

УДК 575.858

МЕТОДЫ ВЫЯВЛЕНИЯ ГИБРИДИЗАЦИИ И ГЕНЕТИЧЕСКОЙ ИНТРОГРЕССИИ ПРИ ФИЛОГЕНЕТИЧЕСКИХ НЕСООТВЕТСТВИЯХ

© 2021 г. В. А. Комарова¹, *, Л. А. Лавренченко¹, **

¹Институт проблем экологии и эволюции им. А.Н. Северцова РАН
Ленинский просп., 33, Москва, 119071 Россия

*E-mail: k_vakka@yahoo.com

**E-mail: llavrenchenko@gmail.com

Поступила в редакцию 30.07.2021 г.

После доработки 04.08.2021 г.

Принята к публикации 26.08.2021 г.

Интрогрессивная гибридизация представляет собой широко распространенное явление в эволюции растений и животных. Заключение о межвидовом потоке генов часто делают при выявлении несогласованности между топологиями генных деревьев. Однако обнаруженное филогенетическое несоответствие может быть вызвано и другими эволюционными процессами, такими как неполная сортировка линий. В таком случае обнаружение и количественная характеристика интрогрессии представляются весьма непростой задачей. В настоящем обзоре дана краткая характеристика основных процессов (гибридизации и интрогрессии, неполной сортировки линий), которые ответственны за несогласованность между топологиями генных деревьев. Обобщены сведения о классических филогенетических и аналитических методах, позволяющих определять и отличать гибридизацию и интрогрессию от неполной сортировки линий при сравнении филогенетических деревьев, построенных на основе отдельных генетических маркеров. Рассмотрены методы, разработанные не только для определения паттернов интрогрессии в присутствии неполной сортировки линий, но и для количественной оценки обнаруженной интрогрессии (доля генетического смешения, направление, время), кратко рассмотрены преимущества и методические ограничения каждого из них. Также рассмотрены методы филогенетических сетей, представляющиеся неотъемлемым инструментом для реконструкции сложных эволюционных историй в присутствии ретикулярных событий (гибридизация, интрогрессия и другие).

DOI: 10.31857/S0044459621060051

Интрогрессивная гибридизация (introgressive hybridization; Anderson, Hubricht, 1938), или интрогрессия (introgression), представляет собой постепенную инфильтрацию генов одного вида в генный пул другого путем повторных возвратных скрещиваний гибридов первого или последующих поколений с одним или обоими родительскими видами (backcrossing). Стоит отметить, что отдельные исследователи нередко по-разному трактовали одни и те же термины, что становилось источником разногласий (например, Anderson, Hubricht, 1938; Anderson, 1949; Heiser, 1973). Так, термин “интрогрессия” нередко понимали в широком смысле, как обмен генами между любыми дифференцированными популяциями, независимо от их таксономического ранга (Anderson, 1949; Schwenk, Spaak, 1995), и применяли при описании различных событий (Rieseberg, Wendel, 1993; Schwenk, Spaak, 1995). В целом феномен интрогрессии можно описать цепочкой последовательных событий. Так, на первом этапе в ходе на-

чального расселения вида возникает неоднородность между его локальными популяциями в пределах ареала. Вторичная экспансия может привести к контакту между этими популяциями, дивергировавшими во время длительной географической изоляции (например, в периоды оледенений, когда ареалы адаптированных к умеренным широтам видов уменьшались; Боркин, Литвинчук, 2013). Если между контактирующими популяциями не выработалась надежная репродуктивная изоляция, то такие контакты могли приводить к гибридизации, одним из последствий которой могла быть интрогрессия отдельных участков ядерного или всего митохондриального генома через серию возвратных скрещиваний в популяции родительских видов (Peters et al., 2007; Arnold, Martin, 2009). Следует добавить, что проницаемость видовых границ различается для разных участков генома. Так, известно, что у организмов, характеризующихся большим уровнем дисперсии гетерогаметного пола, маркеры, свя-

