САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

На правах рукописи

Паженкова Елена Алексеевна

Интегративный анализ сложных в таксономическом отношении групп нимфалоидных чешуекрылых (Lepidoptera, Nymphalidae)

Научная специальность 1.5.14. Энтомология

Диссертация на соискание ученой степени кандидата биологических наук

Научный руководитель:
доктор биологических наук
Лухтанов Владимир Александрович

Санкт-Петербург
2021

Оглавление

Введение 3
Глава 1. История и перспективы применения интегративного анализа в таксономии....10
1.1 Применение ДНК-баркодинга в исследованиях биологического разнообразия 10
1.2 Интегративный подход при делимитации видов....13
1.3 Применение полногеномных данных в филогенетике и таксономии 15
Глава 2. ДНК-баркоды как инструмент для постановки таксономических гипотез....18
2.1 Анализ комплекса *Melitaea ala*....18
2.2. Анализ комплекса *Melitaea didyma* 21
Глава 3. Интегративный таксономический анализ на примере рода *Brenthis* и комплекса *Hyponephele lycaon – H. lupina*....27
3.1 Анализ рода *Brenthis*....27
3.2 Анализ комплекса *Hyponephele lycaon – H. lupina* 35
Глава 4. Использование полных геномов для выявления таксономической структуры и эволюционной истории комплекса *Melitaea acentria-M. persea- M. didyma* 43
4.1. Паттерны генетической дифференциации и геномная интрогрессия 43
4.2 Видовая самостоятельность *Melitaea acentria*, *M. persea* и *M. didyma* 48
4.3 Происхождение вида *M. acentria*....49
Заключение....51
Благодарности 53
Список литературы....54

Введение

Актуальность и степень изученности темы исследования

Таксономия является одной из основных составляющих биологии, поскольку делимитация видов и отнесение их к категориям более высокого ранга (родам, трибам и т.д.) играют ключевую роль в изучении биологического разнообразия. Классификация и идентификация организмов долгое время были основаны на изучении морфологических признаков, которые дополнялись экологическими, этологическими, биогеографическими, эмбриологическими и физиологическими данными. С появлением и развитием молекулярных методов исследователи получили доступ к данным нового типа, которые стали активно внедряться в том числе и в таксономию.

Термин «интегративная таксономия» формально был введен для обозначения комплексного подхода к описанию и делимитации видов с применением различных типов данных и методик сравнительно недавно [1,2], однако преимущества совместного анализа молекулярных и морфологических данных отмечались еще в конце прошлого века (например [3,4]). Интегративный подход позволяет эффективно провести процесс идентификации и делимитации видов в случаях, когда морфологические признаки не показательны [1,5], а также разрешить спорные ситуации при анализе сложных групп [6], обнаружить новые криптические виды [7] и провести их филогеографический анализ, то есть выявить историю возникновения и формирования таксонов [8,9]. Кроме того, зачастую комплексный подход позволяет ускорить процесс видовой идентификации, что особенно важно при исследовании богатых видами, но недостаточно изученных групп организмов и экосистем, например при анализе разнообразия морских [10] и тропических сообществ [11,12].

Необходимо учитывать, что различные типы данных могут поддерживать противоречащие друг другу таксономические гипотезы. Поэтому основным предметом дискуссий в интегративной таксономии является степень соответствия между разными типами признаков, при которой разные популяции или группы популяций могут расцениваться как отдельные виды. Некоторые авторы считают конгруэнтность между различными типами признаков необходимым условием для делимитации видов [1,13], в то время как другие предполагают, что любые несоответствия объясняются с точки зрения эволюционной истории выбранных групп, а для принятия таксономического решения может быть достаточно даже одного набора признаков, если он считается достаточным, чтобы обосновать обособленность таксона [14,15]. Довольно распространенным феноменом является несоответствие филогенетических реконструкций и, соответственно, таксономических заключений, основанных на анализе ядерных и митохондриальных молекулярных маркеров (mito-nuclear

discordance) [16,17]. Особенно часто подобные противоречия появляются при изучении молодых таксонов, где нередко наблюдается ретикулярная эволюция (reticulation events), подразумевающая такие процессы как межвидовая гибридизация и горизонтальный перенос генов. В таких случаях реконструкция эволюционной истории видов на основании отдельных немногочисленных маркеров, как правило, невозможна; однако, эта проблема может быть решена при помощи анализа полных геномов [17].

В настоящее время в исследованиях, ключевые слова которых включают в себя термин «интегративная таксономия», почти в половине случаев использовалось лишь два типа признаков (преимущественно ДНК-баркоды и морфология) [18]. Что примечательно, в 35,4 % случаев анализ проводился с использованием только одного типа признаков (в большинстве случаев, молекулярных). Это может быть связано с тем, что некоторые авторы понимают под термином «интегративная таксономия» использование различных критериев и концепций вида, а не типов данных [18]. Таким образом, разработка терминологии и методологии интегративного анализа является одним из основополагающих направлений в современных исследованиях биологического разнообразия, что говорит об актуальности выполненной диссертации.

Модельными объектами настоящей работы стали дневные бабочки семейства Nymphalidae. Несмотря на сравнительно хорошую изученность этого семейства, таксономические отношения во многих группах недостаточно исследованы в связи с высокой географической или сезонной вариабельностью, наличием криптических видов или межвидовой гибридизацией [17,19,20].

В рамках диссертационного исследования были изучены представители родов *Melitaea*, *Brenthis* и *Hyponephele*, представляющих два подсемейства – Nymphalinae (*Melitaea*, *Brenthis)* и Satyrinae (*Hyponephele*) с использованием трех возможных подходов: (1) анализ ДНК-баркодов (фрагментов гена *COI*) для постановки первичных таксономических гипотез; (2) анализ комбинации ДНК-баркодов, набора ядерных маркеров и морфологических признаков для тестирования первичных таксономических гипотез; (3) анализ полногеномных данных для изучения видового комплекса, в котором обнаружена межвидовая интрогрессия.

Выбор модельных объектов определялся тем, что *Brenthis* – это таксономически один из наиболее изученных родов дневных бабочек, с малым числом видов, причем число и границы видов в пределах рода хорошо известны на основании традиционного морфологического анализа [21,22] и не вызывают сомнения у лепидоптерологов. На примере такого рода, нам казалось интересным проверить результаты делимитации, основанной на молекулярных маркерах. *Melitaea* и *Hyponephele*, напротив, представляют две богатые видами группы бабочек,

число видов и границы видов в пределах которых неясны и являются предметом дискуссий между учеными [20,23,24].

Научная новизна работы

Впервые для дневных бабочек в комплексе *Melitaea didyma* выявлен случай крайне выраженного митохондриального полиморфизма [25,26], когда в пределах одного вида встречается до 11 митохондриальных гаплогрупп, уровень дивергенции между которыми достигает 7,4 %.

Впервые получены нуклеотидные последовательности генов *COI*, *wgl*, *Ca-ATPase, ArgK* и *CAD* для видов рода *Brenthis* и *Hyponephele,* послужившие основой для последующего таксономического анализа [27].

С использованием интегративного подхода выявлена таксономическая структура комплексов видов, близких к *Melitaea ala* [28] и *Hyponephele lycaon* [29].

Отсеквенировано и проанализировано 27 полных геномов бабочек комплекса *M. didyma – M. persea – M. acentria*. Впервые для бабочек рода *Melitaea* показана геномная интрогрессия между неблизкородственными видами – явление, которое крайне редко наблюдается в природе [30].

Теоретическая и практическая значимость работы

Разработка методологии интегративного таксономического анализа имеет ключевое значение для изучения биологического разнообразия. Применение различных подходов к анализу групп дневных бабочек с разной степенью изученности позволяет не только исследовать таксономические взаимоотношения в этих группах, но и апробировать методы, которые впоследствии могут быть использованы при изучении других организмов. Были установлены видовые границы таксонов в комплексе *Hyponephele lycaon – H. lupina* и в роде *Brenthis*. Детальное изучение эволюционной истории таксонов комплекса *Melitaea didyma – M. persea – M. acentria* с применением полногеномных данных показало наличие однонаправленного потока генов от *M. didyma* к *M. acentria*. Установлено, что этот поток генов является дополнительным источником генетического разнообразия в локальной популяции M. *acentria*. Вкладом в эволюционную биологию является открытие геномной интрогрессии между филогенетически далекими видами (*M. didyma* и *M. acentria* дивергировали болеее чем 5 млн лет назад). Значительно пополнена база последовательностей гена *COI*, *ArgK*, *CAD*, *Ca-ATPase*, *wgl*, а также полных геномов для бабочек семейства *Nymphalidae*, что может быть использовано в дальнейших исследованиях. Разработаны специфичные праймеры для генов *ArgK* и *CAD*, специфичные для бабочек семейства *Nymphalidae* [27]. Таксономические исследования имеют

потенциальное природоохранное значение, так как ряд исследованных видов (например, *Melitaea acentria*, *Hyponephele galtcha*) являются малочисленными эндемиками и подлежат охране. Материалы диссертации могут быть использованы при подготовке учебных курсов по энтомологии, генетике и биоинформатике.

Цель и задачи

Целью данной работы является изучение таксономической структуры сложных групп нимфалоидных чешуекрылых родов *Brenthis*, *Melitaea* и *Hyponephele* (Lepidoptera, Nymphalidae) с применением интегративного подхода.
Для достижения цели были поставлены следующие **задачи**:

- Изучить видовую и подвидовую структуру таксонов комплекса видов, близких к *Melitaea ala.*
- Провести первичный таксономический анализ комплекса видов, близких к *Melitaea didyma,* с использованием ДНК-баркодов.
- Проанализировать случай несоответствия филогенетических реконструкций и, соответственно, таксономических заключений, основанных на анализе митохондриальных и ядерных маркеров, на примере рода *Brenthis*.
- Провести таксономический анализ видов группы *Hyponephele lycaon – H. lupina c* использованием мультилокусных генетических маркеров, морфологических признаков и биогеографических данных.
- С использованием полногеномных данных осуществить делимитацию видов и выявить эволюционную историю для видовой группы, в которой подтверждена ретикулярная эволюция (на примере комплекса видов *Melitaea didyma – M. persea – M. acentria*).

Положения, выносимые на защиту

1. Анализ ДНК-баркодов позволил выявить три подвида в составе вида *Melitaea ala* (*M. ala ala*, *M. ala zaisana, M. ala bicolor)*.
2. Комплекс *Melitaea didyma* обладает высоким уровнем митохондриального полиморфизма: в пределах одного вида *M. didyma* sensu stricto обнаруживается 11 дивергировавших митохондриальных генетических линий; 12 линий, выявленных внутри комплекса, соответствуют девяти другим ранее описанным видам.

3. Внутри рода *Brenthis* филогенетические взаимоотношения и группировки особей, обнаруженные с помощью митохондриальных ДНК-баркодов (фрагмент гена *COI*) не соответствуют филогении и таксонам, выявляемым на основании анализа ядерных маркеров и морфологии. В то же время ядерные маркеры и морфологические признаки дают полностью конгруэнтные результаты при выявлении видовых границ.
4. Интегративный анализ ДНК-баркодов, ядерных генов, ареалов, окраски крыльев и строения гениталий самцов позволяет определить границы таксонов в комплексе видов, близких к *Hyponephele lycaon – H. lupina*.
5. Анализ полных геномов бабочек комплекса *Melitaea didyma – M. persea – M. acentria* показал, что они являются отдельными видами, несмотря на неполную репродуктивную изоляцию между *M. didyma* и *M. acentria.*
6. Вид *Melitaea acentria* возник в Леванте как периферический географический изолят *M. persea.* Геномная интрогрессия от неблизкородственного вида *Melitaea didyma* являлась источником генетического разнообразия в локальной популяции *M. acentria*. Половые хромосомы *M. acentria* более устойчивы к интрогрессии, чем аутосомы и играют роль в формировании постзиготических межвидовых репродуктивных барьеров.

Материалы и методы

Основой для исследования послужили коллекционные материалы, хранящиеся в отделении Кариосистематики Зоологического института РАН. Проводились дополнительные сборы свежего материала во время экспедиций по России (Самарская область, Оренбургская область, Башкортостан), Болгарии, Таджикистану, Монголии, Ирану и Израилю. Сбор насекомых проводился при помощи энтомологического сачка, насекомые хранились на ватных матрасиках. Материал для полногеномного анализа фиксировался при помощи Allprotect Tissue reagent (Qiagen), чтобы обеспечить наилучшую сохранность, и долговременно хранился при температуре -80°C. Подробный список использованного материала и точки сбора приводятся в наших публикациях [25–30].

Подготовка препаратов гениталий самцов проводилась в соответствии со следующим протоколом: конец брюшка имаго вываривался в горячем 10% растворе KOH в течение 3-10 минут. Затем брюшко промывалось водой и препарировалось под световым микроскопом. Готовые препараты хранятся в глицерине в плотно закрытых пробирках. Фотографии генитальных структур были сделаны при помощи бинокуляра Leica M205C, оснащенного камерой Leica DFC495. Контуры ункуса и субунций для рода *Hyponephele* были подготовлены в программе GIMP [31]. Геометрическая морфометрия проводилась при помощи библиотеки

Momocs v1.2.9 [32]. Сначала контуры были центрированы, выровнены и приведены к одному размеру, затем применялось дискретное косинусное преобразование, чтобы получить коэффициенты Фурье [33]. Для оценки различий между группами применялся анализ главных компонент (principal component analysis, PCA).

Кладистический анализ бабочек рода *Brenthis* проводился на основании анализа матрицы, содержащей 11 фенотипических признаков с использованием методов максимальной парсимонии в программе PAUP* [34] и Байесова анализа в программе MrBayes v. 3.1.2 [35]. Подробное описание признаков и проведенного анализа приводится в [27].

Подготовка образцов к секвенированию проводилась в отделении Кариосистематики ЗИН РАН и включала в себя экстракцию ДНК, проведение ПЦР, очистку ПЦР-продуктов и анализ полученных результатов при помощи гель-электрофореза (детальное описание см. в наших публикациях [25,27]). Секвенирование проводилось в ресурсном центре СПбГУ «Развитие молекулярных и клеточных технологий». Часть материала для проведения секвенирования была отправлена в Канадский центр ДНК-баркодинга (Canadian Centre for DNA Barcoding, Biodiversity Institute of Ontario, University of Guelph, Canada). Всего было получено 558 последовательностей фрагментов гена *COI* (ДНК-баркодов) и последовательности ядерных генов для 117 образцов (что составляет около 400 последовательностей).

Обработка и выравнивание сиквенсов осуществлялась в программе BioEdit 7.1.7 [36]. Выравнивание проводилось с использованием алгоритма ClustalW [37]. Построение дендрограмм проводилось при помощи Байесова анализа (Bayesian analysis, BI), который осуществлялся в программе MrBayes v. 3.1.2 [35]. Подробнее в [25–27,30].

Подготовка образцов (экстракция ДНК и создание библиотек) и секвенирование полных геномов осуществлялось в компании Макроген, Корея на коммерческой основе. Библиотеки Illumina TruSeq DNA с размером вставки 350 п. н. были отсеквенированы на платформе Illumina HiSeq X Ten с длинной ридов 2x150 п. н. В среднем, было получено 9 млрд п. н. для каждого образца.

Последующий биоинформатический анализ проводился на базе ресурсного центра СПбГУ «Вычислительный центр». Было осуществлено выравнивание прочтений на референсный геном *Melitaea cinxia* v. 1 [38] с последующим поиском однонуклеотидных замен (single nucleotide polymorphism, SNP) при помощи Genome Analysis Toolkit v. 3.7 (GATK) [39]. После фильтрации финальный набор SNP включал в себя 768103 замены высокого качества. Анализ генетической структуры популяции проводился с помощью метода главных компонент (PCA), включенного в библиотеку SNPRelate [40]. Реконструкция событий сетчатой эволюции проводилась с помощью программы SplitsTree 4.16.2 [41]. Были рассчитаны такие показатели, как индекс фиксации F_{ST} и нуклеотидное разнообразие π по всему геному при помощи скрипта

из [42]. Осуществлена оценка направления и уровня интрогрессии с помощью программы dsuite [43]. Проведена реконструкция демографической истории с помощью анализа аллель-частотных спектров в программе δaδi [44]. Филогеографический анализ проводился для маркера *COI* в программе RASP v.3.2 [45]. Филогенетическая сеть для полных митохондриальных геномов была построена в программе Fitchi [46]. Методы полногеномного анализа подробно описаны в [30].

Публикации и апробация результатов

Результаты исследования были изложены на 5 международных конференциях: VI Международная конференция по кариосистематике беспозвоночных животных KARYO-VI (Россия, Саратов, 27-30 августа 2016 г.), VIII International Conference Biology of Butterflies (Bangalore, India, 11-14 June 2018), RECOMB Comparative Genomics 17th (Montpellier, France, 1-4 October 2019), Международный научный форум «Ломоносов-2020» (Москва, Россия 13-17 апреля 2020), the 26th European Meeting of PhD Students in Evolutionary Biology (EMPSEB) (Online, Ireland, 1-4 March 2021).

По теме диссертации было опубликовано 6 научных статей в изданиях, входящих в системы WoS и/или Scopus, из них 3 в журналах первого квартиля SJR, 2 в журнале второго квартиля SJR по генетике или зоологии.

Результаты диссертации были представлены на регулярных заседаниях кафедры энтомологии Санкт-Петербургского государственного университета.

Глава 1. История и перспективы применения интегративного анализа в таксономии

1.1 Применение ДНК-баркодинга в исследованиях биологического разнообразия

Использование молекулярных методов для решения таксономических проблем началось во второй половине прошлого века и получило широчайшее распространение в течение последних десятилетий. Многие исследователи считали перспективным применение митохондриальной ДНК для решения самых разных научных задач по нескольким причинам [47]. Первая — это относительно консервативное устройство митохондриального генома: небольшой размер, константное расположение и состав генов, отсутствие между ними спейсеров. Во-вторых, митохондрии передаются по материнской линии и митохондриальная ДНК не участвует в рекомбинации. В-третьих, считается, что скорость эволюционных изменений митохондриальной ДНК достаточно высока, чтобы можно было использовать ее в качестве маркера в исследованиях в области популяционной экологии, при реконструкции филогении и в филогеографии [9].

Для видовой идентификации ДНК использовали еще до начала 21 века — например, для быстрого определения видов личинок мясных мух *Phormia regina, Phaenicia sericata* и *Lucilia illustris* (Diptera, Calliphoridae), развивающихся в мертвом теле в разное время, что позволяет криминалистам относительно точно установить время смерти [48]. Был разработан метод идентификации для определения видовой принадлежности осетровых рыб, представляющих экономическую ценность [49]. Однако, определение видов, основанное на анализе последовательностей ДНК, приобрело огромную популярность только после публикаций канадского биолога Пола Эбера [50], который не только предложил использовать короткие ДНК-баркоды для видовой идентификации, но и создал постоянно пополняющуюся базу данных, в которой, по замыслу ученого, будут собраны ДНК-баркоды всех живых организмов. Для животных, в том числе насекомых, в качестве ДНК-баркодов используется фрагмент митохондриального гена *COI*, для растений –пластидные гены *matK* и *rbcL*, для грибов – ядерные спейсеры *ITS*. Чтобы подтвердить эффективность предложенного метода, Эбер с коллегами публикует несколько статей, где демонстрирует успешное применение ДНК-баркодинга для различных групп животных — птиц [51], коллембол [52], паукообразных [53], поденок [54], рыб [55].

При работе с последовательностями *COI* в качестве эмпирически установленного порога, позволяющего провести делимитацию видов, используют p-дистанции, превышающие 0,02 (p-дистанция - отношение числа замен в двух сравниваемых сиквенсах к общему числу нуклеотидов в изучаемой последовательности) [50]. Однако, существуют таксоны, у которых внутривидовые различия превосходят пороговое значение, а межвидовые достаточно низки [56]. Таким образом, не всегда можно говорить о наличии межвидового разрыва в ДНК-

баркодах (этот разрыв в англоязычной литературе называется «barcoding gap»), то есть о том, что межвидовая изменчивость превышает уровень внутривидовых различий [56,57]. Альтернативными подходами к разграничению видов являются выявление на филогенетическом дереве групп особей, образующих монофилетические линии, или поиск комбинаций видоспецифичных нуклеотидных замен [58]. Существует несколько эвристических подходов, которые применяются для делимитации видов и направлены на подбор оптимальных параметров для проведения межвидовых границ, таких как Automated barcoding gap discovery (ABGD) [59], General Mixed Yule Coalescent (GMYC) [60] или Poisson Tree Process (PTP) [61]. При наличии большой базы референсных последовательностей применим вероятностный алгоритм PROTAX [62], который не зависит от предустановленного межвидового порога. Этот метод доказал свою эффективность при широкомасштабном анализе европейских дневных бабочек, в ходе которого были получены и проанализированы ДНК-баркоды для 97% видов, зарегистрированных в Европе [63].

Несомненным преимуществом ДНК-баркодинга является то, что для проведения ПЦР и последующего секвенирования достаточно нескольких нанограмм материала [64], и существует опыт выделения ДНК из следов на снегу [65], высушенного яда змей [66], слизи, выделяемой морскими моллюсками на литорали [67] и даже из воздуха [68]. С развитием технологий секвенирования второго поколения стало возможным проводить одновременное прочтение множества образцов на платформе Illumina, что значительно повысило эффективность анализа [69–71]. Секвенирование третьего поколение, в частности, платформа MinION компании Oxford Nanopore позволяет не только прочитывать большое число данных, но и осуществлять анализ в полевых условиях, например, в тропических лесах Эквадора [72] или при изучении коралловых рифов [73], сокращая дистанцию между сбором образца и его лабораторным анализом. Была разработана платформа SEQUEL, использующая технологию одномолекулярного секвенирования в реальном времени (single molecule real-time sequencing, SMRT), которая позволяет в разы сократить стоимость анализа [74]. Это говорит о востребованности метода, к которому адаптируются новейшие технологии прочтения генетической информации - на сегодняшний день в основной базе данных BOLD накоплено около 9.8 млн. последовательностей [75] для 233 тысяч видов животных и 70 тысяч видов растений.

Изначальная задача ДНК-баркодинга – точная видовая идентификация уже описанных видов вне зависимости от внутривидового полиморфизма, пола или стадии онтогенеза. Однако впоследствии Эбер предлагает использовать ДНК-баркодинг для выявления еще не описанных видов. На примере тропической толстоголовки (Lepidoptera, Hesperiidae) *Astraptes fulgerator*, исследователям удалось показать, что этот вид является комплексом из 10 видов, которые были обнаружены как кластеры ДНК-баркодов, коррелирующих с морфологическими и

экологическими различиями [76]. Впоследствии при помощи ДНК-баркодинга неоднократно обнаруживалось скрытое биологическое разнообразие в самых разных группах живых организмов: у мух-тахин [77], олигохет [78], двустворчатых моллюсках [79], пауков [80], рыб [81] и многих других. Скрытое потенциальное биологическое разнообразие среди дневных бабочек также оказалось велико – при масштабном исследовании европейских бабочек было показано, что 27.7% из 299 исследованных видов представлены более чем одной митохондриальной линией [82].

Однако далеко не всегда кластеры, выделяемые на основании анализа гена *COI* соответствуют реальным видовым границам. Множественные митохондриальные линии внутри одного вида могут являться следствием анцестрального полиморфизма [83], митохондриальной интрогрессии [84] или длительной географической изоляции с последующим вторичным контактом [9]. Влияние может оказывать заражение внутриклеточным симбионтом рода *Wolbachia*, который, вызывая цитоплазматическую несовместимость, фактически обеспечивает внутривидовую репродуктивную изоляцию, что приводит к повышению генетического разнообразия митохондриальной ДНК [85]. Более редкая проблема ДНК-баркодинга – гетероплазмия, или присутствие двух типов митохондрий в одном образце [86], которая может быть решена путем привлечения дополнительных маркеров [87]. Сходные трудности вызывает наличие в ядерном геноме псевдогенов митохондриального происхождения (numts), которые могут быть амплифицированы при помощи праймеров, разработанных для ДНК-баркодинга [88].

В недавнем исследовании Махов с соавторами провели детальный анализ возможных причин возникновения митохондриального полиморфизма у пядениц родов *Alcis* и *Thalera* (Lepidoprera, Geometridae) [89]. Авторы не обнаружили связи митохондриальных гаплогрупп у видов *A. deversata* и *Th. chlorosaria* с присутствием симбионта *Wolbachia*, и показали, что наиболее вероятная причина митохондриального полиморфизма – вторичный контакт популяций, долгое время находившихся в аллопатрии. У вида *A. repandata* была обнаружена митохондриальная интрогрессия от *A. extinctaria.*

У распространенной европейской толстоголовки *Thymelicus sylvestis* было обнаружено шесть дифференцированных митохондриальных линий, которые не показывали корреляции с ядерными геномами [16]. Авторами было исключено влияние *Wolbachia*, псевдогенов, интрогрессии и анцестрального полиморфизма, и возможным объяснением является географическая изоляция с последующим слиянием, в результате которого разнообразие ядерного генома элиминировалось при рекомбинации, а митохондриальные линии сохранились [16].

Таким образом, ДНК-баркодинг является удобным инструментом для первичного таксономического анализа, который широко применяется в самых разных группах живых организмов. Тем не менее, несоответствие митохондриальных кластеров границам таксонов в природе встречается довольно часто [90–95]. Поэтому во избежание ошибочных заключений, изучение ДНК-баркодов следует дополнять анализом других генетических маркеров, морфологии, географического распространения, экологических предпочтений и других возможных данных [15,96].

1.2 Интегративный подход при делимитации видов

Применение различных типов данных для определения видовых границ получило формальное название «интегративная таксономия» в 2005 году [1]. Этот подход применяется как при уточнении статуса уже существующих групп, так и при описании новых видов [18,97–99].

За последние десятилетия таксономия претерпела значительные изменения в своей методологии. Наиболее часто используемым типом данных в интегративной таксономии как членистоногих [15], так и других групп животных [18] после традиционных морфологических признаков становятся молекулярные данные, поскольку информация о структуре и функциях генов, протеинов и метаболитов открывает широчайшие возможности для комплексного анализа. С развитием технологий секвенирования происходит переход от филогений, построенных на основании одиночных маркеров к филогеномике, что позволяет делать более тонкие выводы о процессах видообразования [100] и более точные таксономические интерпретации. Интегративный анализ постепенно становится рутинной процедурой для исследователей, что вносит значительный вклад в развитие современной таксономии [18].

Методология интегративного подхода в работах разных исследователей имеет концептуальные различия. Так, Дайрат указывал на необходимость конгруэнтности между несвязанными типами данных [1]. Ожидается, что обнаруженные таким образом эволюционные линии чаще будут соответствовать видам, потому что случайное совпадение эволюционных преобразований маловероятно. Однако, процесс накопления различий на разных уровнях не всегда осуществляется одновременно [101], и велик риск ошибок первого рода, при которых существующие виды не будут обнаружены. Эта проблема особенно актуальна при изучении молодых групп, таких, как например, цихлидовые рыбы, в которых между видами существуют морфологические различия, но слабая репродуктивная изоляция и генетическая дифференциация по нейтральным локусам [102]. Эта проблема также существует при изучении

групп, внутри которых часто происходят случаи т. н. ретикулярной эволюции (reticulation events), включающей такие процессы как межвидовая гибридизация, интрогрессия и горизонтальный перенос генов. При таких событиях эволюционная история разных признаков у одного организма может значительно отличаться, и не все они могут адекватно отражать реальные видовые границы.

Кумулятивный подход к интегративному таксономическому анализу основан на допущении, что, поскольку видообразование – продолжительный процесс, то дивергенция различных признаков может происходить не одновременно, поэтому для выдвижения таксономической гипотезы конгруэнтность разных типов данных желательна, но не является строгим условием [14]. В рамках этого подхода, в некоторых случаях таксономическое решение может быть принято даже по одному набору признаков, если он считается достаточным для обоснования обособленности таксона [14,15]. Чтобы избежать ложноположительных результатов, соответствия и несоответствия между разными типами признаков должны быть объяснены с точки зрения эволюционной истории изучаемых групп, а заключения, основанные на анализе молекулярных маркеров поддержаны анализом морфологии [15].

С методологической проблемой делимитации видов тесно связана концептуальная проблема – определение критериев вида. Существует более 20 концепций вида [103], многие из которых обладают противоречащими друг другу критериями. Чтобы избежать таксономических противоречий, которые возникают при использовании разных определений видовых систем, была предложена унифицированная концепция вида (unified species concept [14]), которая рассматривает вид как самостоятельно эволюционирующую группу метапопуляций. Для того чтобы считать эволюционную линию видом, достаточно соответствие критериям одной из концепций вида, поскольку между молодыми таксонами формирование различий могло не завершиться. Таким образом, подчеркивается, что вид является динамической единицей эволюции, которая изменяется с течением времени, и к молодым таксонам часто неприменимы те же критерии, что и к более древним. При этом, данные о происхождении и эволюционной траектории видов тесно связаны с определением межвидовых границ [96], поэтому понимание процессов видообразования важно для решения таксономических проблем. Следует подчеркнуть, что интегративный подход подразумевает под собой объединение разных типов данных, а не разных критериев вида, которые могут быть применены и к одному типу признаков [18], однако определение вида и его критериев являются ключевыми для таксономии.

1.3 Применение полногеномных данных в филогенетике и таксономии

В настоящее время секвенирование становится все менее и менее дорогостоящей процедурой, и современное оборудование позволяет расшифровывать все более длинные нуклеотидные последовательности, благодаря чему получила развитие филогеномика – ветвь молекулярной филогенетики, основанная на одновременном анализе множественных локусов [104]. Версию филогении насекомых, основанную на анализе последовательностей 1478 генов (2.5 млн пар оснований кодирующей ДНК) 103 вида насекомых, среди которых есть представители всех современных порядков представил в своей работе Мисоф с коллегами [105].

Хорошим примером использования филогеномных данных для интегративной таксономии членистоногих является исследование, включающее анализ полного генома многоножки *Strigamia maritima* [106]. В этой работе была поддержана система, согласно которой насекомые (Insecta) объединяются с раками в таксон Pancrustacea, а многоножки (Myriapoda) являются сестринской кладой по отношению к Pancrustacea (то есть базальной кладой в пределах Mandibulata). Соответственно, в этой работе постулируется независимый выход на сушу для предков Insecta и Myriapoda, что подтверждается тем, что ольфакторные рецепторы этих групп кодируются разными семействами генов. Предложенная классификация также поддержана тем, что многоножки обладают специфическим механизмом обеспечения разнообразия белков иммунного ответа. У дрозофилы и у дафнии ген *Dscam* (Down syndrome adhesion molecule), играющий ключевую роль в иммунной системе, представлен в геноме четырьмя копиями и потенциально способен генерировать около 150 тысяч и 13 тысяч изоформ, соответственно. Таким образом, вероятной синапоморфией насекомых и ракообразных является уникальная система альтернативного сплайсинга. Геном *Strigamia maritima* содержит 80 паралогов *Dscam*, возникших в результате дупликации, и разнообразие белков обеспечивается за счет многокопийности этого гена [106].

Полногеномные данные являются удобным инструментом при описании новых видов. Недавно был описан новый вид орангутангов *Pongo tapanuliensis*, который не только отличается от двух других видов орангутангов *P. pygmaeus* и *P. abelii* по строению нижнечелюстного отдела черепа и зубов, но и является самостоятельной генетической единицей [107]. Датировки, осуществленные на основании анализа полных геномов, показали, что *P. tapanuliensis* представляет из себя наиболее древнюю линию орангутангов, отделившуюся ~3.38 млн лет назад. Этот редкий локальный вид, обитающий около озера Тоба (Суматра) нуждается в охране, поскольку его численность составляет приблизительно 800 особей [107].

Изучение полных геномов позволяет напрямую сопоставить проблему делимитации видов с проблемой видообразования. Поиск регионов генома, отвечающих за видообразование, долгое время привлекал внимание исследователей, и было показано, что в ходе видообразования дивергенция происходит неравномерно и в разных участках генома, выделяются т.н. «геномные островки видообразования» [108]. Эти регионы часто содержат гены, которые либо играют роль в адаптации к условиям среды, либо обеспечивают межвидовую репродуктивную изоляцию [109]. При изучении галапогосских вьюрков (сем. *Geospizinae*), модельного объекта эволюционной биологии, было проведено полногеномное секвенирование для 120 особей всех видов этого эндемичного семейства [110]. Была показана связь гена *ALX1*, кодирующего транскрипционный фактор, отвечающий за развитие черепно-лицевого отдела, с разнообразием морфологии клюва. Таким образом, для этого модельного объекта был найден участок генома, сыгравший ключевую роль в адаптации к пищевым ресурсам, в возникновении экологически обусловленной презиготической репродуктивной изоляции и в видообразовании [110].

Накопление генетических различий и формирование репродуктивных барьеров может происходить неравномерно также на разных хромосомах. Для многих организмов был показан так называемый большой эффект половой X хромосомы (large X-effect), который заключается в том, что гены, ассоциированные с формированием постзиготических репродуктивных барьеров (например, обеспечивающие стерильность гибридов), сконцентрированы скорее в X хромосоме (или в Z хромосоме у видов с ZW системой определения пола), чем в аутосомах [111,112]. Таким образом, половые хромосомы играют непропорционально большую роль в развитии репродуктивной изоляции по сравнению с аутосомами. С этим явлением тесно связан феномен более быстрой эволюции половых X хромосом, называемый faster X-effect (или faster Z-effect), приводящий к тому, что X хромосомы демонстрируют больший уровень дивергенции, чем аутосомы [112–114]. Эти различия могут быть обусловлены действием отбора. Согласно модели Добжанского-Мюллера, в двух независимо эволюционирующих линиях могут возникать такие аллели, которые при вторичном контакте вызывают нежизнеспособность (или сниженную приспособленность) гибридов [115]. Эта модель, как предполагается, объясняет сниженный уровень интрогрессии в половых хромосомах по сравнению с аутосомами [116], который был отмечен у многих организмов: у бабочек рода *Heliconius* [117], у колумбийских ревунов [118], у соловьев [119] и у человека [120]. Анализ генов, расположенных на Z хромосомах бабочек, показал, что они кодируют белки, участвующие в распознавании партнера, развитии крыльев, а также циркадные ритмы, что говорит о существенной роли Z хромосом в формировании признаков, находящихся под влиянием полового отбора [116].

Таким образом, половые хромосомы являются важным регионом генома, участвующим в обеспечении и поддерживании репродуктивной изоляции.

На формирование генетических различий и репродуктивных барьеров между популяциями могут влиять различные факторы. Во-первых, генетические различия могут развиваться под влиянием естественного и полового отбора, действующего на сестринские популяции [121]. Во-вторых, причиной формирования репродуктивной изоляции могуть быть нейтральные мутации, фиксирующиеся в популяциях в результате дрейфа генов, могут служить причиной формирования репродуктивной изоляции [122,123]. В-третьих, интрогрессия генов от несестринского вида к одной из аллопатричных популяций также может играть роль в видообразовании. Такой сценарий был недавно показан при рассмотрении филогенетической истории газели Гранта [124]. Межвидовая интрогрессия более распространена в природе, чем считалось ранее [125–128], и может быть триггером адаптивной эволюции, поскольку она приводит к переносу генов и признаков от одной филогенетической линии к другой [129,130]. Адаптивная интрогрессия у неблизкородственных видов бабочек, обнаруженная при помощи анализа полногеномных данных, была показана на примере рода *Heliconius* (Lepidoptera, Nymphalidae) [131].

Межвидовая гибридизация может значительно затруднять диагностику видов, а также определение видовых границ. Гибридные особи часто демонстрируют промежуточное морфологическое строение между родительскими видами [132], но иногда гибридизация приводит к формированию новых уникальных признаков и видообразованию [133]. В таких группах реконструкция филогении по нескольким генетическим маркерам может приводить к противоречивым результатам. В подобных случаях, увеличение числа данных может помочь при определении видовых границ. Так, при изучении цихлидовых рыб - группы, в которой широко распространена межвидовая гибридизация, использование полногеномных данных, полученных методом RAD, позволило значительно улучшить разрешение филогенетического анализа [134].

Таким образом, анализ полных геномов открывает широкий спектр возможностей для решения самых разнообразных проблем таксономии, оставаясь при этом в зоне относительной доступности для исследователей. В целом, применение современных молекулярных методов служит важным дополнением к традиционным зоологическим подходам и становится неотъемлемой частью интегративного подхода в таксономии.

Глава 2. ДНК-баркоды как инструмент для постановки таксономических гипотез

2.1 Анализ комплекса *Melitaea ala*

Комплекс видов, близких к *Melitaea ala* Staudinger, 1881, распространен в Центральной Азии и включает в себя виды *Melitaea ala* Staudinger, 1881, *M. kotshubeji* Sheljuzhko, 1929, *M. ninae* Sheljuzhko, 1935, *M. chitralensis* Moore, 1901 и *M. enarea* Fruhstorfer, 1917 [24]. Согласно недавней ревизии рода *Melitaea* [20], этот комплекс также включает в себя виды *M. acraeina* Staudinger, 1881, *M. didymina* Staudinger, 1895, *M. bundeli* Kolesnichenko, 1999, *M. sutschana* Staudinger, 1881 и *M. yagakuana* Matsumura, 1927 (последний часто рассматривается в качестве подвида *M. sutschana* [135]). Последующий анализ рода *Melitaea* с использованием молекулярных маркеров показал, что *M. acraeina* - филогенетически далекий вид, образующий сестринскую ветвь к линии (*M. ala* + *M. enarea*), в то время как *M. chitralensis и M. sutschana* принадлежат к комплексу видов, близких к *M. didyma,* который является сестринским к линии ((*M. acraeina* + (*M. ala* + *M. enarea*)) [136].

Нами был проведен анализ ДНК-баркодов (фрагментов митохондриального гена *COI*) для комплекса *M. ala,* который показал, что особи, относимые к *M. ala*, *M. kotshubeji* и *M. enarea* образуют монофилетические линии с высокими поддержками (Рисунок 1) [28]. Вместе эти три вида образуют поддержанную кладу, соответствующую комплексу видов *M. ala*. *Melitaea acraeina* формирует отдельную сестринскую кладу, что согласуется с предыдущими результатами [136]. Таксоны, близкие к *M. didyma*, также кластеризуются вместе, образуя комплекс видов *M. didyma* (Рисунок 1, клада 3), таксономическая структура которого будет детально рассмотрена в следующем разделе. Виды, принадлежащие к группе *M. persea*, выделяются в отдельную кладу, являющуюся сестринской к группе *M. didyma* (Рисунок 1, клада 5).

В составе клады *M. ala* было обнаружено пять поддержанных генетических линий (1) *M. ala ala*, (2) *M. ala irtyshica*, (3) *M. ala zaisana*, (4) *M. ala bicolor* (*b1*) и *M. ala bicolor* (*b2*). Поскольку обнаруженные линии имеют высокие поддержки, они могут теоретически рассматриваться в качестве таксонов с точки зрения филогенетической концепции вида, для которой достаточным критерием является распознаваемость эволюционной единицы независимо от наличия репродуктивной изоляции (о концепции см.: [137,138]).

Чтобы решить вопрос о таксономической интерпретации выявленных линий, мы изучили уровень генетической дивергенции между ними и провели сравнение рисунка крыльев. Сравнение уровня дивергенции гена *COI* внутри и между этими кладами показало, что во всех случаях p-дистанции (отношение числа замен в двух сравниваемых сиквенсах к общему числу

нуклеотидов в изучаемой последовательности) были ниже 2%, значения, которое было предложено считать стандартным порогом для разграничения видов [50].

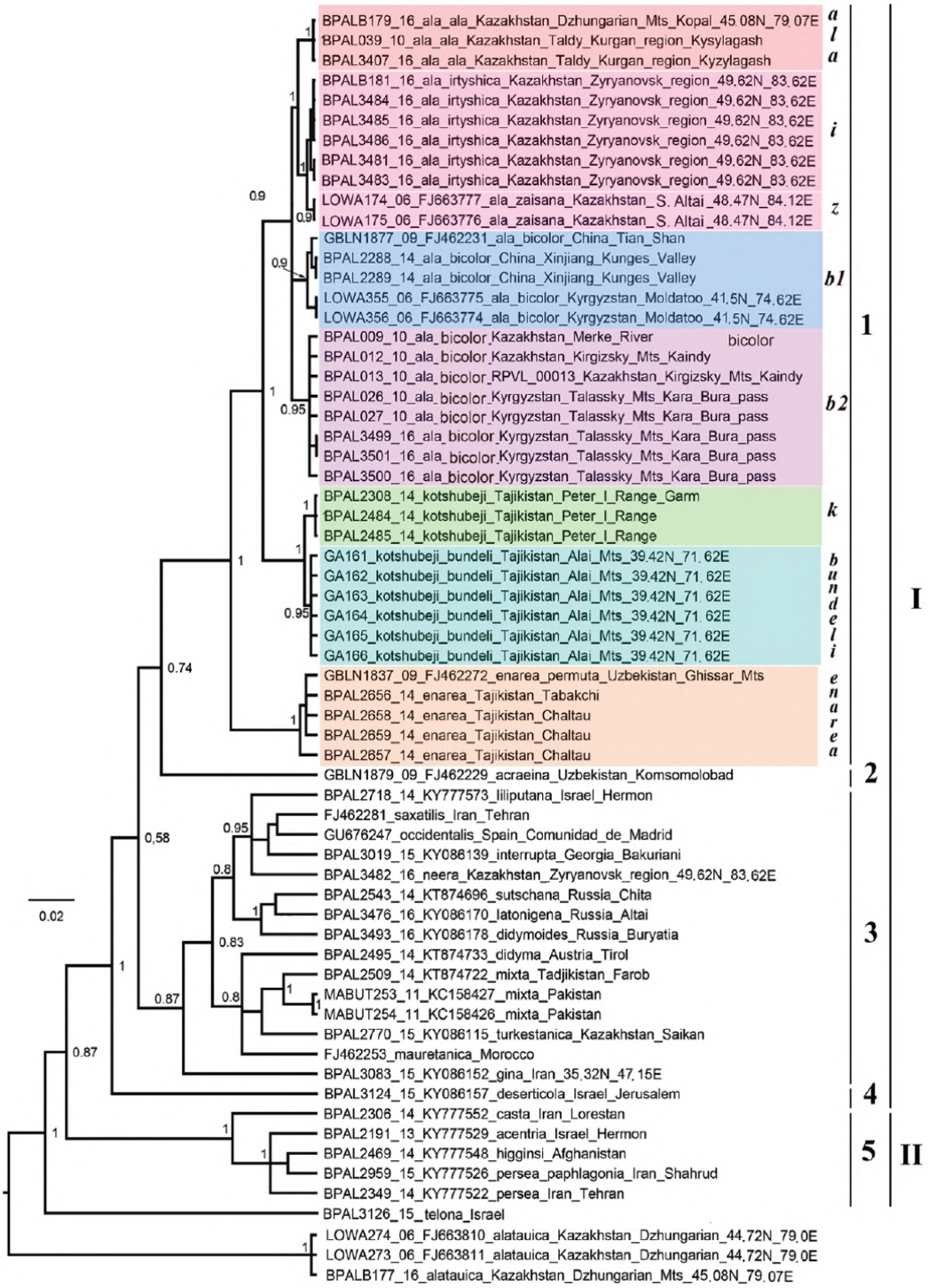

Рисунок 1. Дендрограмма, построенная при помощи метода Байеса, основанная на анализе фрагмента гена *COI*. *Melitaea alatauica* использовалась в качестве внешней группы. На ветвях в качестве поддержек указаны байесовские апостериорные вероятности. Обозначение клад: ***b1*** - *M. ala bicolor* 1, ***b2*** - *M. ala bicolor* 2, ***k*** - *M. kotchubeji kotchubeji*, **i** – *M. ala irtyshica*, **z** – *M. ala zaisana*, 1 – комплекс видов *M. ala*, 2 - *M. acraeina*, 3 – комплекс видов *M. didyma*, 4 - *M. deserticola*, 5 - комплекс видов *M. persea*, I – группа видов *M. didyma*, II – группа видов *M. persea*. Из [28].

Самый низкий уровень дифференциации ДНК-баркодов был обнаружен между кладами *M. ala zaisanica* и *M. ala irtyshica,* в связи с чем предлагается рассматривать их в качестве единой таксономической единицы [28]. В линиях b1и b2 был показан более высокий уровень межвидовых различий (0,3–0,8%), однако внутривидовая вариабельность в этих кладах также высока (до 0,8%), что позволяет рассматривать их в качестве одного таксона *M. ala bicolor*. Таким образом, среди изученных популяций можно выделить три подвида внутри *M. ala: M. ala ala, M. ala bicolor* Seitz, 1908 и *M. ala zaisana* Lukhtanov, 1999 (=*M. ala irtyshica* Lukhtanov, 1999).

Подвид *Melitaea ala ala* распространен в Джунгарском Алатау в Восточном Казахстане и характеризуется затемненными жилками на нижней стороне заднего крыла, которые формируют отчетливые ячейки на срединной перевязи (Рисунок 2).

Подвид *M. ala bicolor* Seitz, 1908 обитает в Северном, Восточном, Центральном и Западном Тянь-Шане (юго-восток Казахстана, северо-запад Китая и Киргизия). У представителей этого подвида жилки на нижней стороне нижнего крыла менее затемнены, ячейки на срединной перевези практически не выделяются, а лишь отмечены темными скобами на внешней стороне перевязи (Рисунок 2).

Подвид *M. ala zaisana* обитает на южном Алтае и морфологически более близок к подвиду *M. ala bicolor*, чем к более близкому географически *M. ala ala* (Рисунок 2) [28].

