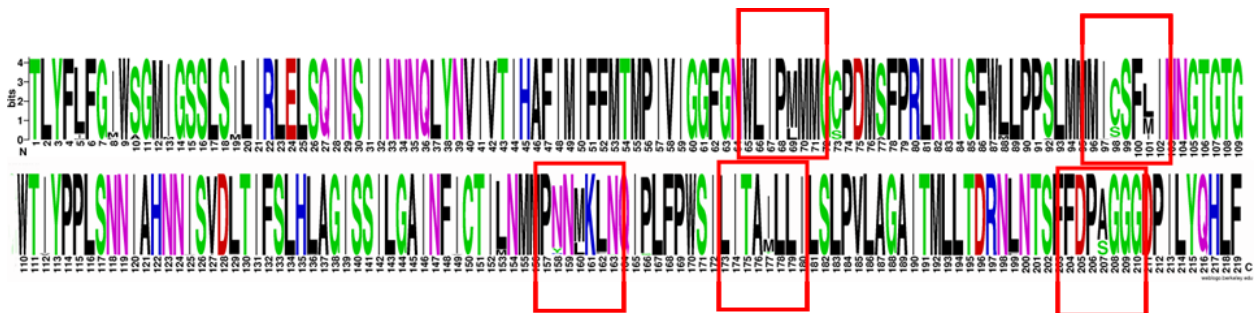


Рисунок 16 – Консенсусная аминокислотная последовательность COI-5 региона тлей 340 видов 16 подсемейств Aphididae

Гетероплазмия или, иначе говоря, присутствие в одной клетке митохондрий с различающимися геномами и наследующихся независимо, встречается как у насекомых, так и, довольно часто, у млекопитающих [48, 49]. Сложности в данном случае связаны с нарушением наблюдаемой дихотомии формирования гаплотипов. То есть, в данном случае, организмы, помеченные синим цветом, будут оцениваться как более близкие друг другу, нежели чем к в действительности близким формам (рисунок 8).

Что касается патеральной передачи митохондрий, то есть от отца к потомкам, то у животных такие случаи описаны [50], в частности у двусторчатых моллюсков – у мидий, частота передачи потомкам отцовских митохондрий сопоставима с частотой передачи материнских. Подобное также описано у некоторых насекомых, это цикады и медоносные пчелы, а также у млекопитающих – мышей, овец и приматов [51]. Патеральная передача представляет собой проблему в случаях, когда существует внутривидовая вариабельность COI, поскольку, также как и при гетероплазмии будет нарушать дихотомию появления признака.

Еще одной проблемой, затрудняющей получение в геносистематике корректных данных, является наличие в ядерном геноме так называемых «Numt» (от англ. Nuclear mitochondrial DNA) – то есть копий митохондриальных генов, встроенных в хромосомы в результате транспозиций (рисунок 9) [52]. Поскольку эти копии представляют собой по сути псевдогены, их изменчивость никак не регулируется. Амплифицируясь под теми же праймерами, что и гены, локализованные в митохондриях, они выдают массу ложных сигналов. В разных группах животных Numt встречаются с различной частотой, широко распространены у ракообразных и жуков, а так же у млекопитающих. В частности, в геноме человека обнаружено 45 ядерных копий COI [53].

У насекомых и филярий есть специфическая особенность. Приблизительно 20% всех видов насекомых в большинстве отрядов и около половины видов филярий инфицированы облигатным внутриклеточным паразитом – грам-отрицательными бактериями рода *Wolbachia* (тип Proteobacteria) (рисунок 10). Эти бактерии оказывают очень сложное влияние на репродукцию своих хозяев, на детерминацию пола, стимулируют партеногенез и, кроме того, оказывают влияние на стабилизацию генома митохондрий [54]. Показано, что у инфицированных видов скорость нуклеотидных замещений в митохондриях во много раз ниже, чем у неинфицированных, даже близкородственных видов. В итоге формируется существенная гетерохрония по отдельным эволюционным линиям насекомых, что серьезно влияет на точность эволюционных расчетов.

Насыщение последовательности, это процесс, при котором нуклеотидные замены происходят многократно в одних и тех же сайтах [55]. В этом случае, через некоторое эволюционное время, может формироваться вторичное сходство между последовательностями, которые на самом деле не являются эволюционно близкими (рисунок 11). Подобный эффект развивается в таксонах, у которых последовательность белка COI высоко консервативна, поскольку в этом случае все наблюдаемые замены сосредоточены в вырожденных сайтах. В последовательности такой длины – 650 нуклеотидов, 254 сайта являются в той или иной степени вырожденными. Это достаточное количество, чтобы обеспечить очень высокую возможность для вариабельности. Однако в действительности мы наблюдаем, что в таких группах при повышении уровня таксона, то есть с увеличением времени дивергенции видов, генетические дистанции между ними не увеличиваются линейно, а выходят на плато по логарифмической кривой (рисунок 12). В наибольшей опасности в отношении насыщения ДНК находятся таксоны с высоко консервативным белком, высокой скоростью генераций и численностью потомства, то есть, прежде всего, беспозвоночные.

