

УДК 575.8

ДЕЛИМИТАЦИЯ ВИДОВ И АНАЛИЗ КРИПТИЧЕСКОГО ВИДОВОГО РАЗНООБРАЗИЯ В XXI ВЕКЕ

© 2019 г. В. А. Лухтанов

¹ Зоологический институт РАН

Университетская наб., 1, С.-Петербург, 199034 Россия

² Санкт-Петербургский государственный университет

Университетская наб., 7/9, С.-Петербург, 199034 Россия

e-mail: lukhtanov@mail.ru

Поступила в редакцию 13.03.2019 г.

После доработки 10.04.2019 г.

Принята к публикации 10.04.2019 г.

В статье обсуждаются возможности и ограничения разных подходов к выявлению видовых границ и описанию криптических видов. Вначале рассматриваются более традиционные методы делимитации видов, основанные преимущественно на анализе морфологических особенностей. Затем дается обзор подходов к делимитации видов с использованием молекулярных маркеров. Коротко рассматриваются также возможности подходов, основанных на машинном обучении компьютерных программ, к распознаванию видов по их дигитализированным изображениям. Делается вывод о том, что молекулярные признаки дают важный материал для разграничения видов, ценность которого многократно возрастает, если полученные молекулярные данные сравниваются с информацией по морфологии, географическому распространению и экологическим предпочтениям изучаемых таксонов. В заключение говорится о том, что зачастую только практикующий систематик, досконально знающий свою группу, может правильно интерпретировать получаемые молекулярные данные и вставить их в систему уже существующего знания для принятия таксономического решения.

Ключевые слова: вид, систематика, насекомые, филогенетическая концепция вида, биологическая концепция вида, генетические дистанции, филогенетический анализ, молекулярный признак, монолокусные генетические данные, мультилокусные генетические данные, полногеномные данные, искусственный интеллект, машинное обучение.

DOI: 10.1134/S0367144519020096

Более 260 лет прошло с момента выхода десятого издания *Systema Naturae* Карла Линнея (Linnaeus, 1758) – отправной точки зоологической систематики, однако поставленная Линнеем задача инвентаризации видов животных еще необозримо далека от решения. Формально описанные на настоящий момент времени 1 600 000 видов животных (Zhang, 2013) составляют, по-видимому, относительно небольшой процент от реально существующего на Земле видового разнообразия, большая часть которого, по крайней мере, среди многоклеточных животных, вне всякого сомнения, относится к классу насекомых (Stork et al., 2015; Larsen et al., 2017). Выявление и описание этого разнообразия еще долго будут оставаться главным вызовом и одной из важнейших задач энтомологии.

Важно подчеркнуть, что «виды», которые появляются в таксономических сводках и ревизиях, – это не обязательно то же самое, что единицы биоразнообразия, реально существующие в природе. Виды систематиков – это скорее теоретические конструкции, представления систематиков о биологических таксонах, т. е., по сути дела, гипотезы, которые пытаются описать действительность. Эти представления могут сильно отличаться от того, что есть на самом деле. Поэтому даже для тех групп организмов, в которых формально виды в основном описаны, работа по выявлению и уточнению существующих межвидовых границ, т. е., работа по делимитации видов, будет продолжаться еще долго.

Неудивительно поэтому, что за последние 30–40 лет в связи с появлением методов секвенирования ДНК и гигантским прогрессом в развитии методов филогенетического анализа появилась масса предложений и подходов, направленных на решение проблемы разграничения видов с использованием молекулярных маркеров (= молекулярных признаков) (Carstens et al., 2013). Кроме того, буквально в последние годы в таксономию стали внедряться методы компьютерного анализа изображений и машинного обучения (искусственного интеллекта). Возможности этих подходов в делимитации еще не описанных видов пока не очевидны, хотя они уже хорошо зарекомендовали себя при решении более простой задачи идентификации уже известных видов (Wäldchen, Mäder, 2018).

В предлагаемом кратком обзоре дается анализ достижений и проблем в области делимитации видов и анализа криптического видового разнообразия, отчасти основанный на конкретном опыте автора этой статьи и являющийся обобщением исследований, проведенных на кафедре энтомологии Санкт-Петербургского университета и в Зоологическом институте РАН.

Концепции видов

Важный вопрос: а что, собственно говоря, мы делимитируем? Понятно, что виды, но что такое вид? Этот вопрос не имеет однозначного ответа, так как существует множество видовых концепций и определений вида, и дискуссии на эту тему бесконечны (De Queiroz, 2007).

Сильно упрощая ситуацию, можно сказать, что в настоящее время наиболее популярны две группы видовых концепций: одна группа имеет дело с видом как эволюционной линией (варианты филогенетических концепций вида), а другая оперирует видом как совокупностью особей, которые репродуктивно совместимы между собой, но репродуктивно изолированы от других подобных совокупностей (варианты биологической концепции вида) (Cracraft, 1989; Coyne, Orr, 2004).

Диагностическая версия филогенетической концепции определяет вид как совокупность популяций, имеющих один или большее число уникальных (диагностических) признаков (Cracraft, 1989; Rosser et al., 2015). Эта версия, с одной стороны, смыкается с более старой типологической концепцией, которая в переводе на язык практической систематики определяет вид как сущность, выявляемую на основании изучения морфологии (Майр, 1968). С другой стороны, она является филогенетической, поскольку уникальные признаки, по крайней мере часть из них, могут быть (син)апорфиями, выявляющими монофилетические группы. В более явной форме использование концепции филогенетического вида предполагает наличие доказательной базы того, что выделяемая группировка (филогенетический вид) является монофилетической группой (Coyne, Orr, 2004).