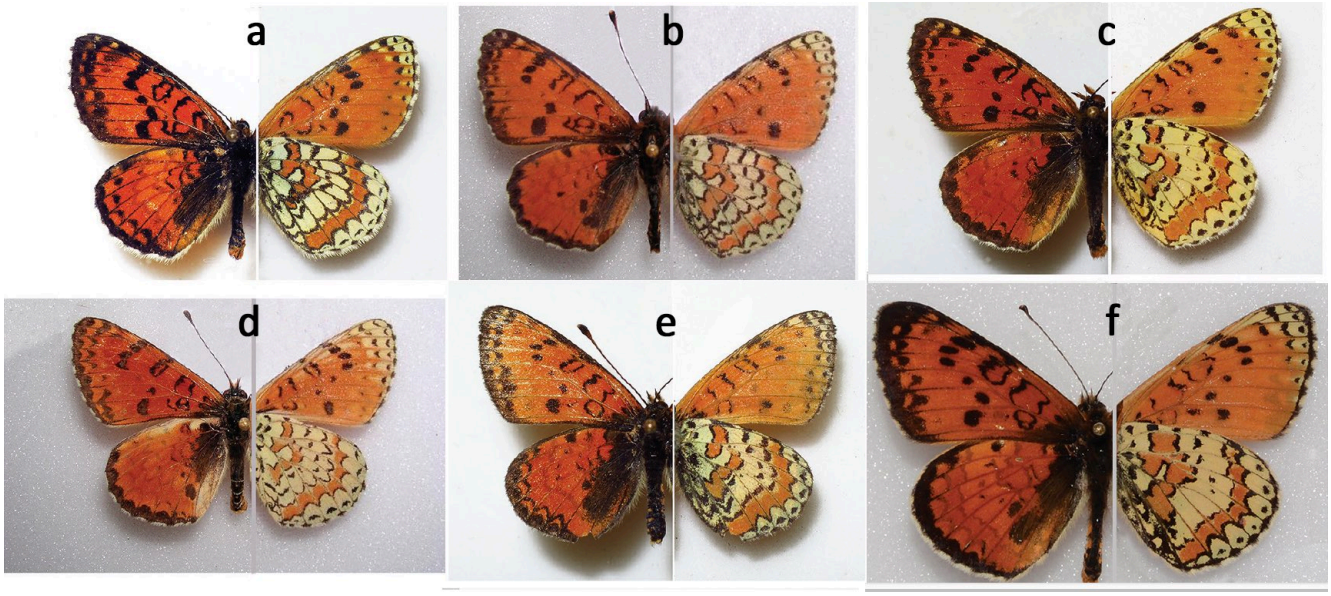

Рисунок 2. Бабочки комплекса *Melitaea ala*. Слева на рисунках показана окраска верхней стороны, справа – испода крыльев. **a** *M. ala ala*, **b** *M. ala bicolor* (гаплогруппа b1), **c** *M. ala bicolor* (гаплогруппа b2), **d** *M. ala zaisana*, **e** *M. kotshubeji bundeli*, **f.** *M. kotshubeji.*

Наш анализ выявил две клады в составе *Melitaea kotshubeji: Melitaea kotshubeji bundeli* и *M. kotshubeji kotshubeji.* Изначально *Melitaea kotshubeji bundeli* был описан как подвид *M. kotshubeji* [24], но в последующих ревизиях либо рассматривался в качестве отдельного вида [20],

либо сводился в синонимы [139]. Проведенный нами анализ ДНК-баркодов показал, что эти два таксона различаются по четырем фиксированным нуклеотидным заменам, что говорит о длительной независимой эволюции этих линий. Также эти таксоны различаются по окраске крыла (Рисунок 2), и, хотя и обитают в аллопатрии, они обнаруживаются в непосредственной близости друг от друга в Киргизии, разделенные только узкой долиной реки Сурхоб [28].

Таким образом, в комплексе *M. ala* обнаружена довольно простая для таксономической интерпретации ситуация, когда уровень обнаруженной генетической дивергенции между аллопатрическими кладами соответствует внутривидовым и межвидовым дистанциям, ожидаемым при анализе ДНК-баркодов, и различия в баркодах в целом скоррелированы с различиями в крыловом рисунке [28]. В следующем разделе будет рассмотрен случай необычно высокого внутривидового митохондриального полиморфизма, который свидетельствует о сложной эволюционной истории рассматриваемой группы и требует комплексного анализа.

2.2. Анализ комплекса *Melitaea didyma*

Представители комплекса таксонов, близких к *Melitaea didyma* Esper, 1779, широко распространены в Палеарктике [20,135]. Внутри комплекса наблюдается высокий уровень внутривидовой географической и сезонной изменчивости, отмечаются различия между особями разных генераций и клинальная вариабельность размеров и окраски крыльев. При этом виды часто слабо различимы по окраске крыльев и морфологии гениталий [20,135]. В подобных группах для выявления криптического таксономического разнообразия часто применяется анализ молекулярных маркеров [7,12]. Проведенный нами анализ ДНК-баркодов показал, что комплекс *M. didyma* является монофилетической группой за исключением трех аномальных по уровню дивергенции гаплогрупп (*turkestanica2*, *turkestanica3* и *gina2*) (Рисунок 3) и характеризуется высоким уровнем митохондриального разнообразия, что может объясняться наличием скрытых видов и/или глубоким внутривидовым полиморфизмом [25]. Чтобы прояснить таксономическую структуру комплекса, был изучен материал из Армении, Болгарии, Греции, Грузии, Израиля, Ирана, Казахстана, Кыргызстана, России, Сирии, Словении и Турции. Анализ показал, что комплекс *M. didyma* представлен 23 дивергировавшими митохондриальными гаплогруппами, которые демонстрируют высокие поддержки на филогенетическом дереве (байесовская апостериорная вероятность >0.95) и, как правило, ассоциированы с определенными территориями (Рисунки 3 и 4) [26].

Для гаплогрупп были выбраны названия, которые соответствуют таксонам из комплекса *M. didyma*, описанным с их территории распространения: *chitralensis* (север Пакистана), *deserticola* (Северная Африка, Израиль, Иордания, Ливан, Сирия), *didyma* (Западная Европа), *didymoides* (Азиатская часть России, Монголия, Северный Китай), *gina* (Западный Иран,

Азербайджан), *gina2* (Иран), *interrupta* (Кавказ, северо-восток Турции), *latonigena* (Азиатская Россия, северо-восток Казахстана, Монголия, северо-запад Китая), *liliputana* (Армения, Турция, Сирия, Ливан), *mauretanica* (юг Испании), *mixta* (Таджикистан, Кыргызстан, Узбекистан, Пакистан, Афганистан), *neera* (Восточная Европа, Северный Кавказ, Западная Сибирь, Северный Казахстан), *neera2* (Словения), *occidentalis* (Испания), *protaeoccidentis* (Северная Африка), *saxatilis* (Северный Иран), *sutschana* (Дальний Восток, Корея, северо-восток Китая), *sutschana2* (Китай), *sutschana3* (Россия, Приморский край), *turkestanica* (Казахстан, Кыргызстан, Узбекистан, Таджикистан, Западный Китай), *turkestanica2* (Казахстан), *turkestanica3* (Казахстан, Узбекистан) и *turkestanica4* (Казахстан) (Рисунки 3 и 4) [25,26]. Показано, что, средние генетические p-дистанции между гаплогруппами, достигают 9,5% [26], что значительно выше «стандартных» дистанций для баркодов 2-3%, которые используются для аллопатричных таксонов в качестве критерия для разграничения видов [50].

Большинство гаплогрупп имеет аллопатричное распространение, тем не менее, было зафиксировано несколько случаев, когда гаплогруппы были обнаружены в симпатрии. Образец M22, относящийся к гаплогруппе *gina2* был отмечен в симпатрии с *gina* на северо-западе Ирана. Различия между кластерами *gina* и *gina2* составляют 6.5%. Гаплогруппы *turkestanica2*, *turkestanica3* и *turkestanica4* обнаруживаются в симпатрии с гаплогруппой t*urkestanica* (Рисунок 5) и обладают различиями по гену *COI* до 7,4%. Два образца, относящихся к гаплогруппе *turkestanica* были обнаружены на территории распространения *neera*: один в Актюбинской области Казахстана, а второй — в Самарской области России (Рисунок 5). При этом, в Казахстане (Карабирюк) два образца с гаплотипом *neera* были отмечены вместе с бабочками, относящимися к гаплогруппам *turkestanica* и *turkestanica 4*.

Поскольку большинство гаплогрупп обнаружено в аллопатрии, учитывая степень различий между митохондриальными генетическими линиями, можно предположить, что их дифференциация происходила в течение длительного периода изоляции, когда их ареалы былиразделены географическими и/или экологическими барьерами. В соответствии с принятой скоростью накопления p-дистанций 1,5-2,3% на миллион лет для гена *COI* для различных групп членистоногих [140,141], период изоляции можно оценить как 0,5–5,0 млн лет. Предположительно, каждая из этих гаплогрупп эволюционировала в одном из основных западно-палеарктических рефугиумов позднего Плиоцена и Плейстоцена в Северной Африке (*protaeoccidentis*, *deserticola*), на Пиренейском полуострове (*occidentalis*, *mauretanica*), Балканском полуострове (*neera*), Ближнем Востоке (*liliputana*, *saxatilis*, *gina*) и Центральной Азии (*turkestanica*, *mixta*, chitralensis) [26]. Наличие малых гаплогрупп *neera2*, *turkestanica2*, *turkestanica3*, t*urkestanica4*, *gina2*, которые могли возникнуть как географические изоляты, но сейчас существуют в условиях вторичной симпатрии с основными гаплогруппами *neera*,

turkestanica и *gina*, согласуется с концепцией рефугиум в рефугиуме (refugia-within-refugia concept) [142].

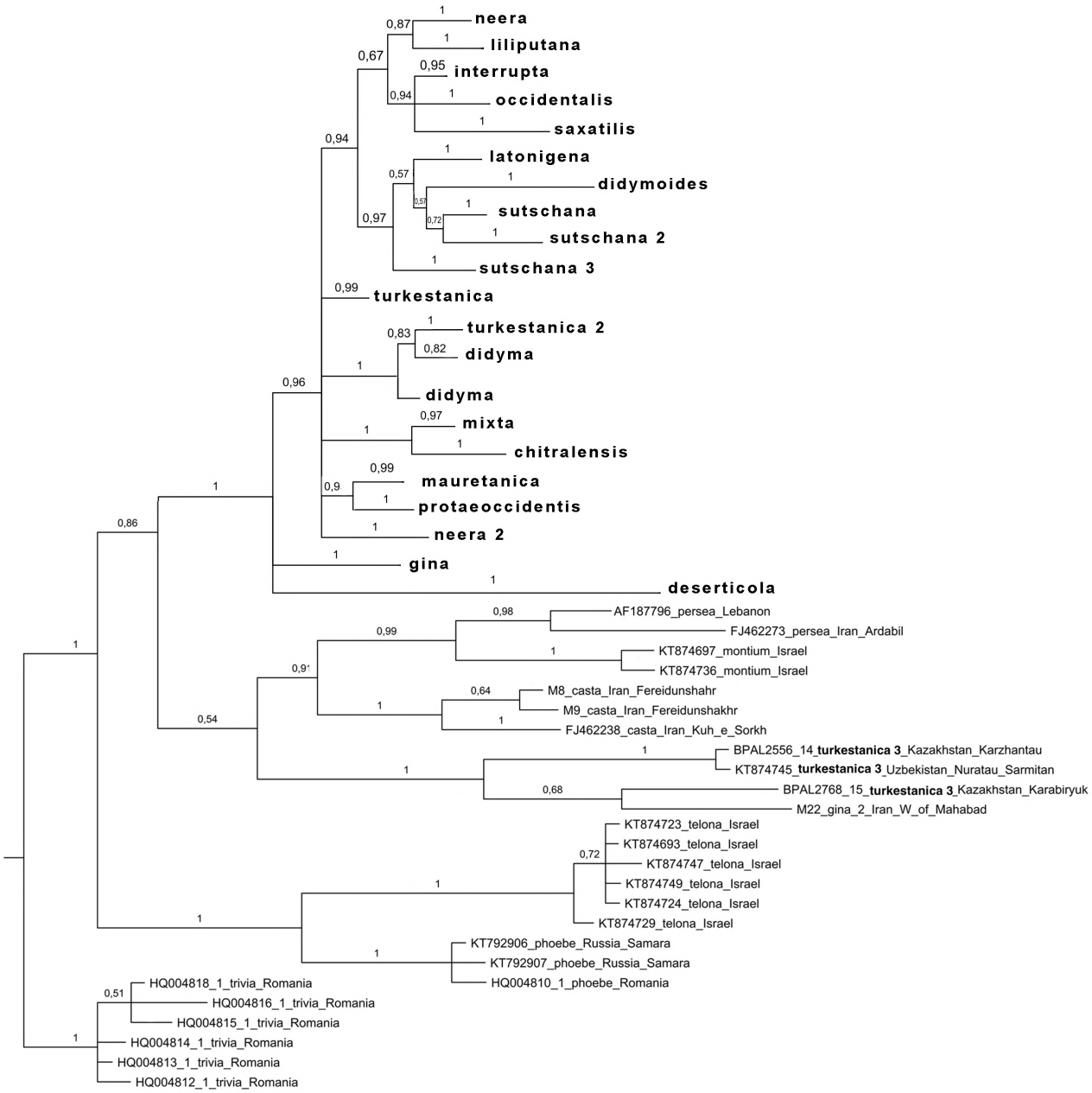

Рисунок 3. Схема, построенная при помощи Байесова метода, отражающая структуру комплекса видов *Melitaea didyma*, выявленные на основании анализа ДНК-баркодов. Выявленные гаплогруппы, относящиеся к комплексу M. didyma, показаны жирным шрифтом. На ветвях указаны апостериорные вероятности. Детальные дендрограммы приведены в [26].

Говоря о таксономической структуре комплекса, следует отметить, что 12 из 23 обрануженных гаплогрупп ассоциированы со следующими девятью традиционно выделяемыми видами: *M. chitralensis* Moore, 1901, *M. deserticola* Oberthür, 1909, *M. didymoides* Eversmann, 1847, *M. gina* Higgins, 1941 (гаплогруппы *gina*+*gina2*), *M. interrupta* Colenati, 1846, *M. latonigena* Eversmann, 1847*, M. mixta* Evans, 1912, *M. saxatilis* Christoph, 1873 и *M. sutschana* Staudinger, 1892 (гаплогруппы *sutschana*+ *sutschana2*+ *sutschana3*) [26].

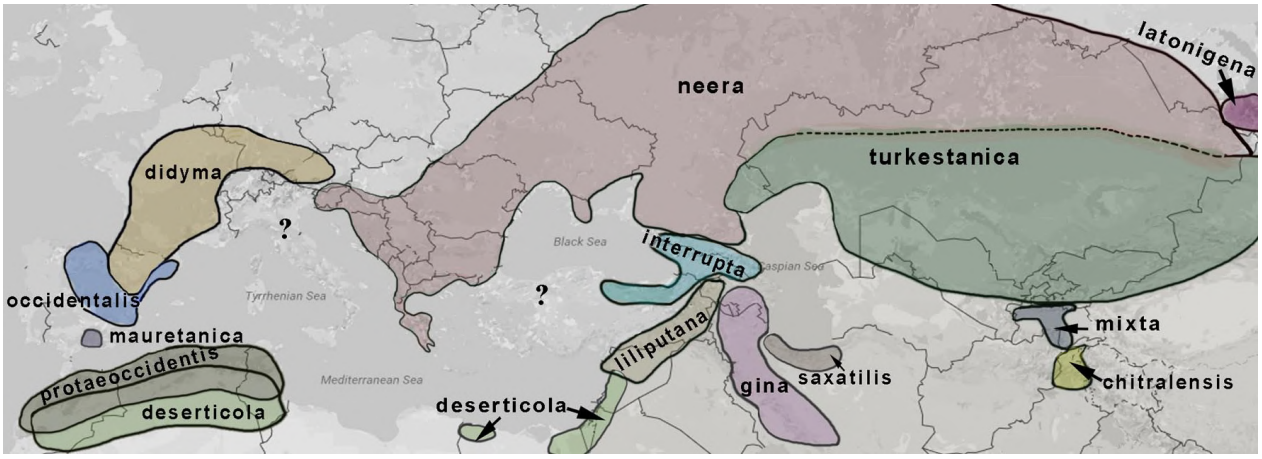

Рисунок 4. Распространение гаплогрупп в составе комплекса видов *Melitaea didyma*. Из [26].

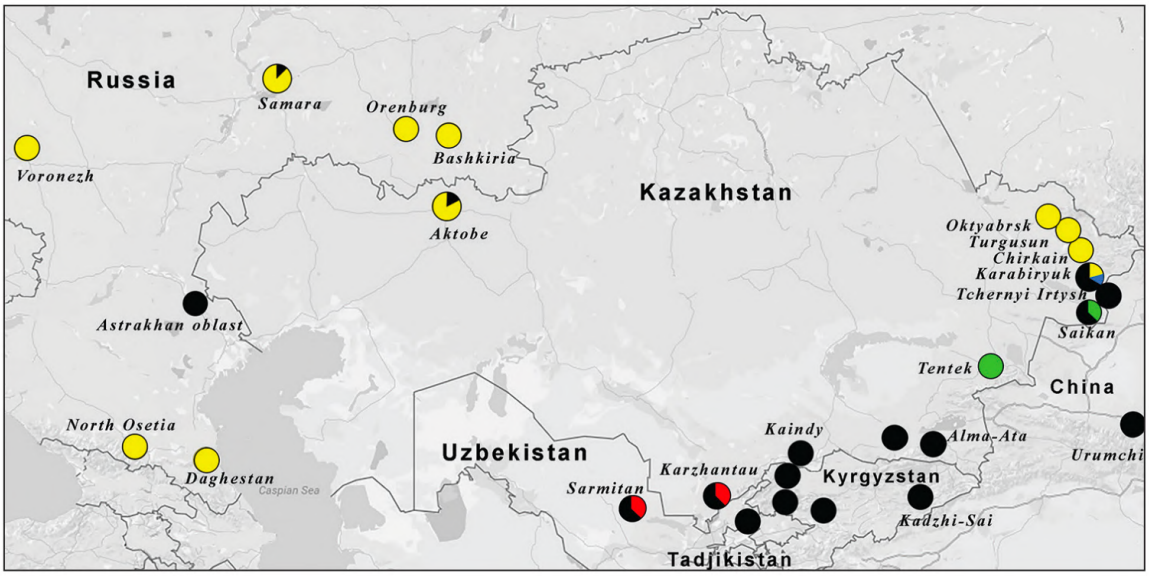

Рисунок 5. Локализация гаплогрупп *neera* и *turkestanica* (желтый – *neera*, черный – *turkestanica*, зеленый – *turkestanica2*, красный – *turkestanica3,* синий – *turkestanica4*).

Оставшиеся 11 гаплогрупп мы предлагаем рассматривать в составе вида *M. didyma* [25,26]. Согласно нашим данным, эти гаплогруппы являются глубоко дивергировавшими внутривидовыми генетическими линиями (p-дистанции до 7,4%), однако мы считаем что было бы преждевременно придавать им видовой ранг. В пользу этой гипотезы говорят следующие данные. Во-первых, согласно последним ревизиям, основанным на изучении морфологии, *M. didyma* не обнаруживает фиксированных морфологических различий между популяциями, в связи с чем все популяции рассматриваются в составе подвида *Melitaea didyma didyma*, кроме популяций из Средней Азии, которые были отнесены к отдельному подвиду *Melitaea didyma turkestanica* [20] или даже разделены на несколько отдельных подвидов [143]. Во-вторых, в

случаях симпатрии гаплогрупп *neera*/*neera2*, *turkestanica/turkestanica*, *turkestanica/turkestanica3* и *turkestanica/turkestanica4* не наблюдается различий в экологических предпочтениях этих бабочек. При этом, когда виды *M. didyma/M. interrupta*, *M. didyma/M. latonigena* и *M. gina/M. saxatilis* обитают симпатрично, они, как правило, занимают различные экологические ниши (*M. didyma* и *M. gina* более ксерофильны, в то время как *M. interrupta, M. latonigena* и *M. saxatilis* более мезофильны). Хотя *M. didyma neera* и *M. didyma turkestanica* дифференцированы экологически [26], в случаях симпатрии представителей разных гаплогрупп различий в экологических предпочтениях отмечено не было. В Самарской области, где преимущественно встречается гаплогруппа *neera*, обе гаплогруппы наблюдаются в биотопе, характерном для *M. didyma neera* (степь), а в Карабирюке, где доминирующей гаплогруппой является *turkestanica*, обе гаплогруппы были отмечены в биотопе, характерном для *M. didyma turkestanica* (пустыня). Эти данные говорят о том, что симпатрично встречающиеся гаплогруппы *M. didyma* представляют случай внутривидовой изменчивости и не являются разными видами [25]. Такое распределение гаплогрупп могло возникнуть вследствие слияния двух линий M. didyma, дивергировавших в аллопатрии, или вследствие межвидовой интрогрессии. Интересно отметить, что гаплогруппа *turkestanica2* не близка гаплогруппе *turkestanica*, а является сестринской к гаплогруппе *didyma* из Западной Европы (Рисунок 3), что, вероятно, является следствием древней интрогрессии. В целом, следы интрогрессии как древней, так и более поздней, свидетельствуют о слабых репродуктивных границах между популяциями *M. didyma*. Таким образом, проведенный анализ демонстрирует феномен полифилии ДНК-баркодов на уровне вида, основанной на глубокой внутривидовой дивергенции. В связи с этим, предлагается рассматривать группы популяций, ассоциированные с основными гаплогруппами *M. didyma*, в качестве подвидов: *M. didyma didyma* Esper, [1779], *M. didyma mauretanica* Oberthür, 1909, *M. didyma occidentalis* Staudinger, 1961*, M. didyma protaeoccidentis* Verity, 1929, *M. didyma liliputana* Oberthür, 1909, *M. didyma neera* Fischer de Waldheim, 1840 и *M. didyma turkestanica* Sheljuzhko, 1929 [26].

Полученные нами для комплекса *M. didyma* [26] таксономические гипотезы фактически были проверены в независимом, более позднем исследовании, проведенном группой испанских исследователей с помощью секвенирования множественных ddRAD маркеров [17]. Это исследование показало неполное соответствие кластеров особей, которые выделяются на основании анализа митохондриальных и ядерных маркеров (mito-nuclear discordance). Полное соответствие было обнаружено лишь для одного кластера из Сицилии вследствие того, что он находится изолированно от остальной популяции и менее подвержен митохондриальной интрогрессии. Данные, полученные Динкой с соавторами [17] не противоречат выдвинутым нами таксономическим гипотезам. Они либо подтверждают их, либо показывают, что два и

более митохондриальных кластера соответствуют одному ядерному кластеру. Последнее указывает на то, что дифференциация по митохондриальным маркерам не всегда является достаточным основанием для выделения видов и подчеркивает важность мультигенного анализа в исследовании биологического разнообразия.

Таким образом, результаты исследований по роду *Melitaea* подтверждают ранее сформулированное положение о том, что анализ одиночных молекулярных маркеров, таких как ДНК-баркоды, может являться лишь инструментом для выдвижения предварительных таксономических гипотез, которые нуждаются в последующей проверке [144]. Эти гипотезы могут затем подтверждаться, как это произошло в нашем анализе видов группы *M. ala* (см. предыдущий раздел) или подтверждаться только частично, как показало тестирование выдвинутых нами гипотез, проведенное Динкой с соавторами [17].

Глава 3. Интегративный таксономический анализ на примере рода *Brenthis* и комплекса *Hyponephele lycaon – H. lupina*

3.1 Анализ рода *Brenthis*

Как было показано выше, ДНК-баркоды являются перспективным маркером, анализ которого позволяет выдвигать первичные таксономические гипотезы. Однако таксономические интерпретации филогений, основанных только на митохондриальных маркерах, не всегда верны. Множественные ядерные гены потенциально более эффективны для решения филогенетических и таксономических задач, но имеют свои особенности [92]. Во-первых, у диплоидных организмов ядерные гены представлены двумя аллелями, которые могут сильно различаться у гетерозиготных особей, что может приводить к несоответствию филогении аллелей и филогении вида. Во-вторых, при выборе маркеров предполагается, что они являются несцепленными и нейтральными, что не всегда соблюдается на практике. В-третьих, невысокая скорость эволюции большинства кодирующих генов в сочетании с рекомбинацией может приводить к слабому филогенетическому сигналу, что не вызывает значительных проблем при анализе высших таксонов (родов, триб, семейств), но накладывает ограничения при анализе межвидовых таксономических отношений. В-четвертых, время коалесценции ядерных аллелей, как правило, значительно больше времени коалесценции митохондриальных аллелей [145]. Поэтому при сравнении эволюционно молодых линий, ядерные гены часто не демонстрируют фиксированных различий по нуклеотидным заменам, хотя по митохондриальным генам различия могут обнаруживаться. В подобных случаях интегративный подход, комбинирующий анализ митохондриальных и ядерных маркеров, а также морфологических признаков, позволяет сделать наиболее точные таксономические заключения [2].

Род *Brenthis* Hübner, 1819 включает в себя четыре вида: *B. hecate* Denis & Schiffermüller, 1775, *B. mofidii* Wyatt, 1969, *B. ino* Rottemburg, 1775, и *B. daphne* Denis & Schiffermüller, 1775 (Рисунок 6). *Brenthis ino* и *B. daphne* широко распространены в Палеарктике (от Западной Европы до Японии). *Brenthis hecate* распространен в западной Палеарктике, в то время как ареал *B. mofidii* ограничен Турцией, Ираком и Ираном. *Brenthis ino*, *B. daphne* и *B. hecate* симпатричны на большей части их ареалов [146]. Таксономическая структура этого маленького рода хорошо изучена, проведен детальный анализ окраски крыльев, гениталий самцов и самок, особенностей преимагинального развития, экологических предпочтений, поведения и распространения [21,22,146,147].

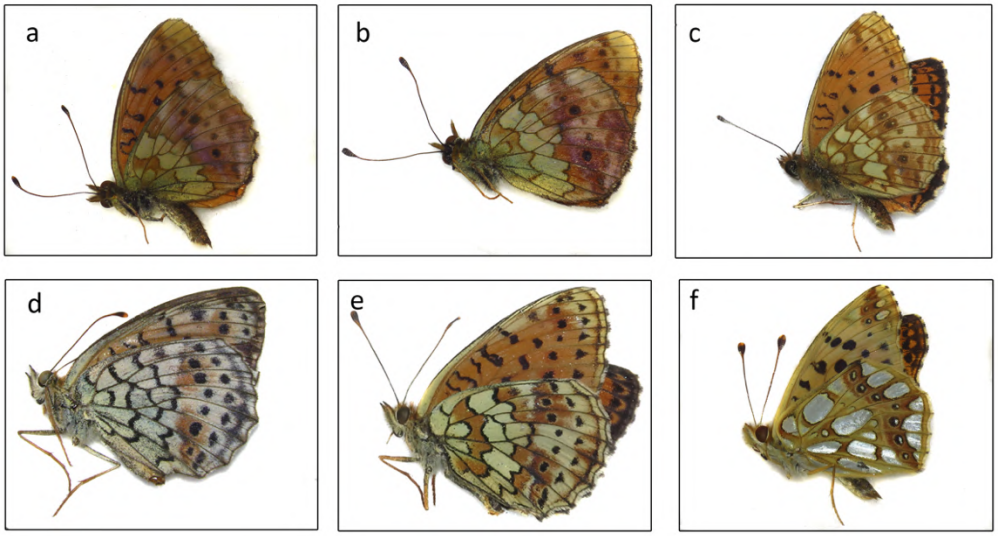

Рисунок 6. Окраска нижней стороны крыла представителей рода *Brenthis*. a – *B. daphne* (D), b – *B. daphne* (D1), c – *B. ino*, d – *B. mofidii*, e – *B. hecate*, f- I*ssoria* lathonia (таксон был использован в качестве внешней группы). Из [27].

Проведенный нами филогенетический анализ, основанный на изучении гена *COI* показал, что в составе рода *Brenthis* выделяются две поддержанные монофилетические группы *B. ino* + *B. daphne* и *B. hecate* + *B. mofidii*, что подтверждает результаты предыдущих исследований [21,22]. Более неожиданным для такой хорошо изученной группы оказалось обнаружение в составе *B. daphne* двух дивергировавших симпатричных митохондриальных линий (гаплогруппы D и D1) (Рисунок 7) [27].

Гаплогруппа D включает в себя образцы *B. daphne*, собранные на участке ареала от Испании до Южного Урала (Россия), в то время как образцы гаплогруппы D1 обнаруживаются от Испании до Японии. Внутри гаплогруппы D1 выделяются три обособленные генетические линии, ассоциированные с географическими регионами: первая - с западной Палеарктикой (от Испании до Урала), вторая - с восточной Палеарктикой (от Алтая до Японии) и третья - с Южной Кореей. Представители гаплогрупп D и D1 были собраны нами одновременно в одних и тех же локациях в Самарской области и в Башкортостане (Россия). Гаплогруппы D и D1 не объединяются в монофилетическую группу, таким образом, *B. daphne* является полифилетическим согласно митохондриальной дендрограмме. Бабочки, принадлежащие к гаплогруппам D и D1 не различаются по окраске крыльев (Рисунок 6 a,b). Минимальная p-дистанция между гаплогруппами D и D1 составляла 0,7%, в то время как минимальная дистанция между гаплогруппой D и *B. ino* также составляла 0,7%, а между D1 и *B. ino* – 0,6% (Таблица 1). Таким образом, уровень дивергенции митохондриального гена *COI* между несомненными видами *B. daphne* и *B. ino* сравним с уровнем дивергенции гаплогрупп D и D1.

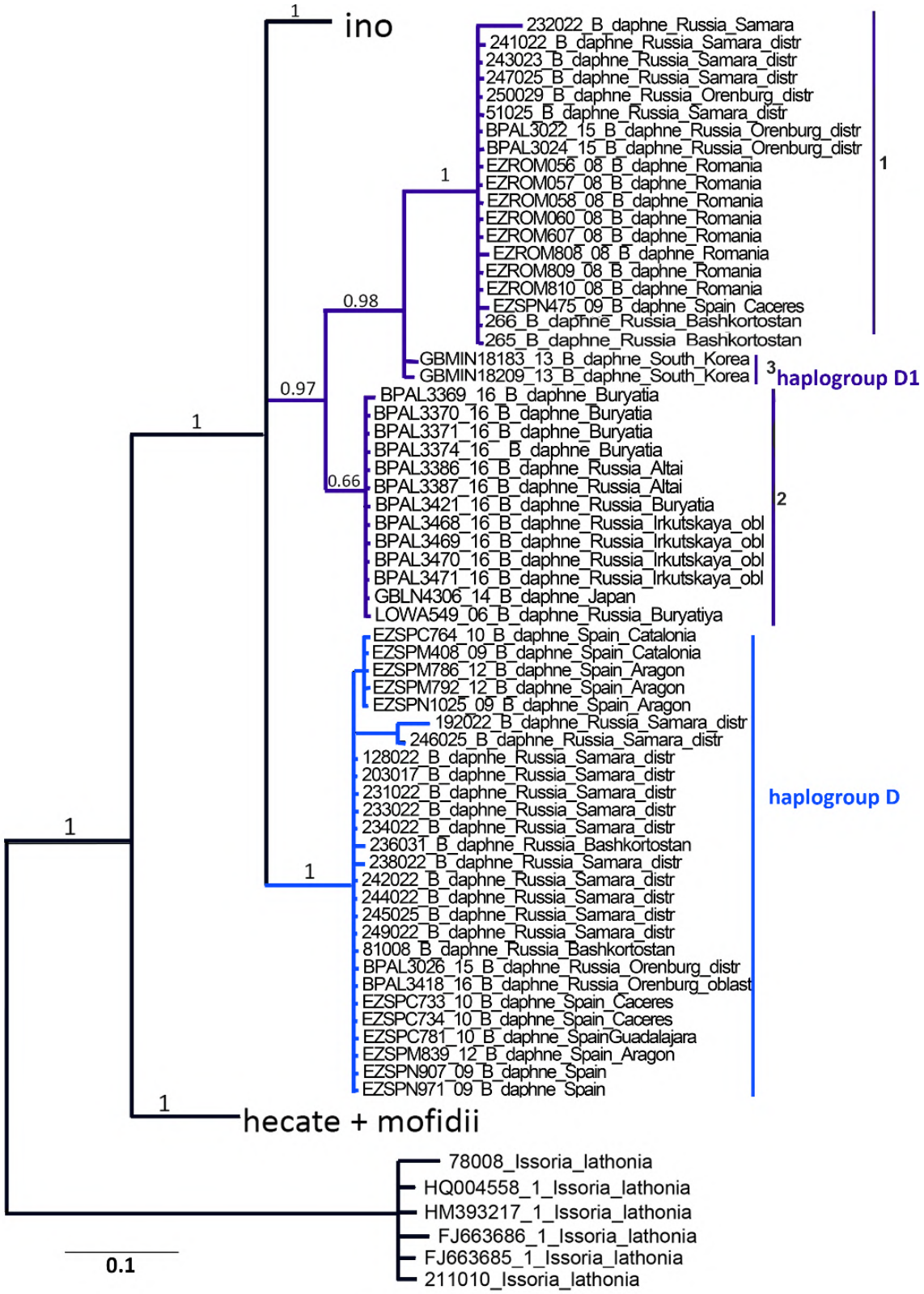

Рисунок 7. Дендрограмма, построенная с помощью Байесова метода на основании анализа фрагмента гена *COI* бабочек рода *Brenthis*. Основные гаплогруппы *B. daphne* (D1 и D) показаны фиолетовым и синим цветом соответственно. 1, 2 и 3 обозначают обособленные генетические линии в пределах гаплогруппы D1. На ветвях обозначены байесовские апостериорные вероятности. Из [27].

Таблица 1. Минимальные и максимальные p-дистанции между ДНК-баркодами бабочек рода *Brenthis* (%). Из [27].

	B. ino		*B. daphne* (D)		*B. daphne* (D1)		*B. hecate*		*B. mofidii*	
	max	min	max	min	max	min	max	min	max	min
B. ino	1.8	0								
B. daphne (D)	2.5	0.7	1.3	0						
B. daphne (D1)	2.1	0.6	2.4	0.7	1.8	0				
B. hecate	3.1	0.9	4.2	1.8	3.7	1.6	2.3	0		
B. mofidii	2.5	1.5	3.3	2	2.9	1.9	1.6	0.2	1.6	0

Что касается остальных видов рода, то следует отметить, что вид *Brenthis ino*, несмотря на свое широкое распространение, не демонстрирует выраженного митохондриального полиморфизма. ДНК-баркоды видов *B. hecate* и *B. mofidii* формируют два гетерогенных кластера, которые не совпадают с традиционными видовыми границами [27]. Внутри вида *B. hecate* выделяется несколько митохондриальных линий с максимальной p-дистанцией между ними 2,3% (Таблица 1).

Чтобы выяснить происхождение и таксономическую принадлежность выявленных митохондриальных линий, нами были проведены два дополнительных филогенетических анализа: 1) с использованием ядерных генов и 2) с использованием морфологических признаков.

В первом дополнительном анализе нами было использовано четыре ядерных гена: *carbamoyl-phosphate synthetase 2, aspartate transcarbamylase and dihydroorotase* (*CAD*), *sarco/endoplasmic reticulum Ca(2 +)-ATPase* (*Ca-ATPase*), *arginine kinase* (*Argk*) и *wingless* (*wg*). Дендрограммы, построенные для отдельных генов с помощью Байесова метода, имели не противоречащую друг другу топологию [27], и, поскольку одиночные маркеры имеют очень слабый филогенетический сигнал, был проведен анализ для конкатенированной матрицы. Разные аллели для одного образца были выделены в результате кодирования двойных пиков на хроматограмме согласно номенклатуре ИЮПАК с последующим выделением гаплотипов в программе PHASE 2.1.1 [148] и обозначены буквами «a» и «b». Сравнение дендрограмм, построенных для ядерного и митохондриального выравниваний представлено на рисунке 8.

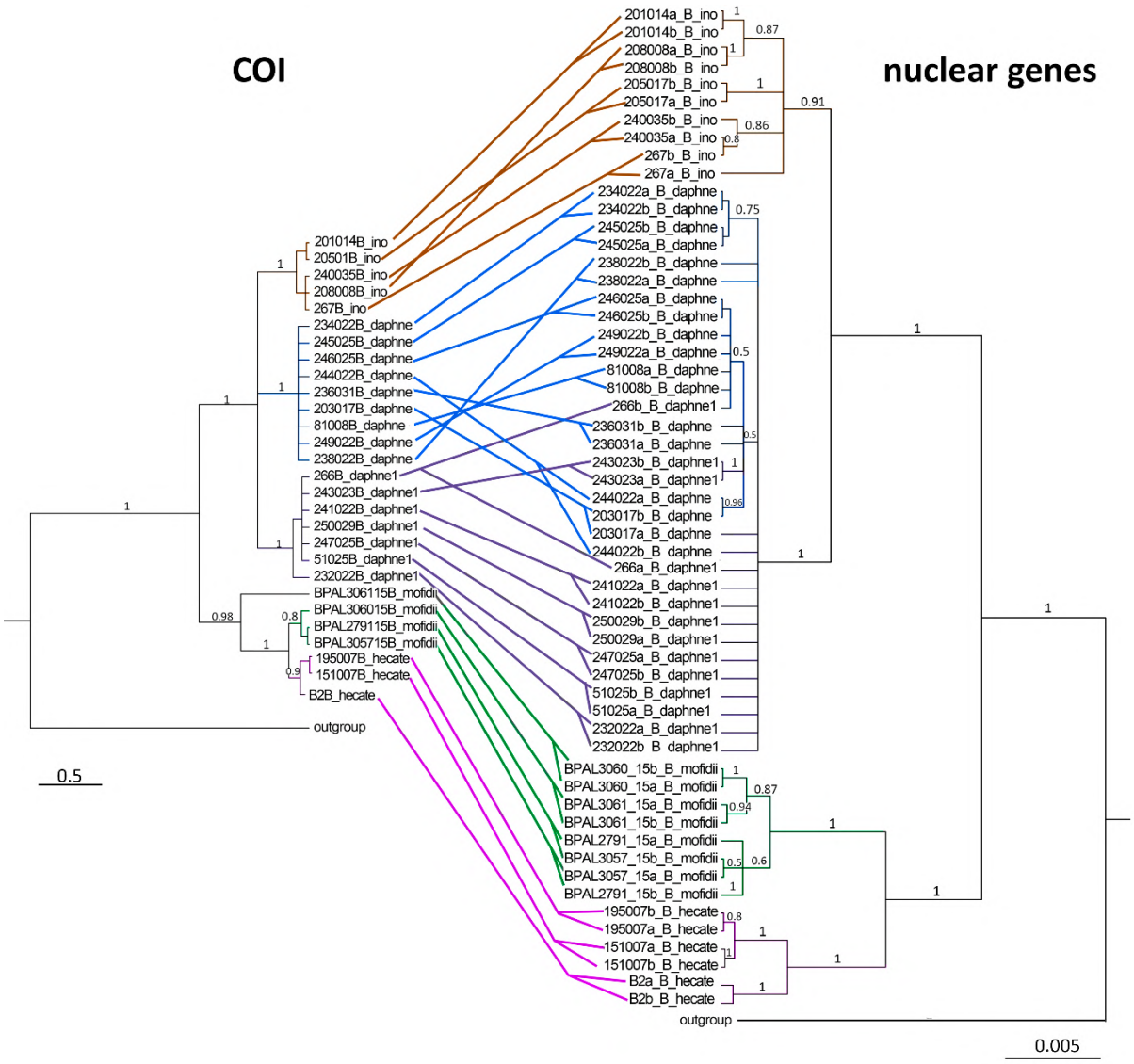

Рисунок 8. Сравнение дендрограмм, полученных при помощи Байесова метода в результате анализа фрагмента гена *COI* (слева) и набора ядерных маркеров (справа) для бабочек рода *Brenthis*. На ветвях обозначены байесовские апостериорные вероятности. Разные гаплотипы одного индивидуума обозначены буквами «a» и «b». Из [27].

Согласно результатам анализа ядерных маркеров, виды *B. daphne* и *B. ino* являются монофилетическими. Представители митохондриальных гаплогрупп D и D1 не образуют отдельных линий на дендрограмме, построенной для ядерных маркеров. Максимальные p-дистанции для конкатенированной матрицы внутри вида *B. daphne* составляли 0,3%, а между видами *B. daphne* и *B. ino* – варьировали между 0,4 и 0,8% (Таблица 2). *Brenthis mofidii* и *B. hecate* формируют отдельные кластеры с высокими поддержками. Образец BPAL306015, соответствующий по окраске крыльев *B. mofidii,* согласно анализу ядерных маркеров, относится к кластеру *mofidii*, но при анализе гена *COI* обнаруживается в качестве базальной ветви в группе *B. hecate* + *B. mofidii* (Рисунок 8).

Таблица 2. Минимальные и максимальные p-дистанции, рассчитанные для набора ядерных маркеров (%). Из [27].

	B. ino		*B. daphne*		*B. hecate*		*B. mofidii*	
	max	min	max	min	max	min	max	min
B. ino	0.4	0						
B. daphne	0.8	0.4	0.3	0				
B. hecate	2.8	1.3	3.3	1.3	1.1	0.1		
B. mofidii	2.5	0.9	2.8	1.1	2.8	0.9	1.2	0.1

Во втором дополнительном анализе были использованы 11 фенотипических признаков (см. полную матрицу признаков в [27]). Кладистический анализ этих признаков выявил, что *B. hecate* + *B. mofidii* и *B. ino* + *B. daphne* формируют монофилетические группы с высокими поддержками (Рисунок 9). Кладистический анализ с использованием меньшего набора морфологических признаков проводился и ранее [21], но была выявлена лишь монофилия рода *Brenthis* и пары видов *B. hecate* + *B. mofidii*, при этом взаимоотношения между оставшимися таксонами оставались неразрешенными.

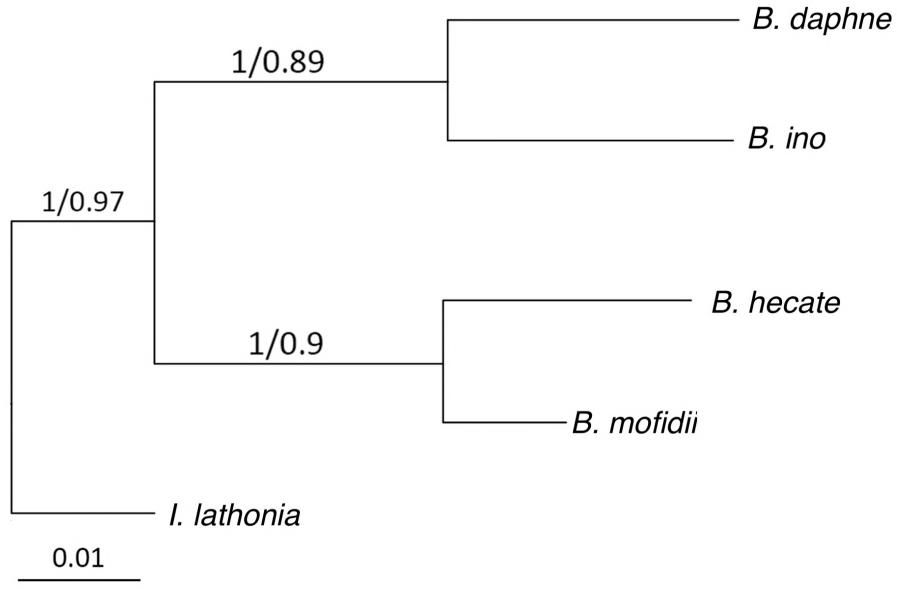

Рисунок 9. Дендргограмма, полученная при помощи Байесова анализа фенотипических признаков бабочек рода *Brenthis*. Вид *Issoria lathonia* был использован в качестве внешней группы. Идентичная топология была выявлена в результате анализа с применением метода Максимальной парсимонии. На ветвях обозначены Байесовские апостериорные вероятности/бутстреп поддержки. Из [27].

Таким образом, таксономические отношения, выявленные в результате анализа фенотипических признаков, полностью согласуются с результатами анализа ядерных маркеров, показавших монофилию всех четырех видов в составе рода *Brenthis,* а также пар видов *B. hecate* + *B. mofidii* и *B. ino* + *B. daphne* (Рисунки 8 и 9). При этом, как показано выше, анализ ДНК-баркодов выявил в пределах рода принципиально другой, противоречивый паттерн распределения митохондриальных гаплогрупп и полифилию таксонов группы (Рисунок 7).

Известно, что подобные несоответствия (mito-nuclear discordance) могут быть вызваны инфицированием эндосимбионтом *Wolbachia,* а также митохондриальной интрогрессией, неполной репродуктивной изоляцией или неполной сортировкой линий [83–85]. Кроме того, в теории, присутствие дивергировавших митохондриальных гаплогрупп в составе *B. daphne* может свидетельствовать о том, что этот таксон представляет из себя комплекс криптических видов. Эта гипотеза может быть проверена при помощи анализа комбинации ядерных генов, дающих большее разрешение на видовом уровне, чем выбранные нами маркеры. Однако, при соответствии митохондриальных кластеров криптическим видам ожидается, они будут демонстрировать привязку к разным географическим территориям (в случае аллопатрии) или экологическим нишам (в случае симпатрии) [26,149]. Поскольку представители двух гаплогрупп в составе *B. daphne* обнаруживались в полной симпатрии и не демонстрировали ни морфологических различий, ни дивергенции по ядерным генам, мы заключаем, что митохондриальный полиморфизм внутри этого таксона, скорее всего, не связан с процессами видообразования [27].