Эффективность использования последовательности COI у тлей

Последние три года на кафедре зоологии и в лаборатории прикладных проблем биохимии БГУ мы проводили исследования поведения последовательности COI у тлей (Aphididae) (таблица 3).

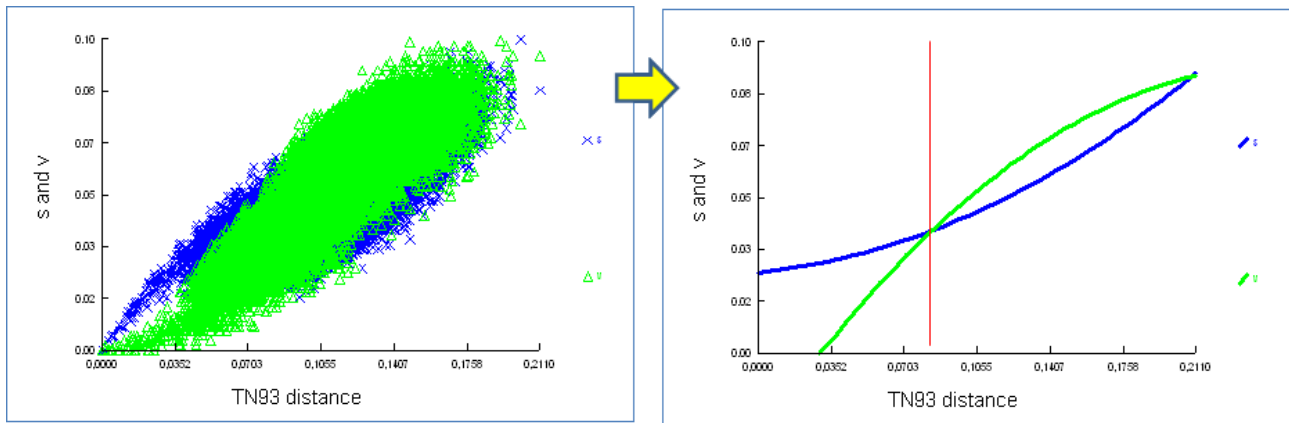


Рисунок 17 – Насыщение последовательности COI у тлей семейства Aphididae

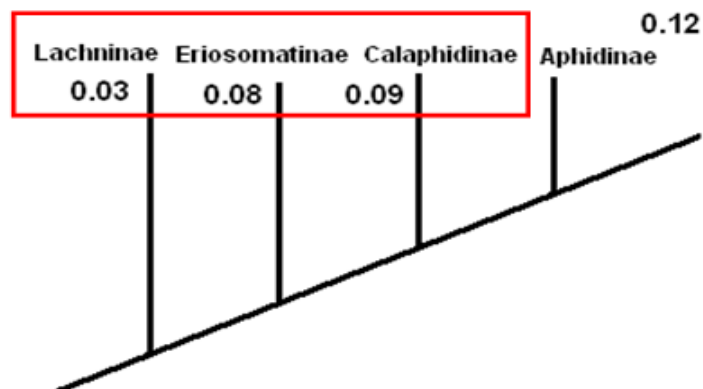
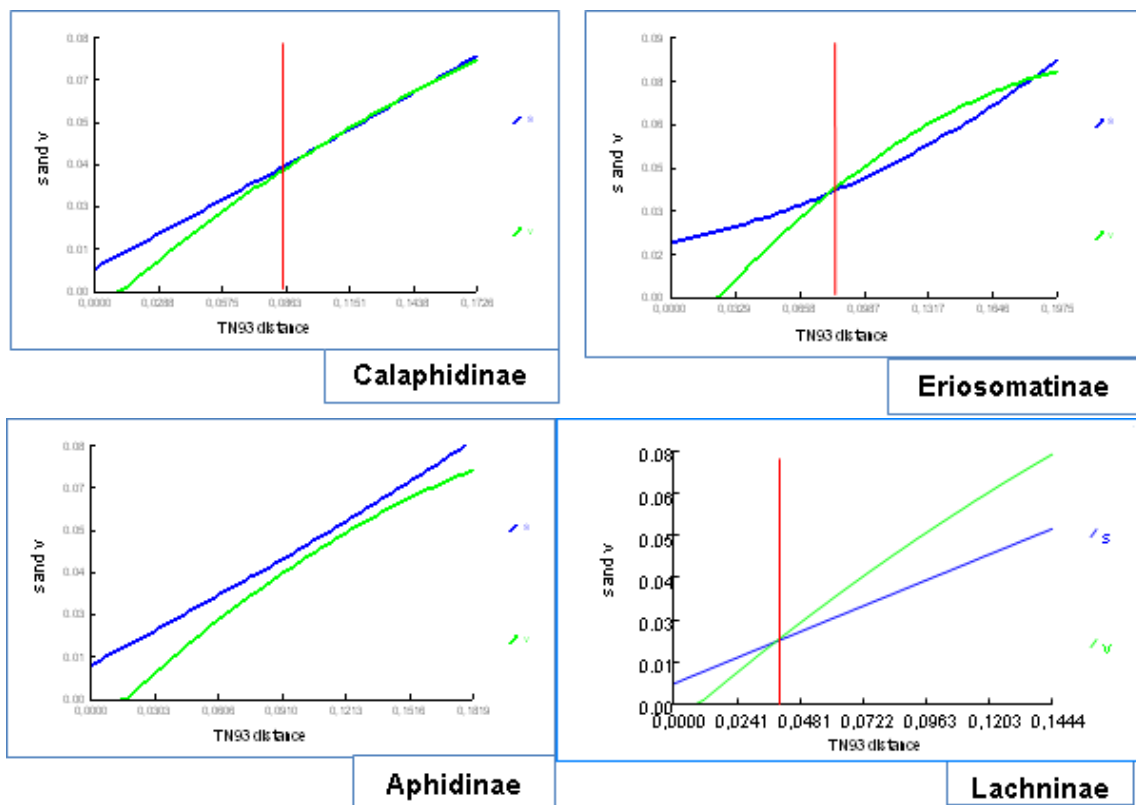