занные с гомогаметным полом (у млекопитающих – митохондриальная ДНК (мтДНК)) значительно легче преодолевают межвидовые барьеры (Petit, Excoffier, 2009). Обнаружены случаи, когда формоспецифические (исходные) гаплотипы мтДНК полностью замещаются чужими гаплотипами при отсутствии интрогрессии ядерных генов (Nevado et al., 2009; Choleva et al., 2014). Против интрогрессии ядерных локусов может действовать выметающий отбор, поскольку заимствование чужеродных ядерных генов может влиять на экологические и морфологические особенности особей (Боркин, Литвинчук, 2013). Тем не менее были обнаружены случаи, свидетельствующие о потоке ядерных генов при отсутствии интрогрессии мтДНК (Di Candia, Routman, 2007).

Межвидовая гибридизация давно привлекала внимание эволюционных биологов. Этому сложному процессу как фактору эволюции большое внимание уделяли ботаники (Anderson, 1949; Heiser, 1973), в то время как отношение зоологов к этому процессу представлялось неоднозначным (Майр, 1968). Однако с интенсивным развитием молекулярных методов в конце XX в. стали накапливаться примеры, свидетельствующие о широком распространении межвидовой интрогрессии среди животных в природе (Arnold, 2006; Toews, Brelsford, 2012). Это неизбежно привело к переоценке роли межвидовой интрогрессии в эволюционной истории живых организмов. В частности, было предположено, что интрогрессия генов может быть источником генетической и фенотипической изменчивости (Arnold, 2006) и способствовать приобретению новых адаптаций (Arnold et al., 2008; Arnold, Martin, 2009; Hedrick, 2013).

Несмотря на то, что межвидовая гибридизация между близкими видами в природе происходит достаточно часто (затрагивает 25% видов растений и около 10% видов животных; Mallet, 2005; Arnold et al., 2008; Mallet et al., 2016), накоплено относительно мало экспериментальных подтверждений межвидовой интрогрессии, в частности, в пользу ее адаптивного характера (приобретение новых адаптаций, увеличение приспособленности таксона-реципиента к изменяющимся условиям среды; например, Whitney et al., 2010). Настоящий прорыв в изучении этого сложного процесса произошел относительно недавно, благодаря прочтению методами высокопроизводительного секвенирования (next generation sequencing, NGS) полных геномов целого ряда “немодельных” видов (Twyford, Ennos, 2012; Hufford et al., 2013; Harrison, Larson, 2014; Mallet et al., 2016; Schaefer et al., 2016; Arnold, Kunte, 2017; Ottenburghs et al., 2017; Rochus et al., 2018; Suarez-Gonzalez et al., 2018; Taylor, Larson, 2019). Так, благодаря использованию геномных данных, в ряде исследований были получены свидетельства и подтверждения интрогрессии среди многих групп организмов: у людей (Reich

et al., 2009; Green et al., 2010; Huerta-Sánchez et al., 2014), насекомых (Lynd et al., 2010; Dasmahapatra et al., 2012; Fontaine et al., 2015), других животных (Shurtliff, 2013; Meier et al., 2017; Palkopoulou et al., 2018; Irisarri et al., 2018), растений (Osborne et al., 2016; Suarez-Gonzalez et al., 2018).