В практике филогенетического анализа часто встречается подход, когда в одну матрицу объединяются митохондриальные и ядерные данные чтобы усилить филогенетический сигнал [150–152]. Можно предполагать, что такая практика ведет к неадекватным результатам в тех случаях, когда анализ митохондриального и ядерного выравниваний приводят к разным топологиям. В нашем исследовании мы протестировали этот подход, объединив митохондриальную и ядерную матрицы для рода *Brenthis* и проведя Байесов анализ. Митохондриальный ген *COI* обладает большим филогенетическим сигналом, чем комбинация ядерных генов (84 и 50 информативных позиции при длине выравниваний 615 и 1779 п.н.), что влияет на результирующую топологию. Два митохондриальных кластера D и D1 поддержаны на комбинированной дендрограмме, однако объединены в кластер, соответствующий *B. daphne* (Рисунок 10). Вид *B. mofidii* соответствует парафилетическому кластеру на комбинированной дендрограмме, что противоречит результатам, полученным при анализе как ядерных, так и митохондриального маркеров. Таким образом, химерная матрица привела к результатам, которые не соответсвуют ни митохондриальной, ни ядерной филогении и фактически не имеют биологического смысла.

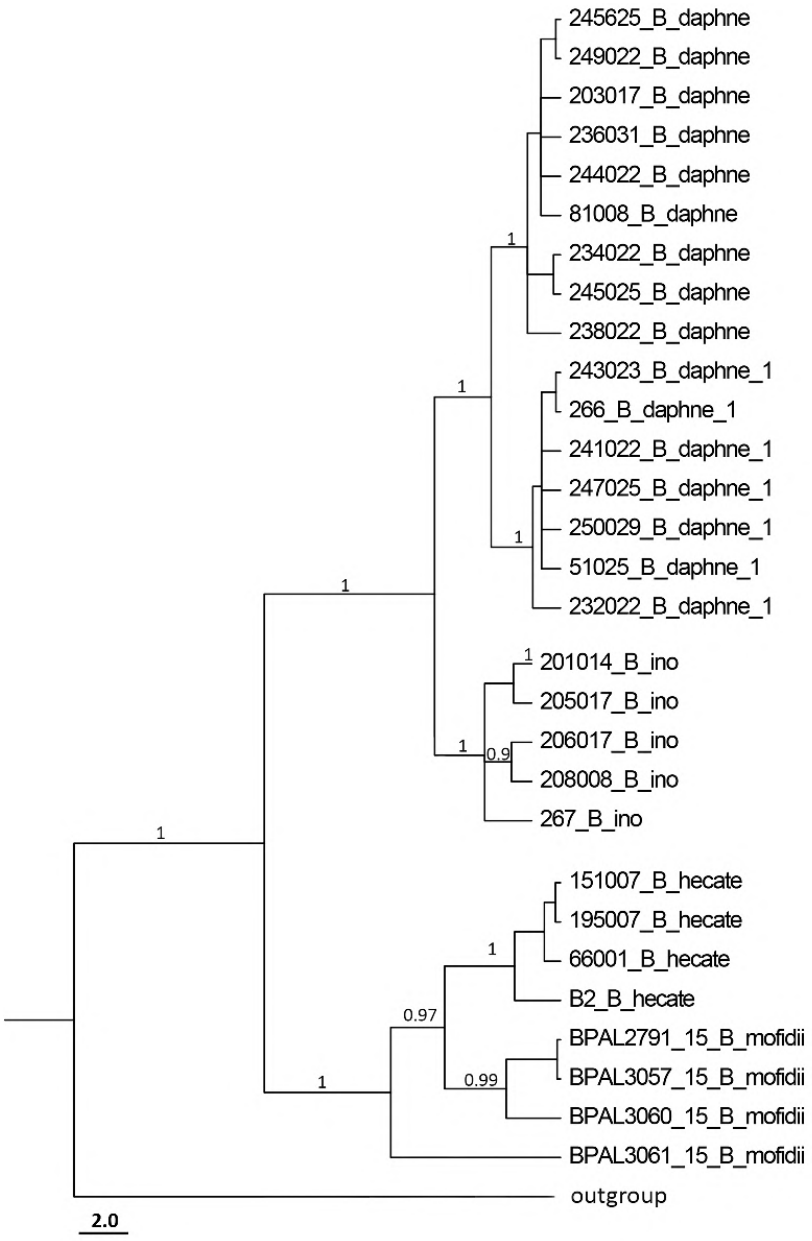

Рисунок 10. Дендрограмма, построенная для бабочек рода *Brenthis* с применением Байесова метода для конкатенированной матрицы, объединяющей выравнивания для гена *COI* и набора ядерных генов. На ветвях обозначены байесовские апостериорные вероятности. Из [27].

В целом, для рода *Brenthis* комбинация ядерных маркеров отражает видовые границы в соответствии с результатами, полученными с помощью анализа морфологических признаков, более точно, чем ДНК-баркоды [27]. В связи с этим, мы предлагаем сопровождать таксономические гипотезы, выдвинутые на основании анализа единичных филогенетических маркеров, с помощью интегративного анализа, а также избегать объединения митохондриальных и ядерных молекулярных данных в единую матрицу, не убедившись предварительно в отсутствии противоречий между данными разного типа.

3.2 Анализ комплекса *Hyponephele lycaon – H. lupina*

Род *Hyponephele* Muschamp, 1915 включает около 40 видов, которые очень сходны морфологически. Межвидовые различия в строении гениталий самцов выражены слабо или отсутствуют, самки практически не различимы по окраске крыльев, самцы часто (но не всегда) различаются по форме андрокониальных пятен на передних крыльях [153]. Выделяемый в пределах рода комплекс таксонов, близких к видам *Hyponephele lycaon* Rottemburg 1775 и *H. lupina* Costa 1863 (Рисунок 11 a, c), представляет из себя чрезвычайно сложную для таксономического анализа группу, о чем свидетельствует большое число противоречивых таксономических гипотез (см. ссылки ниже), предложенных для описания отношений между видами этой группы. До начала 20 века ученые считали, что в этот комплекс входит только один вид *H. lycaon*, представленный несколькими внутривидовыми формами [154]. Затем было установлено, что *Hyponephele lycaon* и *H. lupina* являются разными видами, а не формами одного вида [155].

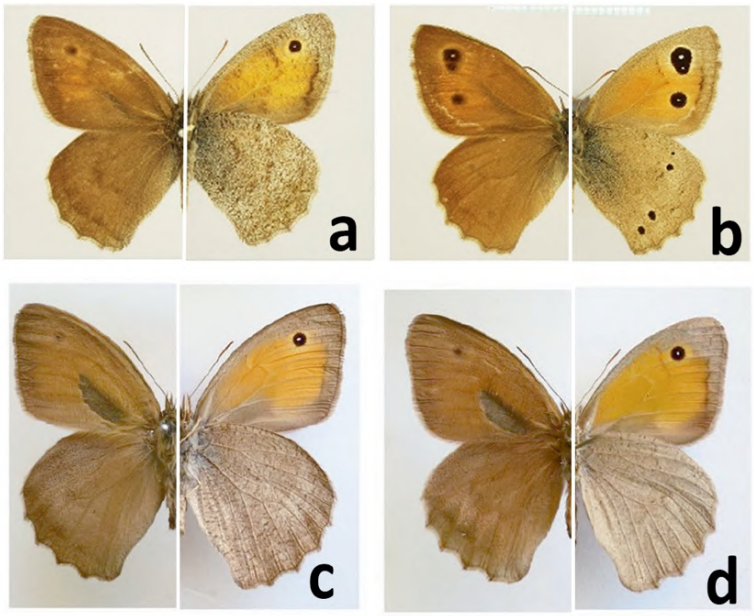

Рисунок 11. Бабочки комплекса *Hyponephele lycaon – lupina*. Слева показана верхняя сторона, справа – испод крыльев. **a** *H. lycaon*, **b** *H. sifanica*, **c** *H. lupina*, **d** *H. mauritanica.* Из [29].

Во второй половине 20 века из Ирана был описан вид *H. lycaonoides* D. Weiss, 1978, предположительно симпатричный с *H. lycaon* в Иране и отличающийся строением гениталий самцов [156]. В недавних работах видовая самостоятельность *H. lycaonoides* была поставлена под сомнение [157].

В обзоре видов рода *Hyponephele,* обитающих на территории бывшего СССР*,* в составе комплекса *H. lycaon* было выделено 6 видов (*H. lycaon*, *H. galtscha* Grum-Grsimailo, 1893, *H.*

pasimelas Staudinger, 1886, *H. dzhungarica* Samodurov, 1996, *H. przhewalskyi* Dubatolov, Sergeev & Zhdanko, 1994, *H. lycaonoides*), из них *H. lycaonoides* симпатричен с *H. lycaon*, а все остальные виды аллопатричны [158]. В обзоре видов рода *Hyponephele*, обитающих на территории Палеарктики, Эквайлер и Бозано предложили классификацию комплекса, в которой большинство ранее описанных популяций и таксонов этой группы относились к виду *H. lycaon*, за исключением аллопатричных видов *H. maroccana*, *H. sifanica*, *H. galtscha* и симпатричного вида *H. lycaonoides* [153].

В нашем исследовании мы протестировали предложенную Самодуровым с соавторами [158] гипотезу о множественных аллопатричных видах в составе *H. lycaon*, основываясь на анализе ядерных и митохондриальных молекулярных маркеров, а также морфологических признаков и анализа экологических предпочтений бабочек в зонах симпатрии. В результате реконструкции филогении, основанной на анализе ДНК-баркодов, было выявлено восемь поддержанных монофилетических групп в составе комплекса *H. lycaon* (Рисунки 12 и 13). Три из них соответствуют видам *H. maroccana*, *H. galtscha* и *H. sifanica*, в то время как остальные включают в себя представителей *H. lycaon* sensu lato: гаплогруппа L1 включает в себя 63 образца из Евразии (определенных как *H. lycaon* s. str., *H. pasimelas*, *H. dzhungarica* и *H. przhewalskiy),* L1b – шесть образцов с Западного Кавказа, L2 – 30 образцов из Израиля, Турции, Северного Ирана и Азербайджана, L3 – 60 образцов из Израиля, Турции, Юго-западного Ирана и Азербайджана (включая *H. lycaon* и *H. lycaonoides*), L4 – 3 образца из Израиля и Ирана.

Гаплогруппа L4 кластеризуется с кладой *H. sifanica*, вместе с ней образуя наиболее базальную группу комплекса *H. lycaon*. Значительный уровень дивергенции группы L4 и остальных представителей *H. lycaon* (p-дистанции 9,3 -11,7%) (Таблица 3) позволяет предполагать, что эта группа является древним изолятом *H. lycaon* или, напротив, произошла в результате быстрой митохондриальной эволюции, спровоцированной эндосимбионтом *Wolbachia* [85].

Таблица 3. Средние p-дистанции между митохондриальными гаплогруппами комплекса *Hypionephele lycaon (%).*

	L1	L1b	*galtscha*	L2	L3	L4	*sifanica*
L1b	3,1						
galtscha	6,7	5,6					
L2	4	2,9	4,9				
L3	5,9	4,6	4,2	3,5			
L4	11,7	10,8	9,2	9,6	9,3		
sifanica	10,6	9,7	8,4	9,2	6,7	6,1	
lupina	8,4	8,1	7,4	7,7	9,5	11,2	9,4

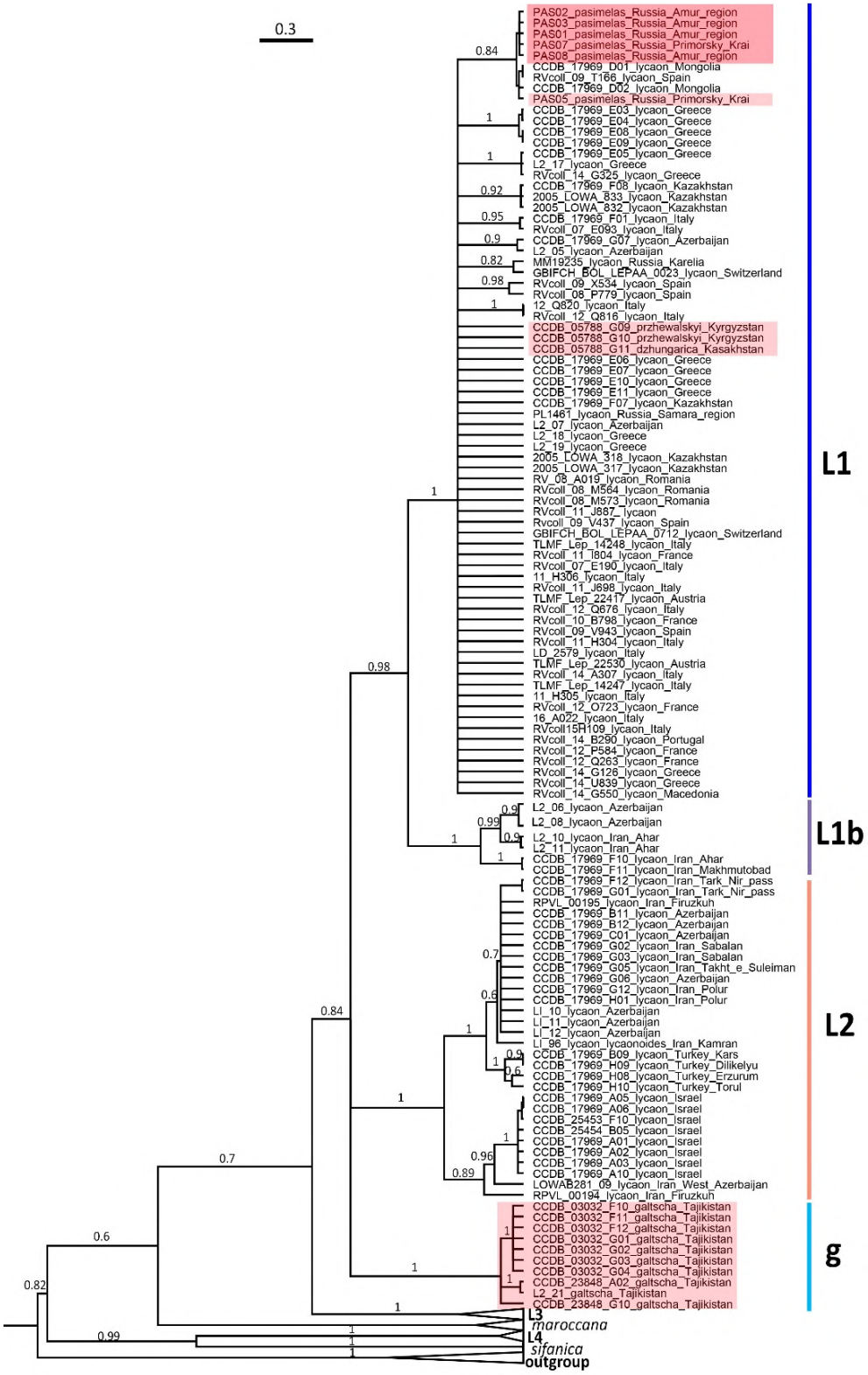

Рисунок 12. Фрагмент дендрограммы, построенной с использованием метода Байеса на основании анализа фрагмента гена *COI* для комплекса видов *Hyponephele lycaon*. Показаны гаплогруппы L1, L1b, L2 и *H. galtscha.* На ветвях показаны байесовские апостериорные вероятности. Красным обозначены образцы, определенные в соответствии с гипотезой о множественных аллопатричных видах как *Hyponephele pasimelas, H. przhewalskyi*, *H. dzhungarica* и *H. galtscha.*

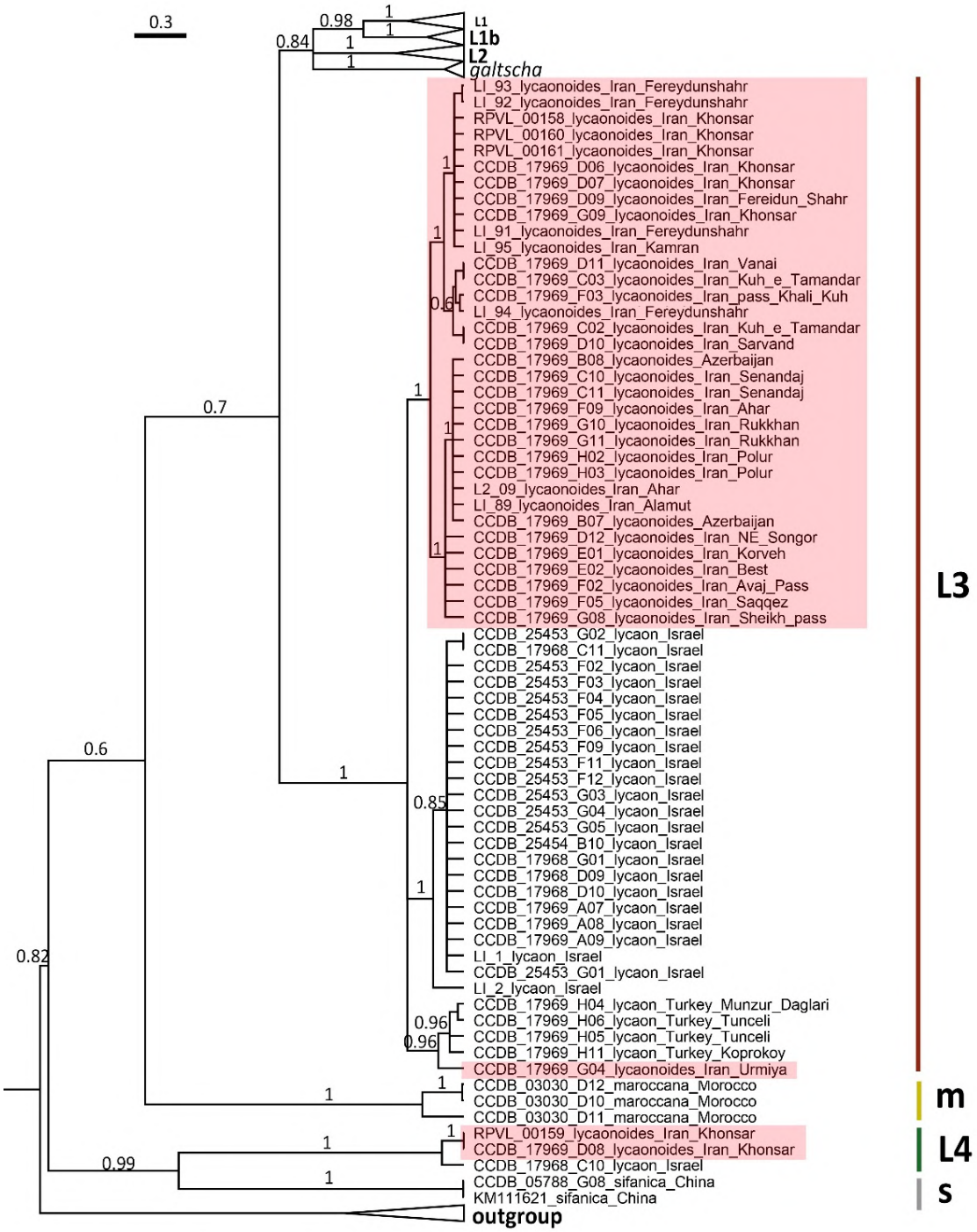

Рисунок 13. Фрагмент дендрограммы, построенной с использованием метода Байеса на основании анализа фрагмента гена *COI* для комплекса видов *Hyponephele lycaon*. Показаны гаплогруппы L3, L4, *H. maroccana* и *H. sifanica.* На ветвях показаны байесовские апостериорные вероятности. Красным обозначены образцы, определенные как *H. lycaonoides*.

Мы провели также анализ трех ядерных маркеров: *glyceraldehyde 3-phosphate dehydrogenase* (*GAPDH*), *sarco/endoplasmic reticulum Ca(2 +)-ATPase* (*Ca-ATPase*) и *wingless* (*wg*) (Рисунок 14). Несмотря на сравнительно низкие поддержки, что говорит о слабом филогенетическом сигнале выбранных маркеров, наблюдается тенденция к образованию тех же клад, что и при анализе ДНК-баркодов. Образцы гаплорупп L1 и L1b (последняя представлена одним

экземпляром) формируют политомию, в то время как образцы гаплогрупп L3 и L2 кластеризуются вместе. Образцы L2 образуют политомию, а L3 формируют поддержанную кладу. Образцы гаплогруппы L4 CCDB-17969 D08 и CCDB-17968 C10 попадают в клады L3 и L2, соответственно (Рисунок 14).

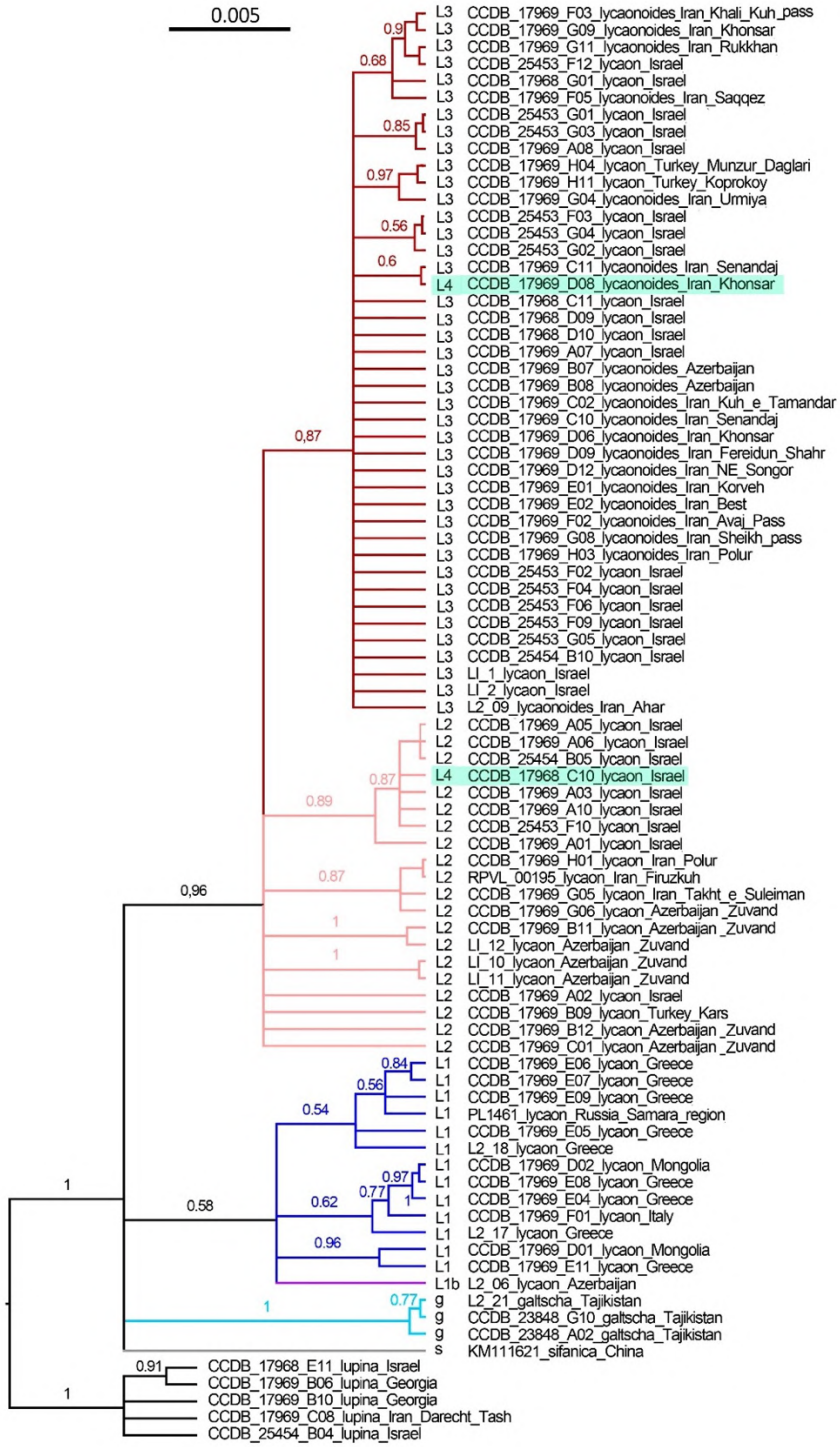

Рисунок 14. Дендрограмма, построенная для комплекса видов *Hyponephele lycaon* с использованием метода Байеса на основании анализа ядерных генов *GAPDH*, *Ca-ATPase* и *wingless.* На ветвях показаны байесовские апостериорные вероятности. Образцы, относящиеся к митохондриальной гаплогруппе L4, выделены зеленым.

Мы применили методы геометрической морфометрии для анализа строения гениталий самцов (Рисунок 15). Представители гаплогруппы L3, обладающие более длинным ункусом и укороченными субунциями, формируют хорошо различимый отдельный кластер, в то время как кластеры образцов гаплогрупп L1 и L2 в значительной степени пересекаются.

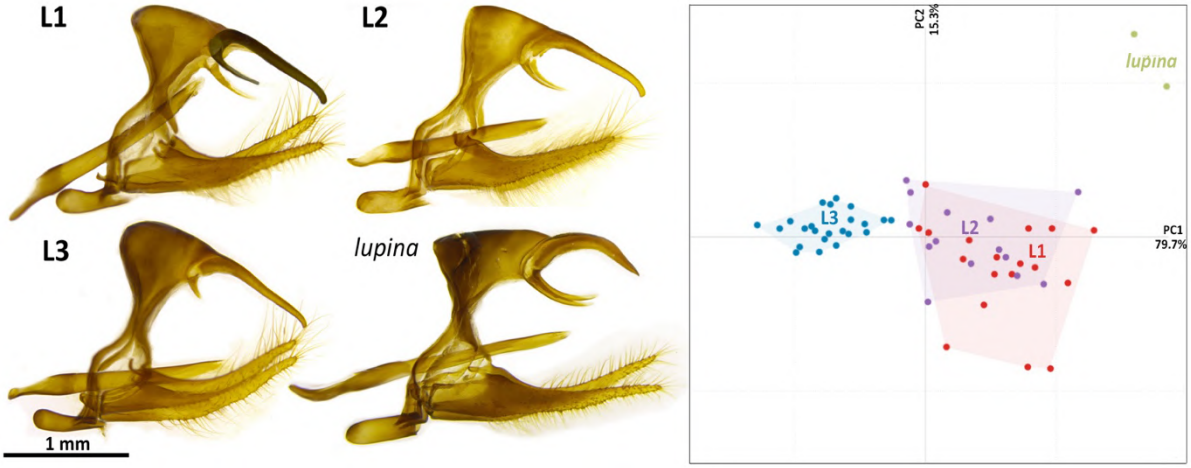

Рисунок 15. Строение и геометрическая морфометрия гениталий самцов рода *Hyponephele*. Слева показаны гениталии самцов *H. lycaon* митохондриальных гаплогрупп L1, L2 и L3, а также *H. lupina*. Черным обозначен пример контура ункуса и субунций, использовавшегося для проведения морфометрического анализа. Справа показаны результаты анализа главных компонент. На осях показаны первые две главные компоненты (PC1 и PC2), объясняющие 79,9% и 15,3% дисперсии формы контуров гениталий, соответственно.

Таким образом, согласно полученным данным, бабочки гаплогруппы L3, имеют обширный ареал, включающий территорию Ирана, Турции, Азербайджана и Израиля, и характеризуются (1) уникальной митохондриальной гаплогруппой, (2) уникальным набором нуклеотидных замен в ядерных генах и (3) уникальным типом генитального аппарата у самцов. На наш взгляд, конгруэнтность этих трех, несвязанных между собой групп признаков является доказательством того, что бабочки гаплогруппы L3 представляют отдельный вид, который в пределах почти всего своего ареала встречается симпатрично с другим видом, представленным гаплогруппами L1 и L2 и отнесенным нами к виду *H. lycaon* s.s. (см. обсуждение ниже). Важно подчеркнуть, что вид с гаплогруппой L3 не идентичен номинальному таксону *H. lycaonoides* и включает как особей, ранее определявшихся как *H. lycaonoides*, так и особей, ранее определявшихся как *H. lycaon* (см. рисунок 13). Визуальный осмотр показывает, что бабочки гаплогруппы L3 обладают, как правило, большим размером крыла и более светлым исподом заднего крыла, чем бабочки гаплогрупп L1и L2 (хотя эти признаки изменчивы и навряд ли могут использоваться для видовой идентификации).

Возникает вопрос о названии, которое может использоваться в качестве валидного для вида с гаплогруппой L3. С территории Турции, Ирана и стран Леванта описано большое число

таксонов из комплекса *H. lycaon*, каждое из которых может представлять как *H. lycaon* (гаплогруппы L1 + L2), так и ранее нераспознанный вид с гаплогруппой L3. Гессельбарт с коллегами предлагают использовать для бабочек с большим размером крыла название *Hyponephele lycaon var. collina* Röber, 1897, поскольку это старейшее пригодное название для бабочек из Турции и вообще из Передней Азии, считая *H. lycaon libanotica* Staudiger, 1901 (страны Леванта) и H. *lycaonoides* D. Weiss, 1978 (Иран) синонимами H. *lycaon var. collina* Röber, 1897 [146]. К сожалению, типовой материал *Hyponephele lycaon var. collina* Röber, 1897 не сохранился [146]. Мы считаем, что правильным решением номенклатурного вопроса будет обозначение в будущем неотипа H. *lycaon var. collina* из типового места этого таксона (юго-восточная Турция, окрестности населенного пункта Gülek), причем в качестве неотипа должен быть выбран самец, несущий молекулярные и морфологические признаки гаплогруппы L3. Поэтому в тексте ниже мы используем название *H. collina* для вида с гаплогруппой L3.

Что касается представителей митохондриальных гаплогрупп L1 и L2, то мы не видим оснований, чтобы относить их к разным видам, поскольку они неразличимы по морфологии и не дифференцированы по изученным ядерным генам. Эти две гаплогруппы аллопатричны и вероятно представляют две линии вида *H. lycaon*, дивергировавшие в условиях географической изоляции. Эти две линии можно на наш взгляд интерпретировать в качестве подвидов. Подвид *H. lycaon lycaon* (гаплогруппа L1) имеет обширный ареал от Западной Европы по Приморья на Дальнем Востоке включительно, захватывающий также Кавказ и север Средней Азии. Поскольку в пределах *H. lycaon lycaon* отдельные популяции слабо дифференцированы по гену COI (средние внутригрупповые p-дистанции по митохондриальному генетическому маркеру *COI* для гаплогруппы L1 составляют 0,7%), а ядерные маркеры мономорфны, мы не можем подтвердить гипотезу Самодурова с соавторами [158] о видовой самостоятельности для аллопатричных таксонов *H. pasimelas* (Дальний Восток), *H. dzhungarica* (Джунгарский Алатау) и *H. przhewalskiy* (Тянь-Шань).

Для подвида с митохондриальной гаплогруппой L2 мы предлагаем использовать название *Hyponephele lycaon zuvandica* Samodurow & Koroljev, 1996, поскольку молекулярные данные были получены для образцов из типовой местности этого таксона (Азербайджан, Талыш). Этот подвид широко распространен в Передней Азии (включая Турцию и Израиль). В Иране он встречается в западной части этой страны, часто симпатрично с *H. collina*, однако отсутствует в южном Загросе, где *H. collina* присутствует.

В процессе полевых исследований мы изучали биотопическую приуроченность и фенологию бабочек комплекса *H. lycaon*-*H. lupina*. Особое внимание было уделено *H. lycaon* и *H. collina* в зоне их симпатрии. Фенология этих видов совпадает, но в зоне их полной симпатрии присутствует разделение по биотопам. По нашим наблюдениям, *H. collina* – более

ксерофильный вид, населяющий сухие остепненные склоны, в то время как *H. lycaon*– более мезофильный вид, предпочитающий лугостепи и опушки леса. В литературе указано, что в провинции Хаккари в Турции *H. lycaon* встречается на больших высотах, чем *H. collina* [153], что подтверждается нашими наблюдениями в Иране. Однако в Израиле на горе Хермон наблюдается обратная картина. Предположительно, выбор местообитания основан не на высоте, а на влажности биотопа. В более сухих районах Ирана и восточной Турции *H. lycaon* летает в верхнем поясе гор ближе к субальпийскому поясу, который является в этих районах более влажным, чем лежащие ниже участки. В Израиле *H. lycaon* приурочен к лесному поясу (1400-1600 м), который влажнее, чем лежащий выше субальпийский пояс. Таким образом, различия в экологических предпочтениях является дополнительным свидетельством в пользу видовой самостоятельности этих таксонов.

Мы провели также филогенетический и таксономический анализ всего комплекса *H. lycaon – H. lupina* [29]. В составе группы *Hyponephele lupina* проведенный нами анализ ДНК-баркодов выявил кластеры, соответствующие таксонам *H. lupina lupina*, *H. lupina mauritanica* Oberthür, 1881, *H. maroccana* Blachier, 1908, *H. sifanica*, *H. cheena* Moore 1865, и *H. interposita* Erschoff, 1874 [29]. Уровень различий между *H. lupina* и *H. mauritanica* варьировал от 3,5 до 4,9%, что превышает «стандартный» уровень межвидовой дивергенции по гену *COI* [50]. Представители клад *lupina* и *mauritanica* также различаются по окраске испода крыла самцов, форме андрокониальных пятен и строению гениталий самцов, в связи с чем предлагается рассматривать *H. mauritanica* в качестве самостоятельного вида [29]. Тем не менее, для подтверждения видового статуса *H. mauritanica.* необходим анализ дополнительных молекулярных маркеров.

Таким образом, на примере комплекса *H. lycaon – H. lupina* мы продемонстрировали, что первоначальные таксономические гипотезы, высказанные на основании анализа митохондриальных ДНК-баркодов, должны быть тщательным образом проверены с использованием интегративного подхода, включающего изучение ядерных генов и морфологии. При этом полученные на выходе классификации могут существенно отличаться как от предварительных таксономических гипотез, основанных на ДНК-баркодах, так и от таксономических интерпретаций, которые можно найти в ранее опубликованных сводках и ревизиях.

Глава 4. Использование полных геномов для выявления таксономической структуры и эволюционной истории комплекса *Melitaea acentria-M. persea- M. didyma*

4.1. Паттерны генетической дифференциации и геномная интрогрессия

Ретикулярная эволюция (reticulation events), включающая в себя такие процессы как межвидовая гибридизация и горизонтальный перенос генов, а также неполная сортировка линий, встречается в природе чаще, чем предполагалось ранее, и вызывает сложности при определении видовых границ [159,160]. Использование большого числа данных, которые возможно получить при анализе полных геномов, позволяет изучить не только степень диверсификации таксонов, но и их эволюционную историю, что в подобных случаях необходимо для постановки корректной таксономической гипотезы.

У бабочек рода *Melitaea* в контактных зонах близких видов иногда отмечаются особи со смешанной морфологией, которые предположительно возникли в результате межвидовой гибридизации [20]. Иногда такие индивидуумы формируют целые популяции, которые были отмечены для *M. didyma* и *M. persea* в Турции и Ливане, а также для *M. persea* и *M. mixta* в Афганистане и Пакистане, *M. athalia* и *M. nevadensis* в Швейцарии [20].

Вид *Melitaea acentria* Lukhtanov, 2017, обнаруженный в Израиле и в Ливане, обладает смешанными чертами строения гениталий самцов и окраски крыльев: с одной стороны, он обладает признаками, встречающимися у неблизкородственного симпатричного вида *M. didyma* Esper, 1777, а с другой стороны он морфологически сходен с близким аллопатричным видом *M. persea* Kolar, 1849. В связи с этим предполагалось, что *M. acentria* появился в результате межвидовой гибридизации [7], однако надо отметить, что эта гипотеза при ее выдвижении не тестировалась с использованием ядерных маркеров.

В нашем исследовании для изучения генетической структуры и выявления взаимоотношений между тремя названными выше таксонами было осуществлено секвенирование 27 полных геномов *M. acentria, M. didyma* и *M. persea* [30]. В результате анализа полученных полных геномов с использованием генома *Melitaea cinxia* в качестве референса [38] были выявлены однонуклеотидные замены (SNP). Анализ полученных данных с использованием метода главных компонент (PCA) выявил, что изученные особи формируют три четких кластера, которые соответствуют трем изученным таксонам; при этом один образец (Mel11) занимает промежуточное положение между ними (Рисунок 16).

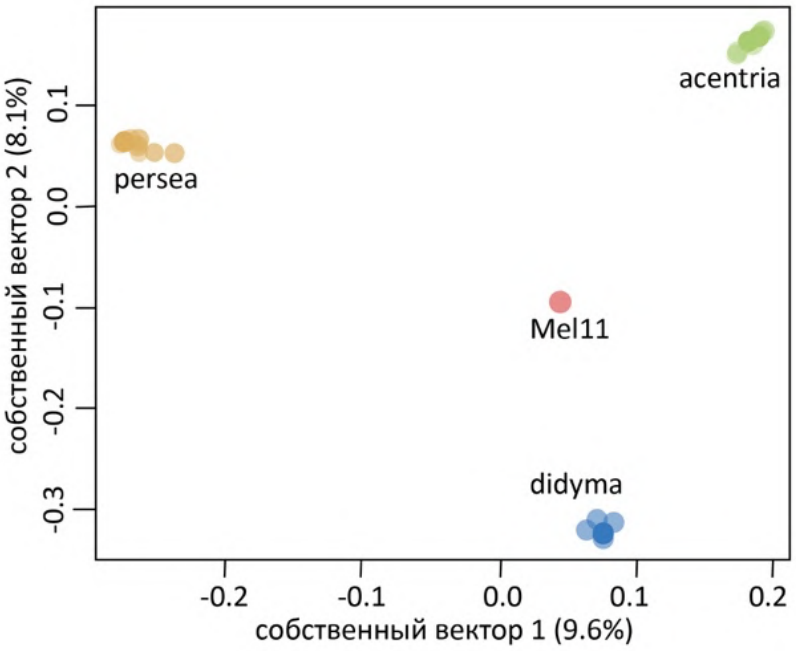

Рисунок 16. Анализ геномных SNP с использованием метода главных компонент. Видоспецифичные кластеры показаны цветом. Образец Mel11 выделен красным. На осях показаны первые два собственных вектора, объясняющие 9.6% и 8.1% дисперсии геномных вариантов. Из [30].

Филогенетическая сеть, построенная в программе Splitstree показывает, что образец Mel11 может иметь гибридное происхождение (*M. acentria* x *M. didyma)* (Рисунок 17). Остальные образцы формируют видоспецифичные кластеры с высокими поддержками.

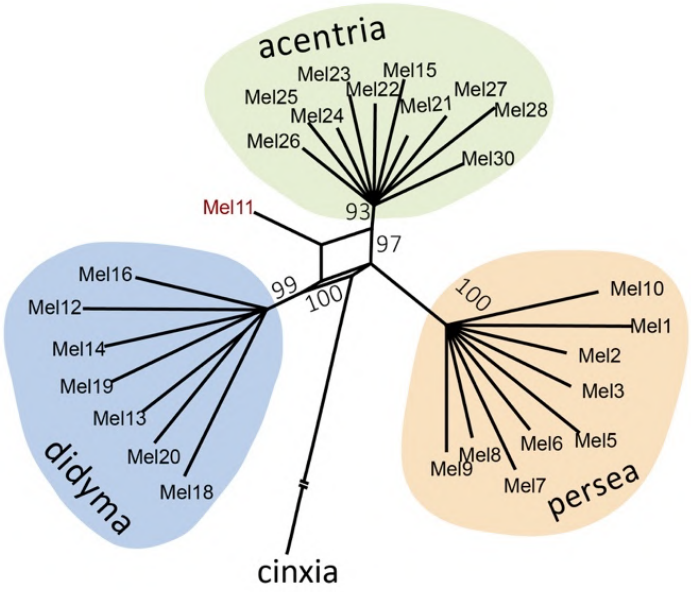

Рисунок 17. Филогенетическая сеть, построенная в программе SplitsTree для полногеномных SNP. Видовые кластеры показаны цветом. Гибридный образец Mel11 выделен красным; видно, что он не включен ни в кластер *M. acentria*, ни в кластер *M. didyma*. На ветвях обозначены бутстреп-поддержки. Из [30].

Можно ожидать, что у гибрида первого поколения все видоспецифичные гомозиготные замены, характерные для *M. didyma* и *M. acentria* будут в гетерозиготном состоянии. Из 104760 видоспецифичных SNP у образца Mel11 только 42% процентов были гетерозиготными, при этом 28% были характерны для *M. didyma и* 30% - для *M. acentri*a. Это говорит о том, что образец Mel11 появился в результате обратного скрещивания гибрида первого поколения с родительскими видами.

Обнаружение гибрида между видами *M. didyma* и *M. acentria* свидетельствует о неполной репродуктивной изоляции между ними, что необычно, так как *M. didyma* и *M. acentria* не являются близкородственными видами и относятся к разным группам видов в пределах рода *Melitaea*. Мы провели дополнительный анализ, чтобы подтвердить или опровергнуть наличие потока генов между этими видами. Образец Mel11 был исключен из дальнейшего анализа, чтобы минимизировать вероятность артефактов.

Для выявления следов геномной интрогрессии применялся ABBA-BABA тест [161,162]. ABBA-BABA тест позволяет обнаружить отклонения от строго дихотомической модели дивергенции видов при помощи подсчета числа участков генома, которые дают топологии, не соответствующие предполагаемой дихотомической схеме. Такие отклонения выражаются через метрику, называемую D-статистика. D-статистика была рассчитана для всех трех теоретически возможных топологий, укорененных по *M. cinxia*. Топология (((*persea*, *acentria*) *didyma*) *cinxia*) являлась наиболее правдоподобной, при этом было установлено, что среднее значение D = 0,21 ± 0.005, а значение Z, характеризующее значимость D-статистики, составляло 18.206. Поскольку Z считается значимым, если оно превышает 3 [162], то мы приходим к выводу о наличии межвидовой интрогрессии между *M. didyma* и *M. acentria*. Пропорция интрогрессии f_d [163] была рассчитана отдельно для каждой хромосомы (Рисунок 18).

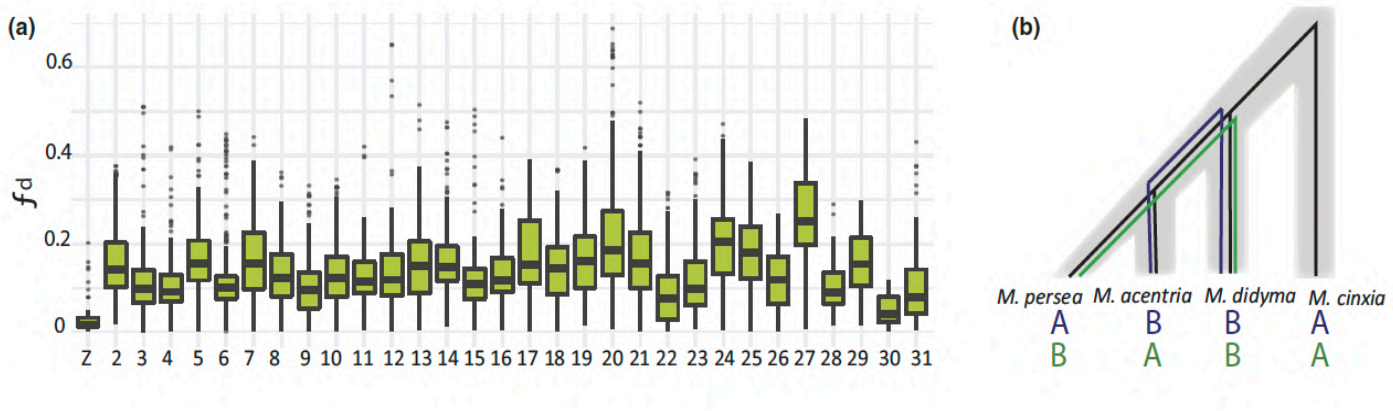

Рисунок 18. Средняя пропорция интрогрессии f_d, рассчитанная отдельно для каждой хромосомы. Справа показана топология, для которой была рассчитана пропорция интрогрессии. Из [30].

Средняя пропорция интрогрессии f_d между *M. acentria* и *M. didyma* составила $f_d = 0.13 \pm 0.035$. Средняя пропорция интрогрессии на половых Z хромосомах была гораздо ниже, чем на аутосомах, и составляла $f_{d_Z} = 0.05 \pm 0.008$, $f_{d_auto} = 0.14 \pm 0.09$. При этом, межвидовая дивергенция, выраженная как индекс фиксации F_{ST}, была примерно одинаковой по всему геному у близких видов *M. acentria* и *M. persea*, но оба вида обладали более высокими дистанциями с *M. didyma* в половых хромосомах, чем в остальных участках генома. Иными словами, различие уровня дивергенции между половыми хромосомами и аутосомами было высоко у *M. acentria* и *M. didyma* ($\Delta F_{ST} = 0{,}44$) и *M. persea* и *M. didyma* ($\Delta F_{ST} = 0{,}39$), но сравнительно низко у *M. acentria* и *M. persea* ($\Delta F_{ST} = 0{,}1$), где $\Delta F_{ST} = F_{ST_Z} - F_{ST_auto}$. Таким образом, было показано, что накопление различий в половых хромосомах и аутосомах происходит неравномерно в течение эволюции – у близких видов половых хромосомы и аутосомы обладают примерно одинаковым F_{ST}, который заметно увеличивается в половых хромосомах у филогенетически более далеких видов (Рисунок 19). Вероятно, разный темп накопления замен может приводить к более быстрому формированию барьеров к интрогрессии на половых хромосомах, чем на аутосомах [30].