Рисунок 18 – Зависимость выраженность насыщения последовательности COI от возраста таксона у тлей

Таблица 3 – Виды тлей фауны Беларуси, для которых были получены ДНК-штрихкоды

Вид	Шифр	Код доступа в GenBank
<i>Amphorophora idaei</i> Börn.	10–530	JF340095
<i>Anuraphis catonii</i> H.L.R.	10–497	JF340106
<i>Anuraphis farfarae</i> (Koch)	09–31	JF340099
<i>Anuraphis subterranea</i> (Walk.)	10–27	JF340091
<i>Aphis idaei</i> (v. d. Goot)	10–454	JF950580
<i>Aphis fabae cirsiacanthoidis</i> (Scop.)	08–34	JF340105
<i>Aphis fabae fabae</i> Scop.	10–43	JF340098
<i>Aphis fabae philadelphi</i> Börn.	08–21	JN004138
<i>Aphis pomi</i> Deg.	09–10	JF950582
<i>Aphis ruborum</i> Börn.	10–504	JF340108
<i>Aphis vaccinii</i> (Börn.)	10–509	JF340094
<i>Brachycaudus aconiti</i> (Mordv.)	9962	JF340110
<i>Brachycaudus lychnidis</i> (L.)	10–489	JF776570
<i>Brachycaudus aff. prunicola</i> (Kalt.)	10132	JF776571
<i>Dysaphis anthrisci anthrisci</i> (Börn.)	10–492	JF340092
<i>Dysaphis apiifolia petroselini</i> (Börn.)	10089	JF776569
<i>Dysaphis newskyi</i> (Börn.)	10084	JF340113
<i>Dysaphis lappae lappae</i> (Koch)	10138	JF340109
<i>Dysaphis lauberti</i> (Börn.)	10–467	JF340090
<i>Dysaphis ossiannilssoni</i> Stroyan	10078	JF776568
<i>Dysaphis ranunkuli</i> Kalt.	10–516	JF340093
<i>Dysaphis sorbi</i> Kalt.	10–42	JF340097
<i>Macrosiphum cholodkovskyi</i> Mordv.	10–50	JF340103
<i>Macrosiphum gei</i> Koch	09–36	JF340102
<i>Macrosiphum gei</i> Koch	09–27	JF340096
<i>Macrosiphum knautiae</i> Holm.	09–54	JF340101
<i>Macrosiphum melampyri</i> Mordv.	10–487	JF340104
<i>Myzus cerasi cerasi</i> (Fab.)	10–503	JF340107
<i>Myzus cerasi pruniavium</i> (Börn.)	10–466	JF340100
<i>Phorodon humuli</i> Schrk.	01–10	JF950579

Мы выяснили, что у тлей COI обладает хорошей стабильностью, и число гаплотипов приближается к числу анализируемых видов (таблица 4).