Интрогрессивная гибридизация приводит к образованию популяций со смешанными геномами и, как следствие, к филогенетическому несоответствиям (phylogenetic discordance, tree incongruence) при анализе различных генетических локусов (Degnan, Rosenberg, 2009). Однако интрогрессия – не единственный процесс, приводящий к подобному филогенетическому несоответствию. Другие процессы, такие как неполная сортировка линий (incomplete lineage sorting, ILS) или дупликация генов (gene duplication), могут также соответствовать аналогичным паттернам (Maddison, 1997; Degnan, Rosenberg, 2009). Таким образом, при обнаружении несоответствий в топологиях генетических деревьев определение причины наблюдаемого паттерна по-прежнему остается нетривиальной задачей (Sang, Zhong, 2000; Fontaine et al., 2015). При использовании традиционных филогенетических подходов и аналитических методов, основанных на независимом исследовании отдельных локусов, возможно предпочтение того или иного эволюционного сценария (Maddison, 1997; Sang, Zhong, 2000; Peters et al., 2007; Joly et al., 2009; Choleva et al., 2014; Poroshina et al., 2020). Однако при использовании данных подходов зачастую невозможна точная оценка процессов, приводящих к несогласованности генных деревьев. Разработанные для этой цели методы, базирующиеся на анализе большого числа генетических локусов и данных всего генома, позволяют эффективно выявлять случаи несоответствия топологии филогенетических деревьев и давать количественную оценку интрогрессии (Pritchard et al., 2000; Alexander et al., 2009; Green et al., 2010; Durand et al., 2011; Huerta-Sánchez et al., 2014; Lohse, Frantz, 2014; Schrider et al., 2018; Edelman et al., 2019; Hamlin et al., 2020), а также определять направление наблюдаемой интрогрессии (Pease, Hahn, 2015). Однако каждый из таких методов задает свои определенные условия для структуры исследуемой группы (например, постоянный размер популяций, большой объем выборки, эволюция последовательностей согласно модели молекулярных часов и др.), что затрудняет возможность их широкого применения.

Несмотря на многообразие накопившихся в последние годы обзоров, посвященных свидетельствам и эволюционному значению межвидовой интрогрессии у различных групп организмов (Toews, Brelsford, 2012; Hedrick, 2013; Harrison, Larson, 2014; Ottenburghs et al., 2017), обобщающих сводок по методам ее обнаружения при наблюдении несоответствий топологий генетиче-

ских деревьев крайне мало (Burgarella et al., 2019). Первая часть настоящего обзора посвящена традиционным филогенетическим и аналитическим методам, разработанным специально для определения и разграничения типов несоответствий филогений на основе отдельных генетических маркеров. Во второй части обобщены накопленные к настоящему моменту сведения о методах, разработанных для поиска и характеристики интрогрессии с использованием геномных данных. В заключение рассмотрены методы филогенетических сетей, представляющиеся неотъемлемым инструментом для реконструкции сложных эволюционных историй в присутствии ретикулярных событий. Кратко рассмотрены преимущества и методические ограничения каждого из них, поскольку понимание основных принципов и ограничений алгоритмов – обязательное условие для грамотной постановки научной задачи.

ПРИЧИНЫ НЕСООТВЕТСТВИЙ ФИЛОГЕНЕТИЧЕСКИХ ДЕРЕВЬЕВ И МЕТОДЫ ИХ РАЗГРАНИЧЕНИЯ НА ОСНОВЕ ИССЛЕДОВАНИЯ ОТДЕЛЬНЫХ ЛОКУСОВ

Нередко на уровне интерпретации молекулярных деревьев (gene tree) приходится сталкиваться с “подводными камнями”, затрудняющими реконструкцию истинной эволюционной истории вида. Наиболее известные из них представляют собой случаи несовпадения результатов, полученных при анализе различных генетических локусов, чаще всего это несоответствие митохондриальной и ядерной филогений (mito-nuclear discordance), а также несоответствие генного дерева видовому (gene tree discordance) (рис. 1). Несовпадение может выражаться не только в отношении времени дивергенции, но и в порядке ветвления (Avice, 2000; Nichols, 2001; Rosenberg, 2002). Прямолинейная интерпретация топологий таких деревьев, без привлечения дополнительных анализов, приводит к серьезным искажениям реконструкции филогении исследуемых таксонов (Абрамсон, 2007, 2009). Подобная проблема несоответствий филогений отдельных генов и видов известна для многих групп организмов (Gompert et al., 2008; Toews, Brelsford, 2012; Weigand et al., 2017; Ivanov et al., 2018; Després, 2019; Edelman et al., 2019; Andersen et al., 2021; Komarova et al., 2021). Интрогрессивную гибридизацию часто рассматривают как основную причину подобных различий (Degnan, Rosenberg, 2009). Однако и другие эволюционные процессы могут приводить к формированию сходных филогенетических паттернов. К наиболее частым причинам, препятствующим реконструкции истинной эволюционной истории вида, относятся предковый полиморфизм (ancestral polymorphism) и неполная сорти-