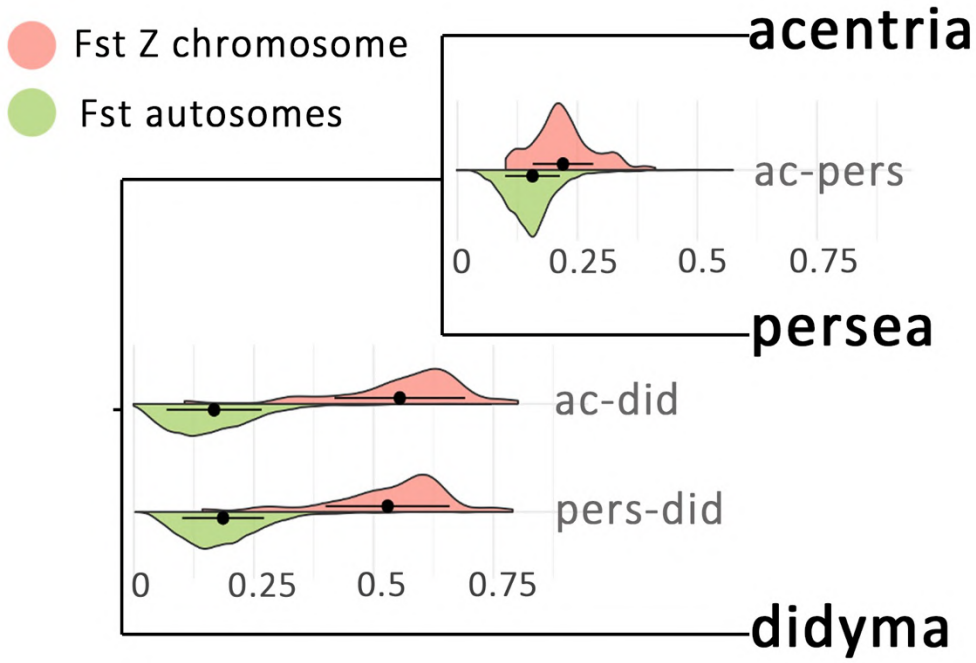

Рисунок 19. Дивергенция половых хромосом и аутосом. Индекс фиксации F_{ST} рассчитан попарно для *M. acentria, M. didyma* и *M. persea* отдельно для половых хромосом и аутосом. Из [30].

Мы оценили уровень нуклеотидного разнообразия (π), и обнаружили, что регионы, обладающие самым высоким нуклеотидным разнообразием, демонстрируют высокую

пропорцию интрогрессии (Рисунок 20). Также было показано, что хромосомы с низкой средней f_d также обладают низким нуклеотидным разнообразием [30].

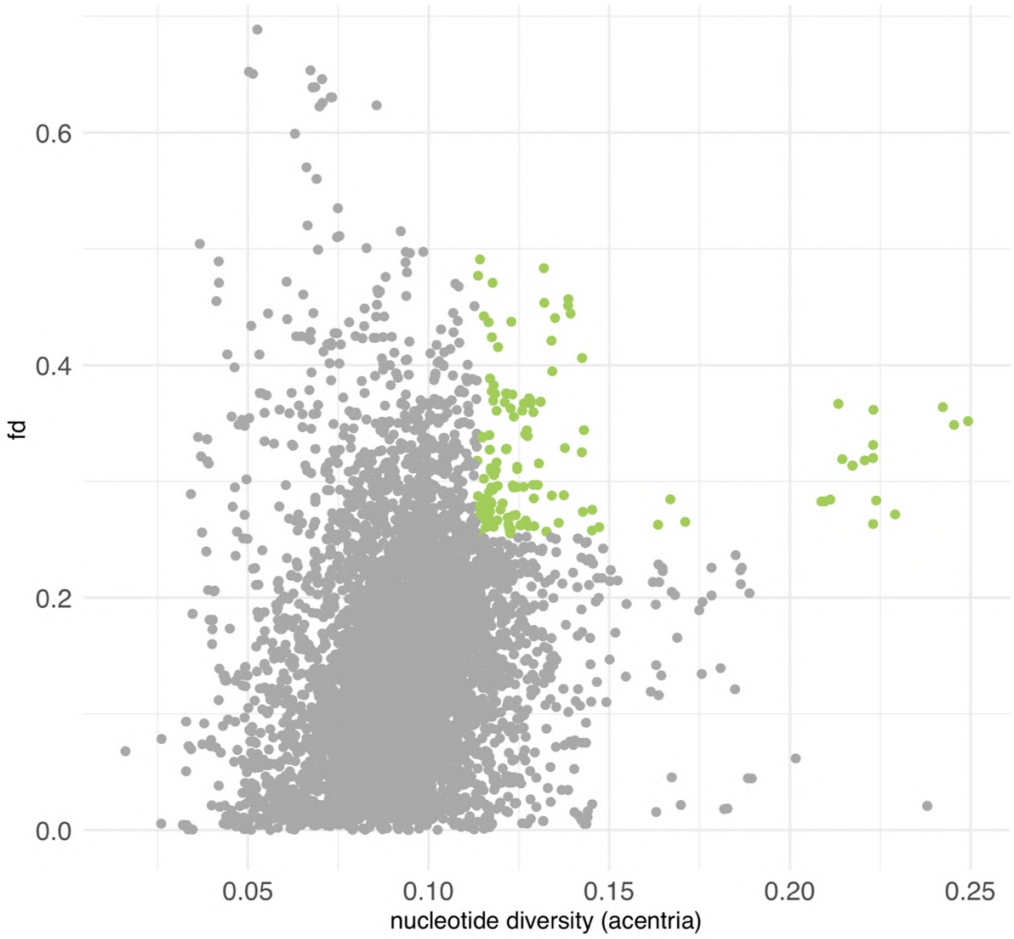

Рисунок 20. Сравнение нуклеотидного разнообразия (π) *Melitaea acentria* и пропорции интрогрессии (f_d) от *M. didyma* к *M. acentria*. Регионы, попадающие в 10% самых высоких значений параметров π и f_d, обозначены зеленым. Из [30].

При анализе возможных демографических сценариев комплекса *M. didyma-M. persea-M. acentria* с помощью анализа аллель-частотных спектров наибольшим правдоподобием обладал сценарий, предполагающий однонаправленный поток генов от *M. didyma* к *M. acentria,* происходивший после расхождения видов *M. persea* и *M. acentria* (Рисунок 21). Согласно параметрам этого сценария, дивергенция *M. didyma* произошла 10,7 млн лет назад, а *M. persea* и *M. acentria* разделились 7,27 млн лет назад, если предположить, что изучаемые виды моновольтинные. Если предположить, что эти виды бивольтинные, то время дивергенции *M. didyma* оценивается как 5,3 млн лет назад, и время дивергенции *M. persea* и *M. acentria* – как 3,6 млн лет назад. Согласно литературным данным, исследуемые виды могут быть как моновольтинными, так и бивольтинными в зависимости от местообитания [146]. Поэтому полученные оценки возрастов можно рассматривать в качестве минимально и максимально возможных.

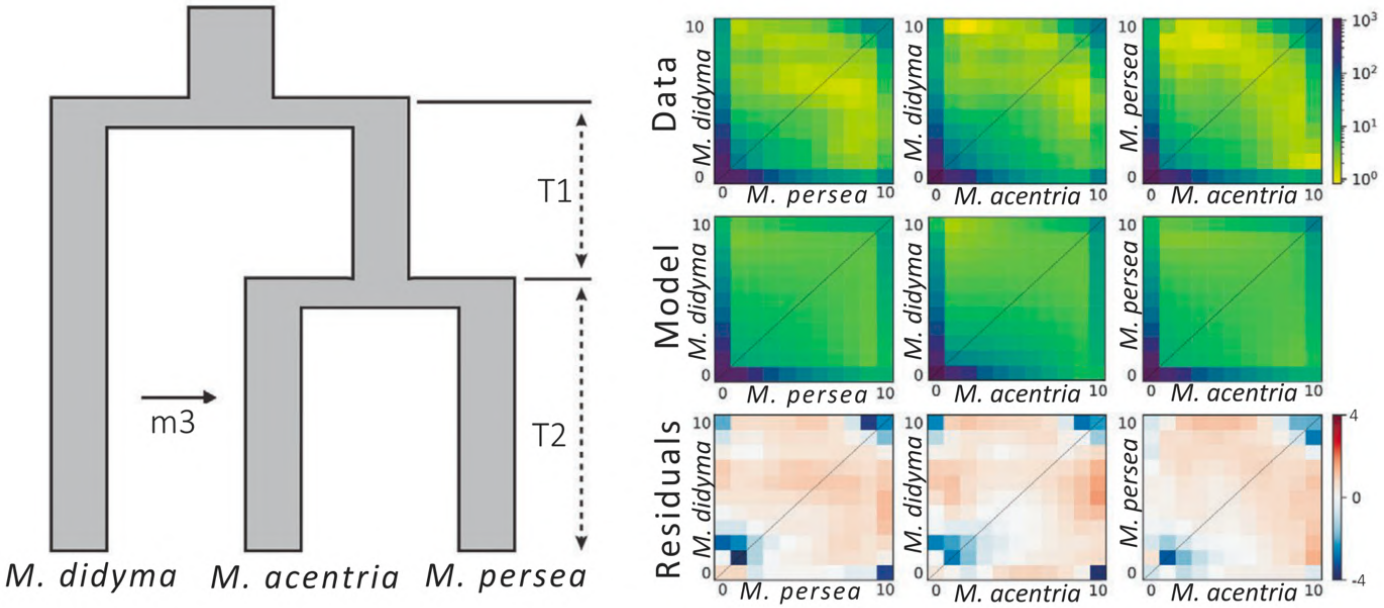

Рисунок 21. Реконструкция демографической истории видов методом анализа аллель-частотных спектров. Слева показана схема наиболее правдоподобной модели. Параметр m3 обозначает однонаправленный поток генов от *M. didyma* к *M. acentria*, параметр T1 показывает время дивергенции *M. didyma* и датирован 5,3-10,7 млн лет назад, T2 обозначает время дивергенции *M. acentria* и *M. persea* и датирован 3.6-7.27 млн лет назад. Справа показаны попарные наблюдаемые и ожидаемые спектры частот аллелей и различия между ними, выраженные через распределение остатков. Из [30].

4.2 Видовая самостоятельность *Melitaea acentria*, *M. persea* и *M. didyma*

Вопрос о видовой самостоятельности всегда зависит от используемого определения вида. В данной работе используется унифицированная концепция вида, которая рассматривает виды как независимые эволюционные линии, используя в качестве видовых характеристик ключевые критерии других концепций (биологической, диагностической, экологической, генеалогической и др.), такие как наличие репродуктивной изоляции между линиями, наличие четких морфологических различий, монофилия, обитание в разных экологических нишах и другие [14].

Рассмотрим комплекс видов с точки зрения **биологической концепции вида**, главным критерием которой является наличие репродуктивной изоляции [138]. *M. persea* и *M. acentria* обитают в аллопатрии, поэтому наличие между ними репродуктивной изоляции нельзя проверить в естественных условиях. В литературе сообщается о спорадической гибридизации между *M. didyma* и *M. acentria* в Турции [20], однако эти данные не были проверены с помощью молекулярного анализа, а наш анализ полногеномных данных не выявил следов интрогрессии между этими видами. Несмотря на поток генов между *M. didyma* и *M. acentria*, анализ полногеномных SNP показал, что все три вида сохраняют генетическую обособленность, что

говорит о наличии изолирующих механизмов, препятствующих их слиянию в единую эволюционную единицу. Одним из механизмов, обеспечивающих репродуктивную изоляцию в комплексе *M. persea -M. didyma - M. acentria,* может являться быстрая дивергенция половых хромосом относительно остального генома.

Поскольку виды *M. persea, M. didyma* и *M. acentria* обладают четкими морфологическими различиями [7,20], они соответствуют критериям диагностической концепции вида [14]. Реконструкция филогении с использованием полных геномов показала, что особи каждого из трех видов формируют монофилетические линии [30], что является критерием генеалогической концепции вида [164]. Также виды комплекса обладают различными экологическими предпочтениями. *M. persea* обитает в открытых сухих биотопах и в случаях симпатрии обычно предпочитает более низкие высоты, чем *M. didyma* [146]. В отличие от *M. persea*, *M. acentria* обитает в субальпийских биотопах, и обычно обнаруживается выше, чем *M. didyma* [7]. Соответственно, эти виды удовлетворяют условиям экологической концепции вида [165,166].

Таким образом, виды *M. persea, M. didyma* и *M. acentria* соответствуют критериям как минимум трех различных концепций, что говорит об их видовой самостоятельности.

4.3 Происхождение вида *M. acentria*

Melitaea acentria и *M. persea* являются полностью аллопатрическими, и реконструкция ареалов, основанная на филогеографическом анализе, показала, что для вида *M. acentria* Левант является наиболее вероятным анцестральным ареалом, в то время как *M. persea* распространен в Турции, Иране и на Кавказе [30]. Предположительно, вид *M. acentria* образовался на территории Леванта в результате географической изоляции от *M. persea* около 3,6-7,27 млн лет назад. Этот период соответствует позднему Миоцену (Мессиний, 7,2 – 5,3 млн лет назад), который был ключевым периодом в колонизации Северной Африки видами рода *Melitaea* благодаря геологическим и климатическим условиям, обеспечивавшим возможность наземных миграций [136].

Анализ полных митохондриальных геномов и ДНК-баркодов выявил две дифференцированные митохондриальные гаплогруппы в составе *M. acentria* (*acentria1* и *acentria2),* которые происходят от *M. persea* и не родственны *M. didyma* [30]. Митохондриальные гаплогруппы не коррелируют ни с кластерами, выделенными на основании анализа полных геномов, ни с морфологическими признаками, и, соответственно, представляют из себя случай внутривидового полиморфизма. В данном случае, наличие двух гаплогрупп может свидетельствовать скорее о двукратном формировании изолятов *M. persea* в Леванте с их

последующим слиянием во время колебаний климата в Четвертичном периоде, чем об интрогрессии между *M. persea* и *M. acentria,* поскольку анализ полных геномов не подтвердил поток генов между этими видами [30].

Melitaea didyma и *M. acentria* не являются близкородственными – согласно результатам анализа аллель-частотных спектров, разделение линий *persea* + *acentria* и *didyma* произошло 5,3-10,7 млн лет назад. Несмотря на то, что гибридизация чаще происходит между близкими таксонами, нами была показана интрогрессия от *M. didyma* к *M. acentria*. Предполагается, что интрогрессия чаще происходит от локального вида к инвазивному [167,168], поскольку второй обладает значительно меньшим размером популяции. Таким образом, при специфических географических и демографических условиях, приводящих к ослабеванию межвидовых репродуктивных барьеров, интрогрессия может происходить даже между неблизкородственными видами. Межвидовая гибридизация может быть дополнительным источником генетического разнообразия, особенно в маленьких популяциях. Наше исследование показало, что у *M. acentria* самое низкое среднее нуклеотидное разнообразие среди изученных видов [30]. Среди регионов генома с наиболее высокой пропорцией интрогрессии от *M. didyma* к *M. acentria* также отмечался самый высокий уровень нуклеотидного разнообразия. Это позволяет предполагать, что поток генов от транспалеарктического вида *M. didyma* являлся источником генетического разнообразия у эндемика Леванта *M. acentria*. Поскольку спорадическая гибридизация между *M. acentria* и *M. didyma* все еще происходит, анализ отдельных молекулярных маркеров отражал бы скорее филогению гена, а не филогению вида, что подчеркивает важность использования полногеномных данных для анализа групп, в которых была отмечена ретикулярная эволюция.

Заключение

Корректная делимитация видов является ключевой проблемой таксономии. Разработка методологии определения границ между видами требует как адаптации концепции вида к современным реалиям, так и изучения применимости различных типов данных для таксономического анализа. Широкая доступность молекулярных данных предоставляет, с одной стороны, возможность обнаружения скрытого биологического разнообразия, но с другой стороны, повышает риск ошибочных таксономических заключений, основанных только на одном типе данных. В настоящем исследовании мы продемонстрировали применимость ДНК-баркодов для постановки таксономических гипотез. Было показано, что вид *Melitaea ala* включает в себя три подвида (*M. ala ala*, *M. ala zaisana, M. ala bicolor*), демонстрирующие морфологические и географические различия, но уровень дивергенции по гену *COI* ниже видового. В составе комплекса *Melitaea didyma* нами были обнаружены 23 глубоко дивергировавшие митохондриальные генетические линии. Двенадцать из них соответствуют видам, 7 – подвидам *M. didyma,* и остальные представляют из себя случай митохондриального полиморфизма.

Несовпадение топологий филогенетических деревьев, построенных с использованием митохондриальных и ядерных маркеров, было продемонстрировано на примере рода *Brenthis*, для которого было показано наличие двух митохондриальных гаплогрупп внутри вида *B. daphne*. Анализ ядерных генов и морфологии не поддержал гипотезу о видовой самостоятельности этих гаплогрупп. Важно подчеркнуть, что объединение противоречащих данных (в конкретном случае, ядерных и митохондриальных генов) в единую матрицу и их совместный анализ приводит к ошибочной филогенетической реконструкции, и мы рекомендуем избегать подобного подхода на практике.

При анализе комплекса *Hyponephele lycaon - H. lupina* также были выявлены митохондриальные линии в составе *H. lycaon*, не соответствующие реальным таксонам. Интегративный анализ, включающий в себя анализ ДНК-баркодов, ядерных генов, ареалов, окраски крыльев и морфометрию гениталий, подтвердил самостоятельность видов *H. lycaon, H. maroccana*, *H. sifanica, H. mauritanica, H. lupina, H. galtscha* и *H. collina,* а также подвида *H. lycaon zuvandica*. Таксоны *H. pasimelas*, *H. dzhungarica* и *H. przhewalskiy* оказались не отличимы от вида *H. lycaon,* что не позволяет поддержать гипотезу о множественных аллопатричных видах внутри этого комплекса.

Секвенирование и последующий анализ полных геномов бабочек комплекса *Melitaea didyma – M. persea – M. acentria* подтвердили видовую самостоятельность этих таксонов, несмотря на неполную репродуктивную изоляцию между *M. didyma* и *M. acentria*. Было

показано, что вид *M. acentria* возник в Леванте как периферический географический изолят *M. persea.* Поток генов от неблизкородственного вида *M. didyma* являлся источником генетического разнообразия в популяции *M. acentria*, что определило эволюционную траекторию этого вида. Половые хромосомы более устойчивы к интрогрессии, чем аутосомы, и играют роль в формировании и поддержании постзиготических межвидовых репродуктивных барьеров.

Таким образом, проведенное диссертационное исследование демонстрирует применимость анализа молекулярных данных всех уровней – от единичных маркеров до полных геномов – для анализа групп нимфалоидных чешуекрылых разной степени изученности. В будущем разработка методологии интегративного анализа в таксономии может продолжена как при изучении других модельных групп живых организмов, так и при исследовании новых типов данных.

Благодарности

В заключение я бы хотела поблагодарить моего научного руководителя В. А. Лухтанова за помощь и поддержку на всех этапах работы над диссертацией. Я глубоко признательна коллективу отделения Кариосистематики ЗИН РАН и кафедры энтомологии СПбГУ за обсуждение работы и помощь в ее реализации. Я благодарю И. А. Махова, А. В. Барабанова, А. В. Новикову, В. В. Тихонова, Н. А. Шаповала, Д. Ф. Шовкуна за собранный материал. Выражаю признательность Анастасии Гагариной за помощь при подготовке образцов к секвенированию. Выражаю благодарность сотрудникам Ресурсного центра СПбГУ «Развитие Молекулярных и Клеточных Технологий» А. Э. Машарскому и Е. Ю. Городиловой и сотрудникам Ресурсного центра «Вычислительный центр» А. Андрееву и Д. Хмель за техническую поддержку проекта.

Исследования по теме диссертации были поддержаны грантами РНФ № 14-14-00541 и 19-14-00202, а также РФФИ № 19-34-90007.

Список литературы

1. Dayrat, B. Towards integrative taxonomy // Biological Journal of the Linnean Society. - 2005. - Vol. 85, № 3. - P. 407–415.
2. Will, K.W., Mishler, B.D., Wheeler, Q.D. The perils of DNA barcoding and the need for integrative taxonomy // Systematic Biology2005. - Vol. 54, № 5. - P. 844–851.
3. Hillis, D.M. Molecular Versus Morphological Approaches to Systematics // Annual Review of Ecology and Systematics. - 1987. - Vol. 18, № 1. - P. 23–42.
4. Hogan, K.M., Hedin, M.C., Koh Hung Sun, et al. Systematic and taxonomic implications of karyotypic, electrophoretic, and mitochondrial-DNA variation in Peromyscus from the Pacific Northwest // Journal of Mammalogy. - 1993. - Vol. 74, № 4. - P. 819–831.
5. DeSalle, R., Egan, M.G., Siddall, M. The unholy trinity: Taxonomy, species delimitation and DNA barcoding // Philosophical Transactions of the Royal Society B: Biological Sciences. - 2005. - Vol. 360, № 1462. - P. 1905–1916.
6. Heethoff, M., Laumann, M., Weigmann, G., et al. Integrative taxonomy: Combining morphological, molecular and chemical data for species delineation in the parthenogenetic Trhypochthonius tectorum complex (Acari, Oribatida, Trhypochthoniidae) // Frontiers in Zoology. - 2011. - Vol. 8, № 1. - P. 1–10.
7. Lukhtanov, V.A. A new species of Melitaea from Israel, with notes on taxonomy, cytogenetics, phylogeography and interspecific hybridization in the Melitaea persea complex (Lepidoptera, Nymphalidae) // Comparative Cytogenetics. - 2017. - Vol. 11, № 2. - P. 325–357.
8. Dincă, V., Wiklund, C., Lukhtanov, V.A., et al. Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex // Journal of Evolutionary Biology. - 2013. - Vol. 26, № 10. - P. 2095–2106.
9. Avise, J.C. Phylogeography: the history and formation of species. - Cambridge, Massachusetts: Harvard University Press, 2000. - 447 p.
10. Schander, C., Willassen, E. What can biological barcoding do for marine biology? // Marine Biology Research. - 2005. - Vol. 1, № 1. - P. 79–83.
11. Condon, M.A., Scheffer, S.J., Lewis, M.L., et al. Hidden neotropical diversity: Greater than the sum of its parts // Science. - 2008. - Vol. 320, № 5878. - P. 928–931.
12. Hebert, P.D.N., Penton, E.H., Burns, J.M., et al. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator // Proceedings of the National Academy of Sciences of the United States of America. - 2004. - Vol. 101, № 41. - P. 14812–14817.
13. Cardoso, A., Serrano, A., Vogler, A.P. Morphological and molecular variation in tiger beetles of the Cicindela hybrida complex: Is an “integrative taxonomy” possible? // Molecular Ecology. -

2009. - Vol. 18, № 4. - P. 648–664.

14. De Queiroz, K. Species Concepts and Species Delimitation // Systematic Biology. - 2007. - Vol. 56, № 6. - P. 879–886.
15. Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., et al. Integrative taxonomy: A multisource approach to exploring biodiversity // Annual Review of Entomology2010. - Vol. 55P. 421–438.
16. Hinojosa, J.C., Koubínová, D., Szenteczki, M.A., et al. A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly Thymelicus sylvestris // Molecular Ecology. - 2019. - Vol. 28, № 17. - P. 3857–3868.
17. Dincă, V., Lee, K.M., Vila, R., et al. The conundrum of species delimitation: A genomic perspective on a mitogenetically super-variable butterfly // Proceedings of the Royal Society B: Biological Sciences. - 2019. - Vol. 286, № 1911.
18. Pante, E., Schoelinck, C., Puillandre, N. From Integrative Taxonomy to Species Description: One Step Beyond // Systematic Biology. - 2015. - Vol. 64, № 1. - P. 152–160.
19. Dincă, V., Montagud, S., Talavera, G., et al. DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity // Scientific Reports 2015 5:1. - 2015. - Vol. 5, № 1. - P. 1–12.
20. van Oorschot, H., Coutsis, J. The genus Melitaea Fabricius, 1807 : taxonomy and systematics with special reference to the male genitalia : (Lepidoptera, Nymphalidae, Nymphalinae). - Pardubice: Tshikolovets Publications, 2014. - 360 p.
21. Simonsen, T.J., Wahlberg, N., Brower, A.V.Z., et al. Morphology, molecules and fritillaries: Approaching a stable phylogeny for Argynnini (Lepidoptera: Nymphalidae) // Insect Systematics and Evolution. - 2006. - Vol. 37, № 4. - P. 405–418.
22. Simonsen, T.J. Fritillary phylogeny, classification, and larval host plants: reconstructed mainly on the basis of male and female genitalic morphology (Lepidoptera: Nymphalidae: Argynnini) // Biological Journal of the Linnean Society. - 2006. - Vol. 89, № 4. - P. 627–673.
23. Higgins, L.G. The Classification of European Butterflies. - London: Collins, 1975. - 320 p.
24. Kolesnichenko, K.A. A review of the East Palaearctic taxa of the Melitaea didyma (Esper, [1779])-group. 1. The M. ala Staudinger, 1881 -M. chitralensis Moore, 1901 -complex. // Atalanta. - 1999. - Vol. 30, № 1. - P. 78–117.
25. Pazhenkova, E.A., Zakharov, E. V., Lukhtanov, V.A. DNA barcoding reveals twelve lineages with properties of phylogenetic and biological species within Melitaea didyma sensu lato (Lepidoptera, Nymphalidae) // ZooKeys. - 2015. - Vol. 538P. 35–46.
26. Pazhenkova, E.A., Lukhtanov, V.A. Chromosomal and mitochondrial diversity in Melitaea Chromosomal and mitochondrial diversity in Melitaea didyma complex (Lepidoptera, Nymphalidae): eleven deeply diverged DNA barcode groups in one non-monophyletic species?

// CompCytogen. - 2016. - Vol. 10, № 4. - P. 697–717.

27. Pazhenkova, E.A., Lukhtanov, V.A. Nuclear genes (but not mitochondrial DNA barcodes) reveal real species: Evidence from the Brenthis fritillary butterflies (Lepidoptera, Nymphalidae) // Journal of Zoological Systematics and Evolutionary Research. - 2019. - Vol. 57, № 2. - .
28. Lukhtanov, V.A., Gagarina, A. V., Pazhenkova, E.A. Chromosomal and DNA barcode analysis of the Melitaea ala Staudinger, 1881 species complex (Lepidoptera, Nymphalidae) // Comparative Cytogenetics. - 2021. - Vol. 15, № 2. - P. 199–216.
29. Lukhtanov, V.A., Pazhenkova, E.A. The Taxa of the Hyponephele lycaon – H. lupina Species Complex (Lepidoptera, Nymphalidae, Satyrinae): Deep DNA Barcode Divergence despite Morphological Similarity // Folia Biologica (Poland). - 2021. - Vol. 69, № 1. - P. 11–21.
30. Pazhenkova, E.A., Lukhtanov, V.A. Genomic introgression from a distant congener in the Levant fritillary butterfly, Melitaea acentria // Molecular Ecology. - 2021.
31. The GIMP Development Team. GIMP. - 2019. URL https://www.gimp.org.
32. Bonhomme, V., Picq, S., Gaucherel, C., et al. Momocs: Outline analysis using R // Journal of Statistical Software. - 2014. - Vol. 56, № 13. - P. 1–24.
33. Dommergues, C.H., Dommergues, J.L., Verrecchia, E.P. The discrete cosine transform, a Fourier-related method for morphometric analysis of open contours // Mathematical Geology. - 2007. - Vol. 39, № 8. - P. 749–763.
34. Swofford, D.L. PAUP*, Phylogenetic analysis using parisimony (*and other methods), version 4.2002. - 2002.
35. Ronquist, F., Teslenko, M., van der Mark, P., et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space // Systematic biology. - . 2012/02/222012. - Vol. 61, № 3. - P. 539–542.
36. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT // Nucleic Acids Symp. - 1999. - Vol. 41. - P. 95–98.
37. Thompson, J.D., Higgins, D.G., Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. // Nucleic Acids Research. - 1994. - Vol. 22, № 22. - P. 4673.
38. Ahola, V., Lehtonen, R., Somervuo, P., et al. The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera // Nature Communications. - 2014. - Vol. 5, № 1. - P. 1–9.
39. McKenna, A., Hanna, M., Banks, E., et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data // Genome Research. - 2010. - Vol. 20, № 9. - P. 1297–1303.

40. Zheng, X., Levine, D., Shen, J., et al. A high-performance computing toolset for relatedness and principal component analysis of SNP data // Bioinformatics. - 2012. - Vol. 28, № 24. - P. 3326–3328.
41. Huson, D.H., Bryant, D. Application of phylogenetic networks in evolutionary studies // Molecular Biology and Evolution. - 2006. - Vol. 23, № 2. - P. 254–267.
42. Martin, S.H., Jiggins, C.D. Interpreting the genomic landscape of introgression // Current Opinion in Genetics and Development. - 2017. - Vol. 47. - P. 69–74.
43. Malinsky, M., Matschiner, M., Svardal, H. Dsuite - Fast D-statistics and related admixture evidence from VCF files // Molecular Ecology Resources. - 2021. - Vol. 21, № 2. - P. 584–595.
44. Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H., et al. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data // PLoS Genetics. - 2009. - Vol. 5, № 10. - P. 1000695.
45. Yu, Y., Harris, A.J., Blair, C., et al. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography // Molecular Phylogenetics and Evolution. - 2015. - Vol. 87. - P. 46–49.
46. Matschiner, M. Fitchi: haplotype genealogy graphs based on the Fitch algorithm // Bioinformatics (Oxford, England). - 2016. - Vol. 32, № 8. - P. 1250—1252.
47. Simon, C., Frati, F., Beckenbach, A., et al. Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers // Annals of the Entomological Society of America. - 1994. - Vol. 87, № 6. - P. 651–701.
48. Sperling, F.A.H., Anderson, G.S., Hickey, D.A. A DNA-Based Approach to the Identification of Insect Species Used for Postmorten Interval Estimation // Journal of Forensic Sciences. - 1994. - Vol. 39, № 2. - P. 13613J.
49. Birstein, V.J., Desalle, R. Molecular Phylogeny of Acipenserinae // Molecular Phylogenetics and Evolution. - 1998. - Vol. 9, № 1. - P. 141–155.
50. Hebert, P.D.N., Ratnasingham, S., DeWaard, J.R. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species // Proceedings of the Royal Society B: Biological Sciences. - 2003. - Vol. 270. - P. 96–99.
51. Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., et al. Identification of birds through DNA barcodes // PLoS Biology. - 2004. - Vol. 2, № 10.
52. Hogg, I.D., Hebert, P.D.N. Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes // Canadian Journal of Zoology. - 2004. - Vol. 82, № 5. - P. 749–754.
53. Barrett, R.D.H., Hebert, P.D.N. Identifying spiders through DNA barcodes // Canadian Journal

of Zoology. - 2005. - Vol. 83, № 3. - P. 481–491.

54. Ball, S.L., Hebert, P.D.N., Burian, S.K., et al. Biological identifications of mayflies (Ephemeroptera) using DNA barcodes // https://doi.org/10.1899/04-142.1. - 2005. - Vol. 24, № 3. - P. 508–524.

55. Ward, R.D., Zemlak, T.S., Innes, B.H., et al. DNA barcoding Australia's fish species // Philosophical Transactions of the Royal Society B: Biological Sciences. - 2005. - Vol. 360, № 1462. - P. 1847–1857.

56. Wiemers, M., Fiedler, K. Does the DNA barcoding gap exist? - A case study in blue butterflies (Lepidoptera: Lycaenidae) // Frontiers in Zoology. - 2007. - Vol. 4, № September 2004. - P. 1–16.

57. Meier, R., Zhang, G., Ali, F. The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification // Systematic Biology. - 2008. - Vol. 57, № 5. - P. 809–813.

58. Zou, S., Li, Q., Kong, L., et al. Comparing the Usefulness of Distance, Monophyly and Character-Based DNA Barcoding Methods in Species Identification: A Case Study of Neogastropoda // PLoS ONE. - 2011. - Vol. 6, № 10.

59. Puillandre, N., Lambert, A., Brouillet, S., et al. ABGD, Automatic Barcode Gap Discovery for primary species delimitation // Molecular Ecology. - 2012. - Vol. 21, № 8. - P. 1864–1877.

60. Pons, J., Barraclough, T.G., Gomez-Zurita, J., et al. Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects // Systematic Biology. - 2006. - Vol. 55, № 4. - P. 595–609.

61. Zhang, J., Kapli, P., Pavlidis, P., et al. A general species delimitation method with applications to phylogenetic placements // Bioinformatics (Oxford, England). - 2013. - Vol. 29, № 22. - P. 2869–2876.

62. Somervuo, P., Koskela, S., Pennanen, J., et al. Unbiased probabilistic taxonomic classification for DNA barcoding // Bioinformatics. - 2016. - Vol. 32, № 19. - P. 2920–2927.

63. Dincă, V., Dapporto, L., Somervuo, P., et al. High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity // Communications Biology. - 2021. - Vol. 4, № 1. - P. 1–11.

64. Hajibabaei, M., Jeremy, R., Ivanova, N. V, et al. Critical factors for assembling a high volume of DNA barcodes. // Philosophical transactions of the Royal society B. - 2005. - Vol. 60, №1462. - P. 1959-1967.

65. Kinoshita, G., Yonezawa, S., Murakami, S., et al. Environmental DNA Collected from Snow Tracks is Useful for Identification of Mammalian Species // Zoological science. - 2019. - Vol. 36, № 3. - P. 198–207.

66. Pook, C.E., McEwing, R. Mitochondrial DNA sequences from dried snake venom: a DNA barcoding approach to the identification of venom samples // Toxicon : official journal of the International Society on Toxinology. - 2005. - Vol. 46, № 7. - P. 711–715.
67. Kawai, K., Shimizu, M., Hughes, R.N., et al. A non-invasive technique for obtaining DNA from marine intertidal snails // Journal of the Marine Biological Association of the United Kingdom. - 2004. - Vol. 84, № 4. - P. 773–774.
68. Clare, E.L., Economou, C.K., Faulkes, C.G., et al. eDNAir: proof of concept that animal DNA can be collected from air sampling // PeerJ. - 2021. - Vol. 9P. - P. e11030.
69. Sonet, G., Pauly, A., Nagy, Z.T., et al. Using next-generation sequencing to improve DNA barcoding: lessons from a small-scale study of wild bee species (Hymenoptera, Halictidae) // Apidologie 2018 49:5. - 2018. - Vol. 49, № 5. - P. 671–685.
70. Sire, L., Gey, D., Debruyne, R., et al. The Challenge of DNA Barcoding Saproxylic Beetles in Natural History Collections—Exploring the Potential of Parallel Multiplex Sequencing With Illumina MiSeq // Frontiers in Ecology and Evolution. - 2019. - P. 495.
71. Shokralla, S., Porter, T.M., Gibson, J.F., et al. Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform // Scientific Reports 2015 5:1. - 2015. - Vol. 5, № 1. - P. 1–7.
72. Pomerantz, A., Peñafiel, N., Arteaga, A., et al. Real-time DNA barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building // GigaScience. - 2018. - Vol. 7, № 4. - P. 1–14.
73. Chang, J.J.M., Ip, Y.C.A., Bauman, A.G., et al. MinION-in-ARMS: Nanopore Sequencing to Expedite Barcoding of Specimen-Rich Macrofaunal Samples From Autonomous Reef Monitoring Structures // Frontiers in Marine Science. - 2020. - P. 448.
74. Hebert, P.D.N., Braukmann, T.W.A., Prosser, S.W.J., et al. A Sequel to Sanger: amplicon sequencing that scales // BMC Genomics 2018 19:1. - 2018. - Vol. 19, № 1. - P. 1–14.
75. Ratnasingham, S., Hebert, P.D.N. BOLD: The Barcode of Life Data System. - 2007. - [http://www.boldsystems.org]
76. Hebert, P.D.N., Penton, E.H., Burns, J.M., et al. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator // Proceedings of the National Academy of Sciences of the United States of America. - 2004. - Vol. 101, № 41. - P. 14812–14817.
77. Smith, M.A., Woodley, N.E., Janzen, D.H., et al. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae) // Proceedings of the National Academy of Sciences. - 2006. - Vol. 103, № 10. - P. 3657–3662.
78. Gustafsson, D.R., Price, D.A., Erséus, C. Genetic variation in the popular lab worm

Lumbriculus variegatus (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation // Molecular phylogenetics and evolution. - 2009. - Vol. 51, № 2. - P. 182–189.

79. Elderkin, C.L., Clewing, C., Wembo Ndeo, O., et al. Molecular phylogeny and DNA barcoding confirm cryptic species in the African freshwater oyster Etheria elliptica Lamarck, 1807 (Bivalvia: Etheriidae) // Biological Journal of the Linnean Society. - 2016. - Vol. 118, № 2. - P. 369–381.
80. Tyagi, K., Kumar, V., Kundu, S., et al. Identification of Indian Spiders through DNA barcoding: Cryptic species and species complex // Scientific Reports 2019 9:1. - 2019. - Vol. 9, № 1. - P. 1–13.
81. Wang, T., Zhang, Y., Yang, Z., et al. DNA barcoding reveals cryptic diversity in the underestimated genus Triplophysa (Cypriniformes: Cobitidae, Nemacheilinae) from the northeastern Qinghai-Tibet Plateau // BMC Evolutionary Biology 2020 20:1. - 2020. - Vol. 20, № 1. - P. 1–15.
82. Dincă, V., Montagud, S., Talavera, G., et al. DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity // Scientific Reports. - 2015. - Vol. 5, № 1 - P. 1–12.
83. Avise, J.C. Molecular markers, natural history and evolution. -Oxford: Oxford University Press, 2004. - 684 p.
84. Cong, Q., Shen, J., Borek, D., et al. When COI barcodes deceive: complete genomes reveal introgression in hairstreaks // Proceedings of the Royal Society B: Biological Sciences. - 2017. - Vol. 284, № 1848. - P. 1 - 9.
85. Cariou, M., Duret, L., Charlat, S. The global impact of Wolbachia on mitochondrial diversity and evolution // Journal of evolutionary biology. - 2017. - Vol. 30, № 12. - P. 2204–2210.
86. Magnacca, K.N., Brown, M.J. Mitochondrial heteroplasmy and DNA barcoding in Hawaiian Hylaeus (Nesoprosopis) bees (Hymenoptera: Colletidae) // BMC Evolutionary Biology. - 2010. - Vol. 10, № 174.
87. Rubinoff, D., Cameron, S., Will, K. A Genomic Perspective on the Shortcomings of Mitochondrial DNA for “Barcoding” Identification // Journal of Heredity. - 2006. - Vol. 97, № 6. - P. 581–594.
88. Song, H., Buhay, J.E., Whiting, M.F., et al. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified // Proceedings of the National Academy of Sciences. - 2008. - Vol. 105, № 36. - P. 13486–13491.
89. Makhov, I.A., Gorodilova, Y.Y.U., Lukhtanov, V.A. Sympatric occurrence of deeply diverged mitochondrial DNA lineages in Siberian geometrid moths (Lepidoptera: Geometridae): cryptic speciation, mitochondrial introgression, secondary admixture or effect of Wolbachia? //

Biological Journal of the Linnean Society. - 2021. - Vol. XXP. 1–24.

90. Ré, F.C. De, Robe, L.J., Wallau, G.L., et al. Inferring the phylogenetic position of the Drosophila flavopilosa group: Incongruence within and between mitochondrial and nuclear multilocus datasets // Journal of Zoological Systematics and Evolutionary Research. - 2017. - Vol. 55, № 3. - P. 208–221.
91. Platt, R.N., Faircloth, B.C., Sullivan, K.A.M., et al. Conflicting Evolutionary Histories of the Mitochondrial and Nuclear Genomes in New World Myotis Bats // Systematic Biology. - 2018. - Vol. 67, № 2. - P. 236–249.
92. Sota, T., Vogler, A.P. Reconstructing species phylogeny of the carabid beetles Ohomopterus using multiple nuclear DNA sequences: heterogeneous information content and the performance of simultaneous analyses // Molecular Phylogenetics and Evolution. - 2003. - Vol. 26, № 1. - P. 139–154.
93. Sun, J.-T., Wang, M.-M., Zhang, Y.-K., et al. Evidence for high dispersal ability and mito-nuclear discordance in the small brown planthopper, Laodelphax striatellus // Scientific Reports. - 2015. - Vol. 5, № 1. - P. 1-10.
94. Toews, D.P.L., Brelsford, A. The biogeography of mitochondrial and nuclear discordance in animals // Molecular Ecology. - 2012. - Vol. 21, № 16. - P. 3907–3930.
95. Tóth, J.P., Bereczki, J., Varga, Z., et al. Relationships within the Melitaea phoebe species group (Lepidoptera: Nymphalidae): New insights from molecular and morphometric information // Systematic Entomology. - 2014. - Vol. 39, № 4. - P. 749–757.
96. Padial, J.M., Miralles, A., De la Riva, I., et al. The integrative future of taxonomy // Frontiers in Zoology. - 2010. - Vol. 7. - P. 1- 14.
97. Bond, J.E., Stockman, A.K. An Integrative Method for Delimiting Cohesion Species: Finding the Population-Species Interface in a Group of Californian Trapdoor Spiders with Extreme Genetic Divergence and Geographic Structuring // Systematic Biology. - 2008. - Vol. 57, № 4. - P. 628–646.
98. Lukhtanov, V.A., Dantchenko, A. V., Vishnevskaya, M.S., et al. Detecting cryptic species in sympatry and allopatry: Analysis of hidden diversity in Polyommatus (Agrodiaetus) butterflies (Lepidoptera: Lycaenidae) // Biological Journal of the Linnean Society. - 2015. - Vol. 116, № 2. - P. 468–485.
99. Garraffoni, A.R.S., Araújo, T.Q., Lourenço, A.P., et al. Integrative taxonomy of a new Redudasys species (Gastrotricha: Macrodasyida) sheds light on the invasion of fresh water habitats by macrodasyids // Scientific Reports 2019 9:1. - 2019. - Vol. 9, № 1. - P. 1–15.
100. Garcia-Erill, G., Kjær, M.M., Albrechtsen, A., et al. Vicariance followed by secondary gene flow in a young gazelle species complex // Molecular Ecology. - 2021. - Vol. 30, № 2. - P. 528–

544.

101. Adams, D.C., Berns, C.M., Kozak, K.H., et al. Are rates of species diversification correlated with rates of morphological evolution? // Proceedings of the Royal Society B: Biological Sciences. - 2009. - Vol. 276, № 1668. - P. 2729–2738.
102. Takahashi K., Y., T., M., N., et al. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons // Molecular biology and evolution. - 2001. - Vol. 18, № 11. - P. 2057–2066.
103. Mayden, R.L. A hierarchy of species concepts: the denouement in the saga of the species problems. - Species. The units of biodiversity. - London: Chapman & Hall.- 1997. - 381-423 p.
104. Delsuc, F., Brinkmann, H., Philippe, H. Phylogenomics and the reconstruction of the tree of life // Nature Reviews Genetics 2005 6:5. - 2005. - Vol. 6, № 5. - P. 361–375.
105. Misof, B., Liu, S., Meusemann, K., et al. Phylogenomics resolves the timing and pattern of insect evolution // Science. - 2014. - Vol. 346, № 6210. - P. 763–767.
106. Chipman, A.D., Ferrier, D.E.K., Brena, C., et al. The First Myriapod Genome Sequence Reveals Conservative Arthropod Gene Content and Genome Organisation in the Centipede Strigamia maritima // PLOS Biology. - 2014. - Vol. 12, № 11. - P. e1002005.
107. Nater, A., Mattle-Greminger, M.P., Nurcahyo, A., et al. Morphometric, Behavioral, and Genomic Evidence for a New Orangutan Species // Current Biology. - 2017. - Vol. 27, № 22. - P. 3487-3498.e10.
108. Harr, B. Genomic islands of differentiation between house mouse subspecies // Genome Research. - 2006. - Vol. 16, № 6. - P. 730–737.
109. Marques, D.A., Lucek, K., Meier, J.I., et al. Genomics of Rapid Incipient Speciation in Sympatric Threespine Stickleback // PLOS Genetics. - 2016. - Vol. 12, № 2. - P. e1005887.
110. Lamichhaney, S., Berglund, J., Almén, M.S., et al. Evolution of Darwin’s finches and their beaks revealed by genome sequencing // Nature 2015 518:7539. - 2015. - Vol. 518, № 7539. - P. 371–375.
111. Masly, J.P., Presgraves, D.C. High-resolution genome-wide dissection of the two rules of speciation in Drosophila // PLoS Biology. - 2007. - Vol. 5, № 9. - P. 1890–1898.
112. Charlesworth, B., Coyne, J.A., Barton, N.H. The relative rates of evolution of sex chromosomes and autosomes. // American Naturalist. - 1987. - Vol. 130, № 1. - P. 113–146.
113. Connallon, T., Singh, N.D., Clark, A.G. Impact of genetic architecture on the relative rates of X versus autosomal adaptive substitution // Molecular Biology and Evolution. - 2012. - Vol. 29, № 8. - P. 1933–1942.
114. Meisel, R.P., Malone, J.H., Clark, A.G. Faster-X Evolution of Gene Expression in Drosophila // PLoS Genetics. - 2012. - Vol. 8, № 10. - P. e1003013–e1003013.