Проблемой филогенетической концепции является то, что на практике ее последовательное применение приводит к чрезмерному дробительству, когда почти каждая

полностью или даже частично изолированная популяция возводится в ранг вида. Это приводит к явлению, которое было названо таксономической инфляцией (Isaac et al., 2004). Другая проблема – существование в природе парафилетических видов, которые возникают в ходе перипатрического видообразования, когда от материнского вида по периферии отпочковываются дочерние виды. Такой способ видообразования превращает материнский таксон в парафилетическую группу, т. е., в ветвь, не включающую всех потомков. Интерпретация таких парафилетических групп в качестве видов может быть головной болью для традиционного кладиста, полностью отрицающего возможность немонофилетических таксонов (Лухтанов, 2013).

Биологическая концепция рассматривает в качестве видов совсем другие сущности: репродуктивно изолированные сообщества (эта изоляция может быть неполной, но все же достаточной для предотвращения слияния двух групп в один генетический пул), не обращая внимания на филогенетическую историю таких сообществ и не оперируя терминами монофилия и (син)апоморфия (Майр, 1968, 1971; Coyne, Orr, 2004; Rosser et al., 2015). Симпатрические биологические виды (безмерные виды, по: Майр, 1971) объективны и делимитируются в подавляющем большинстве случаев совершенно однозначно. К этой концепции близка концепция генотипических кластеров Маллета (Mallet, 1995), которая не делает акцента на репродуктивной изоляции и в качестве видов рассматривает кластеры особей, способные существовать симпатрично. Ограниченность биологической концепции в том, что она плохо работает для аллопатрических таксонов и непригодна для однополых видов.

Важным следствием биологической концепции является допустимость существования слабо дифференцированных по морфологическим признакам (криптических) видов, которые могут обитать совместно благодаря наличию изолирующих поведенческих, экологических или физиологических барьеров. Такие виды трудно или даже почти невозможно выявлять на основании анализа традиционных морфологических маркеров.

Интересно, что, несмотря на описание филогенетическими и биологическими концепциями принципиально разных сущностей, на практике они во многих случаях делимитируют одни и те же единицы. Дело в том, что филогенетические линии, даже если они вначале репродуктивно совместимы, эволюционируя раздельно, рано или поздно накапливают различия, приводящие к репродуктивной изоляции, т. е., становятся биологическими видами. И наоборот, возникающие в ходе симпатрического видообразования репродуктивно изолированные группы поначалу могут не представлять собой монофилетические линии. Однако в силу генетической изоляции такие группы эволюционируют раздельно и поэтому накапливают изменения, приводящие к тому, что на филогенетическом дереве они начинают появляться в виде монофилетических групп, даже если для построения дерева используются только нейтральные признаки, не связанные с появлением и поддержкой репродуктивной изоляции.

Таким образом, если мы сравниваем филогенетически далекие виды, они, как правило, удовлетворяют критериям обеих концепций. Однако проблемы несовпадения видовых границ нередко возникают при применении разных концепций к недавно разошедшимся таксонам.

Делимитация видов в домоллекулярную эпоху

Делимитация видов в домоллекулярную эпоху (которая для многих организмов еще не закончилась) является скорее практическим занятием, чем научной теорией, и ос-

нована на нескольких простых предпосылках, подробно описанных Эрнстом Майром (Майр, 1971).

При использовании типолого-морфологической концепции на основании анализа признаков оценивается уровень внутри- и межгрупповой изменчивости с целью выявления хиатуса между изучаемыми группами. При наличии такого хиатуса группы могут интерпретироваться как разные таксономические виды. В энтомологии при этом традиционно, начиная с начала XX в., большое значение придается изучению признаков строения генитального аппарата.

Делимитация, основанная на концепции биологического вида, выявляет группы особей, морфологические признаки которых не смешиваются при обитании этих групп вместе (симпатрия) или при наличии контакта между их ареалами (парапатрия). Такие группы рассматриваются как виды. Таксономическая интерпретация морфологически дифференцированных аллопатрических популяций (виды или подвиды) с использованием концепции биологического вида затруднена и производится более произвольно. Например, в качестве видов рекомендуется рассматривать географические популяции (группы популяций), морфологические различия между которыми сравнимы с таковыми между симпатрическими видами того же рода либо касаются признаков, потенциально приводящих к репродуктивной изоляции. В редких случаях репродуктивную изоляцию между аллопатрическими формами удается выявить в результате экспериментальной проверки (см., например: Lorković, 1993), и в последнем случае видовой статус аллопатрических форм вряд ли может быть подвергнут сомнению.

Для выявления и количественной оценки хиатуса между потенциальными видами может быть использован анализ данных морфометрии с помощью метода главных компонент (Vrtilek, Reichard, 2016).

Делимитация видов в молекулярную эпоху

Молекулярные подходы, появившиеся в конце XX в., никогда не декларировались в качестве исключительного инструмента для делимитации видов (Pons et al., 2006), однако претендуют на то, что могут:

- 1) ускорить процесс выявления и описания видов;
- 2) унифицировать подходы к выявлению структуры биологического разнообразия на том основании, что одни и те же маркеры могут использоваться для самых разных таксономических групп;
- 3) сделать методы делимитации более формализованными, более прозрачными и воспроизводимыми;
- 4) делать то, что ранее делалось с большим трудом или не делалось вообще: выявлять и описывать криптические виды.