Таблица 4 – Молекулярно-филогенетическая характеристика COI-5 региона тлей

	Семейство Aphididae
Количество проанализированных таксонов	340
Количество гаплотипов (Hap)	334
Модель нуклеотидных замещений	TN93+G+I
Количество вариабельных сайтов (V)	228
Количество парсимонично информативных сайтов (P)	182
Количество 4-вырожденных сайтов	73
Количество 2-вырожденных сайтов	62
Количество транзиций (Ti), %	3,945±0,009
Количество трансверсий (Tv), %	3,168±0,011
Ti/Tv смещение (Ti/Tv)	2,31
G+C содержание, %	24,423
Индекс насыщения нуклеотидных замещений (Iss)	0,136
Статистика насыщения последовательности (Iss/Iss.c)	5,294

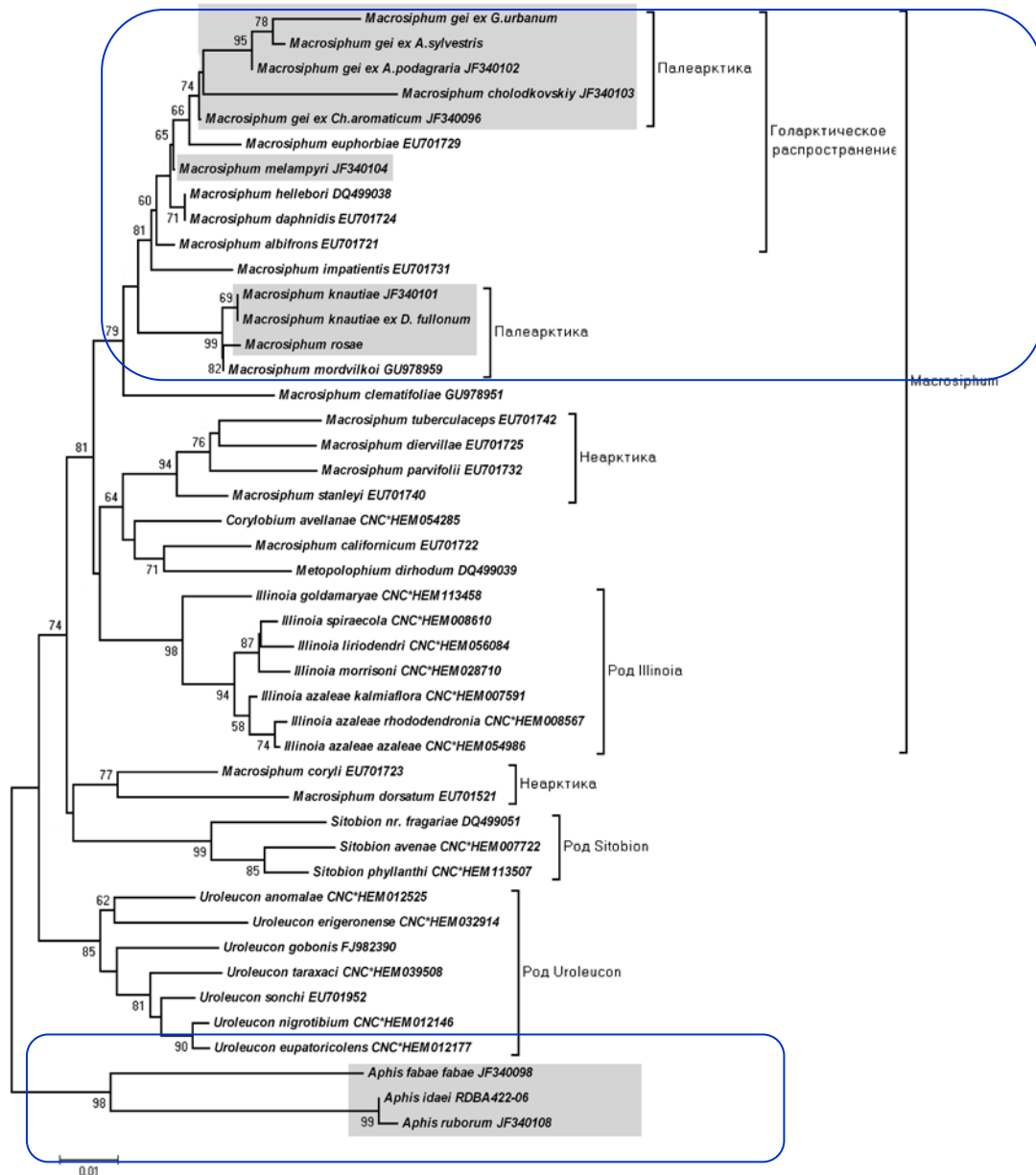


Рисунок 19 – Филогенетические взаимоотношения видов рода *Macrosiphum* Pass., установленные на основе анализа последовательности COI (цветом выделены виды фауны Беларуси)