115. Dobzhansky, T. Studies on hybrid sterility - I. Spermatogenesis in pure and hybrid Drosophila pseudoobscura // Zeitschrift für Zellforschung und Mikroskopische Anatomie. - 1934. - Vol. 21, № 2. - P. 169–223.
116. Cong, Q., Zhang, J., Grishin, N. V. Genomic determinants of speciation // bioRxiv. - 2019. - .
117. Van Belleghem, S.M., Baquero, M., Papa, R., et al. Patterns of Z chromosome divergence among Heliconius species highlight the importance of historical demography // Molecular Ecology. - 2018. - Vol. 27, № 19. - P. 3852–3872.
118. Cortés-Ortiz, L., Nidiffer, M.D., Hermida-Lagunes, J., et al. Reduced Introgression of Sex Chromosome Markers in the Mexican Howler Monkey (Alouatta palliata × A. pigra) Hybrid Zone // International Journal of Primatology. - 2019. - Vol. 40, № 1. - P. 114–131.
119. Storchová, R., Reif, J., Nachman, M.W. Female heterogamety and speciation: Reduced introgression of the z chromosome between two species of nightingales // Evolution. - 2010. - Vol. 64, № 2. - P. 456–471.
120. Sankararaman, S., Mallick, S., Dannemann, M., et al. The genomic landscape of Neanderthal ancestry in present-day humans // Nature. - 2014. - Vol. 507, № 7492. - P. 354–357.
121. Schluter, D. Evidence for ecological speciation and its alternative // Science2009. - Vol. 323, № 5915. - P. 737–741.
122. Nei, M., Maruyama, T., Wu, C. Models of evolution of reproductive isolation // Genetics. - 1983. - Vol. 103, № 3. - P. 557–579.
123. Gavrilets, S. Perspective: Models of speciation - What have we learned in 40 years? // Evolution2003. - Vol. 57, № 10. - P. 2197–2215.
124. Garcia-Erill, G., Kjær, M.M., Albrechtsen, A., et al. Vicariance followed by secondary gene flow in a young gazelle species complex // Molecular Ecology. - 2021. - Vol. 30, № 2. - P. 528–544.
125. Elgvin, T.O., Trier, C.N., Tørresen, O.K., et al. The genomic mosaicism of hybrid speciation // Science Advances. - 2017. - Vol. 3, № 6.
126. Teng, H., Zhang, Y., Shi, C., et al. Population Genomics Reveals Speciation and Introgression between Brown Norway Rats and Their Sibling Species // Molecular Biology and Evolution. - 2017. - Vol. 34, № 9. - P. 2214–2228.
127. Edelman, N.B., Frandsen, P.B., Miyagi, M., et al. Genomic architecture and introgression shape a butterfly radiation // Science. - 2019. - Vol. 366, № 6465. - P. 594–599.
128. Taylor, S.A., Larson, E.L. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature // Nature Ecology and Evolution. - 2019. - Vol. 3, № 2. - P. 170–177.
129. Fontsere, C., de Manuel, M., Marques-Bonet, T., et al. Admixture in mammals and how to

understand its functional implications: On the abundance of gene flow in Mammalian species, its impact on the genome, and roads into a functional understanding // BioEssays. - 2019. - Vol. 41, № 12.

130. Pardo-Diaz, C., Salazar, C., Baxter, S.W., et al. Adaptive Introgression across Species Boundaries in Heliconius Butterflies // PLoS Genetics. - / ed. R. Kronforst M.2012. - Vol. 8, № 6. - P. e1002752.
131. Zhang, W., Dasmahapatra, K.K., Mallet, J., et al. Genome-wide introgression among distantly related Heliconius butterfly species // Genome Biology. - 2016. - Vol. 17, № 1. - P. 1–15.
132. Neri, J., Wendt, T., Palma-Silva, C. Natural hybridization and genetic and morphological variation between two epiphytic bromeliads // AoB PLANTS. - 2018. - Vol. 10, № 1.
133. Lukhtanov, V.A., Shapoval, N.A., Anokhin, B.A., et al. Homoploid hybrid speciation and genome evolution via chromosome sorting // Proceedings of the Royal Society B: Biological Sciences. - 2015. - Vol. 282, № 1807.
134. Wagner, C.E., Keller, I., Wittwer, S., et al. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation // Molecular ecology. - 2013. - Vol. 22, № 3. - P. 787–798.
135. Higgins, L.G. An illustrated catalogue of Palearctic Melitaea (Lep. Rhopalocera) // Transactions of the Royal Entomological Society of London. - 1941. - Vol. 91, № 7. - P. 175–365.
136. Leneveu, J., Chichvarkhin, A., Wahlberg, N. Varying rates of diversification in the genus Melitaea (Lepidoptera: Nymphalidae) during the past 20 million years // Biological Journal of the Linnean Society. - 2009. - Vol. 97, № 2. - P. 346–361.
137. Cracraft, J. Species concepts and speciation analysis. // Current ornithology. Vol. 1. - 1983. - P. 159–187.
138. Coyne, J. A., Orr, H. A. Speciation. – Sanderland, MA: Sinauer Associates, 2004. – 545 p.
139. Tshikolovets, V. V. The Butterflies of Tajikistan. - Brno-Kyiv: Tshikolovets Publications. - 500 p.
140. Brower, A.V.Z., DeSalle, R. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference // Insect Molecular Biology. - 1998. - Vol. 7, № 1. - P. 73–82.
141. Quek, S.-P., Davies, S.J., Itino, T., et al. Codiversification in an ant-plant mutualism: Stem texture and the evolution of host use in Crematogaster (Formicidae: Myrmicinae) inhabitants of Macaranga (Euphorbiaceae) // Evolution. - 2004. - Vol. 58, № 3. - P. 554–570.
142. Gómez, A., Lunt, D.H. Refugia within Refugia: Patterns of Phylogeographic Concordance in the Iberian Peninsula // Phylogeography of Southern European Refugia: Evolutionary

Perspectives on the Origins and Conservation of European Biodiversity. - 2007. - P. 155–188.

143. Kolesnichenko, K.A., Churkin, S. V., Berdiev, R.K. A review of the East Palaearctic taxa of the Melitaea didyma (Esper, [1779])-group. Part II. // Neue Entomologische Nachrichten. - 2011. - Vol. 67. - P. 257–278.
144. Лухтанов, В.А. Делимитация видов и анализ криптического видового разнообразия в XXI веке // Энтомологическое обозрение. - 2019. - Vol. 98, № 2. - P. 358–370.
145. Palumbi, S.R., Cipriano, F., Hare, M.P. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule // Evolution. - 2001. - Vol. 55, № 5. - P. 859–868.
146. Hesselbarth, G., Oorchot, H., Wagener, S. Die Tagfalter der Türkei: unter Berücksichtigung der angrenzenden Länder. - Bocholt: Selbstverlag Siegbert Wagener, 1995. - 2201 p.
147. Zimmermann, K., Fric, Z., Filipová, L., et al. Adult demography, dispersal and behaviour of Brenthis ino (Lepidoptera: Nymphalidae): How to be a successful wetland butterfly // European Journal of Entomology. - 2005. - Vol. 102, № 4. - P. 699–706.
148. Browning, S.R., Browning, B.L. Haplotype phasing: existing methods and new developments // Nature Reviews Genetics 2011 12:10. - 2011. - Vol. 12, № 10. - P. 703–714.
149. Friberg, M., Leimar, O., Wiklund, C. Heterospecific courtship, minority effects and niche separation between cryptic butterfly species // Journal of Evolutionary Biology. - 2013. - Vol. 26, № 5. - P. 971–979.
150. Ramírez, S.R., Nieh, J.C., Quental, T.B., et al. A molecular phylogeny of the stingless bee genus Melipona (Hymenoptera: Apidae) // Molecular Phylogenetics and Evolution. - 2010. - Vol. 56, № 2. - P. 519–525.
151. Silva-Brandão, K.L., Wahlberg, N., Francini, R.B., et al. Phylogenetic relationships of butterflies of the tribe Acraeini (Lepidoptera, Nymphalidae, Heliconiinae) and the evolution of host plant use // Molecular Phylogenetics and Evolution. - 2008. - Vol. 46, № 2. - P. 515–531.
152. van Velzen, R., Wahlberg, N., Sosef, M.S.M., et al. Effects of changing climate on species diversification in tropical forest butterflies of the genus Cymothoe (Lepidoptera: Nymphalidae) // Biological Journal of the Linnean Society. - 2013. - Vol. 108, № 3. - P. 546–564.
153. Eckweiler, W., Bozano, G.C. Guide to the Butterflies of the Palearctic Region. Satyrinae part IV. Tribe Satyrini, Subtribe Maniolina. - Milano: Omnes Artes, 2011. - 102 p.
154. Staudinger, O., Rebel, H. Catalog der Lepidopteren des Palaearctischen Faunengebietes. Famil. Papilionidae-Hepialidae. - Berlin: Friedländer & Sohn, 1901. - 411 p.
155. Turati, E. Revisione delle forme di Epinephele lycaon Rott. // Naturalista Siciliano. - 1909. - Vol. 21, № 1–8. - P. 56–73.
156. Weiss, D. A new species of the genus Hyponephele Muschamp 1915 from West Iran // Atalanta. - 1978. - Vol. 9P. 230–232.

157. Tshikolovets, V. V., Naderi, A., Eckweiler, W. The butterflies of Iran and Iraq. -Pardubice: Tshikolovets Publications, 2014. - 440 p.

158. Samodurov, G.D., Korolev, V.A., Tshikolovets, V.V. Eine Übersicht über die Satyriden der Gattung Hyponephele Muschamp, 1915 VII. Die Arten Hyponephele lycaon (Rottemburg, 1775), H. pasimelas (Staudinger, 1886), H. lycaonoides D. Weiss, 1978, H. przhewalskyi Dubatolov, Sergeev & Zhdanko, 1994, H. dzhungari // Atalanta. - 2001. - Vol. 32, № 1/2. - P. 111–186.

159. Mallet, J., Besansky, N., Hahn, M.W. How reticulated are species? // Bioessays. - 2016. - Vol. 38, № 2. - P. 140-149.

160. Gompert, Z., Fordyce, J.A., Forister, M.L., et al. Homoploid hybrid speciation in an extreme habitat // Science. - 2006. - Vol. 314, № 5807. - P. 1923–1925.

161. Green, R.E., Krause, J., Briggs, A.W., et al. A draft sequence of the neandertal genome // Science. - 2010. - Vol. 328, № 5979. - P. 710–722.

162. Durand, E.Y., Patterson, N., Reich, D., et al. Testing for ancient admixture between closely related populations // Molecular Biology and Evolution. - 2011. - Vol. 28, № 8. - P. 2239–2252.

163. Martin, S.H., Davey, J.W., Jiggins, C.D. Evaluating the use of ABBA-BABA statistics to locate introgressed loci // Molecular Biology and Evolution. - 2015. - Vol. 32, № 1. - P. 244–257.

164. Ball, R.M., Neigel, J.E., Avise, J.C. Gene Genealogies within the Organismal Pedigrees of Random-Mating Populations // Evolution. - 1990. - Vol. 44, № 2. - P. 360-370.

165. Van Valen, L. Ecological species, multispecies, and oaks // Taxon. - 1976. - Vol. 25, № 2–3. - P. 233–239.

166. Andersson, L. The driving force: species concepts and ecology // Taxon. - 1990. - Vol. 39, № 3. - P. 375–382.

167. Currat, M., Ruedi, M., Petit, R.J., et al. The hidden side of invasions: Massive introgression by local genes // Evolution. - 2008. - Vol. 62, № 8. - P. 1908–1920.

168. Valencia-Montoya, W.A., Elfekih, S., North, H.L., et al. Adaptive introgression across semipermeable species boundaries between local helicoverpa zea and invasive helicoverpa armigera moths // Molecular Biology and Evolution. - 2020. - Vol. 37, № 9. - P. 2568–2583.

SAINT-PETERSBURG STATE UNIVERSITY

Printed as manuscript

Pazhenkova Elena Alekseevna

Integrative analysis of taxonomically complex groups of the Nymphalidae butterflies (Lepidoptera)

Scientific specialty 1.5.14. Entomology

Dissertation for a scientific degree of Candidate
of Biological Sciences

Translation from Russian

Supervisor:

Doctor of Biological Sciences,

Lukhtanov Vladimir Aleksandrovich

Saint-Petersburg,
2021

Contents

Introduction....69
Chapter 1. History and perspectives of the application of integrative analysis in taxonomy....75
1.1 Application of DNA barcoding in biodiversity research....75
1.2 Integrative approach to species delimitation....77
1.3 Whole-genome analysis in phylogenetic studies and taxonomy....78
Chapter 2. DNA barcodes as a tool for the formulation of primary taxonomic hypotheses....82
2.1 Analysis of the *Melitaea ala* complex....82
2.2. Analysis of the *Melitaea didyma* complex....85
Chapter 3. Integrative taxonomic analysis of the genus *Brenthis* and the *Hyponephele lycaon – H. lupina* complex....90
3.1 Analysis of the genus *Brenthis*....90
3.2 Analysis of the *Hyponephele lycaon – H. lupina* species complex....97
Chapter 4. Analysis of whole genomes reveals the taxonomic structure and evolutionary history of the *Melitaea acentria-M. persea- M. didyma* species complex....105
4.1. Patterns of genetic differentiation and genomic introgression....105
4.2 *Melitaea acentria*, *M. persea* and *M. didyma* represent different species....109
4.3 Speciation of *M. acentria*....110
Conclusion....112
Acknowledgments....114
References....115

Introduction

Relevance and level of scientific development of research topic

Taxonomy is one of the main components of biology since the delimitation of species and their assignment to higher categories (genera, tribes, etc.) play a key role in the study of biodiversity. Previously, the classification and identification of organisms have been based on morphological traits, which were supplemented by ecological, ethological, biogeographic, embryological, and physiological features. When molecular methods were developed, researchers gained access to a new type of data, which were actively introduced into taxonomy.

The term "integrative taxonomy" was formally introduced to denote an integrated approach to the description and delimitation of species using different types of data and techniques relatively recently [1,2], however, advantages of joint analysis of molecular and morphological data were already noted at the end of the last century (e.g., [3,4]). The integrative approach makes it possible to effectively perform identification and delimitation of species in cases where morphological characters are not indicative [1,5], to resolve controversial taxonomic hypotheses when analyzing complex groups [6], to discover new cryptic species [7] and to carry out their phylogeographical analysis – to reveal the history of emergence and formation of taxa [8,9]. Moreover, an integrative approach often helps to perform species identification faster, which is the most important in research of species-rich and poorly studied groups and ecosystems, e.g., in studies of marine [10] and tropical biodiversity [11,12].

It should be taken into account that different types of data can support contradictory taxonomic hypotheses. Therefore, the main subject of debate in integrative taxonomy is the degree of correspondence between different types of traits which is enough to consider a group of individuals as a separate species. Some authors consider complete congruence between the different types of traits a necessary condition for species delimitation [1,13], while others suggest that any inconsistencies can be explained in terms of the evolutionary history of the selected groups, and even a single set of traits may be sufficient to make a taxonomic decision [14,15]. A relatively common phenomenon is the inconsistency of phylogenetic reconstructions and, accordingly, taxonomic conclusions based on the analysis of nuclear and mitochondrial molecular markers (mito-nuclear discordance) [16,17]. Such contradictions appear particularly often in the study of young taxa, where reticulation events, implying such processes as interspecific hybridization and horizontal gene transfer, are often observed. In such cases, the reconstruction of the evolutionary history of species based on analysis of a few molecular markers is not possible; however, this problem can be solved using analysis of whole genomes [17].

Currently, studies whose keywords include the term "integrative taxonomy" have used only two types of traits (mostly DNA barcodes and morphology) in almost half of the researches [18]. Significantly that in 35.4% of cases the analysis was performed using only one type of traits (in most cases, molecular traits). This may be because some authors understand the term "integrative taxonomy" as the use of different criteria and species concepts rather than data types [18]. Thus, the development of terminology and methodology of integrative analysis is one of the fundamental directions in modern studies of biological diversity, which indicates the relevance of this thesis.

The butterflies of the family Nymphalidae became the model objects of this research. Despite the relatively good knowledge of this family, taxonomic relationships in many groups are insufficiently studied due to its high geographic or seasonal variability, the presence of cryptic species, or interspecific hybridization. [17,19,20].

Butterflies of the genera *Melitaea*, *Brenthis,* and *Hyponephele* representing two subfamilies - Nymphalinae (*Melitaea, Brenthis*) and Satyrinae (*Hyponephele*) were studied within the framework of this study using three approaches: (1) analysis of DNA barcodes (*COI* gene fragments) to make primary taxonomic hypotheses; (2) analysis of a combination of DNA barcodes, a set of nuclear markers, and morphological characters to test primary taxonomic hypotheses; (3) analysis of whole genomic data to examine the species complex in which interspecific introgression was found.

The choice of model objects was determined by the fact that *Brenthis* is taxonomically one of the most studied genera of butterflies, with a small number of species, and the number and boundaries of species within the genus are well known based on traditional morphological analysis [21,22] and are not questioned by lepidopterists. Using such an example, it seemed interesting to verify the results of delimitation based on molecular markers. On the other hand*, Melitaea* and *Hyponephele* represent two species-rich groups of butterflies, the number of species and species boundaries within which are unclear and the subject of debate among scientists [20,23,24].

The scientific novelty of work

For the first time for butterflies, a case of extreme mitochondrial polymorphism was revealed in the *Melitaea didyma* species complex [25,26], when 11 mitochondrial haplogroups were found within one species with the level of divergence up to 7.4 %.

For the first time nucleotide sequences of genes *COI*, *wgl*, *Ca-ATPase, ArgK,* and *CAD* were obtained for *Brenthis* and *Hyponephele,* which was a basis for subsequent taxonomic analysis [27].

Taxonomical structure of species complexes, close *Melitaea ala* [28] and *Hyponephele lycaon* [29] were revealed using an integrative approach.

Twenty-seven whole genomes of butterflies of the complex *M. didyma – M. persea – M. acentria* were sequenced and analyzed. For the first time for genus *Melitaea,* genomic introgression between distantly related species (a phenomenon that is extremely rare in nature) was shown [30].

The theoretical and practical significance of the work

The development of a methodology for integrative taxonomic analysis is of key importance for the study of biodiversity. Application of different approaches to the analysis on well- and least-studied groups of butterflies allows not only to investigate taxonomic relationships in these groups but also to test methods that can be subsequently used in the study of other organisms. Species boundaries of taxa in the complex *Hyponephele lycaon - H. lupina* and in the genus *Brenthis* were explored. A detailed study of the evolutionary history *of Melitaea didyma - M. persea - M. acentria* complex using whole-genome data showed the presence of a unidirectional gene flow from *M. didyma* to *M. acentria*. The gene flow was an additional source of genetic diversity in the local population of *M. acentria*. The important contribustion to evolutionary biology is a discovery of genomic introgression between phylogenetically distant species (the minimum divergence time of *M. didyma* and *M. acentria* was estimated as 5 million years ago). The database of sequences of the genes *COI*, *ArgK*, *CAD*, *Ca-ATPase, wgl,* as well as whole genomes for butterflies of the Nymphalidae family has been significantly replenished, which can be used in further studies. Specific primers for *ArgK* and *CAD* genes specific for Nymphalidae have been developed [27]. Taxonomic studies have potential conservation significance. Some of the studied species (e.g., *Melitaea acentria*, *Hyponephele galtcha*) are endemics and are subject to conservation. The findings of this thesis can be used in the preparation of courses on entomology, genetics, and bioinformatics.

Goals and objectives

The work aims to study the taxonomic structure in complex groups of butterflies of genera *Brenthis*, *Melitaea,* and *Hyponephele* (Lepidoptera, Nymphalidae) using an integrative approach.

Objectives:

- To study the species and subspecies structure of the *Melitaea ala* complex.
- Conduct a primary taxonomic analysis of a complex of species close *to Melitaea didyma* using DNA barcodes.
- Analyze the case of inconsistency of phylogenetic reconstructions and taxonomic conclusions based on the analysis of mitochondrial and nuclear markers, using the example of the genus *Brenthis*.

- Perform a taxonomic analysis of the species of the group *Hyponephele lycaon - H. lupina* using multilocus genetic markers, morphological characters, and biogeographic data.
- Using genome-wide data, carry out species delimitation and reveal the evolutionary history for a species group with reticulation events (using the example of *Melitaea didyma – M. persea – M. acentria* species complex).

Main findings to be defended

1. Analysis of DNA barcodes revealed three subspecies of *Melitaea ala* (*M. ala ala*, *M. ala zaisana, M. ala bicolor)*.
2. The *Melitaea didyma* complex has a high level of mitochondrial polymorphism: within *M. didyma* sensu stricto 11 deeply diverged mitochondrial genetic lines are detected; 12 lineages, found within the complex, corresponding to the previously described species.
3. Within the genus *Brenthis*, phylogenetic relationships and groupings of individuals revealed using mitochondrial DNA barcodes (*COI* gene fragment) do not correspond to phylogeny and taxa identified based on analysis of nuclear markers and morphology. At the same time, nuclear markers and morphological characters are completely congruent in identifying species boundaries.
4. Integrative analysis of DNA barcodes, nuclear genes, geographical ranges, wing coloration, and male genitalia structure allows us to determine the boundaries of the taxa within *Hyponephele lycaon – H. lupina* species complex.
5. Whole-genome analysis of butterflies of the *Melitaea didyma – M. persea – M. acentria* complex confirmed that they represent separate species despite incomplete reproductive isolation between *M. didyma* and *M. acentria.*
6. *Melitaea acentria* originated in the Levant as a peripheral geographic isolate of *M. persea*. Genomic introgression from the nonrelated species *Melitaea didyma* was a source of genetic diversity in the local population of *M. acentria*. Sex Z chromosomes of *M. acentria* are more resistant to gene flow than autosomes and play role in the formation of postzygotic reproductive barriers.

Material and methods

The study was based on the collection materials stored in the Karyosystematics Department of the Zoological Institute of the Russian Academy of Sciences. Additional fresh material was collected during expeditions to Russia (Samara Region, Orenburg Region, Bashkortostan), Bulgaria, Tajikistan,

Mongolia, Iran, and Israel. Insects were collected using an entomological net and were stored on cotton mattresses. Material for whole-genome analysis was fixed using Allprotect Tissue reagent (Qiagen) to ensure the best preservation and stored at -80°C. For a detailed list of the material used and collection points, see our publications [25–30].

Preparation of male genitalia structures was performed according to the following protocol: the end of the abdomen of adults was boiled in a hot 10% KOH solution for 3-10 min. Then the abdomen was washed with water and dissected under a light microscope. Samples were stored in glycerin in tightly sealed tubes. Photographs of genital structures were taken with a Leica M205C binocular camera equipped with a Leica DFC495 camera. The contours of the unci and subunci for the genus *Hyponephele* were prepared using GIMP software [31]. Geometric morphometrics was performed using Momocs v1.2.9 [32]. Firstly, shapes were centered, aligned, and scaled, then a discrete cosine transform was applied to obtain the Fourier coefficients [33]. Principal component analysis (PCA) was used to visualize differences between groups.

Cladistic analysis of butterflies of genus *Brenthis* as performed based on the analysis of a matrix containing 11 phenotypic traits using the methods of maximum parsimony in PAUP* [34] and Bayesian analysis in MrBayes v. 3.1.2 [35]. A detailed description of morphological traits and analysis can be found in [27].

Sample preparation for sequencing was performed in the Karyosystematics Department of the Zoological Institute of the Russian Academy of Sciences and included DNA extraction, PCR, purification of PCR products, and analysis of the results using gel electrophoresis (for a detailed description, see our publications [25,27]). Sequencing was performed at the St. Petersburg State University Resource Center «Development of Molecular and Cellular Technologies". Part of the material was sent for sequencing to the Canadian Centre for DNA Barcoding (Biodiversity Institute of Ontario, University of Guelph, Canada). A total of 558 COI gene fragment sequences (DNA barcodes) and nuclear gene sequences were obtained for 117 samples (about 400 sequences).

Alignment and processing of sequences were performed using BioEdit 7.1.7 software [36]. Alignment was done using the algorithm ClustalW [37]. Phylogenetic trees were obtained using Bayesian analysis in MrBayes v. 3.1.2 [35]. See [25–27,30] for more details.

Sample preparation (DNA extraction and library preparation) and sequencing of whole genomes were performed at Macrogen, Korea on a commercial basis. Illumina TruSeq DNA libraries with an insertion size of 350 bp were sequenced on the Illumina HiSeq X Ten platform with 2x150 bp read lengths. An average of 9 Gb was obtained for each sample.

Subsequent bioinformatic analysis was performed with a support of Resource Center "Computer Center SPbU". Reads were mapped to the reference genome *Melitaea cinxia* v. 1 [38] and single nucleotide polymorphisms (SNPs) were called using Genome Analysis Toolkit v. 3.7 (GATK)

[39]. The final high-quality dataset included 768103 SNP after filtration. Analysis of the genetic structure of populations was performed using PCA, included in SNPRelate library [40]. Reticulation events analysis was conducted in SplitsTree 4.16.2 [41]. Fixation index F_{ST} and nucleotide diversity π were calculated using the script from [42]. Estimation of direction and level of genomic introgression was performed with dsuite software [43]. Demographic modelling using site-frequency spectra was done in δaδi [44]. Phylogeographic analysis using mitochondrial marker *COI* was performed in RASP v.3.2 [45]. A phylogenetic network for mitogenomes was constructed in Fitchi [46]. Detailed description of whole-genome analyses can be found in [30].

Publications and approbation of results

Results of the research were presented at 5 international conferences:
VI International Conference on the Karyosystematics of the Invertebrates KARYO-VI (Saratov, Russia, 27-30 August 2016), VIII International Conference Biology of Butterflies (Bangalore, India, 11-14 June 2018), RECOMB Comparative Genomics 17th (Montpellier, France, 1-4 October 2019), International Youth Scientific Forum «Lomonosov-2020» (Moscow, Russia 13-17 April 2020), the 26th European Meeting of PhD Students in Evolutionary Biology (EMPSEB) (Online, Ireland, 1-4 March 2021).

Based on the results of the research, 6 articles in journals, indexed by WoS and/or Scopus were published, including three articles in journals of the first quartile and two articles in journals of the second quartile of SJR (Genetics or Zoology).

The main findings of this study were regularly presented at meetings of the Department of Entomology of St. Petersburg State University.

Chapter 1. History and perspectives of the application of integrative analysis in taxonomy

1.1 Application of DNA barcoding in biodiversity research

The use of molecular methods to solve taxonomic problems began in the second half of the last century and became widespread during the last decades. Many researchers considered promising the use of mitochondrial DNA to solve a variety of scientific problems for several reasons [47]. The first is the relatively conservative structure of the mitochondrial genome: small size, the constant composition of genes, absence of spacers between them. Secondly, mitochondria are maternally transmitted, and mitochondrial DNA is not involved in recombination. Third, the rate of evolutionary changes of mitochondrial DNA is high enough to be used as a marker in studies of population ecology, for phylogeny reconstruction, and in phylogeography [9].

DNA was used for species identification even before the 21st century - for example, to identify the species of meat fly larvae *Phormia regina, Phaenicia sericata* и *Lucilia illustris* (Diptera, Calliphoridae), that develop in a dead body at different periods, allowing forensics to predict the time of death with relative accuracy [48]. A method of molecular identification was developed to determine the species of economically valuable sturgeon fish [49]. However, species identification based on analysis of DNA sequences became extremely popular only after the publications of Canadian biologist Paul Hebert [50], who suggested not only using DNA barcodes for species identification, but also created a constantly updated database of living organisms, where, according to his plan, DNA barcodes of all living organisms, will be stored. In animals, including insects, a fragment of the mitochondrial gene *COI* is used as DNA barcode, in plants – the plastid genes *matK* и *rbcL*, in fungi – nuclear spacers *ITS*. Hebert and colleagues demonstrated the effectiveness of the proposed method in several groups of animals — in birds [51], Collembola [52], spiders [53], mayflies [54], fishes [55].

For *COI* sequences, p-distances greater than 0.02 are used as an empirically established threshold for species delimitation (p-distance is the ratio of the number of substitutions in two compared sequences to the length of the sequences) [50]. However, in some taxa, intraspecific differences exceed the threshold value and interspecific differences are rather low [56]. Therefore, the barcoding gap does not always exist - interspecific variability exceeds the level of intraspecific variability [56,57]. Alternative approaches to species delimitation are to identify groups of individuals on the phylogenetic tree that form monophyletic lineage or to search for combinations of species-specific nucleotide substitutions [58]. There are several heuristic approaches used for species delimitation that focus on selecting optimal parameters for determination of interspecies boundaries, such as Automated barcoding gap discovery (ABGD) [59], General Mixed Yule Coalescent (GMYC) [60] or Poisson Tree Process (PTP) [61]. Probabilistic algorithm PROTAX [62] can be applied if a large base of reference sequences is available, and it does not depend on the specified interspecific

threshold. This method has proven to be effective in large-scale analysis of European butterflies, in which DNA barcodes were obtained and analyzed for 97% of the species recorded in Europe. [63].

The undoubted advantage of DNA barcoding is that a few nanograms of material are enough for PCR and subsequent sequencing [64], and DNA was successfully extracted from snow tracks [65], dried snake venom [66], marine mollusks mucus [67] and even from the air [68]. With the development of second-generation sequencing technologies, it has become possible to simultaneously read multiple samples on the Illumina platform, which has greatly increased the efficiency of analysis. [69–71]. Third-generation sequencing, in particular Oxford Nanopore's MinION platform, allows not only read a huge amount of data but also perform analysis in the field, such as in the rainforests of Ecuador [72] or during reef monitoring [73], reducing the distance between the sample and the laboratory. The SEQUEL platform, based on molecule real-time sequencing (SMRT), was developed, which significantly reduced the cost of analysis [74]. This indicates the demand for a method to which modern sequencing technologies were adapted - to date, the main BOLD database contains about 9.8 million sequences [75] for 233K animal species and 70K plant species.

The original goal of DNA barcoding is the precise species identification of already described species, regardless of intraspecific polymorphism, sex, or stage of ontogenesis. Subsequently, Hebert suggested using DNA barcoding to identify species that have not yet been described. Using the example of the tropical skipper (Lepidoptera, Hesperiidae) *Astraptes fulgerator*, researchers have shown that this species represents a complex of 10 species, which were detected as clusters of DNA barcodes correlated with morphological and ecological differences [76]. Latter DNA barcoding has repeatedly been used to detect cryptic biodiversity in a wide variety of groups of living organisms: in Tachinid flies [77], Oligochaetes [78], Bivalvia [79], spiders [80], fishes [81] and many others. The hidden biodiversity among butterflies was also high - a large-scale study of European butterflies showed that 27.7% of the 299 species studied represented more than one mitochondrial lineage [82].

However, the clusters identified based on *COI* gene analysis do not always correspond to the actual species boundaries. Multiple mitochondrial lineages may be a consequence of ancestral polymorphism [83], mitochondrial introgression [84] or long-time geographic isolation with following secondary contact [9]. Infection with an intracellular symbiont of the genus *Wolbachia*, which, by causing cytoplasmic incompatibility, actually provides intraspecific reproductive isolation, resulting in increased genetic diversity of mitochondrial DNA, can affect species delimitation [85]. A rarer problem of DNA barcoding is heteroplasmy, or the presence of two types of mitochondria in the same sample [86], which can be solved by involving additional genetic markers [87]. Similar difficulties are caused by the presence of mitochondrial pseudogenes (numts) in the nuclear genome, which can be amplified using primers developed for DNA barcoding [88].

In a recent study, Makhov with colleagues conducted a detailed analysis of possible causes of mitochondrial polymorphism in the moths of the genera *Alcis* and *Thalera* (Lepidoptera, Geometridae) [89]. Researchers did not find a connection between mitochondrial haplogroups of species *A. deversata* and *Th. chlorosaria* with *Wolbachia*, and revealed, that, most probably, the mitochondrial polymorphism was caused by secondary contact of allopatric populations. Also, mitochondrial introgression from *A. extinctaria* to *A. repandata* was detected.

Within small skipper *Thymelicus sylvestis* (Lepidoptera, Hesperiida) six differentiated mitochondrial genetic lines were detected, which did not correlate with nuclear genomes [16]. The authors excluded the influence of *Wolbachia*, pseudogenes, introgression, and ancestral polymorphism, and proposed as a possible explanation geographical isolation with subsequent fusion, as a result of which the diversity of the nuclear genome was eliminated during recombination, and mitochondrial lines were preserved. [16].

Thus, DNA barcoding is a convenient tool for primary taxonomic analysis, which is widely used in various groups of living organisms. Nevertheless, the mismatch of mitochondrial clusters with the boundaries of taxa is quite common. [90–95]. Therefore, the taxonomical studies, based on DNA barcodes, should be supplemented with the analysis of other genetic markers, morphology, geographic distribution, ecological preferences, and other possible types of data to avoid erroneous conclusions [15,96].

1.2 Integrative approach to species delimitation

The approach to the definition of species boundaries using different types of data was formally called "integrative taxonomy" in 2005 [1]. This approach is used both in clarifying the status of existing taxa and in describing new species [18,97–99].

Taxonomy has undergone significant changes in its methodology over the past decades. The most commonly used data type in integrative taxonomic analysis of arthropod [15] and other animals [18] after traditional morphological traits become molecular data because information about structure and functions of genes, proteins, and metabolites opens up the broadest possibilities for comprehensive analysis. With the development of sequencing technologies, phylogenetics, which is based on single markers, is slowly being replaced by phylogenomic, which allows more subtle conclusions about the processes of speciation [100] and more accurate taxonomic interpretation. The integrative analysis is gradually becoming a routine procedure for researchers, which makes a significant contribution to the development of modern taxonomy. [18].

The methodology of the integrative approach has conceptual difference in the works of different researches. For example, Dayrat pointed out the need for congruence between unrelated data types [1]. It is expected that evolutionary lineages detected in this way are more likely to correspond to

species because the random coincidence of evolutionary changes is unlikely. However, the process of the accumulation of specific traits at different levels is not always carried out simultaneously [101], and there is a high risk of errors of Type I when existing species will not be detected. This problem is especially relevant when studying young groups, such as cichlid fish, in which there are morphological differences between species, but weak reproductive isolation and genetic differentiation by neutral loci [102]. This problem also exists when studying groups, where reticulation events were detected, including such processes as interspecific hybridization, introgression, and horizontal gene transfer. In such events, the evolutionary history of different traits in one organism may significantly differ, and not all of them can adequately reflect real species boundaries.

The cumulative approach to integrative taxonomic analysis is based on the assumption that, since speciation is a long process, the divergence of different traits may not occur simultaneously, so congruence of different data types is desirable but not a strict condition for making a taxonomic hypothesis [14]. In some cases, even a single set of characters may be sufficient to make a taxonomic conclusion [14,15]. To avoid false positive assignments, matches, and inconsistencies between different types of data need to be explained in terms of the evolutionary history of the groups studied, for example, conclusions based on analysis of molecular markers should be supported by analysis of morphology [15].

The definition of species criteria is another conceptual problem, which is closely related to the methodological problem of species delimitation. There are more than 20 species concepts [103], many of which have conflicting criteria. To avoid the taxonomic contradictions that arise when using different definitions of species systems, a unified species concept has been proposed, which considers a species as an independently evolving group of metapopulations [14]. Since between young taxa the formation of differences may not have been completed, an evolutionary lineage can be considered a species in compliance with the criteria of one of the species concepts. Thus, it is emphasized that a species is a dynamic unit of evolution that changes over time, and the same criteria are often not applicable to young taxa as to more ancient ones. At the same time, data on the origin and evolutionary trajectory of species are closely related to the determination of interspecific boundaries [96], so understanding the processes of speciation is important for solving taxonomic problems. It should be emphasized that the integrative approach implies combining different types of data, rather than different species criteria, which can also be applied to the same type of traits [18], however, the definition of species and its criteria are key to taxonomy.

1.3 Whole-genome analysis in phylogenetic studies and taxonomy

Recently, sequencing has become less expensive, and modern equipment makes it possible to read increasingly longer nucleotide sequences, so phylogenomics, based on the simultaneous analysis

of multiple loci, has developed [104]. A version of insect phylogeny based on sequence analysis of 1,478 genes (2.5 Mb of coding DNA) of 103 species of insects, which include representatives of all modern orders, was published recently [105].

A good example of the use of phylogenomic data for the integrative taxonomy of arthropods is a study that included analysis of the whole genome of centipede *Strigamia maritima* [106]. This work supported the system according to which insects are combined with crustaceans in the taxon Pancrustacea, and centipedes (Myriapoda) forms a sister clade to Pancrustacea (i.e., a basal clade within Mandibulata). Accordingly, the authors postulate an independent adaptation to the terrestrial environment for the ancestors of Insecta and Myriapoda, which is confirmed by the fact that the olfactory receptors of these groups are encoded by different gene families. The proposed classification is also supported by the fact that centipedes have a specific mechanism for providing a variety of immune response proteins. In *Drosophila* and *Daphnia*, the *Dscam* gene (Down syndrome adhesion molecule), which plays a key role in the immune system, is represented in the genome by four copies and is capable to generate about 150k and 13k isoforms, respectively. Thus, the probable synapomorphy of insects and crustaceans is a unique system of alternative splicing. The genome of *Strigamia maritima* contains 80 *Dscam* paralogs resulting from duplication, and the diversity of proteins is provided due to the presence of multiple copies of this gene [106].

Whole-genome data can be a useful tool for describing new species. Recently, a new species of orangutans, *Pongo tapanuliensis*, was discovered, which differs from two other species of orangutans *P. pygmaeus* and *P. abelii* not only morphologically in the structure of the mandibular region of the skull and teeth, but it also represents an independent genetic unit [107]. Dating based on analysis of whole genomes showed that *P. tapanuliensis* represents the oldest orangutan lineage, having separated ~3.38 million years ago. This species is an endemic living near Lake Toba (Sumatra) with a population size of about 800 animals, thus, it needs to be under protection. [107].

The study of whole genomes allows to directly relate the problem of species delimitation to the problem of speciation. The searching for the regions responsible for speciation has long attracted the attention of researchers, and when it has been shown that divergence occurs unevenly in different parts of the genome, so-called "genomic islands of speciation" were distinguished [108]. These regions often contain genes that play a key role in adaptation to the environment or provide interspecies reproductive isolation [109]. In the study on the Galapagos finches (*Geospizinae*), the whole-genome sequencing of 120 individuals of all species of these endemic group was performed [110]. The association of the *ALX1* gene, encoding the transcription factor responsible for craniofacial development, with the diversity of beak morphology was shown. Thus, the genome region that played a role in adaptation to food resources, emergence of ecological prezygotic reproductive isolation and speciation in this model object was found [110].

The accumulation of genetic differences and the formation of reproductive barriers can also occur unevenly on different chromosomes. For many organisms, the large X-effect has been shown, i.e. genes associated with the formation of postzygotic reproductive barriers (for example, hybrids sterility) are concentrated rather on the X chromosome (or on the Z chromosome in species with a ZW sex determination system) than in autosomes [111,112]. Thus, sex chromosomes play a disproportionately greater role in the development of reproductive isolation than autosomes. Closely related to this phenomenon is the phenomenon called faster X-effect (or faster Z-effect), indicating that X chromosomes show a higher level of divergence than autosomes [112–114]. These differences may be caused by selection. According to the Dobzhansky-Müller model, in two independently evolving lineages, such alleles can arise that, upon secondary contact, cause the non-viability (or reduced fitness) of the hybrids [115]. This model explains the reduced level of introgression in sex chromosomes compared to autosomes[116], which was detected in many animals: *Heliconius* butterflies [117], Mexican Howler monkeys [118], nightingales [119] and human [120]. Analysis of genes located on the Z chromosomes of butterflies showed that they encode proteins involved in partner recognition, wing development, and circadian rhythms, which indicates a significant role of Z chromosomes in the formation of traits influenced by sexual selection [116]. Thus, sex chromosomes are an important genomic region involved in promoting and maintaining reproductive isolation.

Various factors can influence the formation of genetic divergence and reproductive barriers between populations. First, genetic differences may develop under the influence of natural and sexual selection acting on sister populations [121]. Second, neutral mutations fixed in population due to genetic drift can cause the formation of reproductive isolation [122,123]. Third, the introgression of genes from a non-sister species to one of the allopatric populations may also play a role in speciation. Such a scenario was recently shown when studying the phylogenetic history of the Grant's gazelle [124]. Interspecies introgression is more common in nature than previously thought [125–128], and can play role in adaptation because it leads to the transfer of genes and traits from one phylogenetic lineage to another [129,130]. Adaptive introgression in non-closely related species of butterflies revealed by analysis of whole-genome data was shown in *Heliconius* (Lepidoptera, Nymphalidae) [131].

Interspecific hybridization can significantly complicate species identification as well as the determination of species boundaries. Hybrids often show an intermediate morphological structure between parental species [132], however, in some cases hybridization may lead to the formation of unique traits and speciation [133]. In such groups, phylogeny reconstruction based on several genetic markers can lead to contradictory results and increasing the number of data can help in determining species boundaries. For example, in a study of cichlid fishes, a well-known example of a group with

common interspecific hybridization, analysis of whole-genome data obtained by RAD has greatly improved the resolution of phylogeny [134].

Thus, the analysis of complete genomes opens up a wide range of possibilities for solving taxonomic problems, while remaining available for researchers. The use of modern molecular techniques serves as an important addition to traditional zoological approaches and becomes a part of an integrative approach to taxonomy.

Chapter 2. DNA barcodes as a tool for the formulation of primary taxonomic hypotheses

2.1 Analysis of the *Melitaea ala* complex

A complex of species, close to *Melitaea ala* Staudinger, 1881 is widespread in Central Asia and includes species *Melitaea ala* Staudinger, 1881, *M. kotshubeji* Sheljuzhko, 1929, *M. ninae* Sheljuzhko, 1935, *M. chitralensis* Moore, 1901 and *M. enarea* Fruhstorfer, 1917 [24]. According to the recent revision of genus *Melitaea* [20], the complex includes species *M. acraeina* Staudinger, 1881, *M. didymina* Staudinger, 1895, *M. bundeli* Kolesnichenko, 1999, *M. sutschana* Staudinger, 1881 and *M. yagakuana* Matsumura, 1927 (the last is sometimes considered as a subspecies of *M. sutschana* [135]). Subsequent analysis of *Melitaea* using molecular markers shown that *M. acraeina* is a phylogenetically distant species, forming a sister clade to the lineage (*M. ala* + *M. enarea*) when *M. chitralensis* and *M. sutschana* belong to *M. didyma* species complex, which is a sister branch to ((*M. acraeina* + (*M. ala* + *M. enarea*)) [136].

We analyzed DNA barcodes (fragments of mitochondrial gene *COI*) of *M. ala* and showed, that samples, identified as *M. ala*, *M. kotshubeji,* and *M. enarea* were highly supported monophyletic lines (Fig. 1) [28]. These three species formed a clade, corresponding to *M. ala* species complex. *Melitaea acraeina* formed a separate clade, which is consistent with previous results [136]. Taxa close to *M. didyma* were clustered together, forming *M. didyma* complex (Fig. 1, clade 3), which will be discussed in detail in the next section. Species of *M. persea* group were placed in a separate clade, which was sister to the *M. didyma* group (Fig. 1, clade 5).

M. ala included five supported genetic lineages (1) *M. ala ala*, (2) *M. ala irtyshica*, (3) *M. ala zaisana*, (4) *M. ala bicolor* (*b1*), and *M. ala bicolor* (*b2*). Since the discovered lines had high supports, they can be considered as taxa from the point of view of the phylogenetic concept of the species, in which diagnosable unit can be considered as taxa regardless of the presence of reproductive isolation (about the concept see: [137,138]).

To resolve the issue of the taxonomic interpretation of the identified lines, we studied the level of genetic divergence between them and compared the wing patterns. We compared a level of divergence of COI gene within and between the clades, and in all the cases p-distances were lower than the standard interspecific barcoding threshold of 2% [50]. The lowest level of differentiation of DNA barcodes was between *M. ala zaisanica* and *M. ala irtyshica,* and we suggest considering them as a single taxonomical unit [28]. Haplogroups b1and b2 had a slightly higher level of divergence (0.3–0.8%), however, p-distances within a group were also high (up to 0.8%), so they possibly represent a case of mitochondrial polymorphism within *M. ala bicolor*. Thus, among the studied individuals, we can distinguish three subspecies within *M. ala: M. ala ala, M. ala bicolor* Seitz, 1908, and *M. ala zaisana* Lukhtanov, 1999 (=*M. ala irtyshica* Lukhtanov, 1999).