Методы, основанные на дистанциях

Эволюция преимущественно дивергентна: в ее ходе между эволюционирующими линиями накапливаются генетические различия, и уровень этих различий можно использовать для видовой делимитации. Эмпирически показано, что так называемый ДНК-баркод, фрагмент митохондриального гена *COI* длиной 658 нуклеотидных пар (т. е. ничтожная по размеру часть генома) позволяет различать до 95 % видов во многих группах живых организмов (Hebert et al., 2003, 2004; Najibabaei et al., 2006; Hebert, Gregory, 2005, Lukhtanov et al., 2009). При этом различия в 2 и более процентов почти всегда соответствуют уровню различий между «хорошими» видами (но есть многочисленные исключения, см.: Wiemers, Fiedler, 2007). На основании молекулярного сход-

ства по этому фрагменту особи могут объединяться в «корзины» (BINs, barcode index numbers), которые можно рассматривать в качестве операциональных таксономических единиц, т. е. в качестве потенциальных обладателей видового статуса (Ratnasingham, Hebert, 2013). Такие «корзины» вполне оправдывают свое название («bin» в английском языке означает не просто емкость, но и мусорное ведро): они пригодны лишь для предварительной оценки и не могут быть серьезным основанием для принятия таксономических решений. Однако они полезны в качестве предварительных таксономических гипотез и могут тестироваться с использованием других подходов.

Методы, основанные на филогенетическом анализе молекулярных признаков

Виды возникают в природе в ходе дихотомической или ретикулярной эволюции, и реконструкция их истории с использованием методов филогенетического анализа может выявлять не только крупные эволюционные линии, но и терминальные клады, которые можно интерпретировать в качестве видов. На практике это означает, что при использовании этого метода делимитации в филогенетический анализ включаются многочисленные особи каждого из предполагаемых видов. Эти особи формируют на филогенетическом дереве кластеры, которые и являются объектом таксономической интерпретации. При первичной оценке таких кластеров большое значение имеет уровень статистической поддержки выявляемых кластеров в виде бутстреп-поддержки при проведении анализов с использованием методов максимального правдоподобия и максимальной парсимонии и в виде постериорной вероятности при использовании метода Байеса. Уровень этой поддержки должен быть более 90 %, чтобы придавать серьезное значение выявленному кластеру.

Такой анализ может проводиться на основании секвенирования одного или нескольких генов. При этом никогда нельзя верить анализу, основанному на изучении одного гена, так как филогенетическая история любого одного гена может сильно отличаться от истории организма (Nichols, 2001). (То же самое можно сказать о традиционных домолекулярных подходах: серьезная таксономическая работа не может основываться на одном признаке).

Тем не менее, филогенетическая реконструкция, основанная на изучении одного гена, может использоваться вместе с данными по морфологии, экологии, кариологии и географическому распространению изучаемых таксонов и давать ценную дополнительную информацию для принятия таксономического решения (Lukhtanov et al., 2015a, 2015b, 2016).

Использование митохондриальных ДНК-баркодов для видовой делимитации. ДНК-баркод (фрагмент митохондриального гена *COI* длиной 658 нуклеотидных пар) является наиболее распространенным молекулярным маркером для видовой делимитации у животных, в том числе у насекомых (Hebert et al., 2003, 2004). Чаще всего его использование приводит к правильным таксономическим заключениям, однако существует опасность получения ложноположительных и ложноотрицательных результатов, т. е., этот маркер может быть дифференцирован и давать статистически значимые разные кластеры в ситуациях, где нет разных видов, и наоборот, дифференциация по этому маркеру может отсутствовать при сравнении разных видов (Toews, Brelsford, 2012).

Первый случай (митохондриальный ДНК-баркод дифференцирован там, где нет разных видов) у насекомых чаще всего определяется наличием внутриклеточного

паразита *Wolbachia*, наличие которого в части популяции может индуцировать внутри-видовую дивергенцию по митохондриальным генам (Ritter et al. 2013).

Второй случай (виды не дифференцированы по митохондриальной ДНК) может быть связан с реальной эволюционной молодостью таксонов или же с так называемой митохондриальной интрогрессией (Toews, Brelsford, 2012), когда в результате спорадических гибридизаций митохондрии одного вида замещают митохондрии другого вида, приводя к тому, что два вида (иногда более чем два) (Lukhtanov et al., 2015b) имеют идентичные митохондриальные геномы. Хотя последнее явление кажется экзотическим, эмпирические данные показывают, что митохондриальная интрогрессия – довольно широко распространенное явление.

Еще одной проблемой использования митохондриальных баркодов являются то, что в некоторых случаях вследствие генных дупликаций и последующей эволюции могут возникать модифицированные ядерные копии этих последовательностей (псевдогены) (Bensasson et al., 2001). Хотя по сути дела, получаемые псевдогены – это другие маркеры, которые находятся в другой части генома и эволюционируют независимо, они могут амплифицироваться теми же праймерами, что и митохондриальные гены. Использование митохондриальных генов (для одних видов) с псевдогенами (для других видов) в рамках одного филогенетического анализа неправомерно и приводит к ошибочным заключениям (Galtier, 2009; Cameron, 2014). Хотя в общем случае различить митохондриальные и ядерные копии не так легко, часто помогает простой анализ: попытка сделать трансляцию. Ядерные копии (псевдогены) не кодируют белки, и поэтому в их последовательностях могут содержаться нуклеотиды, нарушающие правила генетических кодов, а также стоп-кодоны, в то время как кодирующие белок митохондриальные гены, такие как COI и CytB, должны транслироваться идеально.

Использование множественных генов для видовой делимитации. Более адекватным методом межвидовой делимитации является филогенетический анализ, основанный на использовании нескольких, желательно несцепленных (находящихся на разных хромосомах) ядерных генов. Филогенетический анализ можно провести для каждого из этих генов по отдельности, и факт конгруэнтности получаемых результатов (разные гены дают одни и те же кластеры особей) можно интерпретировать как свидетельство того, что выявленные межвидовые границы являются истинными. Недостаток такого подхода заключается в том, что ядерные гены обычно эволюционируют медленнее митохондриальных. Поэтому филогении, основанные на отдельных ядерных генах, часто имеют низкое разрешение и низкую статистическую поддержку выявляемых клад. Эта проблема может быть решена путем конкатенирования – объединения нескольких генов в одну матрицу. Однако при проведении конкатенирования обязательной процедурой является предварительная проверка отдельных генов на филогенетическую непротиворечивость. В противном случае может получиться химера, которая не отражает ни историю таксонов, ни историю отдельных генов, и не имеет биологического смысла (Pazhenkova, Lukhtanov, 2019).