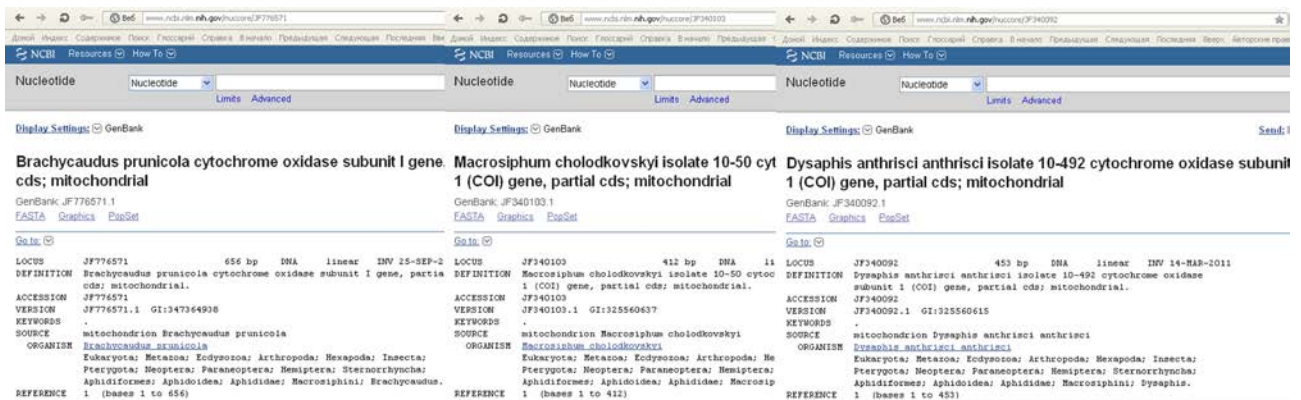


Рисунок 20 – ДНК-паспорта тлей фауны Беларуси

В некоторых случаях наблюдается внутривидовая изменчивость COI, однако, как правило, это виды, в которых имеются описанные подвиды. Нам удалось связать некоторые из этих гаплотипов с подвидами тлей, в частности в видах *Myzus cerasi* s.l., *Aphis fabae* s.l. и *Dyzaphis newskyi* s.l. (рисунки 13, 15) [56].

Нуклеотидная последовательность COI тлей вариабельна на всем протяжении COI-5-секции, однако абсолютное большинство позиций в последовательности обладают высокой определенностью (рисунок 14). Аминокислотная последовательность COI у тлей чрезвычайно консервативна. При сравнении последовательностей 340 видов из 16 подсемейств мы обнаружили только 14 вариабельных позиций в последовательности белка длиной 206 аминокислот. Причем лишь по 6-ти из них вариабельность можно не считать крайне низкой (рисунок 16).

Однако такая консервативность белка в эволюционно древней группе, а возраст семейства не менее 150 млн. лет, привела к насыщению последовательности и разрушению филогенетического сигнала при работе с далекими таксонами тлей. Выражается насыщение последовательности графически как пересечение кривых, отражающих частоту наблюдаемых транзиций и трансверсий (рисунок 17).

Интересно, что у тлей степень насыщения последовательности хорошо коррелирует с возрастом подсемейства. То есть в более молодом подсемействе последовательность не насыщена, в наиболее старом – последовательность насыщена максимально (рисунок 18).

Тем не менее, на уровне родов тлей COI оказалась высоко информативной. В частности нам удалось установить филогенетические взаимоотношения видов рода *Macrosiphum* Pass., используя в качестве единственного филогенетического маркера последовательность COI (рисунок 19) [57]. А также получить генетические паспорта 32 видов и подвидов тлей из числа вредителей сельскохозяйственных культур, позволяющие идентифицировать эти виды при помощи системы BLAST со 100% вероятностью (рисунок 20).

Выводы

Техника геносистематики требует корректного применения маркеров, поскольку использование маркеров, не способных предоставить требуемый филогенетический сигнал, неизбежно приведет к получению невалидных результатов. Как показывает опыт, невозможно а priori предсказать характер поведения COI даже в достаточно близких таксонах, таких как семейства одного отряда. В свете вышесказанного, изучение поведения маркера в каждом конкретном таксоне становится важнейшей частью проведения филогенетических исследований. Мы надеемся, что такие работы будут появляться, и это, в конечном итоге, позволит синтезировать данные о COI в систему знаний об эволюции самого гена, что может быть крайне интересно не только для геносистематики, но и для понимания механизмов поддержания функций столь важных для самых разных организмов белков на протяжении миллионов лет эволюции.

Благодарности

Работа выполнена при частичной поддержке БРФФИ проекта Б11АРМ-015.