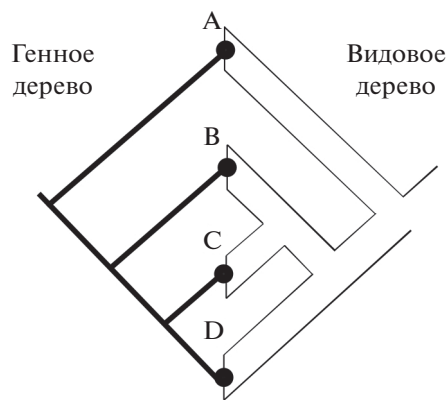


Рис. 1. Несовпадение генных и видовых деревьев. Справа представлено видовое дерево для четырех видов: А, В, С и D; слева – генное дерево. Виды В и С представляют собой сестринские виды, но их генные копии таковыми не являются.

ровка генеалогических линий (ILS) (Avice, 2000; Rosenberg, 2002; Degnan, Rosenberg, 2009; Degnan, 2018). Показано, что феномен ILS широко распространен среди живых организмов в природе (например, Fontaine et al., 2015; Pease et al., 2016; Chan et al., 2020). Вследствие всех перечисленных выше эволюционных сценариев мы сталкиваемся с полифилией (polyphyly) на генных деревьях, когда генетические последовательности одного вида расположены внутри клад другого (сходство в генетических последовательностях “сестринских” видов в случае интрогрессии будет связано с полным замещением общих для одного вида аллелей аллелями другого вида). В таком случае распознать и отличить полифилию, возникшую вследствие интрогрессии, от полифилии двух других вышеописанных сценариев весьма непросто (Holder et al., 2001), особенно в случае “молодых” видов (события интрогрессии у таких видов могут происходить в присутствии ILS) (Avice, 2000; Degnan, Rosenberg, 2009; Degnan, 2018). Следовательно, филогенетическое несоответствие следует рассматривать не как свидетельство интрогрессии, а скорее как отправную точку для дальнейшего анализа. Существует ряд традиционных филогенетических и аналитических подходов, позволяющих отдавать предпочтение тому или иному эволюционному сценарию при сравнении филогенетических деревьев, реконструированных на основе анализа отдельных генов.

Согласно базовым положениям молекулярной филогенетики количество нуклеотидных замен в последовательностях ДНК двух форм прямо пропорционально времени дивергенции от общего предка, а филогения отдельных генов соответствует филогении видов (Nichols, 2001). Принимая это во внимание, следует учитывать, что время дивергенции генетических последовательностей от об-

шего предка может быть больше времени дивергенции видов, другими словами, разделение популяций происходит позже, чем разделение генов (Nichols, 2001). Поэтому одним из путей, способных помочь при разграничении ILS и интрогрессии, может быть оценка предполагаемого возраста генетического обособления вида. Учитывая тот факт, что выраженность ILS обратно пропорциональна эволюционному возрасту вида, можно предполагать, что чем “моложе” вид, тем более вероятно, что различие в топологиях генетических деревьев объясняется скорее ILS, чем интрогрессией (Avice, 2000). В случае “молодых” видов сортировка предковых линий еще не успеет завершиться за этот промежуток времени (Avice, 2000). В случае интрогрессии, наоборот, степень ее выраженности пропорциональна эволюционному возрасту вида (Avice, 2000). Стоит также добавить, что при анализе митохондриальных генов наиболее вероятно, что наблюдаемая полифилия будет вызвана скорее интрогрессией, чем ILS, поскольку сортировка линий по митохондриальным локусам происходит быстрее, чем по ядерным, благодаря своей гаплоидности (наследование только по материнской линии) и отсутствию рекомбинации (Hudson, Turelli, 2003).