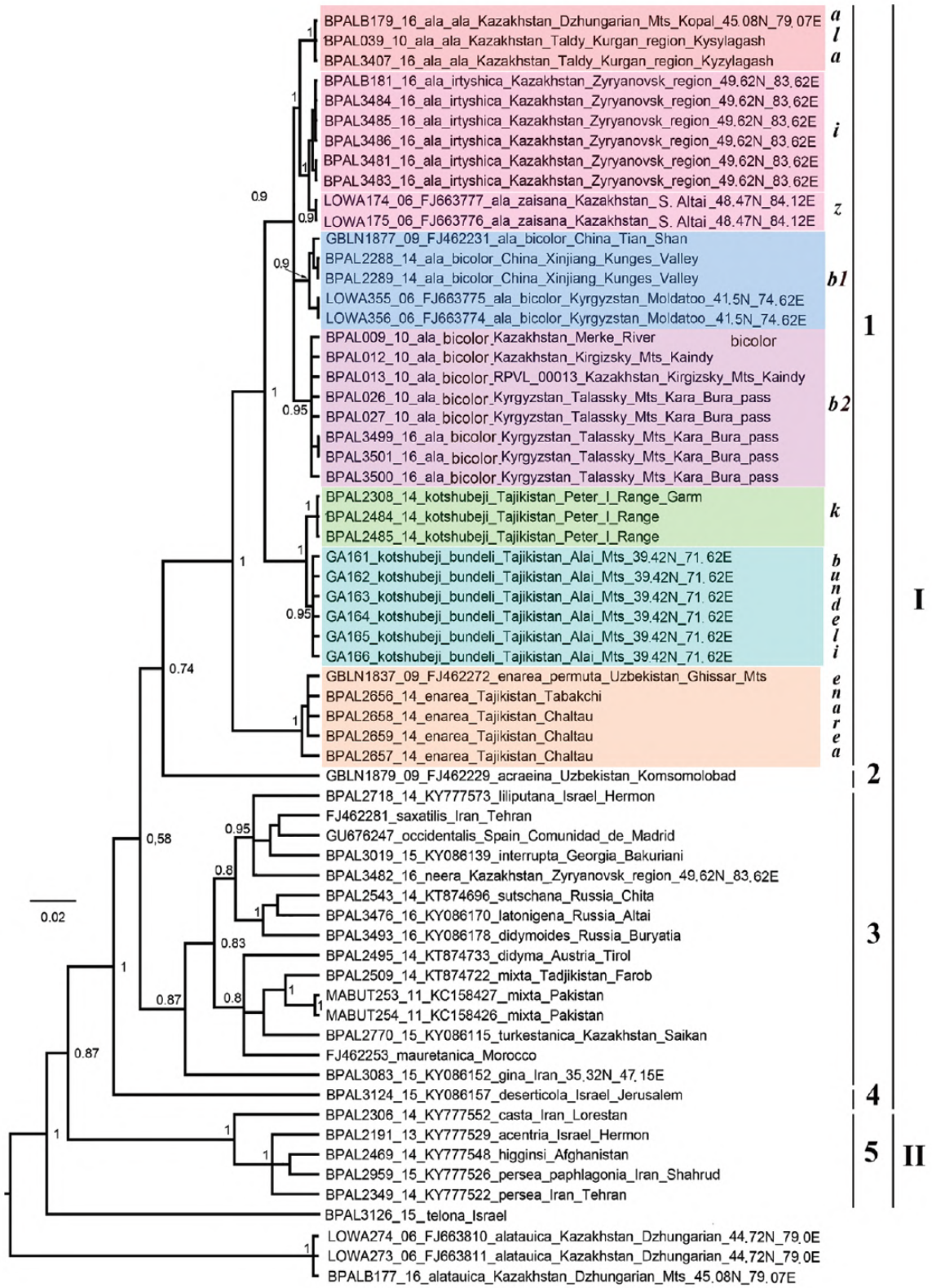

Figure 1. Bayesian phylogenetic tree inferred from sequences of a fragment of *COI* gene. *Melitaea alatauica* was used as an outgroup. Bayesian posterior probabilities are shown on branches. Clades: ***b1*** - *M. ala bicolor* 1, ***b2*** - *M. ala bicolor* 2, ***k*** - *M. kotchubeji kotchubeji*, **i** – *M. ala irtyshica*, **z** – *M. ala zaisana*, 1 –*M. ala* complex, 2 - *M. acraeina*, 3 –*M. didyma* complex, 4 - *M. deserticola*, 5 - *M. persea* complex, I –*M. didyma* group, II –*M. persea* group. From [28].

Subspecies *Melitaea ala ala* is found in East Kazakhstan and the Dzungarian Alatau and has darkened veins on the underside of the hind wing, which form clear cells on the median band (Figure 2). Subspecies *M. ala bicolor* Seitz, 1908 is distributed in Tian-Shan (except for the southern part), including southeast Kazakhstan, northwest China, and Kyrgyzstan. These butterflies have less highlighted veins on the underside of the hind wing and the median band is marked only by dark brackets on its outer side (Figure 2). *M. ala zaisana* is distributed in South Altai and morphologically more similar to subspecies *M. ala bicolor* than to geographically closer *M. ala ala* (Figure 2) [28].

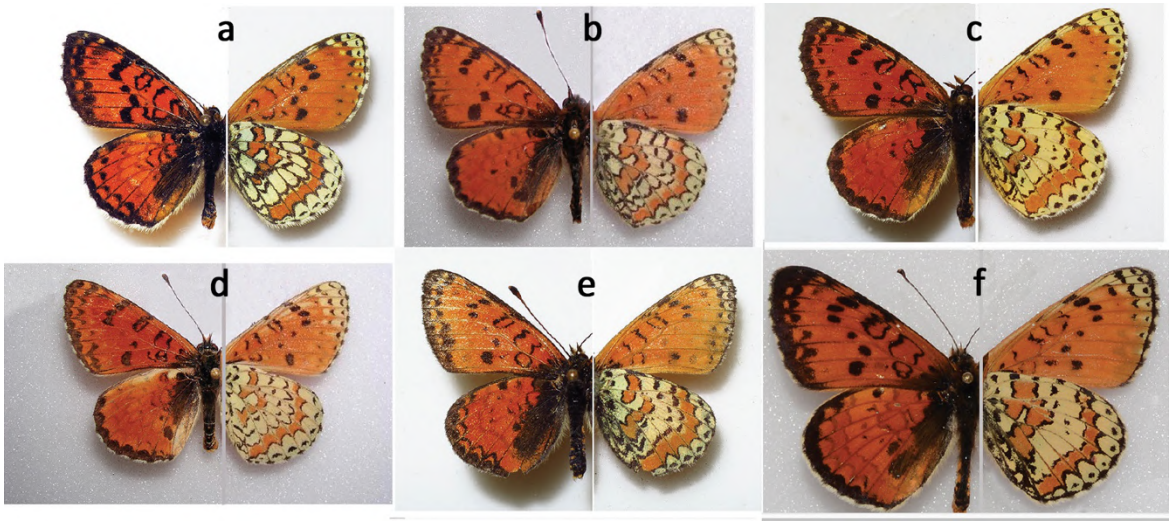

Figure 2. Butterflies of *Melitaea ala* species complex, upperside (right) and underside (left). **a** *M. ala ala*, **b** *M. ala bicolor* (haplogroup b1), **c** *M. ala bicolor* (haplogroup b2), **d** *M. ala zaisana*, **e** *M. kotshubeji bundeli*, **f.** *M. kotshubeji.* Based on [28].

Our analysis revealed two clades within *Melitaea kotshubeji: Melitaea kotshubeji bundeli* and *M. kotshubeji kotshubeji.* Originally *Melitaea kotshubeji bundeli* was described as a subspecies of *M. kotshubeji* [24], but the latter was considered as a separate species [20], or as a synonym [139]. Our analysis of DNA barcodes showed that these two taxa differ in four fixed nucleotide substitutions, indicating a long independent evolution of these lines. Also, these taxa differ in wing color pattern (Fig. 2), and, although they are allopatric, they were found close to each other in Kyrgyzstan, separated only by the narrow valley of the Surkhob river [28].

Thus, in the *M. ala* complex, a rather simple situation for taxonomic interpretation was found when the level of detected genetic divergence between allopatric clades corresponds to the intraspecific and interspecific distances expected in DNA barcode analysis, and differences in DNA barcodes are generally correlated with differences in wing pattern [28]. The next section will consider the case of unusually high intraspecific mitochondrial polymorphism, which indicates the complex evolutionary history of the group studied and requires a comprehensive analysis.

2.2. Analysis of the *Melitaea didyma* complex

Butterflies of the complex of taxa close to *Melitaea didyma* Esper, 1779 are widely distributed in Palearctic [20,135]. Within the complex, a high level of intraspecific geographic and seasonal variability is observed, including differences between individuals of different generations and clinal variability in the size and the color pattern of the wings. At the same time, the species are often poorly distinguished by the coloration of wings and the morphology of the genital structure [20,135]. In such groups, analysis of molecular markers is often used to discover a cryptic taxonomic diversity [7,12]. Our analysis of DNA barcodes showed that the *M. didyma* complex formed a monophyletic group except for three haplogroups with an abnormally high level of divergence (*turkestanica2*, *turkestanica3*, and *gina2*) (Figure 3) and was characterized by a high level of mitochondrial diversity, which can be explained by the presence of cryptic species and/or deep intraspecific polymorphism [25]. To clarify the taxonomic structure of the complex, material from Armenia, Bulgaria, Greece, Georgia, Israel, Iran, Kazakhstan, Kyrgyzstan, Russia, Syria, Slovenia, and Turkey was studied. The analysis showed that the *M. didyma* complex was represented by 23 divergent mitochondrial haplogroups, which had high supports on the phylogenetic tree (Bayesian posterior probability > 0.95) and were associated with certain areas (Figures 3 and 4) [26].

For these haplogroups we chose names, that correspond to taxa of the *M. didyma* complex described from their distribution areas: *chitralensis* (northern Pakistan), *deserticola* (North Africa, Israel, Jordan, Lebanon, Syria), *didyma* (Western Europe), *didymoides* (Asian part of Russia, Mongolia, Northern China), *gina* (Western Iran, Azerbaijan), *gina2* (Iran), *interrupta* (Caucasus, North East Turkey), *latonigena* (Asian Russia, North East Kazakhstan, Mongolia, North West China), *liliputana* (Armenia, Turkey, Syria, Lebanon), *mauretanica* (southern Spain), *mixta* (Tajikistan, Kyrgyzstan, Uzbekistan, Pakistan, Afghanistan), *neera* (Eastern Europe, North Caucasus, Western Siberia, Northern Kazakhstan), *neera2* (Slovenia), *occidentalis* (Spain), *protaeoccidentis* (North Africa), *saxatilis* (Northern Iran), *sutschana* (Far East, Korea, northeast China), *sutschana2* (China), *sutschana3* (Russia, Primorsky Krai), *turkestanica* (Kazakhstan, Kyrgyzstan, Uzbekistan, Tajikistan, Western China), *turkestanica2* (Kazakhstan), *turkestanica3* (Kazakhstan, Uzbekistan) and *turkestanica4* (Kazakhstan) (Figures 3 and 4) [25,26]. Average genetic p-distances between the haplogroups reached 9,5% [26], which was significantly higher than the "standard" threshold for DNA barcodes of 2-3%, which is used for allopatric taxa as a criterion for species delimitation [50].

Most of the haplogroups had an allopatric distribution, however, we detected several cases when butterflies from different haplogroups have been found in sympatry. An M22 sample (*gina2* haplogroup) was observed in sympatry with *gina* in northwestern Iran. The DNA barcodes divergence between the *gina* and *gina2* clusters was 6.5%.

Figure 3. Bayesian phylogenetic tree of *Melitaea didyma* species complex inferred from sequences of a fragment of *COI* gene. Haplogroups of *M. didyma* are shown by bold font. Bayesian posterior probabilities are shown on branches. For detailed dendrograms see [26].

The haplogroups *turkestanica2*, *turkestanica3,* and *turkestanica4* were found in sympatry with haplogroup *turkestanica* (Figure 5) and demonstrated differences in the COI gene up to 7.4%. Two samples of t*urkestanica* haplogroup were found in the area of *neera*: one in the Aktobe region of Kazakhstan, and the other in the Samara region of Russia (Figure 5). At the same time, in Kazakhstan (Karabiryuk) two samples with the *neera* haplotype were observed together with the butterflies, which belong to the *turkestanica* and *turkestanica4* haplogroups.

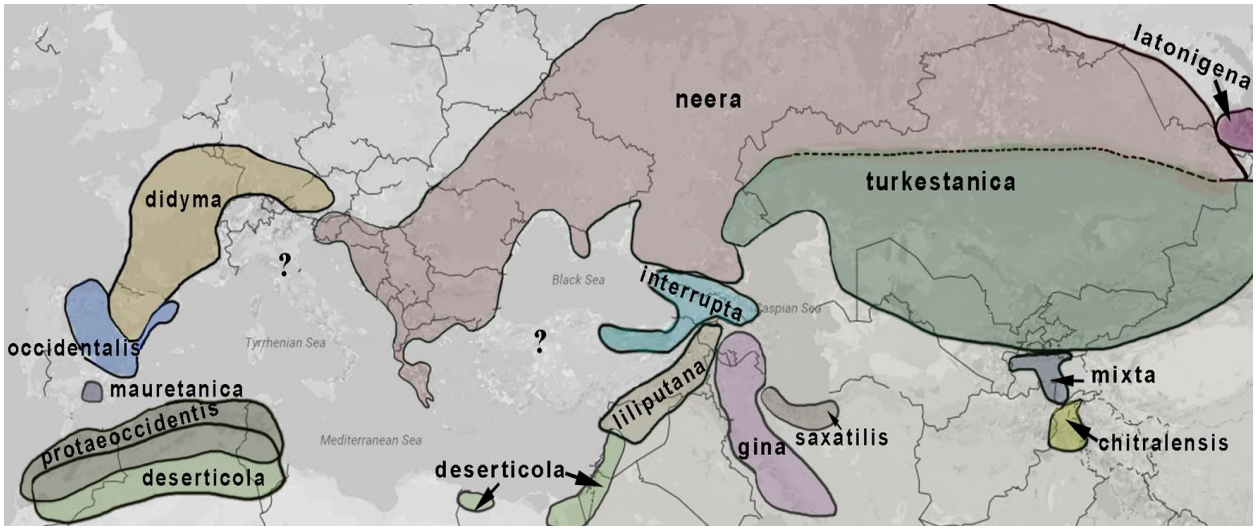

Figure 4. Distribution of mitochondrial haplogroups of *Melitaea didyma* species complex. From [26].

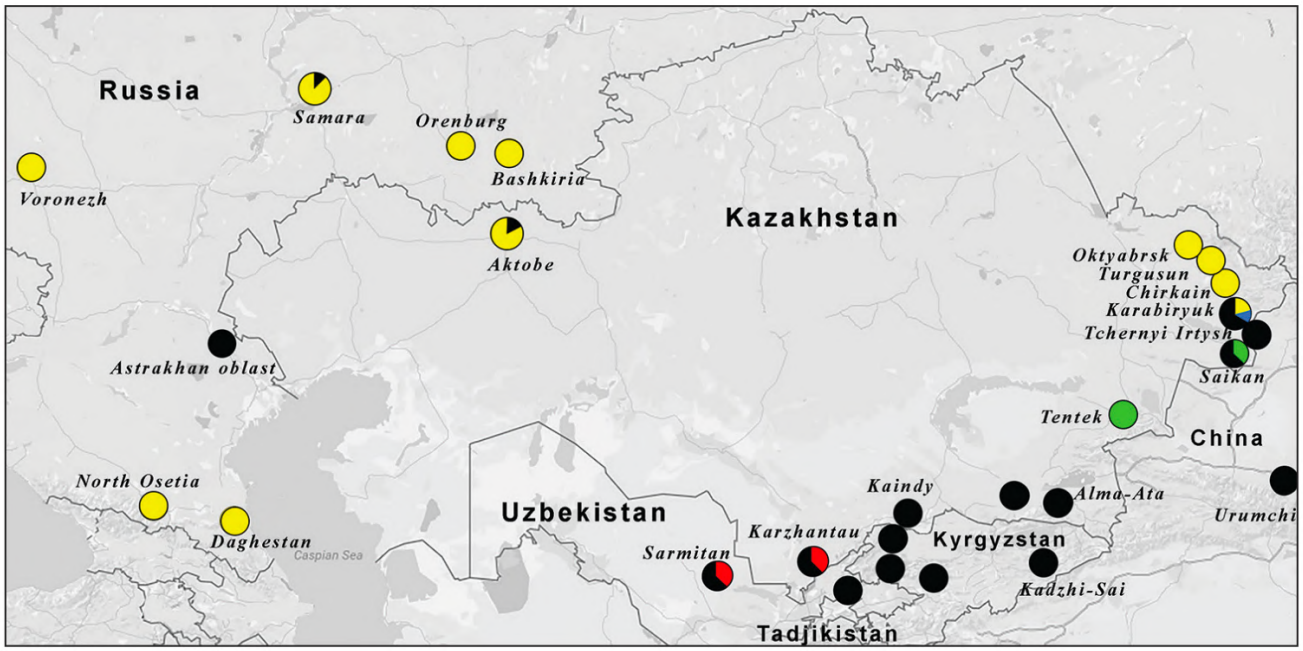

Figure 5. Localization of *neera* and *turkestanica* haplogroups (yellow – *neera*, black – *turkestanica*, green – *turkestanica2*, red – *turkestanica3,* blue– *turkestanica4*).

Since most haplogroups were found in allopatry, given the degree of differences between mitochondrial genetic lineages, it can be assumed that their differentiation occurred during a long period of isolation when their ranges were separated by geographic and/or ecological barriers. According to the accepted p-distribution accumulation rate of 1.5-2.3% per million years for the *COI* gene for different groups of Arthropods [140,141], the isolation between haplogroups can be estimated as approximately 0,5–5,0 million years. Presumably, each of these haplogroups evolved in one of the main West Palaearctic refugia of the Late Pliocene and Pleistocene in North Africa

(*protaeoccidentis*, *deserticola*), on the Iberian Peninsula (*occidentalis*, *mauretanica*), the Balkan Peninsula (*neera*), Middle East (*liliputana*, *saxatilis*, *gina*) and Central Asia (*turkestanica*, *mixta*, chitralensis) [26]. The presence of small haplogroups *neera2*, *turkestanica2*, *turkestanica3*, *turkestanica4*, *gina2*, which may have arisen as geographical isolates but now exist in secondary sympatry with the major haplogroups *neera*, *turkestanica,* and *gina* is consistent with the refugia-within-refugia concept [142].

Regarding the taxonomic structure of the complex, 12 out of 23 identified haplogroups were associated with the following nine traditionally distinguished species: *M. chitralensis* Moore, 1901, *M. deserticola* Oberthür, 1909, *M. didymoides* Eversmann, 1847, *M. gina* Higgins, 1941 (haplogroups *gina*+*gina2*), *M. interrupta* Colenati, 1846, *M. latonigena* Eversmann, 1847*, M. mixta* Evans, 1912, *M. saxatilis* Christoph, 1873 and *M. sutschana* Staudinger, 1892 (haplogroups *sutschana*+ *sutschana2*+ *sutschana3*) [26].

We propose to consider the remaining 11 haplogroups within *M. didyma* sensu stricto [25,26]. According to our results, these haplogroups represent deeply divergent intraspecific genetic lineages (with p-distances up to 7.4%), however, we believe that it would be premature to assign them a species rank. This decision is supported by the following data. First, according to the latest revisions based on the study of morphology, *M. didyma* did not show fixed morphological differences between populations, and therefore all populations are considered as a subspecies *Melitaea didyma didyma*, except for populations from Central Asia, which were assigned to a separate subspecies *Melitaea didyma turkestanica* [20] or even split to several subspecies [143]. Second, in cases of sympatry of haplogroups *neera*/*neera2*, *turkestanica/turkestanica*, *turkestanica/turkestanica3,* and *turkestanica/turkestanica4* no difference in ecological preferences was observed. At the same time, when species *M. didyma/M. interrupta*, *M. didyma/M. latonigena,* and *M. gina/M. saxatilis* were found in sympatry, they occupied separate ecological niches (*M. didyma* and *M. gina* are more xerophilic, whereas *M. interrupta, M. latonigena* and *M. saxatilis* are more mesophilic). Although *M. didyma neera* and *M. didyma turkestanica* are ecologically differentiated [26], in case of sympatry they were found in the same ecological niches. In the Samara region (Russia), where the *neera* haplogroup is predominant, both haplogroups were observed in the typical biotope of *M. didyma neera* (steppe), and in Karabiryuk (Kazakhstan), where the dominant haplogroup is *turkestanica*, both haplogroups were observed in the typical biotope of *M. didyma turkestanica* (desert). These data indicate that the sympatric haplogroups of *M. didyma* represent rather a case of intraspecific variation than different species [25]. This distribution of haplogroups could have resulted from the fusion of two *didyma* lineages that diverged in the allopatry or from interspecies introgression. Interestingly, the haplogroup t*urkestanica2* is not closely related to the haplogroup *turkestanica* but forms a sister clade to the haplogroup *didyma* from Western Europe (Figure 3), which is probably a consequence of ancient

introgression. In general, traces of both ancient and recent introgression indicate weak reproductive barriers between *M. didyma* populations. Thus, our results demonstrate the phenomenon of polyphyly of DNA barcodes at the species level based on deep intraspecific divergence. We propose to consider groups of populations associated with the main haplogroups *M. didyma* as subspecies: *M. didyma didyma* Esper, [1779], *M. didyma mauretanica* Oberthür, 1909, *M. didyma occidentalis* Staudinger, 1961*, M. didyma protaeoccidentis* Verity, 1929, *M. didyma liliputana* Oberthür, 1909, *M. didyma neera* Fischer de Waldheim, 1840 and *M. didyma turkestanica* Sheljuzhko, 1929 [26].

The taxonomic hypotheses we obtained for the *M. didyma* complex [26] were tested latter in an independent study conducted by a group of Spanish researchers using sequencing of ddRAD markers [17]. This study showed an incomplete correspondence between the clusters of individuals, which are distinguished based on the analysis of mitochondrial and nuclear markers (mito-nuclear discordance). A complete match was found for only one cluster from Sicily because it is isolated from the rest of the population and is less prone to mitochondrial introgression. The results, obtained by Dincă et al [17] do not contradict our taxonomic hypotheses. They either confirm them or show that two or more mitochondrial clusters correspond to a single nuclear cluster. The latter indicates that differentiation by mitochondrial markers is not always sufficient for species delimitation and emphasizes the importance of the analysis of multiple genetic markers in the study of biodiversity.

Thus, the results of studies on the genus *Melitaea* confirm the previously formulated statement that the analysis of single molecular markers, such as DNA barcodes, can only be a tool for proposing preliminary taxonomic hypotheses, which need further verification [144]. These hypotheses may then be confirmed, as showed by our analysis of *M. ala* complex (see the previous section), or confirmed only partially, as demonstrated by the testing of our hypotheses by Dincă et al [17].

Chapter 3. Integrative taxonomic analysis of the genus *Brenthis* and the *Hyponephele lycaon – H. lupina* complex

3.1 Analysis of the genus *Brenthis*

As shown above, analysis of such a promising marker as DNA barcodes allows making primary taxonomic hypotheses. However, taxonomic interpretations of phylogenies based only on mitochondrial markers are not always correct. Multiple nuclear genes are potentially more efficient for solving phylogenetic and taxonomic problems, but they have their own specificities [92]. First, in diploid organisms, nuclear genes are represented by two alleles, which can greatly differ in heterozygous individuals, leading to a discrepancy between the phylogeny of alleles and the phylogeny of the species. Secondly, chosen markers are assumed to be unlinked and neutral, which is not always possible to ensure in practice. Third, the relatively low rate of evolution of most of the coding genes together with recombination can lead to a weak phylogenetic signal, which does not cause significant problems in the analysis of higher taxa (genera, tribes, families), but limits the analysis of interspecific taxonomic relationships. Fourth, the time of coalescence of nuclear alleles usually is much longer than the time of coalescence of mitochondrial alleles [145]. Therefore, when comparing evolutionarily young lineages, nuclear genes often may not show fixed nucleotide substitutions, although differences can be found in mitochondrial genes. In such cases, an integrative approach, combining analysis of mitochondrial and nuclear markers, as well as morphological characters leads to more accurate taxonomic conclusions [2].

Genus *Brenthis* Hübner, 1819 includes four species: *B. hecate* Denis & Schiffermüller, 1775, *B. mofidii* Wyatt, 1969, *B. ino* Rottemburg, 1775, and *B. daphne* Denis & Schiffermüller, 1775 (Fig 6).

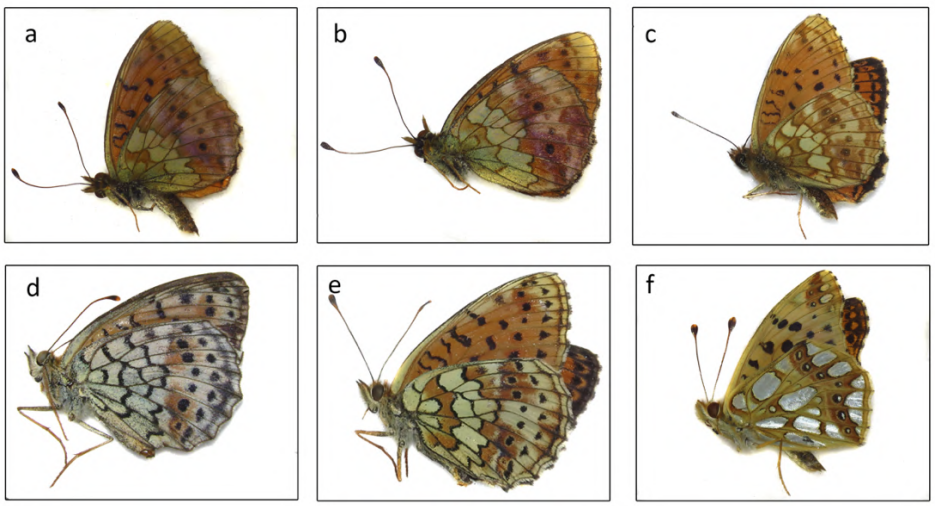

Figure 6. The coloration of the wing underside of *Brenthis*. a – *B. daphne* (D), b – *B. daphne* (D1), c – *B. ino*, d – *B. mofidii*, e – *B. hecate*, f - I*ssoria* lathonia (used as an outgroup). From [27].

Brenthis ino and *B. daphne* are widely distributed in Palearctic (from Western Europe to Japan). *Brenthis hecate* is distributed in the western Palaearctic, and the range of *B. mofidii* is more local (Turkey, Iran, and Iraq). *Brenthis ino*, *B. daphne,* and *B. hecate* are sympatric on most of their distributions [146]. The taxonomic structure of this small genus has been well studied, with detailed analysis of wing coloration, male and female genitalia, preimaginal development, ecological preferences, behavior, and distribution [21,22,146,147].

We performed a phylogenetic analysis based on the study of the *COI* gene, which showed that within the genus *Brenthis* there were two supported monophyletic groups: *B. ino* + *B. daphne* and *B. hecate* + *B. mofidii*, which is consistent with the results of the previous studies [21,22]. More surprising for such a well-studied group was our discovery of two divergent sympatric mitochondrial lineages within *B. daphne* (haplogroups D and D1) based on analysis of gene *COI* (Figure 7) [27].

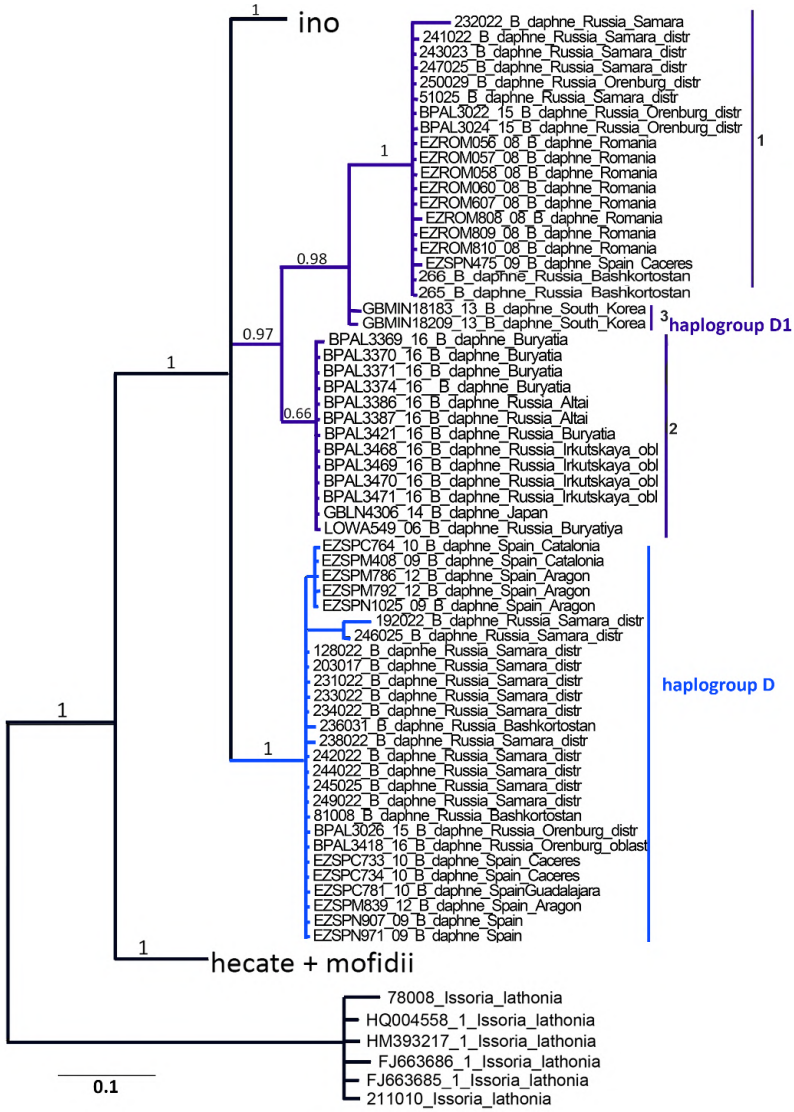

Figure 7. Bayesian phylogenetic tree of genus *Brenthis* inferred from sequences of a fragment of *COI* gene. Main haplogroups of *B. daphne* (D1 and D) are shown by purple and blue, respectively. Numbers 1, 2 and 3 represent distinct genetic lines within haplogroup D1. Bayesian posterior probabilities are shown on branches. From [27].

Haplogroup D included samples of *B. daphne*, collected from Spain to the South Ural (Russia) when samples of haplogroups D1 were detected from Spain to Japan. There were three separate genetic lineages within haplogroup D1, which were associated with geographic regions: the first - with the western Palaearctic (from Spain to the Ural), the second - with the eastern Palearctic (from Altai to Japan), and the third - with South Korea. Butterflies with haplotypes D and D1 were collected by us simultaneously in the same locations in the Samara region and Bashkortostan (Russia). Haplogroups D and D1 did not form a monophyletic group, thus, according to mitochondrial dendrogram, *B. daphne* is a polyphyletic assemblage. Butterflies of D and D1 haplogroups did not differ by the wings color pattern (Figure 6 a,b). The minimal p-distance between these haplogroups was 0.7% when minimal p-distance between haplogroup D and *B. ino* was also 0.7%, and between haplogroup D1 and *B. ino* – 0.6% (Table 1). Therefore, the level of divergence of mitochondrial gene *COI* between undeniable species *B. daphne* and *B. ino* was similar to the level of divergence between haplogroups D and D1.

Table 1. Minimal and maximal p-distances between DNA barcodes of *Brenthis* butterflies (%). From [27].

	B. ino		*B. daphne* (D)		*B. daphne* (D1)		*B. hecate*		*B. mofidii*	
	max	min	max	min	max	min	max	min	max	min
B. ino	1.8	0								
B. daphne (D)	2.5	0.7	1.3	0						
B. daphne (D1)	2.1	0.6	2.4	0.7	1.8	0				
B. hecate	3.1	0.9	4.2	1.8	3.7	1.6	2.3	0		
B. mofidii	2.5	1.5	3.3	2	2.9	1.9	1.6	0.2	1.6	0

As for the other species of the genus, it should be noted that the species *Brenthis ino*, despite its wide distribution, did not demonstrate intraspecific mitochondrial polymorphism. The DNA barcodes of *B. hecate* and *B. mofidii* formed two heterogeneous clusters that did not correlate with the traditional species boundaries [27]. Within *B. hecate* we detected several mitochondrial lineages with a maximum p-distance between them of 2.3% (Table 1).

To clarify the origin and taxonomic status of the identified mitochondrial lines, we performed two additional phylogenetic analyses: 1) using nuclear genes and 2) using morphological characters.

The first additional analysis was performed using four nuclear genes: *carbamoyl-phosphate synthetase 2, aspartate transcarbamylase and dihydroorotase* (*CAD*), *sarco/endoplasmic reticulum*

Ca(2 +)-ATPase (*Ca-ATPase*), *arginine kinase* (*Argk*), and *wingless* (*wg*). Bayesian trees for single nuclear genes had a noncontradictory topology [27], but because single markers had a weak phylogenetic signal, we performed an analysis for the concatenated nuclear matrix. Different alleles in heterozygous states were coded according to the standard IUPAC nucleotide code after inspection of chromatograms for double peaks, haplotypes were estimated in PHASE 2.1.1 software [148] and marked by letters «a» and «b». A comparison of the dendrograms constructed for the nuclear and mitochondrial alignments is shown in Figure 8.

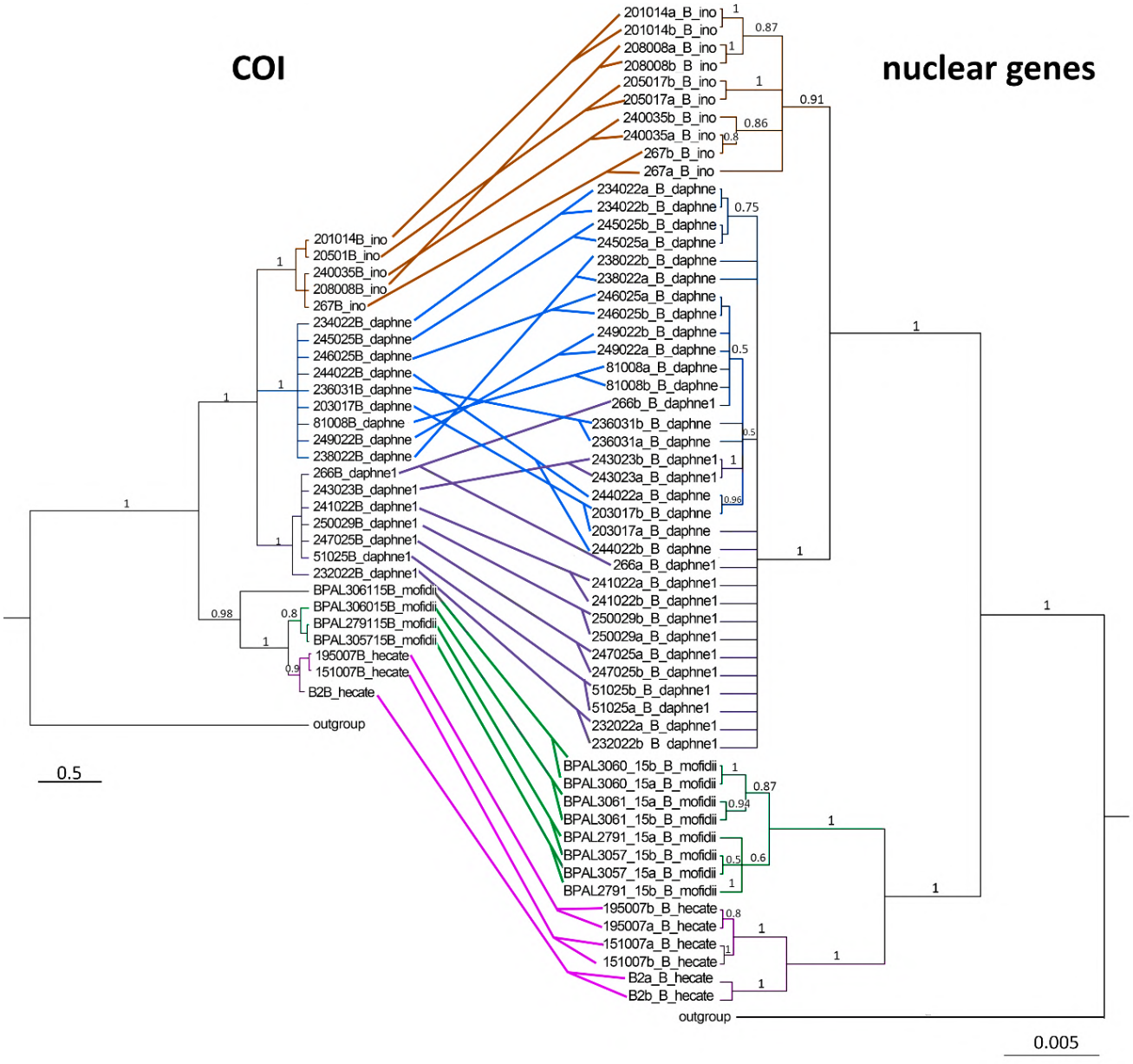

Figure 8. The Bayesian analysis of a fragment of gene *COI* (left) and a set of nuclear genes (right) of *Brenthis* butterflies. Numbers on branches show Bayesian posterior probability. Different nuclear haplotypes of the same individual were marked by the letters «a» and «b». From [27].

According to the results of nuclear marker analysis, the species *B. daphne* and *B. ino* are monophyletic. Samples with mitochondrial haplogroups D and D1 did not form separate lineages on

the nuclear tree. The maximum p-distances for the concatenated nuclear matrix within B. *daphne* were 0.3%, and between *B. daphne* and *B. ino* were in the range of 0.4 - 0.8% (Table 2). *Brenthis mofidii* and *B. hecate* formed separate clades with high support. Specimen BPAL306015 (identified as *B. mofidii*) belonged to the *mofidii* cluster according to the analysis of nuclear markers but formed a basal branch in the group *B. hecate* + *B. mofidii* according to the analysis of the *COI* gene (Figure 8).

Table 2. Minimal and maximal p-distances were calculated for a set of nuclear genes of *Brenthis* butterflies (%). From [27].

	B. ino		*B. daphne*		*B. hecate*		*B. mofidii*	
	max	min	max	min	max	min	max	min
B. ino	0.4	0						
B. daphne	0.8	0.4	0.3	0				
B. hecate	2.8	1.3	3.3	1.3	1.1	0.1		
B. mofidii	2.5	0.9	2.8	1.1	2.8	0.9	1.2	0.1

The second additional analysis included 11 phenotypic traits (for the complete matrix of the trait see [27]). Cladistic analysis of these traits showed, that *B. hecate* + *B. mofidii* and *B. ino* + *B. daphne* are highly supported monophyletic groups (Figure 9). Cladistic analysis of a smaller set of morphological traits has been performed in previous studies [21], but it revealed monophyly of genus *Brenthis* and *B. hecate* + *B. mofidii* species pair, while relationships between the rest of taxa remained unresolved.

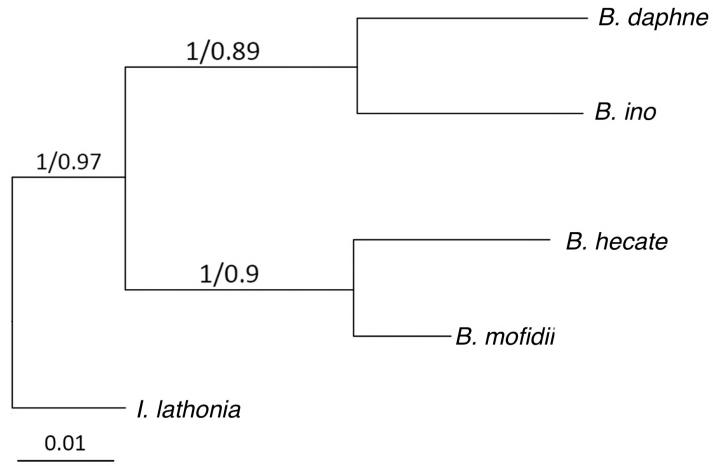

Figure 9. Bayesian tree of *Brenhtis* butterflies based on analysis of phenotypic traits. *Issoria lathonia* was used as an outgroup. The identical topology was revealed with the Maximum Parsimony (MP) method. Numbers on branches show Bayesian posterior probabilities/bootstrap values of MP. From [27].

Thus, analysis of phenotypic traits revealed taxonomic structure, which was fully consistent with the results of the analysis of nuclear markers and showed the monophyly of all four species within the genus *Brenthis,* as well as species pairs *B. hecate* + *B. mofidii* and *B. ino* + *B. daphne* (Figures 8, 9). At the same time, as shown above, the analysis of DNA barcodes revealed a completely different contradictory pattern of distribution of mitochondrial haplogroups and polyphyly of the studied species within the genus (Figure 7).

It is known that such mismatches (mito-nuclear discordance) can be caused by infection with the endosymbiont *Wolbachia*, as well as by mitochondrial introgression, incomplete reproductive isolation, or incomplete lineage sorting[83–85]. In addition, the presence of divergent mitochondrial haplogroups within *B. daphne* may indicate that it represents a complex of cryptic species. This hypothesis can be tested by analyzing the combination of nuclear genes that give better resolution at the species level than the markers we have chosen. However, if mitochondrial clusters correspond to cryptic species, it is expected that they will show attachment to different geographic areas (in allopatry) or ecological niches (in sympatry) [26,149]. Since samples of two haplogroups of *B. daphne* were found in complete sympatry and did not demonstrate either morphological differences or divergence by nuclear genes, we conclude that mitochondrial polymorphism within this taxon is not associated with speciation [27].

In the practice of phylogenetic analysis mitochondrial and nuclear data are often combined into one matrix in order to increase phylogenetic signal. [150–152]. We can assume that this practice may lead to inadequate results when the analysis of mitochondrial and nuclear alignments leads to different topologies. In our study, we tested this method by combining mitochondrial and nuclear matrices for the genus *Brenthis* and performed Bayesian analysis. The mitochondrial *COI* gene has a stronger phylogenetic signal than the combination of nuclear genes (84 and 50 informative positions and alignment lengths of 615 and 1779 bp, respectively), which affected the resulting topology. Two mitochondrial clusters D and D1 were represented on the dendrogram but were combined into a cluster corresponding to *B. daphne* (Figure 10). *B. mofidii* species corresponded to a paraphyletic cluster, which contradicts the results obtained in the analysis of both nuclear and mitochondrial markers. Thus, the chimeric matrix led to results that did not correspond to either mitochondrial or nuclear phylogeny and had no biological sense.

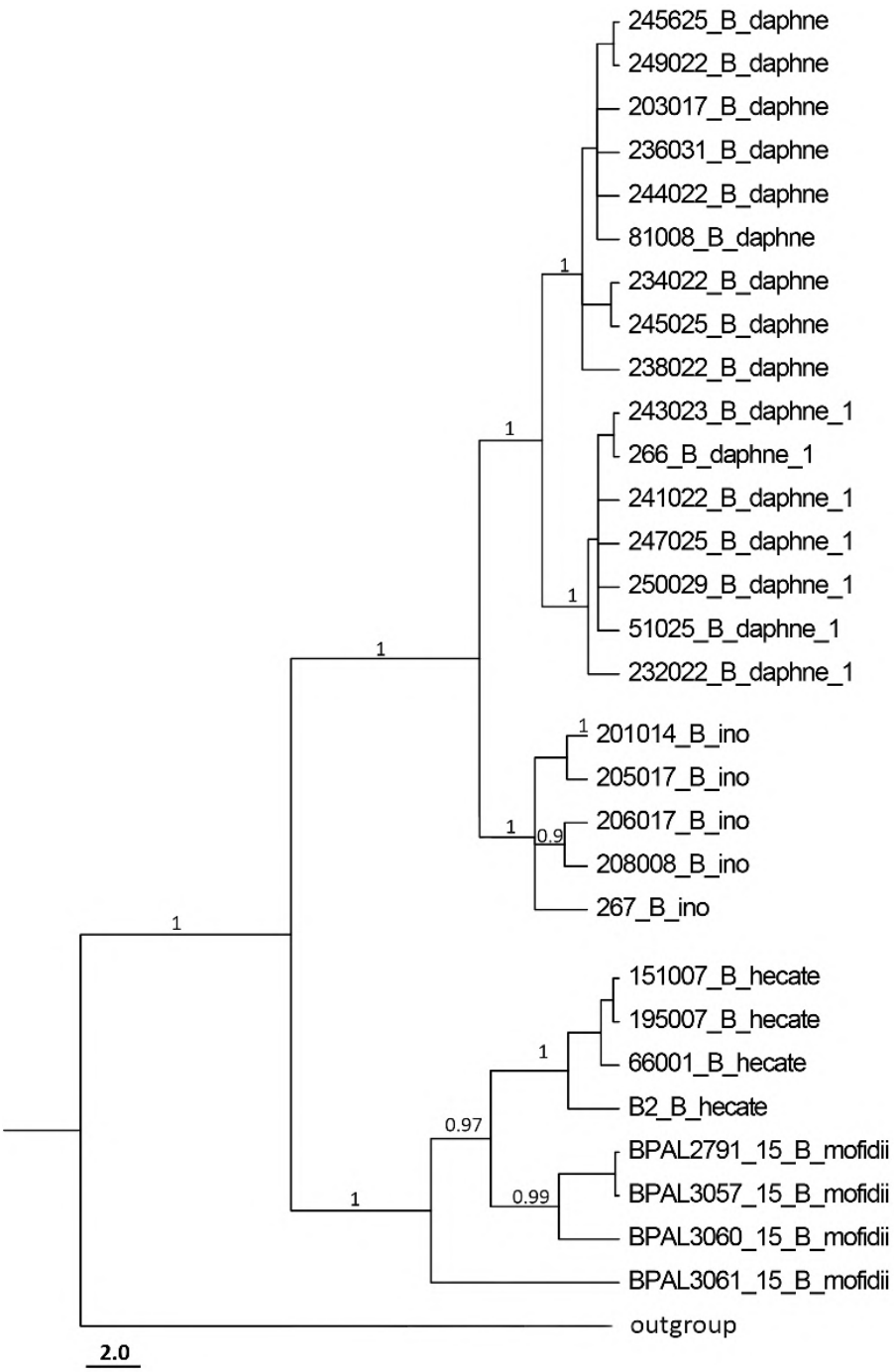

Figure 10. The Bayesian tree based on analysis of the concatenated matrix of *Brenthis* butterflies, ccombining alignments for COI sequences and a set of nuclear markers. Numbers on branches show Bayesian posterior probability. From [27].

In general, for the genus *Brenthis,* the combination of nuclear markers reflects the species boundaries in accordance with the results obtained using the analysis of morphological characters more accurately than the DNA barcodes. [27]. Therefore, we propose to accompany taxonomic hypotheses based on the analysis of single phylogenetic markers with an integrative analysis, as well as to avoid combining mitochondrial and nuclear markers into a single matrix without first checking for contradictions between these data.