Список литературы

1. Pfenninger, M. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions / M. Pfenninger, K. Schwenk // BMC Evolutionary Biology. – 2007. – Vol. 7, № 121. – P. 1–6.
2. Review on global co-transmission of human *Plasmodium* species and *Wuchereria bancrofti* by *Anopheles* mosquitoes / S. Manguin [et al.] // Infection, Genetics and Evolution. – 2010. – № 10. – P. 159–177.
3. Biological identifications through DNA barcodes / D. Paul [et al.] // Proc. R. Soc. Lond. B. – 2003. – № 270. – P. 313–321.

4. Ratnasingham, S. BOLD: the barcode of life data system (www.barcodinglife.org) / S. Ratnasingham, P.D.N. Hebert // *Mol. Ecol. Notes.* – 2007. – № 7. – P. 355–364.
5. Stoeckle, M. Barcode of Life / M. Stoeckle, P.D.N. Hebert // *Scientific American.* – 2008. – № 299. – P. 65–71.
6. Ratnasingham, S. BOLD's role in barcode data management and analysis / S. Ratnasingham, P.D.N. Hebert // *Mol. Ecol. Res.* – 2011. – № 11. – P. 941–942.
7. Smith, V.S. DNA barcoding: perspectives from a «Partnerships for enhancing expertise in taxonomy» (PEET) debate / V.S. Smith // *Syst. Biol.* – 2005. – Vol. 54, № 5. – P. 841–844.
8. Moritz, C. DNA barcoding: promise and pitfalls / C. Moritz, C. Cicero // *PLoS Biology.* – 2004. – Vol. 2, Iss. 10. – P. e279–e354.
9. Pires, A.C. DNA barcoding and traditional taxonomy unified through Integrative Taxonomy: a view that challenges the debate questioning both methodologies / A.C. Pires, L. Marinoni // *Biota Neotrop.* – 2010. – Vol. 10, № 2. – P. 339–346.
10. DNA Barcoding the Native Flowering Plants and Conifers of Wales / de Vere [et al.] // *PLoS ONE.* – 2012. – Vol. 7, Iss. 6. – P. e37945.
11. ITS as an environmental DNA barcode for fungi: an *in silico* approach reveals potential PCR biases / E. Bellemain [et al.] // *BMC Microbiology.* – 2010. – Vol. 10, № 189. – P. 1–9.
12. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates / O. Folmer [et al.] // *Molecular Marine Biology and Biotechnology.* – 1994. – Vol. 3, № 5. – P. 294–299.
13. The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies / D.H. Lunt [et al.] // *Insect Molecular Biology.* – 1996. – Vol. 5, № 3. – P. 153–165.
14. Countering criticisms of single mitochondrial DNA gene barcoding in birds / A.J. Baker [et al.] // *Molecular Ecology Resources.* – 2009. – Vol. 9. – P. 257–268.
15. DNA barcoding and the mediocrity of morphology / L. Packer [et al.] // *Molecular Ecology Resources.* – 2009. – Vol. 9. – P. 42–50.
16. DNA barcoding identifies *Eimeria* species and contributes to the phylogenetics of coccidian parasites (Eimeriorina, Apicomplexa, Alveolata) / J.D. Ogedengbe [et al.] // *International Journal for Parasitology.* – 2011. – Vol. 41. – P. 843–850.
17. Struder-Kypke, M.C. Comparative analysis of the mitochondrial cytochrome c oxidase subunit I (COI) gene in ciliates (Alveolata, Ciliophora) and evaluation of its suitability as a biodiversity marker / M.C. Struder-Kypke, D.H. Lynn // *Systematics and Biodiversity.* – 2010. – Vol. 8, № 1. – P. 131–148.
18. COI gene and ecological data suggest size-dependent high dispersal and low intra-specific diversity in free-living terrestrial protists (Euglyphida: Assulina) / E. Lara [et al.] // *Journal of Biogeography.* – 2011. – Vol. 38. – P. 640–650.
19. Neotypification of *Drawida hattamimizu* Hatai, 1930 (Annelida, Oligochaeta, Megadrili, Moniligastridae) as a model linking mtDNA (COI) sequences to an earthworm type, with a response to the 'Can of Worms' theory of cryptic species / R.J. Blakemore [et al.] // *ZooKeys.* – 2010. – Vol. 41. – P. 1–29.
20. Sequence amplification and gene rearrangement in parasitic nematode mitochondrial DNA / B.C. Hyman [et al.] // *Genetics.* – 1998. – Vol. 140. – P. 707–712.
21. Analysis of cytochrome c-oxidase (COI) gene of mitochondrial DNA from the *Trichinella* spp. in China / Y. Yang [et al.] // *Parasitol. Res.* – 2008. – Vol. 103. – P. 1355–1359.
22. Kornobis, E. Discordance in variation of the ITS region and the mitochondrial COI gene in the subterranean amphipod *Crangonyx islandicus* / E. Kornobis, S. Palsson // *J. Mol. Evol.* – 2011. – Vol. 73. – P. 34–44.
23. Mitochondrial cytochrome oxidase I gene sequences support an Asian origin for the Portuguese oyster *Crassostrea angulate* / D.O.A. Foighil [et al.] // *Marine Biology.* – 1998. – Vol. 131. – P. 497–503.