Некоторые существующие методы основаны на подсчете времени дивергенции исследуемых таксонов, так, один тест (Sang, Zhong, 2000) разработан специально для разграничения процессов гибридизации и интрогрессии от событий ILS, вызывающих несоответствие между генетическими деревьями. Если предположить сценарий с существованием трех таксонов А, В и С, где таксон В будет сестринским по отношению к таксону С на одном филогенетическом дереве, в то время как на другом дереве этот же таксон В будет сестринским таксону А, то при наблюдаемом несоответствии филогений можно ожидать, что таксон В имеет гибридное происхождение (рис. 2а). Для подтверждения данного предположения было предложено оценивать время дивергенции согласно модели молекулярных часов для таксонов А и С на каждом из исследуемых деревьев. Если обе временные оценки практически не отличаются, то предпочтение отдается гибридизации, интрогрессии или горизонтальному переносу генов, в противном случае – ILS (рис. 2а). Однако этот тест не учитывает, что время “слияния” для разных генов, как правило, всегда различается независимо от наличия ILS. Таким образом, тест, основанный на данной теоретической модели, ненадежен при разграничении этих процессов при сравнении филогений на основе отдельных генов, что было наглядно продемонстрировано (Holder et al., 2001).

Согласно результатам этого исследования (Holder et al., 2001), процессы гибридизации и неполной сортировки линий формируют разные

прогнозы относительно топологии и длины ветвей генных деревьев, которые эволюционируют в соответствии с филогенией видов. Если предположить сценарий полифилии, при котором сходные последовательности “сливаются” до диверсификации двух видов, то наблюдаемое несоответствие можно объяснить неполной сортировкой линий. Следовательно, при ILS время, прошедшее с момента видообразования, представляет собой нижнюю границу минимального времени расхождения между этими последовательностями. При гибридизации сходные последовательности могут “сливаться” как до, так и после расхождения видов. Если “слияние” происходит после расхождения видов, следует ожидать, что минимальное время дивергенции между последовательностями будет меньше, чем при неполной сортировке линий (Joly, Bruneau, 2006). В некоторых статистических методах данный подход применяется, чтобы отличить гибридизацию и интрогрессию от неполной сортировки линий. Например, статистический параметрический метод (Joly et al., 2009), предполагая ILS в качестве нулевой гипотезы, вычисляет минимальные попарные дистанции между последовательностями исследуемых видов при помощи коалесцентного моделирования. Данный подход основан на предположении о том, что минимальная генетическая дистанция между последовательностями двух видов должна быть меньше в случае событий гибридизации, чем ILS.

Еще один метод, основанный на оценке скорректированных эволюционных дистанций (d) между общим предком каждого из исследуемых видов (L) и общим предком для обоих видов (CA), позволяет отличать интрогрессию от ILS (Pogoshina et al., 2020). Алгоритм данного подхода довольно прост и включает ряд последовательных действий: использование файлов с построенными филогенетическими деревьями в качестве входных данных, проведение теста на взаимную монофилию (подсчет коэффициента P ; при $P = 0$ группа монофилетична) исследуемых групп и подсчет набора эволюционных дистанций (d , D , $E\Sigma$) (рис. 2б). Так, в случае анцестрального полиморфизма разница D между da и db (где a и b соответствуют исследуемым видам, d – дистанция от CA до L соответствующего вида) будет равно нулю ($D = 0$) и незначительно отличаться от последнего при неполной сортировке линий. В случае же интрогрессии, наоборот, значение D будет значительно отличаться от нуля (рис. 2б). Параметр $E\Sigma$, представляющий сумму эволюционных дистанций всех последовательностей внутри одного вида до ближайшего общего предкового узла дерева (внешней группы), также важен при определении интрогрессии. Так, значение разницы между Ea и Eb , посчитанное для митохондриального гена и значительно отличное от нуля, соответствует