Методы делимитации, основанные на моделях

Методы этой группы основаны на филогениях, но учитывают то, что кладогенез (филогения таксонов) и филогения гена (генов) – разные явления (Nichols, 2001). В этих методах делается попытка смоделировать точку перехода от филогении аллелей к кладогенезу на основании предположения о том, что коалесценция аллелей (переход к монофилии аллелей изучаемого гена) происходит быстрее, чем видообразование (Carstens et al., 2013).

Так же, как в предыдущем случае, при использовании этих методов делимитации в филогенетический анализ включаются многочисленные особи каждого из предполагаемых видов, а сами методы могут оперировать одним или альтернативно несколькими генами. При использовании одного гена, в частности ДНК-баркодов, большое распространение получила модель GMYC (The General Mixed Yule Coalescent model) (Pons et al., 2006).

Хотя в теоретическом отношении методы делимитации с использованием одного гена, основанные на модели GMYC, более совершенны, чем алгоритмы, описанные в предыдущем разделе, все монолокусные подходы страдают одним и тем же существенным недостатком: неопределенностью получаемых таксономических заключений, связанной с тем, что один признак недостаточен для решения таксономических задач. Поэтому все кластеры, получаемые с помощью монолокусных подходов, являются не более чем предварительными гипотезами.

Существуют также мультилокусные алгоритмы выявления видовых кластеров, основанные на модели GMYC (O'Meara, 2010), однако они менее популярны, чем методы, основанные на валидации предварительно выдвинутых видовых гипотез (см. ниже).

Методы, основанные на валидации предварительно выдвинутых видовых гипотез (построение «деревьев видов»)

В настоящее время это один из наиболее совершенных и популярных подходов к делимитации видов. При использовании этого подхода вначале выдвигаются гипотезы о видах (о том, что тот или иной набор особей составляет особый вид). Эти гипотезы могут быть альтернативными, т. е., для одной и той же полной совокупности особей можно предлагать разные варианты разбиения на виды (Yang, Rannala, 2010; Camargo et al., 2012; Carstens et al., 2013). Предварительные гипотезы могут выдвигаться на основании ранее существовавших таксономических интерпретаций, например, на основании изучения морфологии, строения митохондриальной или ядерной ДНК, особенностей кариотипов или произвольно (Talavera et al., 2013; Todisco et al., 2018; Pazhenkova, Lukhtanov, 2019).

Затем с использованием методов Байесовой филогенетики для разных вариантов разбиения полной совокупности особей на гипотетические виды строятся филогенетические деревья. Это делается на основании анализа нуклеотидных последовательностей одного или, предпочтительнее, нескольких генов (Carstens et al., 2013) или даже на основании анализа данных полногеномного секвенирования (Leache et al., 2014; O'Connell, Smith, 2018). При этом в алгоритм анализа закладывается ограничение для получаемых топологий: совокупности особей, которые были выбраны в качестве видовых гипотез, должны обязательно появиться на получаемых деревьях в виде клад. Однако статистическая поддержка получаемых клад (уровень их правдоподобия) будет разной в разных вариантах разбиения исходной совокупности. В качестве решения задачи принимается топология (топологии), обеспечивающая максимальный уровень статистической поддержки.

К бесспорным преимуществам этого можно отнести то, что

1) существует возможность прямого сравнения альтернативных гипотез разбиения совокупности особей на виды;

2) дерево видов строится на основании совокупности нуклеотидных последовательностей по всем изученным генам без применения такой теоретически малообоснован-

ной процедуры как конкатенирование (механическое объединение генов), т. е., с учетом индивидуальных особенностей коалесценции аллелей каждого анализируемого гена;

3) заключение о каждом виде делается на основании анализа распределения большого числа деревьев, что позволяет учитывать филогенетическую неопределенность (уровень ненадежности получаемых реконструкций) при принятии таксономического решения.

Анализ криптического видового разнообразия (выявление видов-двойников)

При использовании молекулярных маркеров эта процедура принципиально не отличается от стандартной делимитации видов, поскольку морфологические маркеры не используются. Таким образом, при поиске криптических видов вначале на основании анализа молекулярных маркеров выявляются кластеры генетически сходных особей, а затем делается интерпретация этих кластеров в качестве видов на основании использования дополнительных критериев (Lukhtanov et al., 2006, 2014; Лухтанов, Шаповал, 2008; Лухтанов, Кузнецова, 2009; Verшинина, Lukhtanov, 2010; Vila et al., 2010; Dinca et al., 2011; Hernández-Roldán et al., 2016; Vishnevskaya et al., 2016; Lukhtanov, Dantchenko, 2017b; Lukhtanov, Shapoval, 2017a).