24. Evolutionary relationships among squids of the family Gonatidae (Mollusca: Cephalopoda) inferred from three mitochondrial loci / A.R. Lindgren [et al.] // *Molecular Phylogenetics and Evolution*. – 2005. – Vol. 36. – P. 101–111.
25. Limitations of mitochondrial gene barcoding in Octocorallia / C.S. McFadden [et al.] // *Molecular Ecology Resources*. – 2011. – Vol. 11. – P. 19–31.
26. *Ixodes philipi* (Acari: Ixodidae): phylogenetic status inferred from mitochondrial cytochrome oxidase subunit I gene sequence comparison / H. Mitani [et al.] // *J. Parasitol.* – 2007. – Vol. 93, № 3. – P. 719–722.
27. Phylogenetic relationship between *Dermanyssus gallinae* populations in European countries based on mitochondrial COI gene sequences / M. Marangi [et al.] // *Exp. Appl. Acarol.* – 2009. – Vol. 48. – P. 143–155.
28. Gantenbeina, B. The phylogeographic importance of the Strait of Gibraltar as a gene flow barrier in terrestrial arthropods: a case study with the scorpion *Buthus occitanus* as model organism / B. Gantenbeina, C.R. Lariader // *Molecular Phylogenetics and Evolution*. – 2003. – Vol. 28. – P. 119–130.
29. An integrative approach to species discovery in odonates: from character-based DNA barcoding to ecology / S. Damm [et al.] // *Molecular Ecology*. – 2010. – Vol. 19. – P. 3881–3893.
30. Using COI gene sequence to barcode two morphologically alike species: the cotton bollworm and the oriental tobacco budworm (Lepidoptera: Noctuidae) / Q.Q. Li [et al.] // *Mol. Biol. Rep.* – 2011. – Vol. 38. – P. 5107–5113.
31. Evolution of the hymenopteran megaradiation / J. Heraty [et al.] // *Molecular Phylogenetics and Evolution*. – 2011. – Vol. 60. – P. 73–88.
32. Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA / J.M. Cornuet [et al.] // *Genetics*. – 1991. – Vol. 1128. – P. 393–403.
33. Magnacca, K.N. Mitochondrial heteroplasmy and DNA barcoding in Hawaiian *Hylaeus (Nesoprosopis)* bees (Hymenoptera: Colletidae) / K.N. Magnacca, M.J.F. Brown // *BMC Evolutionary Biology*. – 2010. – Vol. 10, № 174. – P. 1–16.
34. Geographic variation in host selection behaviour and reproductive success in the stemborer parasitoid *Cotesia flavipes* (Hymenoptera: Braconidae) / R.P.J. Potting [et al.] // *Bulletin of Entomological Research*. – 1997. – Vol. 87, Iss. 5. – P. 515–524.
35. Geographic homogeneity and high gene flow of the pear psylla, *Cacopsylla pyricola* (Hemiptera: Psyllidae), detected by mitochondrial COI gene and nuclear ribosomal internal transcribed spacer 2 / A.R. Kanga [et al.] // *Animal Cells and Systems*. – 2012. – Vol. 16, № 2. – P. 145–153.
36. COI barcoding of true bugs (Insecta, Heteroptera) / S. Jang [et al.] // *Mol. Ecol. Resour.* – 2011. – Vol. 11, № 2. – P. 266–270.
37. A new mtDNA COI gene lineage closely related to *Anopheles janconnae* of the Albitarsis complex in the Caribbean region of Colombia / L.A. Gutierrez [et al.] // *Mem. Inst. Oswaldo Cruz, Rio de Janeiro*. – 2010. – Vol. 105, № 8. – P. 1019–1025.
38. Identification of forensically important sarcophagid flies (Diptera: Sarcophagidae) based on COI gene in China / G. Yadong [et al.] // *Rom. J. Leg. Med.* – 2010. – Vol. 18. – P. 217–224.
39. Phylogenetic relationships and the evolution of mimicry in the *Chauliognathus* yellow-black species complex (Coleoptera: Cantharidae) inferred from mitochondrial COI sequences / V. Machado [et al.] // *Genetics and Molecular Biology*. – 2004. – Vol. 27, № 1. – P. 55–60.
40. COI (cytochrome oxidase-I) sequence based studies of Carangid fishes from Kakinada coast, India / M. Persis [et al.] // *Mol. Biol. Rep.* – 2009. – Vol. 36. – P. 1733–1740.
41. Tavares, E.S. Single mitochondrial gene barcodes reliably identify sister-species in diverse clades of birds / E.S. Tavares, A.J. Baker // *BMC Evolutionary Biology*. – 2008. – Vol. 8, № 81. – P. 1–14.
42. Molecular identification of three *Ompok* species using mitochondrial COI gene / A.K. Malakar [et al.] // *Mitochondrial DNA*. – 2012. – Vol. 23, № 1. – P. 20–24.