3.2 Analysis of the *Hyponephele lycaon – H. lupina* species complex

The genus *Hyponephele* Muschamp, 1915 includes about 40 species, which are morphologically very similar. Interspecific differences in the structure of males genitalia are slightly expressed or absent, females practically cannot be distinguished by the coloration of wings, males can often (but not always) be distinguished by the form of androconial patch shape on forewings [153]. The complex of taxa close to *Hyponephele lycaon* Rottemburg 1775 and *H. lupina* Costa 1863 (Figure 11 a, c) is an extremely difficult group for taxonomic analysis, as evidenced by a large number of contradictory taxonomic hypotheses (see references below) proposed to describe relationships between species of this group. Until the early 20th century, scientists believed that this complex included only one species of *H. lycaon*, represented by several intraspecific forms [154]. It was subsequently discovered, that *Hyponephele lycaon* and *H. lupina* represent different species, not interspecific forms [155].

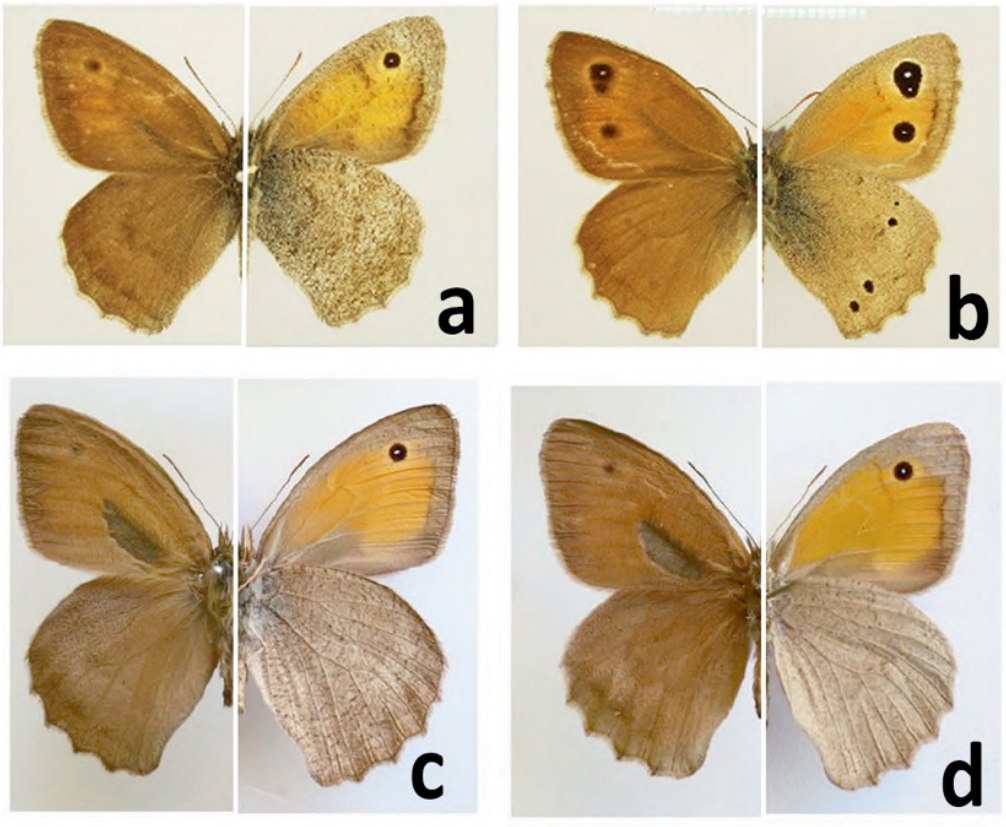

Figure 11. Butterflies of *Hyponephele lycaon – lupina* species complex, upperside (left) and underside (right). **a** *H. lycaon*, **b** *H. sifanica*, **c** *H. lupina*, **d** *H. mauritanica*. From [29].

In the second half of the 20th century, *H. lycaonoides* D. Weiss, 1978 was described from Iran and separated from *H. lycaon* based on supposed sympatry and differences in male genitalia structure [156]. In recent works, the species independence of *H. lycaonoides* has been questioned [157].

In a review of *Hyponephele* species from the territory of the former USSR, *H. lycaon* was divided into 6 species (*H. lycaon*, *H. galtscha* Grum-Grsimailo, 1893, *H. pasimelas* Staudinger, 1886, *H. dzhungarica* Samodurov, 1996, *H. przhewalskyi* Dubatolov, Sergeev & Zhdanko, 1994, *H. lycaonoides*). *H. lycaonoides* is sympatric with *H. lycaon*, and all other species are allopatric [158]. In a review of Palearctic *Hyponephele* species, Eckweiler and Bozano made a compromise classification by assigning most of the populations and described taxa to the *H. lycaon*, but retaining the allopatric *H. maroccana*, *H. sifanica*, *H. galtscha,* and sympatric *H. lycaonoides* as distinct species [153].

In our research, we have tested a multiple allopatric species hypothesis proposed by Samodurov et al. [158] using an analysis of the mitochondrial (COI) and nuclear (*GAPDH*, *Ca-ATPase*, *wingless*) genes, study of morphology and field observation on butterfly ecological preferences, especially in zones of sympatry between putative species.

Reconstruction of phylogeny based on the *COI* barcode fragment analysis using the Bayesian method revealed eight strongly supported monophyletic groups within *H. lycaon* complex (Figures 12, 13). Three of these clades corresponded to *H. maroccana*, *H. galtscha,* and *H. sifanica* species, and the other five represented haplogroups within *H. lycaon* sensu lato: L1 including 63 specimens collected in Eurasia (identified as *H. lycaon* s. str., *H. pasimelas*, *H. dzhungarica,* and *H. przhewalskiy),* L1b including 6 specimens from East Caucasus, including 30 specimens from Israel, Turkey, North Iran, and Azerbaijan, L3 including 60 samples from Israel, Turkey, Southwest Iran and Azerbaijan (comprising *H. lycaon* and *H. lycaonoides)*, and anomalous strongly diverged haplogroup L4 including three samples: one is from Israel and two are from Iran.

Haplogroup L4 was clustered together with *H. sifanica* and together they formed the most basal clade within *lycaon* complex. High level of divergence of L4 and other groups within *H. lycaon* complex (p-distances 9.3 -11.7%) (Table 3) was observed, and this haplogroup possibly represented an ancient isolate of *lycaon* ancestor or a result of fast mitochondrial evolution driven, for example, by *Wolbachia* infection [85].

Table 3. Average p-distances between mitochondrial haplogroups of *Hypionephele lycaon* complex (%).

	L1	L1b	*galtscha*	L2	L3	L4	*sifanica*
L1b	3,1						
galtscha	6,7	5,6					
L2	4	2,9	4,9				
L3	5,9	4,6	4,2	3,5			
L4	11,7	10,8	9,2	9,6	9,3		
sifanica	10,6	9,7	8,4	9,2	6,7	6,1	
lupina	8,4	8,1	7,4	7,7	9,5	11,2	9,4

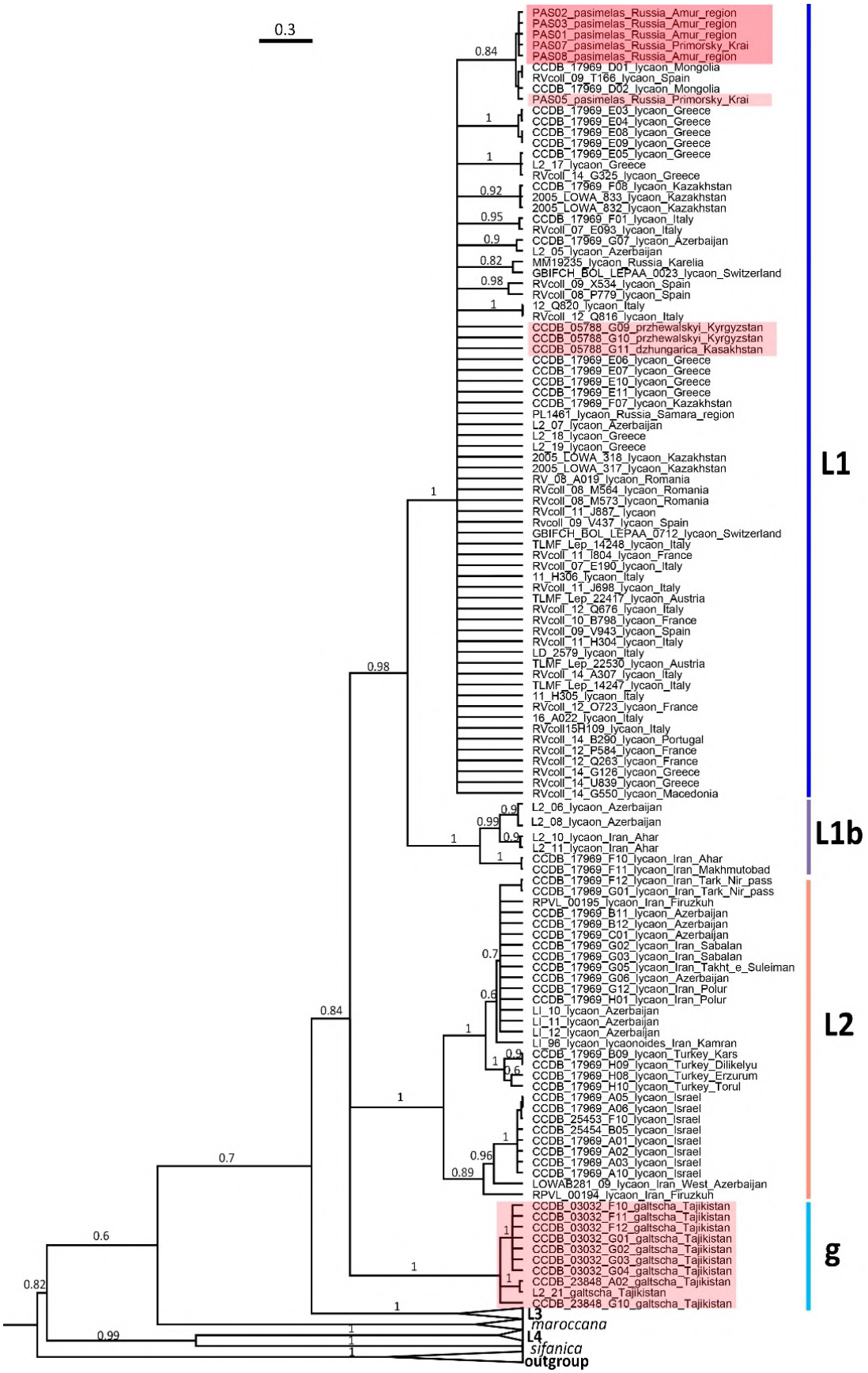

Figure 12. The fragment of the Bayesian tree of *Hyponephele lycaon* complex based on analysis of the *COI* gene. Haplogroups L1, L1b, L2, and *H. galtscha* species are shown. Numbers at branches indicate Bayesian posterior probabilities. The samples identified as *Hyponephele pasimelas, H. przhewalskyi,* and *H. dzhungarica* according to multiple allopatric species concept, are marked red.

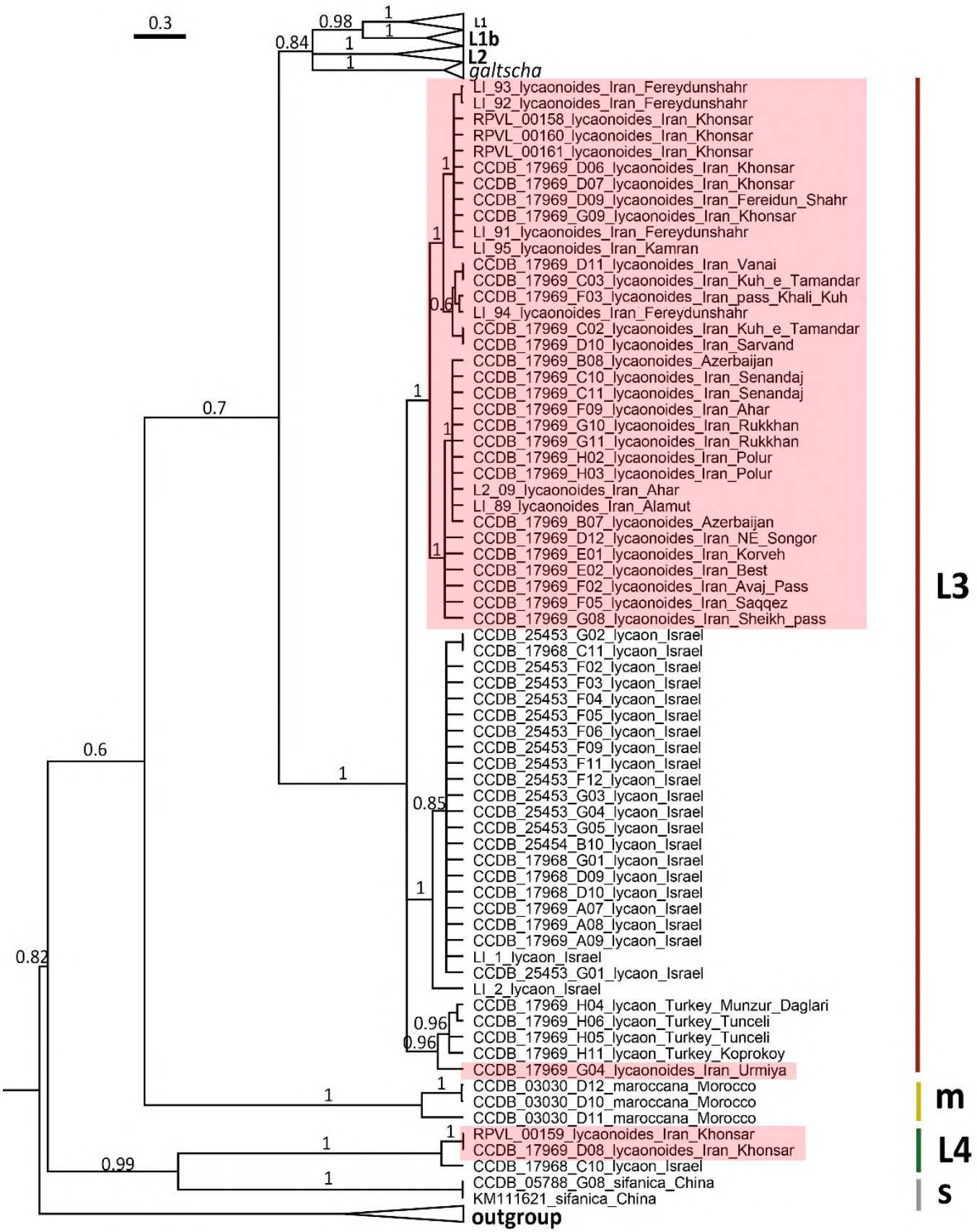

Figure 13. The fragment of the Bayesian tree of *Hyponephele lycaon* complex based on analysis of the *COI* gene. Haplogroups L3, L4 and species *H. maroccana* and *H. sifanica* are shown. Numbers at branches indicate Bayesian posterior probabilities. Specimens, identified as *H. lycaonoides*, are marked red.

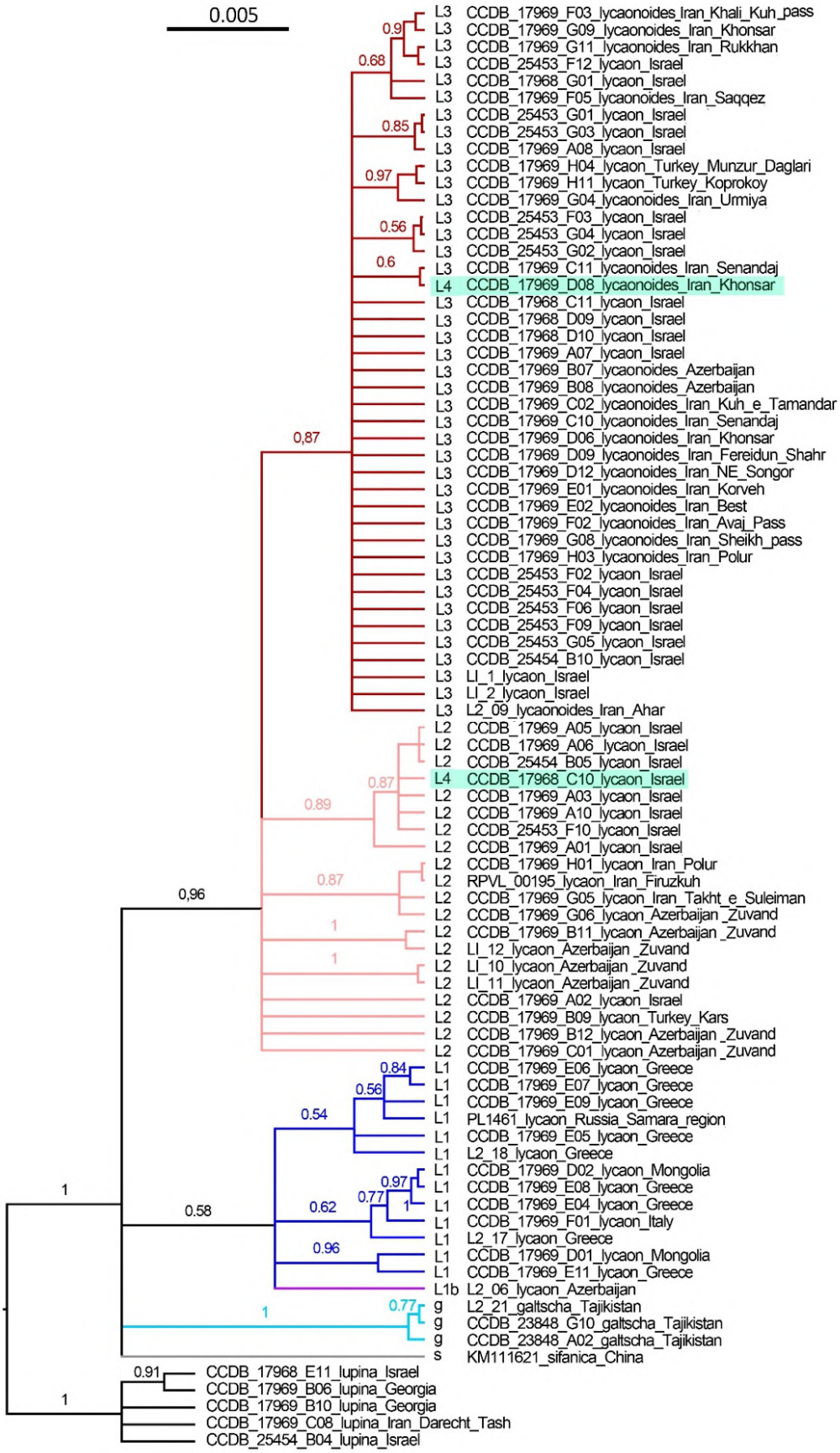

Figure 14. Bayesian tree of *Hyponephele lycaon* complex based on analysis of the concatenated matrix of nuclear genes *GAPDH*, *Ca-ATPase,* and *wingless*. Numbers at branches indicate Bayesian posterior probabilities. Samples of mitochondrial haplogroup L4 are marked green.

We applied geometric morphometrics to analyze male genitalia structure (Figure 15). Specimens of haplogroup L3 have longer uncus and shortened subunci were clearly separated from clusters L1 and L2, which considerably overlapped.

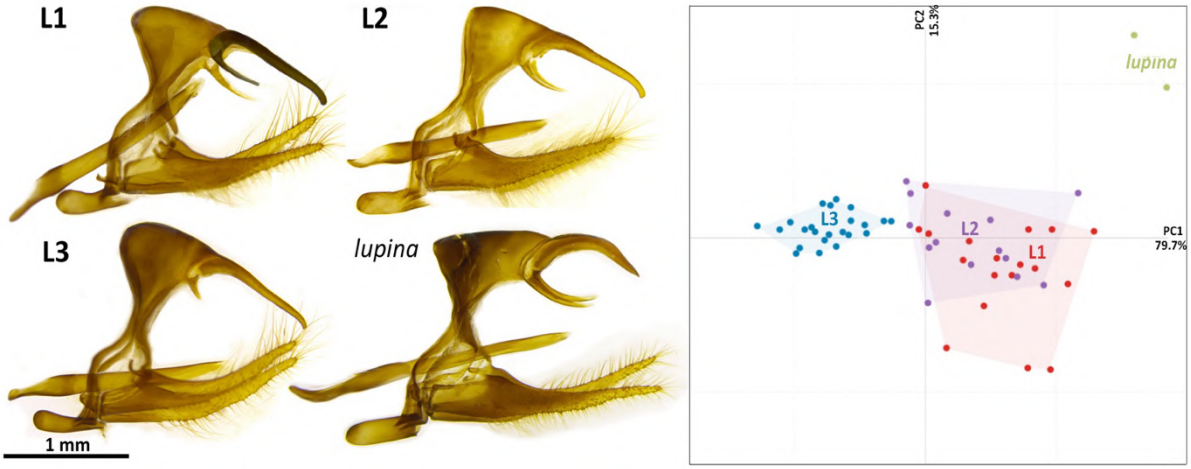

Figure 15. Structure and geometric morphometric of the male genitalia of *Hyponephele* butterflies. Male genitalia of *H. lycaon* with mitochondrial haplogroups L1, L2 and L3 as well as of *H. lupina* are shown on the left. An example of the uncus and subunci contour used for morphometrics is shown by transparent black color on the L1 haplogroup. PCA results are shown on the right. The first two principal components are shown (PC1 and PC2), explaining 79.9% and 15.3% of the variance of shapes, respectively.

Thus, according to the obtained data, butterflies of the L3 haplogroup have an extensive range including the territory of Iran, Turkey, Azerbaijan, and Israel, and are characterized by (1) a unique mitochondrial haplogroup, (2) unique nucleotide substitutions in nuclear genes, and (3) a unique structure of males genitalia. In our opinion, the congruence of these three, unrelated groups of characters is a piece of evidence that butterflies of haplogroup L3 represent a separate species, which occurs in sympatry within almost its entire range with another species represented by haplogroups L1 and L2 which we assign to *H. lycaon* sensu stricto. (see discussion below). It is important to emphasize that the species with haplogroup L3 is not identical to the *H. lycaonoides* and includes both individuals identified as *H. lycaonoides* and *as H. lycaon* (see Figure 13). Visual inspection shows that butterflies of haplogroup L3 usually possess a larger wing size and a lighter hind wing margin than butterflies of haplogroups L1 and L2 (although these characters are variable and hardly can be used for species identification).

The next question to be discussed is a valid name for the species with haplogroup L3. A large number of taxa from the *H. lycaon* complex have been described from Turkey, Iran, and the Levant, each of which may represent both *H. lycaon* (haplogroups L1 + L2) and a previously unrecognized species with haplogroup L3. Hesselbart et al. (1995) suggested to use the name *Hyponephele lycaon var. collina* Röber, 1897 for butterflies with larger wing size, since this was the oldest available

scientific name for butterflies from Turkey and generally from West Asia. These authors considered *H. lycaon libanotica* Staudinger, 1901 (the Levant) and *H. lycaonoides* D. Weiss, 1978 (Iran) as synonyms of *H. lycaon var. collina* Röber, 1897 [146]. Unfortunately, the type material of *Hyponephele lycaon var. collina* Röber, 1897 was lost [146]. We believe that the correct solution to the nomenclatural question in the future would be to designate the neotype *of H. lycaon var. collina* from the type locality (southeastern Turkey, the vicinity of Gülek locality), with a male which has the molecular and morphological features of haplogroup L3. Therefore, we will use the name *H. collina* for the species with haplogroup L3 below.

Regarding butterflies of mitochondrial haplogroups L1 and L2, we see no reason to assign them to different species, because they are morphologically indistinguishable and not differentiated by the studied nuclear genes. These two haplogroups are allopatric and probably represent two lineages of the *H. lycaon* species that diverged in geographic isolation. In our opinion, these two lineages can be interpreted as subspecies. The subspecies *H. lycaon lycaon* (haplogroup L1) has an extensive range from Western Europe to the Far East, including the Caucasus and northern Central Asia. Since within *H. lycaon lycaon* populations are weakly differentiated by the *COI* gene (the average intragroup p-distances by the mitochondrial COI genetic marker for haplogroup L1 are 0.7%), and the nuclear markers are monomorphic, we cannot confirm the hypothesis of Samodurov et al. [158] about allopatric species *H. pasimelas* (Russian Far East), *H. dzhungarica* (Dzhungarian Alatau), and *H. przhewalskiy* (Tian Shan).

We propose to use the name *Hyponephele lycaon zuvandica* Samodurow & Koroljev, 1996 for the subspecies with mitochondrial haplogroup L2 because molecular data were obtained for samples from the type locality of this taxon (Azerbaijan, Talysh). This subspecies is widespread in West Asia (including Turkey and Israel). In Iran, *H. lycaon zuvandica* can be found in the western part of that country, often sympatrically with *H. collina*, but it is absent in the southern Zagros, where *H. collina* is present.

During our field expeditions, we studied the biotopic confinement and phenology of the *H. lycaon-H. lupina* complex. Special attention was paid to *H. lycaon* and *H. collina* in regions of their sympatry. The phenology of these species coincides, but in the zone of their complete sympatry, they differ in biotopic preferences. According to our observations, *H. collina* is a more xerophilic species inhabiting dry steppe slopes, while *H. lycaon* is a more mesophilic species preferring grasslands and forest meadows. *Hyponephele lycaon* was found at higher altitudes than *H. collina* in Hakkari province of Turkey [153], which is confirmed by our observations in Iran. However, in Israel, the opposite pattern is observed on Mount Hermon. Presumably, the choice of habitat is based not on the altitude but the humidity of the biotope. In the drier areas of Iran and eastern Turkey, *H. lycaon* flies in the

upper zone of the mountains closer to the subalpine ecoregion, which is more humid in these areas. In Israel, *H. lycaon* was detected in the forest region (1400-1600 m), which is wetter than the subalpine ecoregion above. Thus, the differences in ecological preferences are additional evidence in favor of the species independence of these taxa.

We have also performed phylogenetic and taxonomic analysis of the entire *H. lycaon – H. lupina* complex [29]. Within the group *Hyponephele lupina* our analysis of DNA barcodes revealed clusters corresponding to taxa *H. lupina lupina*, *H. lupina mauritanica* Oberthür, 1881, *H. maroccana* Blachier, 1908, *H. sifanica*, *H. cheena* Moore 1865, and *H. interposita* Erschoff, 1874 [29]. The level of divergence between *H. lupina* and *H. mauritanica* varied in a range of 3,5-4,9%, which was higher than the "standard" level of interspecies divergence for the *COI* gene [50]. Specimens of the *lupina* and *mauritanica* clades also differed in the color of the underside of the wing of males, the shape of androconial patches, and the structure of the male genitalia, thus, we proposed to consider *H. mauritanica* as an independent species [29]. However, to confirm the status of *H. mauritanica*, analysis of additional molecular markers is required.

Thus, using the example of the *H. lycaon* - *H. lupina* complex, we demonstrated that the primary taxonomic hypotheses based on the analysis of mitochondrial DNA barcodes should be thoroughly tested using an integrative approach, including the study of nuclear genes and morphology. At the same time, the resulting classifications may differ significantly both from preliminary taxonomic hypotheses based on DNA barcodes and from taxonomic interpretations that can be found in previously published revisions.

Chapter 4. Analysis of whole genomes reveals the taxonomic structure and evolutionary history of the *Melitaea acentria-M. persea- M. didyma* species complex

4.1. Patterns of genetic differentiation and genomic introgression

Reticulation events, including interspecific hybridization and horizontal gene transfer, and also incomplete lineage sorting, occur in nature more frequently than it was previously thought and cause difficulties in the determination of species boundaries [159,160]. The use of a large amount of data that can be obtained from the analysis of complete genomes makes it possible to study not only the degree of diversification of taxa but also their evolutionary history, which sometimes is necessary to formulate a correct taxonomic hypothesis.

In butterflies of the genus *Melitaea*, individuals with mixed morphology were observed in the contact zone of the two closely related species, which could arise as a result of interspecific hybridization [20]. Sometimes such individuals form entire populations, which have been described for *M. didyma* and *M. persea* in Turkey and Lebanon, for *M. persea* and *M. mixta* in Afghanistan and Pakistan, *M. athalia* and *M. nevadensis* in Switzerland [20].

Species *Melitaea acentria* Lukhtanov, 2017, distributed in Israel and Lebanon, has mixed traits of the male genitalia structure and the wing color: on the one hand, M. acentria demonstrates traits of an unrelated sympatric species *M. didyma* Esper, 1777, and on the other hand, it is morphologically similar to a closely related allopatric species *M. persea* Kolar, 1849. Therefore, it was assumed that *M. acentria* originated as a result of interspecific hybridization [7], however, this hypothesis has not been tested using nuclear markers previously.

In our study, sequencing of 27 whole genomes of *M. acentria, M. didyma,* and *M. persea* was performed to study the genetic structure and reveal the relationships between these three taxa [30]. We have analyzed the whole genomes and obtained a set of SNPs using the genome of *Melitaea cinxia* as a reference [38]. Analysis of the SNP set using the Principal component analysis (PCA) revealed non-overlapping species-specific clusters, with one sample (Mel11) occupying an intermediate position between them (Figure 16).

The phylogenetic network constructed in the Splitstree software showed that the sample Mel11 may have a hybrid origin (*M. acentria* x *M. didyma*) (Figure 17). The rest of the samples formed species-specific clusters with high bootstrap supports. In the first-generation hybrid, homozygous substitutions species-specific for *M. didyma* and *M. acentria* are expected to be in a heterozygous state. Of 104,760 species-specific SNPs in the Mel11 sample, only 42% were heterozygous, 28% of the rest were related to *M. didyma* and 30% - to *M. acentria*. Therefore, specimen Mel11 emerged from backcrossing of the first-generation hybrid with the parental species.

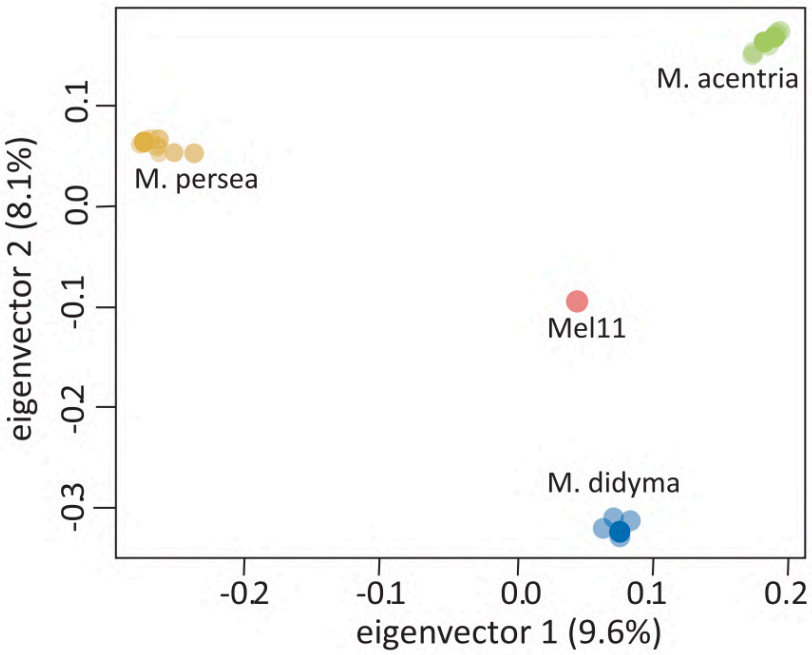

Figure 16. Principle component analysis of the whole-genome SNP. Species-specific clusters are shown by color. Sample Mel11 is marked red. The axes show the first two eigenvectors explaining 9.6% and 8.1% of the variance of genomic SNPs. From [30].

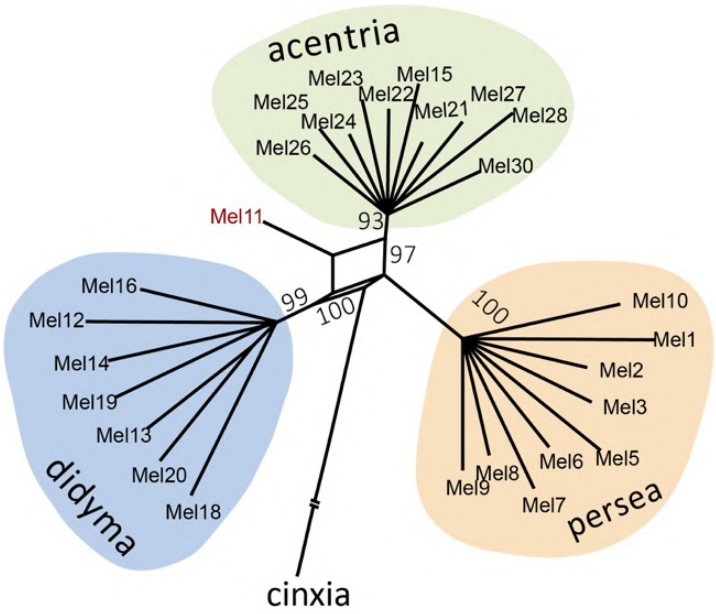

Figure 17. Phylogenetic network calculated using SplitsTree software for the whole-genome SNP. Species-specific clusters are shown by color. Hybrid sample Mel11 is marked red and excluded from both *M. acentria* and *M. didyma* clusters. Numbers on branches represent bootstrap values. From [30].

The finding of a hybrid between *M. didyma* and *M. acentria* indicates incomplete reproductive isolation, which was surprising because *M. didyma* and *M. acentria* are not closely related species and belong to different species groups within the genus *Melitaea*. We performed additional analysis to detect a possible gene flow between these species. The Mel11 sample was excluded from further analysis to minimize the possibility of artifacts.

We used ABBA-BABA test to detect signs of genomic introgression [161,162]. ABBA-BABA test detects deviations from the strictly dichotomic model of the species divergence by counting the number of genome regions with topologies that do not correspond to the original dichotomic model. Such deviations are expressed through a metric called the D-statistics. D-statistics was calculated for all three theoretically possible rooted topologies (*M. cinxia* was used as an outgroup). The topology (((*persea*, *acentria*) *didyma*) *cinxia*) had the highest likelihood; and it had an average D = 0.21 ± 0.005 and Z-score, calculated to assess the significance of D-statistics, was 18.206. Since Z-score is considered significant if it is bigger than 3 [162], we can assume an interspecific introgression between *M. didyma* and *M. acentria*. The proportion of introgression f_d [163] was calculated separately for each chromosome (Figure 18).

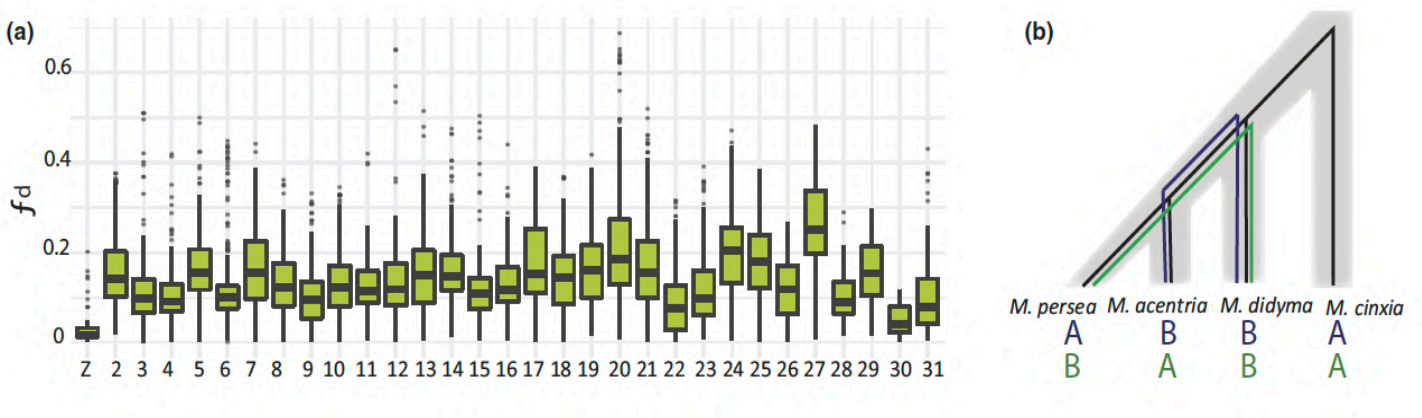

Figure 18. The average proportion of introgression f_d, calculated separately for each chromosome. The topology used for f_d estimation is shown on the right side. From [30].

An average proportion of introgression f_d between *M. acentria* and *M. didyma* was f_d = 0.13 ± 0.035. An average proportion of introgression in sex Z chromosomes was much lower than in autosomes (f_{d_Z} = 0.05 ± 0.008, f_{d_auto} = 0.14 ± 0.09). At the same time, fixation index F_{ST} was approximately the same along the whole genome between closely related species *M. acentria* and *M. persea*, but both of them had higher divergence with *M. didyma* in sex chromosomes than in autosomes. In other words, the difference in the level of divergence between sex chromosomes and autosomes was high in *M. acentria* -*M. didyma* (ΔF_{ST} = 0,44) and *M. persea* - *M. didyma* (ΔF_{ST} = 0,39) species pairs, but relatively low in *M. acentria* - *M. persea* (ΔF_{ST} = 0,1), where $\Delta F_{ST} = F_{ST_Z} - F_{ST_auto}$. Thus, sex chromosomes and autosomes accumulate interspecific differences unevenly during evolution – closely related species have rather similar F_{ST} in sex chromosomes and autosomes, when in distantly related species F_{ST} significantly increased on sex chromosomes (Figure 19). Therefore, the different rate of accumulation of substitutions may lead to faster formation of barriers to introgression on sex chromosomes than on autosomes [30].

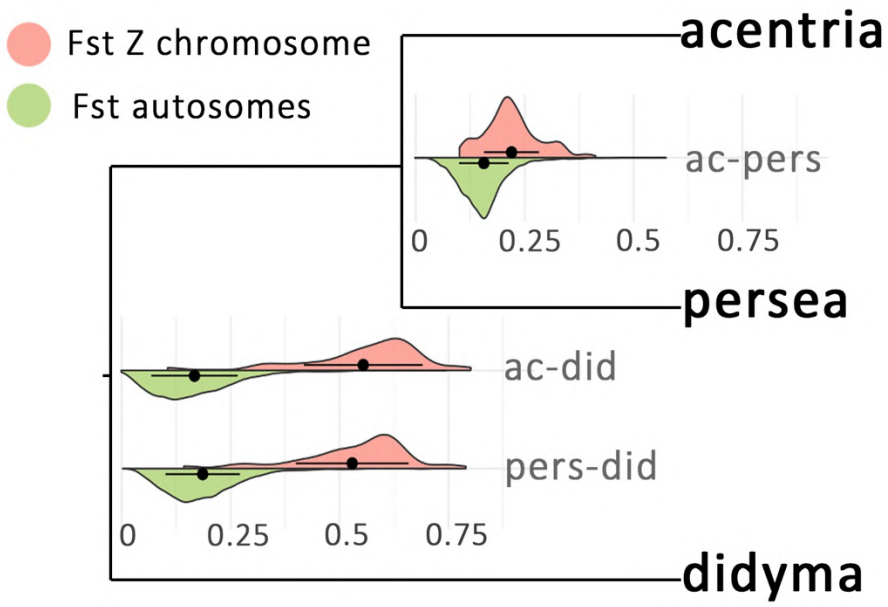

Figure 19. Level of divergence in sex chromosomes and autosomes. The fixation index F_{ST} was calculated between all the studied species pairs separately for sex chromosomes and autosomes. From [30].

We calculated nucleotide diversity (π) among the genomes, and found out, that genomic regions with the highest π also demonstrated a high proportion of introgression (Figure 20). Also, we demonstrated that chromosomes with low average f_d also had low average nucleotide diversity [30].

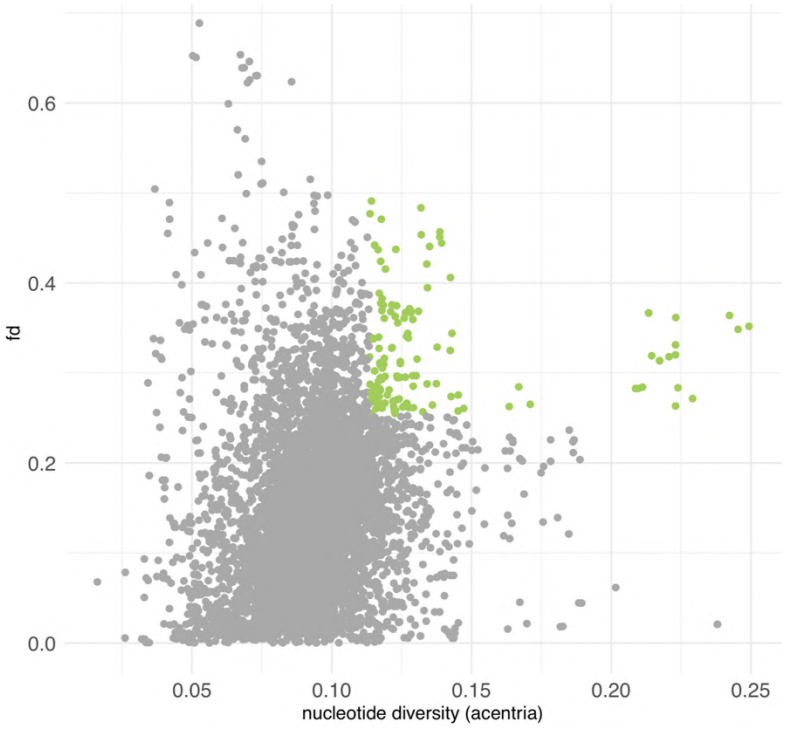

Figure 20. Comparison of nucleotide diversity (π) of *Melitaea acentria* and proportion of introgression (f_d) from *M. didyma* to *M. acentria*. 10% outliers of π and f_d, are marked green. From [30].

Analysis of possible demographic scenarios of the *M. didyma-M. persea-M. acentria* complex using site-frequency spectra revealed, that the highest value of likelihood had a scenario assuming a unidirectional gene flow from *M. didyma* to *M. acentria* after the divergence of *M. persea* and *M. acentria* (Figure 21). According to the parameters of this model, *M. didyma* separated 10.7 Mya, while *M. persea* and *M. acentria* diverged 7.27 Mya, assuming that the studied species are monovoltine. If we assume that these species are bivoltine, then the time of divergence of *M. didyma* was estimated as 5.3 Mya, and the time of divergence of *M. persea* and *M. acentria* - as 3.6 Mya. According to literature, the studied species can be both monovoltine and bivoltine, depending on the habitat [146]. Hence, the obtained estimates of ages can be considered as a minimum and maximum.

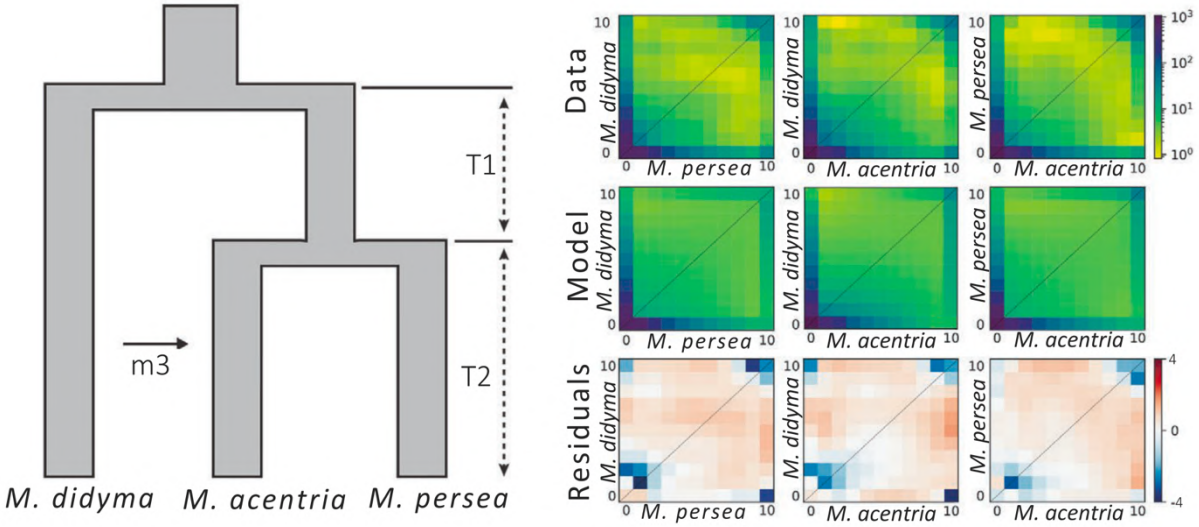

Figure 21. Reconstruction of demographic history using analysis of site frequency spectra. The scheme of the most probable model is shown on the left. Parameter m3 represents a unidirectional gene flow from *M. didyma* to *M. acentria*, parameter T1 is a time of divergence of *M. didyma* (5,3-10,7 Mya), T2 is a time of divergence of *M. acentria* and *M. persea* (3.6-7.27 Mya). The pairwise observed and expected allele frequency spectra and distributions of residues (difference between expected and observed spectra) are shown on the right. From [30].

4.2 *Melitaea acentria*, *M. persea* and *M. didyma* represent different species

The basic question in species delimitation is a definition of species. In this study we used a unified species concept that defines species as independent evolutionary lineages, using as species properties the key criteria of other species concepts (biological, diagnostic, genealogical, ecological), such as reproductive isolation from other lineages, morphological differentiation, inhabiting of separate ecological niches, monophyly, and others [14].