При анализе двух кластеров, обнаруженных симпатрично, вывод об их неконспецифичности может быть сделан на основании:

1) конгруэнтности деревьев, полученных при использовании разных, генетически несцепленных (то есть находящихся на разных хромосомах) маркеров (иными словами, если разные гены выявляют одни и те же кластеры особей) (Avice, 2000, 2004);

2) бимодальности в распределении молекулярных маркеров в исследуемой выборке (Mallet, Willmott, 2003);

3) отсутствия или дефицита гетерозигот по дифференцированным видовым ядерным маркерам, которые могли бы свидетельствовать о гибридизации. Этот метод чаще применяется для анализа хромосомных признаков, так как выявление гетерозигот по нуклеотидным заменам является сложной процедурой и требует применения методов клонирования (Shapoval, Lukhtanov, 2015) или, в крайнем случае, методов фэйзинга (Pazhenkova, Lukhtanov, 2019);

4) выявления картины, которая имитирует сцепление изученных, заведомо несцепленных маркеров (Лухтанов, Шаповал, 2008; Lukhtanov et al., 2008)

При анализе двух морфологически неразличимых (или близких) кластеров, обнаруженных аллопатрично, заключение или предположение об их неконспецифичности может быть сделано на основании:

1) прямых экспериментальных доказательств наличия репродуктивной изоляции (Dincă et al., 2013);

2) косвенных свидетельств возможной репродуктивной изоляции. В качестве таких могут выступать параметры, которые потенциально играют роль в возникновении презиготической репродуктивной изоляции (например, феромоны и звуковые сигналы) (Lukhtanov et al., 2005) или постзиготической изоляции (например, разные кариотипы) (Lukhtanov et al., 2015a, 2015b, 2018);

3) высокого уровня дифференциации по изученным молекулярным маркерам (см. выше Методы, основанные на дистанциях).

Таксономическая интуиция и возможности искусственного интеллекта

Под таксономической интуицией здесь понимается способность распознавать потенциально новые таксоны «на глаз», без приведения формализованного вербального списка признаков, на основании которых распознавание было выполнено.

Вряд ли кто-то будет отрицать важность таксономической интуиции в повседневной работе систематика: зачастую именно она является триггером, запускающим процесс выявления неописанных таксонов. Однако таксономическая интуиция может также играть отрицательную роль в работе систематика, поскольку может быть основой для некачественного описания таксонов без приведения их реальных диагностических признаков.

Интуиция основана на многолетнем опыте, в ходе которого мозг систематика научается принимать во внимание множество мелких деталей, учет которых позволяет принять таксономическое решение, логику которого трудно описать словами. В принципе алгоритм, на основании которого действует интуиция (анализ и систематизация большого числа мелких деталей у большого числа особей), может быть основой компьютерной программы. В настоящее время уже существуют компьютерные методы видовой идентификации, основанные на машинном обучении (то, что иначе называется искусственным интеллектом) компьютерных программ распознавать виды по их дигитализированным изображениям (Wäldchen, Mäder, 2018).

Можно предполагать, что появление таких или аналогичных систем для решения более сложной и творческой задачи делимитации видов – дело недалекого будущего.

ЗАКЛЮЧЕНИЕ

В настоящее время с проблемой делимитации видов сложилась парадоксальная ситуация. С одной стороны, есть много работ по теории делимитации на основании использования молекулярных маркеров. Однако авторы этих работ, как правило, сами практической делимитацией не занимаются, в том смысле, что они не дают реальной таксономической интерпретации полученных молекулярных кластеров (Hebert et al., 2004; Brower, 2006, 2010; Carstens et al., 2013).

С другой стороны, имеется масса практикующих систематиков, которые ежедневно занимаются этой работой, и многие даже используют для этого молекулярные маркеры, чаще всего в виде митохондриальных ДНК-баркодов. Однако они в подавляющем большинстве случаев не только не используют эти разработанные методы делимитации, но и даже не подозревают об их существовании.

Таксономическая пассивность первой группы ученых понятна и оправдана. Во-первых, описание новых видов требует навыков и квалификации, для получения которых нужны годы, и которых, как правило, нет у теоретиков молекулярной делимитации. Во-вторых, и это самое главное, все алгоритмы молекулярной делимитации дают исходно клады (кластеры особей), которые не обязательно идентичны реальным видам, существующим в природе. Пассивность второй группы ученых в освоении новых теорий кажется совершенно неоправданной, так как она лишает их мощного инструмента, которым они (и только они!) могут эффективно пользоваться.

Создается впечатление, что молекулярные работы по делимитации видов, даже те, которые основаны на полногеномных данных и оперируют сотнями и тысячами маркеров, не дают окончательного ответа на вопрос о видовых границах и о статусах таксонов (вид или подвид) (O'Connell, Smith, 2018). Но они дают необычайно важный

материал, ценность которого многократно возрастает, если полученные данные коррелируют (или не коррелируют) с информацией по географическому распространению, морфологии и экологическим предпочтениям.

Зачастую только практикующий систематик, досконально знающий свою группу, может правильно интерпретировать эти молекулярные данные и вставить их в систему уже существующего знания для принятия таксономического решения.

БЛАГОДАРНОСТИ

Я выражаю благодарность А. О. Вершининой (University of California, Santa Cruz, California, U.S.A.), М. С. Вишневецкой (СПбГУ), А. В. Данченко (МГУ), Е. А. Паженовой (СПбГУ) и Н. А. Шаповалу (ЗИН) за помощь в проведении исследований и обсуждение статьи.

ФИНАНСИРОВАНИЕ

Работа выполнена в рамках государственной темы АААА-А19-119020790106-0 (разделы «Делимитация видов в домолекулярную эпоху» и «Делимитация видов в молекулярную эпоху») и при финансовой поддержке проектов Российского фонда фундаментальных исследований № 18-04-00263а (разделы «Методы, основанные на дистанциях» и «Методы, основанные на филогенетическом анализе молекулярных признаков»), № 17-04-00828 (разделы «Методы делимитации, основанные на моделях» и «Методы, основанные на валидации предварительно выдвинутых видовых гипотез»), № 17-04-00754 (раздел «Таксономическая интуиция и возможности искусственного интеллекта») и Российского научного фонда № 19-14-00202 (разделы «Концепции видов» и «Анализ критического видового разнообразия»).