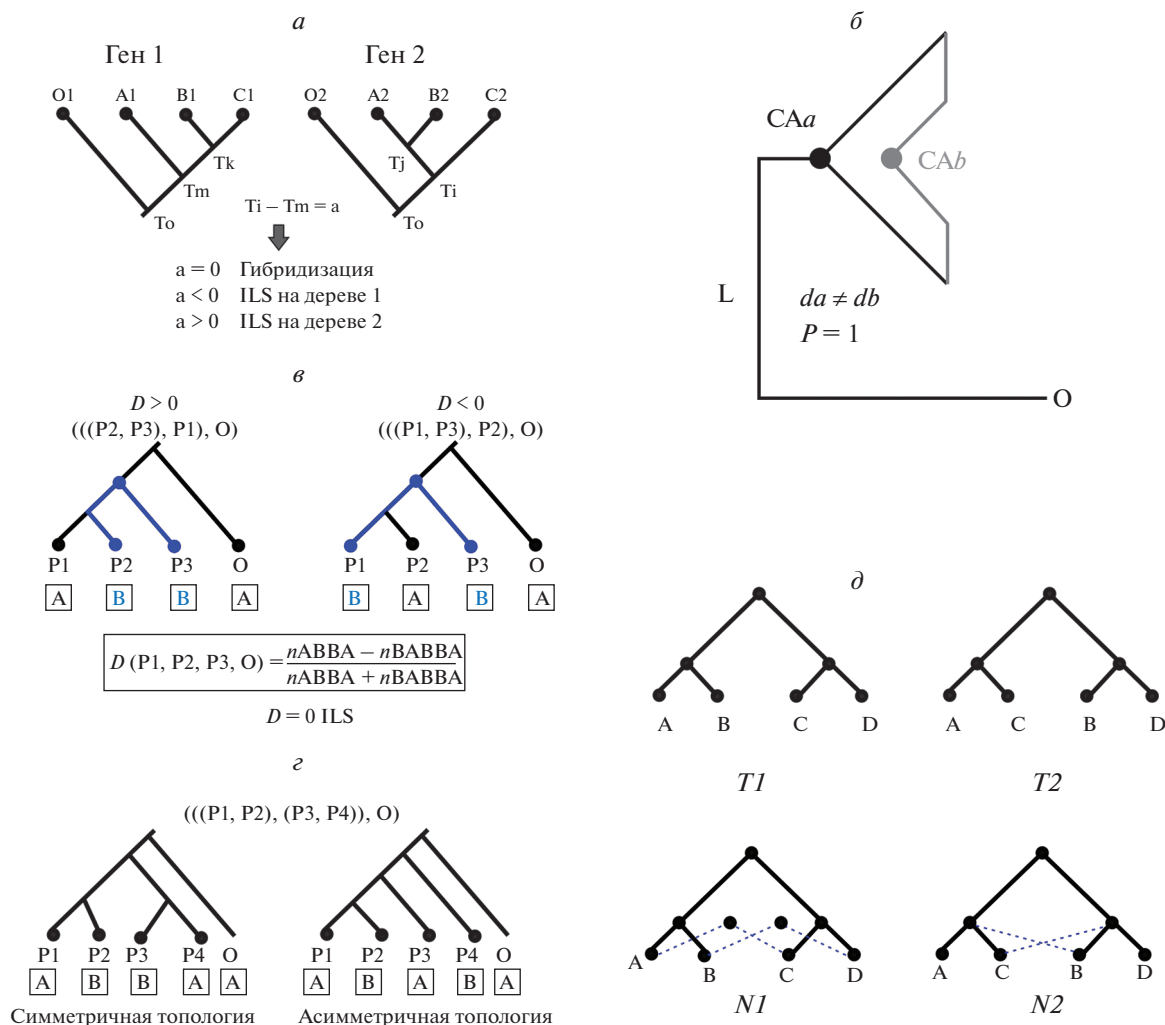


Рис. 2. Схематические иллюстрации отдельных методов, упоминаемых в тексте. *a* – модель для филогении четырех таксонов, где A, B и C – исследуемые виды, O – внешняя группа. A1, B1, C1 и O1 представляют собой последовательности гена 1 для видов A, B, C и O соответственно. A2, B2, C2 и O2 представляют собой последовательности гена 2 для видов A, B, C и O соответственно. Tm, Tk, Ti и Tj представляют время дивергенции между последовательностями A1 и (B1, C1), B1 и C1, C2 и (A2, B2), а также A2 и B2 соответственно. *б* – схематическое представление полной интрогрессии. *в* – схема филогении трех таксонов (P1–P3) и одной внешней группы (O). D-статистика учитывает родительские (“A”) и дочерние (“B”) аллели для филогении четырех таксонов. По сценарию ILS паттерны “ABBA” и “BABA” должны встречаться равновероятно (D = 0). Превышение “ABBA” или “BABA”, приводящее к значению D, значительно отличного от нуля, свидетельствует о потоке генов между двумя таксонами. Превышение сайтов “ABBA” свидетельствует о потоке генов между P2 и P3. Превышение сайтов “BABA” свидетельствует о потоке генов между P1 и P3 (выделено синим цветом). *г* – схема филогении четырех таксонов (P1–P4) и одной внешней группы (O). Показаны симметричная и асимметричная топологии с примерами паттернов аллельных сайтов “ABBA” и “BABA” соответственно. *д* – укорененные бифуркационные деревья T1 и T2 и две филогенетические сети N1 и N2, объясняющие оба дерева.