43. Mitochondrial cytochrome c subunit 1 sequence variation in prostate cancer / T.A. Scott [et al.] // *Scientifica*. – 2012. – ID 701810. – P. 1–7.
44. Segregation of point mutation heteroplasmy in the control region of dog mtDNA studied systematically in deep generation pedigrees / C.F. Klutsch [et al.] // *International Journal of Legal Medicine*. – 2011. – Vol. 125, Iss. 4. – P. 527–535.
45. Wiemers, M. Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae) / M. Wiemers, K. Fiedler // *Frontiers in Zoology*. – 2007. – Vol. 4, № 8. – P. 1–16.
46. Funk, D.J. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA / D.J. Funk, K.E. Omland // *Annu. Rev. Ecol. Evol. Syst.* – 2003. – Vol. 34. – P. 397–423.
47. Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores / A.R. Hoelzel [et al.] // *J. Mol. Evol.* – 1994. – Vol. 39. – P. 191–199.
48. Mitochondrial heteroplasmy and the evolution of insecticide resistance: Non-Mendelian inheritance in action / T. Van Leeuwen [et al.] // *PNAS*. – 2008. – Vol. 105, № 16. – P. 5980–5985.
49. Meusel, M.S. Transfer of paternal mitochondrial DNA during fertilization of honeybee (*Apis mellifera* L.) eggs / M.S. Meusel, R.F. Moritz // *Curr. Genet.* – 1993. – Vol. 24, Iss. 6. – P. 539–543.
50. St. John, J.C. Paternal mitochondrial DNA transmission during nonhuman primate nuclear transfer / J.C. St. John, G. Schatten // *Genetics*. – 2004. – Vol. 167. – P. 897–905.
51. Detection of mitochondrial insertions in the nucleus (NuMts) of Pleistocene and modern muskoxen / S.O. Kolokotronis [et al.] // *BMC Evolutionary Biology*. – 2007. – Vol. 7, № 67. – P. 1–10.
52. Ancient mtDNA sequences in the human nuclear genome: a potential source of errors in identifying pathogenic Mutations / D.C. Wallace [et al.] // *PNAS*. – 1997. – Vol. 94, № 26. – P. 14900–14905.
53. Comparisons of host mitochondrial, nuclear and endosymbiont bacterial genes reveal cryptic fig wasp species and the effects of *Wolbachia* on host mtDNA evolution and diversity / X.J. Sun [et al.] // *BMC Evolutionary Biology*. – 2011. – Vol. 11, № 86. – P. 1–8.
54. Characterizing the time dependency of human mitochondrial DNA mutation rate estimates / B.M. Henn [et al.] // *Mol. Biol. Evol.* – 2009. – Vol. 26. – P. 217–230.
55. Воронова, Н.В. Выявление криптических подвидов в эволюционно молодых группах тлей (Rhynchota: Homoptera: Aphididae) с использованием гена субъединицы I цитохромоксидазы *c* (*COI*) в качестве филогенетического маркера / Н.В. Воронова, С.В. Буга, В.П. Курченко // Доклады НАН Беларуси. – 2011. – Т. 55, № 3. – С. 97–102.
56. Воронова, Н.В. Филогения тлей рода *Macrosiphum* Pass. (Rhynchota: Homoptera: Aphididae) на основе анализа нуклеотидной последовательности гена субъединицы 1 цитохромоксидазы *c* (*COI*) / Н.В. Воронова, С.В. Буга, В.П. Курченко // Труды Белорусского государственного университета. – 2011. – Т. 6, ч. 1. – С. 193–200.

COI-5 REGION AS A MARKER FOR MOLECULAR TAXONOMY OF ANIMALS: APPROACHES, OUTCOMES AND CONSTRAINTS

N.V. Voronova, S.V. Buga, V.P. Kurchenko

Belarusian State University, Minsk, Belarus

In this paper we described the examples of using of sequences of the mitochondrial COI gene in molecular taxonomic studies. The efficiency of COI as a phylogenetic marker in 42 taxa belonging to different types of the kingdom Animalia was analyzed. We discussed the reasons which may lead to poor performance of COI in some animal taxa. The data on efficiency of COI in aphids are given.