СПИСОК ЛИТЕРАТУРЫ

- Лухтанов В. А. 2013. Принципы реконструкции филогенезов: признаки, модели эволюции и методы филогенетического анализа. В кн.: А. Ф. Алимов, С. Д. Степаньянц (ред.). Современные проблемы биологической систематики. СПб.: Зоологический институт РАН, с. 39–52. (Труды Зоологического института РАН, приложение № 2).
- Лухтанов В. А., Кузнецова В. Г. 2009. Молекулярно-генетические и цитогенетические подходы к проблемам видовой диагностики, систематики и филогенетики. Журнал общей биологии **70** (5): 415–437.
- Лухтанов В. А., Шаповал Н. А. 2008. Выявление симпатрично обитающих видов-двойников бабочек из комплекса *Agrodiaetus kendevani* (Lepidoptera, Lycaenidae) с помощью популяционного анализа несцепленных генетических маркеров. Доклады Академии наук **423** (3): 421–426.
- Майр Э. 1968. Зоологический вид и эволюция. М.: Мир, 598 с.
- Майр Э. 1971. Принципы зоологической систематики. М.: Мир, 454 с.
- Avise J. C. 2000. Phylogeography: the History and Formation of Species. Cambridge–London: Harvard University Press, 447 p.
- Avise J. C. 2004. Molecular Markers, Natural History and Evolution. Sunderland; Massachusetts: Sinauer Associates, 684 p.
- Bensasson D., Zhang D. X., Hartl D. L., Hewitt G. M. 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. Trends in Ecology and Evolution **16** (6): 314–321. doi: 10.1016/S0169-5347(01)02151-6
- Brower A. V. Z. 2006. Problems with DNA barcodes for species delimitation: 'ten species' of *Astraptus fulgerator* reassessed (Lepidoptera: Hesperidae). Systematics and Biodiversity **4**: 127–132.
- Brower A. V. Z. 2010. Alleviating the taxonomic impediment of DNA barcoding and setting a bad precedent: names for ten species of '*Astraptus fulgerator*' (Lepidoptera: Hesperidae: Eudaminae) with DNA-based diagnoses. Systematics and Biodiversity **8** (4): 485–491.
- Camargo A., Morando M., Avila L. J., Sites J. W. 2012. Species delimitation with ABC and other coalescent-based methods: a test of accuracy with simulations and an empirical example with lizards of the *Liolaemus darwini* complex (Squamata: Liolaemidae). Evolution **66** (9): 2834–2849.
- Cameron S. 2014. How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. Systematic Entomology **39** (3): 400–411. doi: 10.1111/syen.12071

- Carstens B. C., Pelletier T. A., Reid N. M., Satler J. D. 2013. How to fail at species delimitation. *Molecular Ecology* **22**: 4369–4383. doi: 10.1111/mec.12413
- Coyne J. A., Orr H. A. 2004. *Speciation*. Sunderland, Massachusetts: Sinauer Associates, 545 p.
- Cracraft J. 1989. Speciation and its ontology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In: D. Otte, J. Endler (eds). *Speciation and its Consequences*. Sunderland: Sinauer Associates, pp. 28–59.
- De Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56** (6): 879–886. doi: 10.1080/10635150701701083
- Dincă V., Lukhtanov V. A., Talavera G., Vila R. 2011. Unexpected layers of cryptic diversity in wood white *Leptidea* butterflies. *Nature Communications* **2**, 324. doi: 10.1038/ncomms1329
- Dincă V., Wiklund C., Lukhtanov V. A., Kodandaramaiah U., Norén N., Dapporto L., Wahlberg N., Vila R., Friberg M. 2013. Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex. *Journal of Evolutionary Biology* **26** (10): 2095–2106. doi: 10.1111/jeb.12211
- Galtier N., Nabholz B., Glémin S., Hurst G. D.D. 2009. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology* **18** (22): 4541–4550. doi: 10.1111/j.1365-294X.2009.04380.x
- Hajibabaei M., Janzen D. H., Burns J. M., Hallwachs W., Hebert P. D. N. 2006. DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 968–971.
- Hebert P. D. N., Cywinska A., Ball S. L., deWaard J. R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* **270**: 313–321.
- Hebert P. D. N., Gregory T. R. 2005. The promise of DNA barcoding for taxonomy. *Systematic Biology* **54**: 852–859. <https://doi.org/10.1080/10635150500354886>
- Hebert P. D. N., Penton E. H., Burns J. M., Janzen D. H., Hallwachs W. 2004. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrartes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* **101** (41): 14812–14817. doi: 10.1073/pnas.0406166101
- Hernández-Roldán J. L., Dapporto L., Dincă V., Vicente J. C., Hornett E. A., Šichová J., Lukhtanov V. A., Talavera G., Vila R. 2016. Integrative analyses unveil speciation linked to host plant shift in *Spialia* butterflies. *Molecular Ecology* **25** (17): 4267–4284. doi: 10.1111/mec.13756
- Isaac N. J. B., Mallet J., Mace G. M. 2004. Taxonomic inflation: its influence on macroecology and conservation. *Trends in Ecology and Evolution* **19**: 464–469.
- Larsen B. B., Miller E. C., Rhodes M. K., Wiens J. J. 2017. Inordinate fondness multiplied and redistributed: The number of species on earth and the new pie of life. *Quarterly Review of Biology* **92** (3): 229–265. <https://doi.org/10.1086/693564>
- Leache A. D., Fujita M. K., Minin V. N., Bouckaert R. R. 2014. Species delimitation using genome-wide SNP data. *Systematic Biology* **63** (4): 534–542. doi:10.1093/sysbio/syu018
- Linnaeus C. 1758. *Systema Naturae*. Tomus I. Editio Decima, Reformata. Holmiae: Laurentii Salvii, 824 p.
- Lorković Z. 1993. *Leptidea reali* Reissinger, 1989 (= *lorkovicii* Real, 1988), a new European species (Lepid., Pieridae). *Natura Croatica* **2**: 1–26.
- Lukhtanov V. A., Kandul N. P., Plotkin J. B., Dantchenko A. V., Haig D., Pierce N. E. 2005. Reinforcement of prezygotic isolation and karyotype evolution in *Agrodiaetus* butterflies. *Nature* **436**: 385–389.
- Lukhtanov V. A., Dantchenko A. V. 2017. A new butterfly species from south Russia revealed through chromosomal and molecular analysis of the *Polyommatus (Agrodiaetus) damonides* complex (Lepidoptera, Lycaenidae). *Comparative Cytogenetics* **11** (4): 769–795. <https://doi.org/10.3897/CompCytogen.v11i4.20072>
- Lukhtanov V. A., Shapoval N. A. 2017. Chromosomal identification of cryptic species sharing their DNA barcodes: *Polyommatus (Agrodiaetus) antidolus* and *P. (A.) morgani* in Iran (Lepidoptera, Lycaenidae). *Comparative Cytogenetics* **11** (4): 759–768. doi: 10.3897/compcytogen.v11i4.20876
- Lukhtanov V. A., Vila R., Kandul N. P. 2006. Rearrangement of the *Agrodiaetus dolus* species group (Lepidoptera, Lycaenidae) using a new cytological approach and molecular data. *Insect Systematics and Evolution* **37** (3): 325–334.
- Lukhtanov V. A., Dantchenko A. V., Vishnevskaya M. S., Saifitdinova A. F. 2015. Detecting cryptic species in sympatry and allopatry: analysis of hidden diversity in *Polyommatus (Agrodiaetus)* butterflies (Lepidoptera: Lycaenidae). *Biological Journal of the Linnean Society* **116** (2): 468–485. doi: 10.1111/bij.12596
- Lukhtanov V. A., Dincă V., Friberg M., Šichová J., Olofsson M., Vila R., Marec F., Wiklund C. 2018. Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. *Proceedings of the National Academy of Sciences of the United States of America* **115** (41): E9610–E9619. <https://doi.org/10.1073/pnas.1802610115>
- Lukhtanov V. A., Shapoval N. A., Anokhin B. A., Saifitdinova A. F., Kuznetsova V. G. 2015b. Homoploid hybrid speciation and genome evolution via chromosome sorting. *Proceedings of the Royal Society B: Biological Sciences* **282** (1807): 20150157. doi:10.1098/rspb.2015.0157