Let us consider this species complex from the view point of **biological species concept**, the main criterion of which is the presence of reproductive isolation [138]. *M. persea* and *M. acentria* are completely allopatric, so the hypothesis of the presence of reproductive isolation between them cannot

be tested in nature. Previously was reported that *Melitaea persea* and *M. didyma* sporadically hybridize in Turkey [20], but this data was not tested using molecular data, and our whole-genome analysis did not reveal any signs of introgression between these species. Despite a gene flow between *M. didyma* and *M. acentria*, analysis of whole-genome SNP showed, that all the three species are genetically distinct due to isolating mechanisms that prevent them from merging into a single evolutionary unit. One of such mechanisms, maintaining reproductive isolation within *M. persea -M. didyma - M. acentria* species complex can be a faster divergence of sex chromosomes compared to the rest of the genome.

Species *M. persea, M. didyma,* and *M. acentria* are morphologically distinct [7,20], therefore, they correspond to the criteria of the **diagnostic species concept** [14]. Phylogenetic reconstruction using wholes showed, that specimens of each three studied species formed monophyletic units [30], which is a key criterion of the **genealogical species concept** [164]. Also, the species occupy different ecological niches. *M. persea* lives in open dry biotopes and in cases of sympatry usually prefers lower altitudes than *M. didyma* [146]. Unlike *M. persea*, *M. inhabits* subalpine habitats and is usually found higher than *M. didyma* [7]. Thus, they represent species from the viewpoint of the **ecological species concept** [165,166]. Hence, *M. persea, M. didyma,* and *M. acentria* meet at least three species criteria, which allows us to consider them as different species.

4.3 Speciation of *M. acentria*

Melitaea acentria and *M. persea* are completely allopatric, and ancestral area reconstruction, based on phylogeographic analysis, showed, that the most probable ancestral distribution area for *M. acentria* was the Levant, whereas for *M. persea* it was Turkey, Iran, and Caucasus [30]. Presumably, species *M. acentria* originated in the Levant as a result of geographic isolation from *M. persea* about 3.6-7.27 Mya. This period corresponds to the Late (the Messinian age, 7,2 – 5,3 Mya), which was supposed to be a key period in the Northern Africa colonization by *Melitaea* butterflies due to its geological and climatic conditions, which made possible terrestrial migrations [136].

Analysis of whole mitochondrial genomes and DNA barcodes revealed two differentiated mitochondrial haplogroups within *M. acentria* (*acentria1* and *acentria2),* which originated from *M. persea* and are distantly related to *M. didyma* [30]. Mitochondrial haplogroups do not correlate either with the clusters identified based on the analysis of complete genomes or with morphological differences and represent a case of intraspecific polymorphism. The presence of the haplogroups more likely occurs due to two independent invasions of *M. persea* to the Levant with their subsequent fusion during climate fluctuations in the Quaternary, rather than due to mitochondrial introgression between

M. persea and *M. acentria*, since the analysis of complete genomes did not confirm the gene flow between these species [30].

Melitaea didyma and *M. acentria* are distantly related –analysis of allele frequency spectra predicted the separation of branches *persea* + *acentria* and *didyma* of 5.3-10.7 Mya. Although hybridization occurs more often between closely related taxa, we have detected introgression from *M. didyma* to *M. acentria*. Introgression is thought to be more likely to occur from local to invasive species [167,168], because the second has a significantly smaller population size. Thus, under specific geographic and demographic conditions, leading to a weakening of interspecific reproductive barriers, introgression can occur even between non-closely related species. Interspecies hybridization can be an additional source of genetic diversity, especially in small populations with low levels of genetic diversity. Our study showed that *M. acentria* has the lowest average nucleotide diversity among the studied species [30]. Among the regions of the genome with the highest proportion of introgression from *M. didyma* to *M. acentria*, the highest level of nucleotide diversity was also detected. This suggests that the gene flow from the Trans-Palaearctic species *M. didyma* was the source of genetic diversity in the Levant endemic *M. acentria*. Since sporadic hybridization between *M. acentria* and *M. didyma* still occurs, in this case, the analysis of individual molecular markers would rather reflect the phylogeny of the gene than the phylogeny of the species, which underlines the importance of using whole-genome data to analyze groups with reticular evolution.

Conclusion

The correct delimitation of species is a key problem of taxonomy. The development of a methodology for determining the boundaries between species requires both adapting the species concepts to modern realities and studying the applicability of various types of data for taxonomic analysis. The wide availability of molecular data provides, on the one hand, the possibility of detecting hidden biological diversity, but on the other hand, increases the risk of erroneous taxonomic conclusions based on only one type of characters. In this study, we have demonstrated the applicability of DNA barcodes for formulating primary taxonomic hypotheses. We showed, that *Melitaea ala* includes three subspecies (*M. ala ala*, *M. ala zaisana, M. ala bicolor)*, which demonstrate morphological and geographical differences but have the level of divergence of *COI* gene lower than the standard species-specific threshold. We have detected 23 deeply divergent mitochondrial genetic lines within the *Melitaea didyma* complex. Twelve of them correspond to the previously described species, 7 - to subspecies of *M. didyma*, and the rest represents cases of mitochondrial polymorphism.

The inconsistency of topologies of phylogenetic trees based on mitochondrial and nuclear markers was demonstrated in the butterflies of genus *Brenthis*, for which two mitochondrial haplogroups were shown within the species *B. daphne*. Analysis of nuclear genes and morphological traits did not support the hypothesis that the haplogroups represent different species. It is important to emphasize that combining contractionary data (in this case, nuclear and mitochondrial genes) into a single matrix and their joint analysis leads to an erroneous phylogenetic reconstruction, and we recommend avoiding such an approach in practice.

Primary analysis of *Hyponephele lycaon- H. lupina* species complex also revealed mitochondrial lineages within *H. lycaon*, which did not correspond to the real taxa. Integrative analysis, including analysis of DNA barcode, nuclear genes, geographical distributions, wing color patterns, and morphometry of male genitalia structure, confirmed that *H. lycaon, H. maroccana*, *H. sifanica, H. mauritanica, H. lupina, H. galtscha,* and *H. collina* represent a separate species*,* and *H. lycaon zuvandica* represents a subspecies. Taxa *H. pasimelas*, *H. dzhungarica,* and *H. przhewalskiy* are indistinguishable from *H. lycaon,* so multiple allopatric species hypothesis was not supported.

Sequencing and following analysis of whole genomes of *Melitaea didyma – M. persea – M. acentria* complex confirmed, that they are different species despite incomplete reproductive isolation between *M. didyma* and *M. acentria*. We showed, *M. acentria* originated in the Levant as a peripheral geographic isolate of *M. persea*. Gene flow from the not closely related species *M. didyma* was a source of genetic diversity in *M. acentria* population, which determined the evolutionary trajectory of this species. Sex chromosomes are more resistant to introgression than autosomes and play a role in the formation and maintenance of postzygotic interspecific reproductive barriers.

Thus, the dissertation demonstrates the applicability of molecular data analysis of all levels (from single markers to whole genomes) for the analysis of well- and least-studied groups of Nymphalidae. In the future, the development of the methodology of integrative analysis in taxonomy can be continued both in the study of other model groups of living organisms and in the study of new types of data. In the future, the development of a methodology for integrative analysis in taxonomy can be continued both in the study of other model groups of living organisms and in the study of new types of data.

Acknowledgments

In conclusion, I would like to thank my scientific advisor V.A.Lukhtanov for his help and support at all stages of work on the dissertation. I am deeply grateful to the staff of the Karyosystematics Department of the Zoological Institute of the Russian Academy of Sciences and the Department of Entomology of St. Petersburg State University for the discussion of the work and assistance in its implementation. I thank I. A. Makhov, A. V. Barabanov, A. V. Novikova, V. V. Tikhonov, N. A. Shapoval, D. F. Shovkun for sharing the collected material. I would like to express my gratitude to Anastasia Gagarina for her help in preparing samples for sequencing in the lab. I would like to express my gratitude to the staff of the Resource Center of St. Petersburg State University "Development of Molecular and Cellular Technologies" A.E. Masharsky and E. Yu. Gorodilova and the staff of the Resource Center " Computer Center SPbU " A. Andreev and D. Khmel for technical support of the project.

This work was supported by RSF grants № 14-14-00541, 19-14-00202; RFBR grant № 19-34-90007.

References

1. Dayrat, B. Towards integrative taxonomy // Biological Journal of the Linnean Society. - 2005. - Vol. 85, № 3. - P. 407–415.
2. Will, K.W., Mishler, B.D., Wheeler, Q.D. The perils of DNA barcoding and the need for integrative taxonomy // Systematic Biology2005. - Vol. 54, № 5. - P. 844–851.
3. Hillis, D.M. Molecular Versus Morphological Approaches to Systematics // Annual Review of Ecology and Systematics. - 1987. - Vol. 18, № 1. - P. 23–42.
4. Hogan, K.M., Hedin, M.C., Koh Hung Sun, et al. Systematic and taxonomic implications of karyotypic, electrophoretic, and mitochondrial-DNA variation in Peromyscus from the Pacific Northwest // Journal of Mammalogy. - 1993. - Vol. 74, № 4. - P. 819–831.
5. DeSalle, R., Egan, M.G., Siddall, M. The unholy trinity: Taxonomy, species delimitation and DNA barcoding // Philosophical Transactions of the Royal Society B: Biological Sciences. - 2005. - Vol. 360, № 1462. - P. 1905–1916.
6. Heethoff, M., Laumann, M., Weigmann, G., et al. Integrative taxonomy: Combining morphological, molecular and chemical data for species delineation in the parthenogenetic Trhypochthonius tectorum complex (Acari, Oribatida, Trhypochthoniidae) // Frontiers in Zoology. - 2011. - Vol. 8, № 1. - P. 1–10.
7. Lukhtanov, V.A. A new species of Melitaea from Israel, with notes on taxonomy, cytogenetics, phylogeography and interspecific hybridization in the Melitaea persea complex (Lepidoptera, Nymphalidae) // Comparative Cytogenetics. - 2017. - Vol. 11, № 2. - P. 325–357.
8. Dincă, V., Wiklund, C., Lukhtanov, V.A., et al. Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex // Journal of Evolutionary Biology. - 2013. - Vol. 26, № 10. - P. 2095–2106.
9. Avise, J.C. Phylogeography: the history and formation of species. - Cambridge, Massachusetts: Harvard University Press, 2000. - 447 p.
10. Schander, C., Willassen, E. What can biological barcoding do for marine biology? // Marine Biology Research. - 2005. - Vol. 1, № 1. - P. 79–83.
11. Condon, M.A., Scheffer, S.J., Lewis, M.L., et al. Hidden neotropical diversity: Greater than the sum of its parts // Science. - 2008. - Vol. 320, № 5878. - P. 928–931.
12. Hebert, P.D.N., Penton, E.H., Burns, J.M., et al. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator // Proceedings of the National Academy of Sciences of the United States of America. - 2004. - Vol. 101, № 41. - P. 14812–14817.
13. Cardoso, A., Serrano, A., Vogler, A.P. Morphological and molecular variation in tiger beetles of

the Cicindela hybrida complex: Is an “integrative taxonomy” possible? // Molecular Ecology. - 2009. - Vol. 18, № 4. - P. 648–664.

14. De Queiroz, K. Species Concepts and Species Delimitation // Systematic Biology. - 2007. - Vol. 56, № 6. - P. 879–886.
15. Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., et al. Integrative taxonomy: A multisource approach to exploring biodiversity // Annual Review of Entomology2010. - Vol. 55P. 421–438.
16. Hinojosa, J.C., Koubínová, D., Szenteczki, M.A., et al. A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly Thymelicus sylvestris // Molecular Ecology. - 2019. - Vol. 28, № 17. - P. 3857–3868.
17. Dincă, V., Lee, K.M., Vila, R., et al. The conundrum of species delimitation: A genomic perspective on a mitogenetically super-variable butterfly // Proceedings of the Royal Society B: Biological Sciences. - 2019. - Vol. 286, № 1911.
18. Pante, E., Schoelinck, C., Puillandre, N. From Integrative Taxonomy to Species Description: One Step Beyond // Systematic Biology. - 2015. - Vol. 64, № 1. - P. 152–160.
19. Dincă, V., Montagud, S., Talavera, G., et al. DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity // Scientific Reports 2015 5:1. - 2015. - Vol. 5, № 1. - P. 1–12.
20. van Oorschot, H., Coutsis, J. The genus Melitaea Fabricius, 1807 : taxonomy and systematics with special reference to the male genitalia : (Lepidoptera, Nymphalidae, Nymphalinae). - Pardubice: Tshikolovets Publications, 2014. - 360 p.
21. Simonsen, T.J., Wahlberg, N., Brower, A.V.Z., et al. Morphology, molecules and fritillaries: Approaching a stable phylogeny for Argynnini (Lepidoptera: Nymphalidae) // Insect Systematics and Evolution. - 2006. - Vol. 37, № 4. - P. 405–418.
22. Simonsen, T.J. Fritillary phylogeny, classification, and larval host plants: reconstructed mainly on the basis of male and female genitalic morphology (Lepidoptera: Nymphalidae: Argynnini) // Biological Journal of the Linnean Society. - 2006. - Vol. 89, № 4. - P. 627–673.
23. Higgins, L.G. The Classification of European Butterflies. - London: Collins, 1975. - 320 p.
24. Kolesnichenko, K.A. A review of the East Palaearctic taxa of the Melitaea didyma (Esper, [1779])-group. 1. The M. ala Staudinger, 1881 -M. chitralensis Moore, 1901 -complex. // Atalanta. - 1999. - Vol. 30, № 1. - P. 78–117.
25. Pazhenkova, E.A., Zakharov, E. V., Lukhtanov, V.A. DNA barcoding reveals twelve lineages with properties of phylogenetic and biological species within Melitaea didyma sensu lato (Lepidoptera, Nymphalidae) // ZooKeys. - 2015. - Vol. 538P. 35–46.
26. Pazhenkova, E.A., Lukhtanov, V.A. Chromosomal and mitochondrial diversity in Melitaea Chromosomal and mitochondrial diversity in Melitaea didyma complex (Lepidoptera,

Nymphalidae): eleven deeply diverged DNA barcode groups in one non-monophyletic species? // CompCytogen. - 2016. - Vol. 10, № 4. - P. 697–717.

27. Pazhenkova, E.A., Lukhtanov, V.A. Nuclear genes (but not mitochondrial DNA barcodes) reveal real species: Evidence from the Brenthis fritillary butterflies (Lepidoptera, Nymphalidae) // Journal of Zoological Systematics and Evolutionary Research. - 2019. - Vol. 57, № 2. - .
28. Lukhtanov, V.A., Gagarina, A. V., Pazhenkova, E.A. Chromosomal and DNA barcode analysis of the Melitaea ala Staudinger, 1881 species complex (Lepidoptera, Nymphalidae) // Comparative Cytogenetics. - 2021. - Vol. 15, № 2. - P. 199–216.
29. Lukhtanov, V.A., Pazhenkova, E.A. The Taxa of the Hyponephele lycaon – H. lupina Species Complex (Lepidoptera, Nymphalidae, Satyrinae): Deep DNA Barcode Divergence despite Morphological Similarity // Folia Biologica (Poland). - 2021. - Vol. 69, № 1. - P. 11–21.
30. Pazhenkova, E.A., Lukhtanov, V.A. Genomic introgression from a distant congener in the Levant fritillary butterfly, Melitaea acentria // Molecular Ecology. - 2021.
31. The GIMP Development Team. GIMP. - 2019. URL https://www.gimp.org.
32. Bonhomme, V., Picq, S., Gaucherel, C., et al. Momocs: Outline analysis using R // Journal of Statistical Software. - 2014. - Vol. 56, № 13. - P. 1–24.
33. Dommergues, C.H., Dommergues, J.L., Verrecchia, E.P. The discrete cosine transform, a Fourier-related method for morphometric analysis of open contours // Mathematical Geology. - 2007. - Vol. 39, № 8. - P. 749–763.
34. Swofford, D.L. PAUP*, Phylogenetic analysis using parisimony (*and other methods), version 4.2002. - 2002.
35. Ronquist, F., Teslenko, M., van der Mark, P., et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space // Systematic biology. - . 2012/02/222012. - Vol. 61, № 3. - P. 539–542.
36. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT // Nucleic Acids Symp. - 1999. - Vol. 41. - P. 95–98.
37. Thompson, J.D., Higgins, D.G., Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. // Nucleic Acids Research. - 1994. - Vol. 22, № 22. - P. 4673.
38. Ahola, V., Lehtonen, R., Somervuo, P., et al. The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera // Nature Communications. - 2014. - Vol. 5, № 1. - P. 1–9.
39. McKenna, A., Hanna, M., Banks, E., et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data // Genome Research. - 2010. -

Vol. 20, № 9. - P. 1297–1303.

40. Zheng, X., Levine, D., Shen, J., et al. A high-performance computing toolset for relatedness and principal component analysis of SNP data // Bioinformatics. - 2012. - Vol. 28, № 24. - P. 3326–3328.
41. Huson, D.H., Bryant, D. Application of phylogenetic networks in evolutionary studies // Molecular Biology and Evolution. - 2006. - Vol. 23, № 2. - P. 254–267.
42. Martin, S.H., Jiggins, C.D. Interpreting the genomic landscape of introgression // Current Opinion in Genetics and Development. - 2017. - Vol. 47. - P. 69–74.
43. Malinsky, M., Matschiner, M., Svardal, H. Dsuite - Fast D-statistics and related admixture evidence from VCF files // Molecular Ecology Resources. - 2021. - Vol. 21, № 2. - P. 584–595.
44. Gutenkunst, R.N., Hernandez, R.D., Williamson, S.H., et al. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data // PLoS Genetics. - 2009. - Vol. 5, № 10. - P. 1000695.
45. Yu, Y., Harris, A.J., Blair, C., et al. RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography // Molecular Phylogenetics and Evolution. - 2015. - Vol. 87. - P. 46–49.
46. Matschiner, M. Fitchi: haplotype genealogy graphs based on the Fitch algorithm // Bioinformatics (Oxford, England). - 2016. - Vol. 32, № 8. - P. 1250—1252.
47. Simon, C., Frati, F., Beckenbach, A., et al. Evolution, Weighting, and Phylogenetic Utility of Mitochondrial Gene Sequences and a Compilation of Conserved Polymerase Chain Reaction Primers // Annals of the Entomological Society of America. - 1994. - Vol. 87, № 6. - P. 651–701.
48. Sperling, F.A.H., Anderson, G.S., Hickey, D.A. A DNA-Based Approach to the Identification of Insect Species Used for Postmorten Interval Estimation // Journal of Forensic Sciences. - 1994. - Vol. 39, № 2. - P. 13613J.
49. Birstein, V.J., Desalle, R. Molecular Phylogeny of Acipenserinae // Molecular Phylogenetics and Evolution. - 1998. - Vol. 9, № 1. - P. 141–155.
50. Hebert, P.D.N., Ratnasingham, S., DeWaard, J.R. Barcoding animal life: Cytochrome c oxidase subunit 1 divergences among closely related species // Proceedings of the Royal Society B: Biological Sciences. - 2003. - Vol. 270. - P. 96–99.
51. Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., et al. Identification of birds through DNA barcodes // PLoS Biology. - 2004. - Vol. 2, № 10.
52. Hogg, I.D., Hebert, P.D.N. Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes // Canadian Journal of Zoology. - 2004. - Vol. 82, № 5. - P. 749–754.

53. Barrett, R.D.H., Hebert, P.D.N. Identifying spiders through DNA barcodes // Canadian Journal of Zoology. - 2005. - Vol. 83, № 3. - P. 481–491.

54. Ball, S.L., Hebert, P.D.N., Burian, S.K., et al. Biological identifications of mayflies (Ephemeroptera) using DNA barcodes // https://doi.org/10.1899/04-142.1. - 2005. - Vol. 24, № 3. - P. 508–524.

55. Ward, R.D., Zemlak, T.S., Innes, B.H., et al. DNA barcoding Australia's fish species // Philosophical Transactions of the Royal Society B: Biological Sciences. - 2005. - Vol. 360, № 1462. - P. 1847–1857.

56. Wiemers, M., Fiedler, K. Does the DNA barcoding gap exist? - A case study in blue butterflies (Lepidoptera: Lycaenidae) // Frontiers in Zoology. - 2007. - Vol. 4, № September 2004. - P. 1–16.

57. Meier, R., Zhang, G., Ali, F. The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification // Systematic Biology. - 2008. - Vol. 57, № 5. - P. 809–813.

58. Zou, S., Li, Q., Kong, L., et al. Comparing the Usefulness of Distance, Monophyly and Character-Based DNA Barcoding Methods in Species Identification: A Case Study of Neogastropoda // PLoS ONE. - 2011. - Vol. 6, № 10.

59. Puillandre, N., Lambert, A., Brouillet, S., et al. ABGD, Automatic Barcode Gap Discovery for primary species delimitation // Molecular Ecology. - 2012. - Vol. 21, № 8. - P. 1864–1877.

60. Pons, J., Barraclough, T.G., Gomez-Zurita, J., et al. Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects // Systematic Biology. - 2006. - Vol. 55, № 4. - P. 595–609.

61. Zhang, J., Kapli, P., Pavlidis, P., et al. A general species delimitation method with applications to phylogenetic placements // Bioinformatics (Oxford, England). - 2013. - Vol. 29, № 22. - P. 2869–2876.

62. Somervuo, P., Koskela, S., Pennanen, J., et al. Unbiased probabilistic taxonomic classification for DNA barcoding // Bioinformatics. - 2016. - Vol. 32, № 19. - P. 2920–2927.

63. Dincă, V., Dapporto, L., Somervuo, P., et al. High resolution DNA barcode library for European butterflies reveals continental patterns of mitochondrial genetic diversity // Communications Biology. - 2021. - Vol. 4, № 1. - P. 1–11.

64. Hajibabaei, M., Jeremy, R., Ivanova, N. V, et al. Critical factors for assembling a high volume of DNA barcodes. // Philosophical transactions of the Royal society B. - 2005. - Vol. 60, №1462. - P. 1959-1967.

65. Kinoshita, G., Yonezawa, S., Murakami, S., et al. Environmental DNA Collected from Snow Tracks is Useful for Identification of Mammalian Species // Zoological science. - 2019. - Vol.

36, № 3. - P. 198–207.

66. Pook, C.E., McEwing, R. Mitochondrial DNA sequences from dried snake venom: a DNA barcoding approach to the identification of venom samples // Toxicon : official journal of the International Society on Toxinology. - 2005. - Vol. 46, № 7. - P. 711–715.
67. Kawai, K., Shimizu, M., Hughes, R.N., et al. A non-invasive technique for obtaining DNA from marine intertidal snails // Journal of the Marine Biological Association of the United Kingdom. - 2004. - Vol. 84, № 4. - P. 773–774.
68. Clare, E.L., Economou, C.K., Faulkes, C.G., et al. eDNAir: proof of concept that animal DNA can be collected from air sampling // PeerJ. - 2021. - Vol. 9P. - P. e11030.
69. Sonet, G., Pauly, A., Nagy, Z.T., et al. Using next-generation sequencing to improve DNA barcoding: lessons from a small-scale study of wild bee species (Hymenoptera, Halictidae) // Apidologie 2018 49:5. - 2018. - Vol. 49, № 5. - P. 671–685.
70. Sire, L., Gey, D., Debruyne, R., et al. The Challenge of DNA Barcoding Saproxylic Beetles in Natural History Collections—Exploring the Potential of Parallel Multiplex Sequencing With Illumina MiSeq // Frontiers in Ecology and Evolution. - 2019. - P. 495.
71. Shokralla, S., Porter, T.M., Gibson, J.F., et al. Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform // Scientific Reports 2015 5:1. - 2015. - Vol. 5, № 1. - P. 1–7.
72. Pomerantz, A., Peñafiel, N., Arteaga, A., et al. Real-time DNA barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building // GigaScience. - 2018. - Vol. 7, № 4. - P. 1–14.
73. Chang, J.J.M., Ip, Y.C.A., Bauman, A.G., et al. MinION-in-ARMS: Nanopore Sequencing to Expedite Barcoding of Specimen-Rich Macrofaunal Samples From Autonomous Reef Monitoring Structures // Frontiers in Marine Science. - 2020. - P. 448.
74. Hebert, P.D.N., Braukmann, T.W.A., Prosser, S.W.J., et al. A Sequel to Sanger: amplicon sequencing that scales // BMC Genomics 2018 19:1. - 2018. - Vol. 19, № 1. - P. 1–14.
75. Ratnasingham, S., Hebert, P.D.N. BOLD: The Barcode of Life Data System. - 2007. - [http://www.boldsystems.org]
76. Hebert, P.D.N., Penton, E.H., Burns, J.M., et al. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator // Proceedings of the National Academy of Sciences of the United States of America. - 2004. - Vol. 101, № 41. - P. 14812–14817.
77. Smith, M.A., Woodley, N.E., Janzen, D.H., et al. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae) // Proceedings of the National Academy of Sciences. - 2006. - Vol. 103, № 10. - P. 3657–3662.

78. Gustafsson, D.R., Price, D.A., Erséus, C. Genetic variation in the popular lab worm Lumbriculus variegatus (Annelida: Clitellata: Lumbriculidae) reveals cryptic speciation // Molecular phylogenetics and evolution. - 2009. - Vol. 51, № 2. - P. 182–189.
79. Elderkin, C.L., Clewing, C., Wembo Ndeo, O., et al. Molecular phylogeny and DNA barcoding confirm cryptic species in the African freshwater oyster Etheria elliptica Lamarck, 1807 (Bivalvia: Etheriidae) // Biological Journal of the Linnean Society. - 2016. - Vol. 118, № 2. - P. 369–381.
80. Tyagi, K., Kumar, V., Kundu, S., et al. Identification of Indian Spiders through DNA barcoding: Cryptic species and species complex // Scientific Reports 2019 9:1. - 2019. - Vol. 9, № 1. - P. 1–13.
81. Wang, T., Zhang, Y., Yang, Z., et al. DNA barcoding reveals cryptic diversity in the underestimated genus Triplophysa (Cypriniformes: Cobitidae, Nemacheilinae) from the northeastern Qinghai-Tibet Plateau // BMC Evolutionary Biology 2020 20:1. - 2020. - Vol. 20, № 1. - P. 1–15.
82. Dincă, V., Montagud, S., Talavera, G., et al. DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity // Scientific Reports. - 2015. - Vol. 5, № 1 - P. 1–12.
83. Avise, J.C. Molecular markers, natural history and evolution. -Oxford: Oxford University Press, 2004. - 684 p.
84. Cong, Q., Shen, J., Borek, D., et al. When COI barcodes deceive: complete genomes reveal introgression in hairstreaks // Proceedings of the Royal Society B: Biological Sciences. - 2017. - Vol. 284, № 1848. - P. 1 - 9.
85. Cariou, M., Duret, L., Charlat, S. The global impact of Wolbachia on mitochondrial diversity and evolution // Journal of evolutionary biology. - 2017. - Vol. 30, № 12. - P. 2204–2210.
86. Magnacca, K.N., Brown, M.J. Mitochondrial heteroplasmy and DNA barcoding in Hawaiian Hylaeus (Nesoprosopis) bees (Hymenoptera: Colletidae) // BMC Evolutionary Biology. - 2010. - Vol. 10, № 174.
87. Rubinoff, D., Cameron, S., Will, K. A Genomic Perspective on the Shortcomings of Mitochondrial DNA for “Barcoding” Identification // Journal of Heredity. - 2006. - Vol. 97, № 6. - P. 581–594.
88. Song, H., Buhay, J.E., Whiting, M.F., et al. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified // Proceedings of the National Academy of Sciences. - 2008. - Vol. 105, № 36. - P. 13486–13491.
89. Makhov, I.A., Gorodilova, Y.Y.U., Lukhtanov, V.A. Sympatric occurrence of deeply diverged mitochondrial DNA lineages in Siberian geometrid moths (Lepidoptera: Geometridae): cryptic

speciation, mitochondrial introgression, secondary admixture or effect of Wolbachia? // Biological Journal of the Linnean Society. - 2021. - Vol. XXP. 1–24.

90. Ré, F.C. De, Robe, L.J., Wallau, G.L., et al. Inferring the phylogenetic position of the Drosophila flavopilosa group: Incongruence within and between mitochondrial and nuclear multilocus datasets // Journal of Zoological Systematics and Evolutionary Research. - 2017. - Vol. 55, № 3. - P. 208–221.
91. Platt, R.N., Faircloth, B.C., Sullivan, K.A.M., et al. Conflicting Evolutionary Histories of the Mitochondrial and Nuclear Genomes in New World Myotis Bats // Systematic Biology. - 2018. - Vol. 67, № 2. - P. 236–249.
92. Sota, T., Vogler, A.P. Reconstructing species phylogeny of the carabid beetles Ohomopterus using multiple nuclear DNA sequences: heterogeneous information content and the performance of simultaneous analyses // Molecular Phylogenetics and Evolution. - 2003. - Vol. 26, № 1. - P. 139–154.
93. Sun, J.-T., Wang, M.-M., Zhang, Y.-K., et al. Evidence for high dispersal ability and mito-nuclear discordance in the small brown planthopper, Laodelphax striatellus // Scientific Reports. - 2015. - Vol. 5, № 1. - P. 1-10.
94. Toews, D.P.L., Brelsford, A. The biogeography of mitochondrial and nuclear discordance in animals // Molecular Ecology. - 2012. - Vol. 21, № 16. - P. 3907–3930.
95. Tóth, J.P., Bereczki, J., Varga, Z., et al. Relationships within the Melitaea phoebe species group (Lepidoptera: Nymphalidae): New insights from molecular and morphometric information // Systematic Entomology. - 2014. - Vol. 39, № 4. - P. 749–757.
96. Padial, J.M., Miralles, A., De la Riva, I., et al. The integrative future of taxonomy // Frontiers in Zoology. - 2010. - Vol. 7. - P. 1- 14.
97. Bond, J.E., Stockman, A.K. An Integrative Method for Delimiting Cohesion Species: Finding the Population-Species Interface in a Group of Californian Trapdoor Spiders with Extreme Genetic Divergence and Geographic Structuring // Systematic Biology. - 2008. - Vol. 57, № 4. - P. 628–646.
98. Lukhtanov, V.A., Dantchenko, A. V., Vishnevskaya, M.S., et al. Detecting cryptic species in sympatry and allopatry: Analysis of hidden diversity in Polyommatus (Agrodiaetus) butterflies (Lepidoptera: Lycaenidae) // Biological Journal of the Linnean Society. - 2015. - Vol. 116, № 2. - P. 468–485.
99. Garraffoni, A.R.S., Araújo, T.Q., Lourenço, A.P., et al. Integrative taxonomy of a new Redudasys species (Gastrotricha: Macrodasyida) sheds light on the invasion of fresh water habitats by macrodasyids // Scientific Reports 2019 9:1. - 2019. - Vol. 9, № 1. - P. 1–15.
100. Garcia-Erill, G., Kjær, M.M., Albrechtsen, A., et al. Vicariance followed by secondary gene

flow in a young gazelle species complex // Molecular Ecology. - 2021. - Vol. 30, № 2. - P. 528–544.

101. Adams, D.C., Berns, C.M., Kozak, K.H., et al. Are rates of species diversification correlated with rates of morphological evolution? // Proceedings of the Royal Society B: Biological Sciences. - 2009. - Vol. 276, № 1668. - P. 2729–2738.
102. Takahashi K., Y., T., M., N., et al. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons // Molecular biology and evolution. - 2001. - Vol. 18, № 11. - P. 2057–2066.
103. Mayden, R.L. A hierarchy of species concepts: the denouement in the saga of the species problems. - Species. The units of biodiversity. - London: Chapman & Hall.- 1997. - 381-423 p.
104. Delsuc, F., Brinkmann, H., Philippe, H. Phylogenomics and the reconstruction of the tree of life // Nature Reviews Genetics 2005 6:5. - 2005. - Vol. 6, № 5. - P. 361–375.
105. Misof, B., Liu, S., Meusemann, K., et al. Phylogenomics resolves the timing and pattern of insect evolution // Science. - 2014. - Vol. 346, № 6210. - P. 763–767.
106. Chipman, A.D., Ferrier, D.E.K., Brena, C., et al. The First Myriapod Genome Sequence Reveals Conservative Arthropod Gene Content and Genome Organisation in the Centipede Strigamia maritima // PLOS Biology. - 2014. - Vol. 12, № 11. - P. e1002005.
107. Nater, A., Mattle-Greminger, M.P., Nurcahyo, A., et al. Morphometric, Behavioral, and Genomic Evidence for a New Orangutan Species // Current Biology. - 2017. - Vol. 27, № 22. - P. 3487-3498.e10.
108. Harr, B. Genomic islands of differentiation between house mouse subspecies // Genome Research. - 2006. - Vol. 16, № 6. - P. 730–737.
109. Marques, D.A., Lucek, K., Meier, J.I., et al. Genomics of Rapid Incipient Speciation in Sympatric Threespine Stickleback // PLOS Genetics. - 2016. - Vol. 12, № 2. - P. e1005887.
110. Lamichhaney, S., Berglund, J., Almén, M.S., et al. Evolution of Darwin’s finches and their beaks revealed by genome sequencing // Nature 2015 518:7539. - 2015. - Vol. 518, № 7539. - P. 371–375.
111. Masly, J.P., Presgraves, D.C. High-resolution genome-wide dissection of the two rules of speciation in Drosophila // PLoS Biology. - 2007. - Vol. 5, № 9. - P. 1890–1898.
112. Charlesworth, B., Coyne, J.A., Barton, N.H. The relative rates of evolution of sex chromosomes and autosomes. // American Naturalist. - 1987. - Vol. 130, № 1. - P. 113–146.
113. Connallon, T., Singh, N.D., Clark, A.G. Impact of genetic architecture on the relative rates of X versus autosomal adaptive substitution // Molecular Biology and Evolution. - 2012. - Vol. 29, № 8. - P. 1933–1942.
114. Meisel, R.P., Malone, J.H., Clark, A.G. Faster-X Evolution of Gene Expression in Drosophila //

PLoS Genetics. - 2012. - Vol. 8, № 10. - P. e1003013–e1003013.

115. Dobzhansky, T. Studies on hybrid sterility - I. Spermatogenesis in pure and hybrid Drosophila pseudoobscura // Zeitschrift für Zellforschung und Mikroskopische Anatomie. - 1934. - Vol. 21, № 2. - P. 169–223.
116. Cong, Q., Zhang, J., Grishin, N. V. Genomic determinants of speciation // bioRxiv. - 2019. - .
117. Van Belleghem, S.M., Baquero, M., Papa, R., et al. Patterns of Z chromosome divergence among Heliconius species highlight the importance of historical demography // Molecular Ecology. - 2018. - Vol. 27, № 19. - P. 3852–3872.
118. Cortés-Ortiz, L., Nidiffer, M.D., Hermida-Lagunes, J., et al. Reduced Introgression of Sex Chromosome Markers in the Mexican Howler Monkey (Alouatta palliata × A. pigra) Hybrid Zone // International Journal of Primatology. - 2019. - Vol. 40, № 1. - P. 114–131.
119. Storchová, R., Reif, J., Nachman, M.W. Female heterogamety and speciation: Reduced introgression of the z chromosome between two species of nightingales // Evolution. - 2010. - Vol. 64, № 2. - P. 456–471.
120. Sankararaman, S., Mallick, S., Dannemann, M., et al. The genomic landscape of Neanderthal ancestry in present-day humans // Nature. - 2014. - Vol. 507, № 7492. - P. 354–357.
121. Schluter, D. Evidence for ecological speciation and its alternative // Science2009. - Vol. 323, № 5915. - P. 737–741.
122. Nei, M., Maruyama, T., Wu, C. Models of evolution of reproductive isolation // Genetics. - 1983. - Vol. 103, № 3. - P. 557–579.
123. Gavrilets, S. Perspective: Models of speciation - What have we learned in 40 years? // Evolution2003. - Vol. 57, № 10. - P. 2197–2215.
124. Garcia-Erill, G., Kjær, M.M., Albrechtsen, A., et al. Vicariance followed by secondary gene flow in a young gazelle species complex // Molecular Ecology. - 2021. - Vol. 30, № 2. - P. 528–544.
125. Elgvin, T.O., Trier, C.N., Tørresen, O.K., et al. The genomic mosaicism of hybrid speciation // Science Advances. - 2017. - Vol. 3, № 6.
126. Teng, H., Zhang, Y., Shi, C., et al. Population Genomics Reveals Speciation and Introgression between Brown Norway Rats and Their Sibling Species // Molecular Biology and Evolution. - 2017. - Vol. 34, № 9. - P. 2214–2228.
127. Edelman, N.B., Frandsen, P.B., Miyagi, M., et al. Genomic architecture and introgression shape a butterfly radiation // Science. - 2019. - Vol. 366, № 6465. - P. 594–599.
128. Taylor, S.A., Larson, E.L. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature // Nature Ecology and Evolution. - 2019. - Vol. 3, № 2. - P. 170–177.

129. Fontsere, C., de Manuel, M., Marques-Bonet, T., et al. Admixture in mammals and how to understand its functional implications: On the abundance of gene flow in Mammalian species, its impact on the genome, and roads into a functional understanding // BioEssays. - 2019. - Vol. 41, № 12.

130. Pardo-Diaz, C., Salazar, C., Baxter, S.W., et al. Adaptive Introgression across Species Boundaries in Heliconius Butterflies // PLoS Genetics. - / ed. R. Kronforst M.2012. - Vol. 8, № 6. - P. e1002752.

131. Zhang, W., Dasmahapatra, K.K., Mallet, J., et al. Genome-wide introgression among distantly related Heliconius butterfly species // Genome Biology. - 2016. - Vol. 17, № 1. - P. 1–15.

132. Neri, J., Wendt, T., Palma-Silva, C. Natural hybridization and genetic and morphological variation between two epiphytic bromeliads // AoB PLANTS. - 2018. - Vol. 10, № 1.

133. Lukhtanov, V.A., Shapoval, N.A., Anokhin, B.A., et al. Homoploid hybrid speciation and genome evolution via chromosome sorting // Proceedings of the Royal Society B: Biological Sciences. - 2015. - Vol. 282, № 1807.

134. Wagner, C.E., Keller, I., Wittwer, S., et al. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation // Molecular ecology. - 2013. - Vol. 22, № 3. - P. 787–798.

135. Higgins, L.G. An illustrated catalogue of Palearctic Melitaea (Lep. Rhopalocera) // Transactions of the Royal Entomological Society of London. - 1941. - Vol. 91, № 7. - P. 175–365.

136. Leneveu, J., Chichvarkhin, A., Wahlberg, N. Varying rates of diversification in the genus Melitaea (Lepidoptera: Nymphalidae) during the past 20 million years // Biological Journal of the Linnean Society. - 2009. - Vol. 97, № 2. - P. 346–361.

137. Cracraft, J. Species concepts and speciation analysis. // Current ornithology. Vol. 1. - 1983. - P. 159–187.

138. Coyne, J. A., Orr, H. A. Speciation. – Sanderland, MA: Sinauer Associates, 2004. – 545 p.

139. Tshikolovets, V. V. The Butterflies of Tajikistan. - Brno-Kyiv: Tshikolovets Publications. - 500 p.

140. Brower, A.V.Z., DeSalle, R. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of wingless as a source of characters for phylogenetic inference // Insect Molecular Biology. - 1998. - Vol. 7, № 1. - P. 73–82.

141. Quek, S.-P., Davies, S.J., Itino, T., et al. Codiversification in an ant-plant mutualism: Stem texture and the evolution of host use in Crematogaster (Formicidae: Myrmicinae) inhabitants of Macaranga (Euphorbiaceae) // Evolution. - 2004. - Vol. 58, № 3. - P. 554–570.

142. Gómez, A., Lunt, D.H. Refugia within Refugia: Patterns of Phylogeographic Concordance in the Iberian Peninsula // Phylogeography of Southern European Refugia: Evolutionary Perspectives on the Origins and Conservation of European Biodiversity. - 2007. - P. 155–188.
143. Kolesnichenko, K.A., Churkin, S. V., Berdiev, R.K. A review of the East Palaearctic taxa of the Melitaea didyma (Esper, [1779])-group. Part II. // Neue Entomologische Nachrichten. - 2011. - Vol. 67. - P. 257–278.
144. Лухтанов, В.А. Делимитация видов и анализ криптического видового разнообразия в XXI веке // Энтомологическое обозрение. - 2019. - Vol. 98, № 2. - P. 358–370.
145. Palumbi, S.R., Cipriano, F., Hare, M.P. Predicting nuclear gene coalescence from mitochondrial data: the three-times rule // Evolution. - 2001. - Vol. 55, № 5. - P. 859–868.
146. Hesselbarth, G., Oorchot, H., Wagener, S. Die Tagfalter der Türkei: unter Berücksichtigung der angrenzenden Länder. - Bocholt: Selbstverlag Siegbert Wagener, 1995. - 2201 p.
147. Zimmermann, K., Fric, Z., Filipová, L., et al. Adult demography, dispersal and behaviour of Brenthis ino (Lepidoptera: Nymphalidae): How to be a successful wetland butterfly // European Journal of Entomology. - 2005. - Vol. 102, № 4. - P. 699–706.
148. Browning, S.R., Browning, B.L. Haplotype phasing: existing methods and new developments // Nature Reviews Genetics 2011 12:10. - 2011. - Vol. 12, № 10. - P. 703–714.
149. Friberg, M., Leimar, O., Wiklund, C. Heterospecific courtship, minority effects and niche separation between cryptic butterfly species // Journal of Evolutionary Biology. - 2013. - Vol. 26, № 5. - P. 971–979.
150. Ramírez, S.R., Nieh, J.C., Quental, T.B., et al. A molecular phylogeny of the stingless bee genus Melipona (Hymenoptera: Apidae) // Molecular Phylogenetics and Evolution. - 2010. - Vol. 56, № 2. - P. 519–525.
151. Silva-Brandão, K.L., Wahlberg, N., Francini, R.B., et al. Phylogenetic relationships of butterflies of the tribe Acraeini (Lepidoptera, Nymphalidae, Heliconiinae) and the evolution of host plant use // Molecular Phylogenetics and Evolution. - 2008. - Vol. 46, № 2. - P. 515–531.
152. van Velzen, R., Wahlberg, N., Sosef, M.S.M., et al. Effects of changing climate on species diversification in tropical forest butterflies of the genus Cymothoe (Lepidoptera: Nymphalidae) // Biological Journal of the Linnean Society. - 2013. - Vol. 108, № 3. - P. 546–564.
153. Eckweiler, W., Bozano, G.C. Guide to the Butterflies of the Palearctic Region. Satyrinae part IV. Tribe Satyrini, Subtribe Maniolina. - Milano: Omnes Artes, 2011. - 102 p.
154. Staudinger, O., Rebel, H. Catalog der Lepidopteren des Palaearctischen Faunengebietes. Famil. Papilionidae-Hepialidae. - Berlin: Friedländer & Sohn, 1901. - 411 p.
155. Turati, E. Revisione delle forme di Epinephele lycaon Rott. // Naturalista Siciliano. - 1909. - Vol. 21, № 1–8. - P. 56–73.

156. Weiss, D. A new species of the genus Hyponephele Muschamp 1915 from West Iran // Atalanta. - 1978. - Vol. 9P. 230–232.
157. Tshikolovets, V. V., Naderi, A., Eckweiler, W. The butterflies of Iran and Iraq. -Pardubice: Tshikolovets Publications, 2014. - 440 p.
158. Samodurov, G.D., Korolev, V.A., Tshikolovets, V.V. Eine Übersicht über die Satyriden der Gattung Hyponephele Muschamp, 1915 VII. Die Arten Hyponephele lycaon (Rottemburg, 1775), H. pasimelas (Staudinger, 1886), H. lycaonoides D. Weiss, 1978, H. przhewalskyi Dubatolov, Sergeev & Zhdanko, 1994, H. dzhungari // Atalanta. - 2001. - Vol. 32, № 1/2. - P. 111–186.
159. Mallet, J., Besansky, N., Hahn, M.W. How reticulated are species? // Bioessays. - 2016. - Vol. 38, № 2. - P. 140-149.
160. Gompert, Z., Fordyce, J.A., Forister, M.L., et al. Homoploid hybrid speciation in an extreme habitat // Science. - 2006. - Vol. 314, № 5807. - P. 1923–1925.
161. Green, R.E., Krause, J., Briggs, A.W., et al. A draft sequence of the neandertal genome // Science. - 2010. - Vol. 328, № 5979. - P. 710–722.
162. Durand, E.Y., Patterson, N., Reich, D., et al. Testing for ancient admixture between closely related populations // Molecular Biology and Evolution. - 2011. - Vol. 28, № 8. - P. 2239–2252.
163. Martin, S.H., Davey, J.W., Jiggins, C.D. Evaluating the use of ABBA-BABA statistics to locate introgressed loci // Molecular Biology and Evolution. - 2015. - Vol. 32, № 1. - P. 244–257.
164. Ball, R.M., Neigel, J.E., Avise, J.C. Gene Genealogies within the Organismal Pedigrees of Random-Mating Populations // Evolution. - 1990. - Vol. 44, № 2. - P. 360-370.
165. Van Valen, L. Ecological species, multispecies, and oaks // Taxon. - 1976. - Vol. 25, № 2–3. - P. 233–239.
166. Andersson, L. The driving force: species concepts and ecology // Taxon. - 1990. - Vol. 39, № 3. - P. 375–382.
167. Currat, M., Ruedi, M., Petit, R.J., et al. The hidden side of invasions: Massive introgression by local genes // Evolution. - 2008. - Vol. 62, № 8. - P. 1908–1920.
168. Valencia-Montoya, W.A., Elfekih, S., North, H.L., et al. Adaptive introgression across semipermeable species boundaries between local helicoverpa zea and invasive helicoverpa armigera moths // Molecular Biology and Evolution. - 2020. - Vol. 37, № 9. - P. 2568–2583.