- Lukhtanov V. A., Shapoval N. A., Dantchenko A. V. 2008. *Agrodiaetus shahkhusensis* sp. n. (Lepidoptera, Lycaenidae), a cryptic species from Iran discovered by using molecular and chromosomal markers. *Comparative Cytogenetics* **2** (2): 99–114.
- Lukhtanov V. A., Shapoval N. A., Dantchenko A. V. 2014. Taxonomic position of several enigmatic *Polyommatus* (*Agrodiaetus*) species (Lepidoptera, Lycaenidae) from Central and Eastern Iran: insights from molecular and chromosomal data. *Comparative Cytogenetics* **8** (4): 313–322. doi: 10.3897/CompCytogen.v8i4.8939
- Lukhtanov V. A., Sourakov A., Zakharov E. V. 2016. DNA barcodes as a tool in biodiversity research: testing pre-existing taxonomic hypotheses in Delphic Apollo butterflies (Lepidoptera, Papilionidae). *Systematics and Biodiversity* **14** (6): 599–613. doi: 10.1080/14772000.2016.1203371
- Lukhtanov V. A., Sourakov A., Zakharov E. V., Hebert P. D. N. 2009. DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification. *Molecular Ecology Resources* **9**: 1302–1310. doi: 10.1111/j.1755-0998.2009.02577.x
- Mallet J. A. 1995. Species definition for the modern synthesis. *Trends in Ecology and Evolution* **10**: 294–299.
- Mallet J., Willmott K. 2003. Taxonomy: Renaissance or Tower of Babel? *Trends in Ecology and Evolution* **18**: 57–59.
- Nichols R. 2001. Gene trees and species trees are not the same. *Trends in Ecology and Evolution* **16**: 358–364.
- O'Connell K. A., Smith E. N. 2018. The effect of missing data on coalescent species delimitation and a taxonomic revision of whipsnakes (Colubridae: *Masticophis*). *Molecular Phylogenetics and Evolution* **127**: 356–366. <https://doi.org/10.1016/j.ympev.2018.03.018>
- O'Meara B. C. 2010. New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology* **59**: 59–73.
- Pazhenkova E. A., Lukhtanov V. A. 2019. Nuclear genes (but not mitochondrial DNA barcodes) reveal real species: evidence from the *Brenthis fritillaria* butterflies (Lepidoptera, Nymphalidae). *Journal of Zoological Systematics and Evolutionary Research* **57** (2): 298–313. doi: 10.1111/jzs.12252
- Pons J., Barraclough T. G., Gomez-Zurita J., Cardoso A., Duran D. P., Hazell S., Kamoun S., Sumlin W. D., Vogler A. P. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55** (4): 595–609. doi: 10.1080/10635150600852011
- Ratnasingham S., Hebert P. D. N. 2013. A DNA-based registry for all animal species: The Barcode Index Number (BIN) system. *PLoS ONE* **8** (8): e66213. doi: 10.1371/journal.pone.0066213
- Ritter S., Michalski S. G., Settele J., Wiemers M., Fric Z. F., Sielezniew M., Šašić M., Rozier Y., Durka W. 2013. *Wolbachia* infections mimic cryptic speciation in two parasitic butterfly species, *Phengaris teleius* and *P. nausithous* (Lepidoptera: Lycaenidae). *PLoS ONE* **8** (11), e78107. <https://doi.org/10.1371/journal.pone.0078107>
- Rosser N., Kozak K. M., Phillimore A. B., Mallet J. 2015. Extensive range overlap between heliconiine sister species: evidence for sympatric speciation in butterflies? *BMC Evolutionary Biology* **15**, 125. doi: 10.1186/s12862-015-0420-3
- Shapoval N. A., Lukhtanov V. A. 2015. Intragenomic variations of multicopy *ITS2* marker in *Agrodiaetus* blue butterflies (Lepidoptera, Lycaenidae). *Comparative Cytogenetics* **9** (4): 483–497. doi: 10.3897/CompCytogen.v9i4.5429
- Stork N. E., McBroom J., Gely C., Hamilton A. J. 2015. New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 7519–7523.
- Talavera G., Lukhtanov V. A., Rieppel L., Pierce N. E., Vila R. 2013. In the shadow of phylogenetic uncertainty: the recent diversification of *Lysandra* butterflies through chromosomal change. *Molecular Phylogenetics and Evolution* **69**: 469–478. <http://dx.doi.org/10.1016/j.ympev.2013.08.004>
- Todisco V., Grill A., Fiedler K., Gottsberger B., Dincă V., Vodă R., Lukhtanov V., Letsch H. 2018. Molecular phylogeny of the Palaearctic butterfly genus *Pseudophilotes* (Lepidoptera: Lycaenidae) with focus on the Sardinian endemic *P. barbaggiae*. *BMC Zoology* **3**: 4. <https://doi.org/10.1186/s40850-018-0032-7>
- Toews D., Brelford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21** (16): 3907–3930. doi: 10.1111/j.1365-294X.2012.05664.x
- Vershinina A. O., Lukhtanov V. A. 2010. Geographical distribution of the cryptic species *Agrodiaetus alcestis alcestis*, *A. alcestis karacetinae* and *A. demavendi* (Lepidoptera, Lycaenidae) revealed by cytogenetic analysis. *Comparative Cytogenetics* **4** (1): 1–11.
- Vila R., Lukhtanov V. A., Talavera G., Gil-T F., Pierce N. E. 2010. How common are dot-like distribution ranges? Taxonomical oversplitting in Western European *Agrodiaetus* (Lepidoptera, Lycaenidae) revealed by chromosomal and molecular markers. *Biological Journal of Linnean Society* **101**: 130–154.
- Vishnevskaya M. S., Saifitdinova A. F., Lukhtanov V. A. 2016. Karyosystematics and molecular taxonomy of the anomalous blue butterflies (Lepidoptera, Lycaenidae) from the Balkan Peninsula. *Comparative Cytogenetics* **10** (5): 1–85. doi: 10.3897/CompCytogen.v10i5.10944.

- Vrtilek M., Reichard M. 2016. Patterns of morphological variation among populations of the widespread annual killifish *Nothobranchius orthonotus* are independent of genetic divergence and biogeography. *Journal of Zoological Systematics and Evolutionary Research* **54** (4): 289–298. doi: 10.1111/jzs.12134
- Wäldchen J., Mäder P. 2018. Machine learning for image based species identification. *Methods in Ecology and Evolution* **9**: 2216–2225. doi: 10.1111/2041-210X.13075
- Wiemers M., Fiedler K. 2007. Does the DNA barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology* **4**, 8. doi: 10.1186/1742-9994-4-8
- Yang Z., Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 9264–9269.
- Zhang Z.-Q. 2013. (ed.). *Animal biodiversity: an outline of higher-level classification and survey of taxonomic richness* (Addenda 2013). *Zootaxa* **3703**: 1–82.

SPECIES DELIMITATION AND ANALYSIS OF CRYPTIC SPECIES DIVERSITY IN THE 21ST CENTURY

V. A. Lukhtanov

Key words: species, taxonomy, insects, phylogenetic species concept, biological species concept, genetic distance, phylogenetic analysis, molecular character, monolocus data, multi-locus data, whole genome data, machine learning.

SUMMARY

At present, there is a paradoxical situation with the problem of species delimitation. On the one hand, there are many papers on the theory of delimitation based on the use of molecular markers. However, the authors of these works, as a rule, do not engage in practical delimitation, in the sense that they do not provide a real taxonomic interpretation of the molecular clusters obtained.

On the other hand, there are practicing taxonomists who do this work every day, and many even use molecular markers for this, most often in the form of mitochondrial DNA barcodes. However, in the overwhelming majority of cases, they not only do not use these developed methods of delimitation, but they are not even aware of their existence.

Taxonomic passivity of the first group of scientists is clear. First, the description of new species requires skills and qualifications, and theorists of molecular delimitation do not usually have these skills. Second, and most importantly, all molecular delimitation algorithms result initially in clades (clusters of individuals), which are not necessarily identical to the real species that exist in nature. But the passivity of the second group of scientists in the using the molecular delimitation algorithms seems completely unjustified, since it deprives them of a powerful tool that they (and only they!) can effectively use.

It seems that molecular markers, even those based on whole-genome data and operating with hundreds and thousands of SNPs, do not provide a definitive answer to the question of species boundaries and taxon status (species / subspecies). However, they provide very important data, the value of which increases manifold if these data correlate (or do not correlate) with information on geographical distribution, morphology and ecological preferences.

Often, only a practicing taxonomist who knows his or her group thoroughly can correctly interpret this molecular information and insert it into the system of already existing knowledge to make a taxonomic decision.