полной интрогрессии мтДНК. При этом для ядерного(ых) гена(ов): $Ea = Eb$. Следует отметить, что в случае полной интрогрессии мтДНК стоит ожидать, что филогенетические линии всех исследуемых видов будут обособлены друг от друга и монофилетичны на ядерном генетическом дереве, в то время как на митохондриальном дереве гаплотипы одного из видов будут расположены внутри клады другого вида (полифилетический паттерн). При этом группа предполагаемого “донора” будет всегда проявлять полифилетиче-

ский характер и образовывать внутреннюю кладу по отношению к виду-реципиенту, в то время как предполагаемый “реципиент” может быть как монофилетичным, так и полифилетичным. Учитывая это, данный метод включает тест определения филогенетического статуса исследуемых групп (монофилия, парафилия или полифилия). Программное обеспечение для данного алгоритма находится в открытом доступе (<https://github.com/dysh/MRDR>). Описанный выше подход был разработан и впервые приме-

нен для исследования брюхоногих моллюсков из р. *Baicalia*, для представителей которого характерны несоответствия между митохондриальной и ядерной филогениями (Poroshina et al., 2020).

Рассмотренные выше алгоритмы специализируются на распознавании гибридизации и интрогрессии при филогенетическом несоответствии и разграничении таковых от неполной сортировки линий, не подразумевая, что они могут происходить одновременно (Degnan, 2018). Для определения вышеописанных типов ретикулярных процессов в присутствии ILS был разработан ряд аналитических методов, позволяющих также проводить дополнительную характеристику интрогрессии (оценка доли заимствований, направление и временная оценка наблюдаемой интрогрессии).

СОВРЕМЕННЫЕ МЕТОДЫ ОБНАРУЖЕНИЯ И ХАРАКТЕРИСТИКИ СЛЕДОВ ИНТРОГРЕССИИ НА ОСНОВЕ ГЕНОМНЫХ ДАННЫХ И ИХ ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Развитие новых молекулярно-генетических методов и прочтение полных геномов ряда современных и вымерших видов значительно изменили наше представление об эволюционной истории человека и других видов животных и увеличило шансы определения даже незначительных следов интрогрессии (Racimo et al., 2015; Schaefer et al., 2016; Rochus et al., 2018). Был разработан ряд методов, направленных на поиск генетического смешения (genetic admixture; формирование современной популяции за счет потока ядерных генов между предковыми популяциями) в масштабе всего генома (genome-wide test). Целью таких методов является поиск популяций-доноров и популяций-реципиентов с “чужим” генетическим материалом, количественная оценка наблюдаемой интрогрессии, а также определение “чужеродных” геномных сегментов, проникших через видовые границы (Burgarella et al., 2019). Как было отмечено выше, определение следов интрогрессии затруднено в присутствии ILS среди диверсифицирующихся видов, поскольку оба процесса создают схожие паттерны несоответствия генетических деревьев (Holder et al., 2001; Meleshko et al., 2021). Широкий спектр методов учитывает это, предполагая ILS в качестве нулевой гипотезы (Pease, Hahn, 2015). В целом методы, о которых пойдет речь, могут применяться и к отдельным генам, но их надежность значительно возрастает, когда они применяются к данным всего генома (Patterson et al., 2012). Однако следует принять во внимание, что ни один из них не дает надежных доказательств интрогрессии, поэтому необходимо применять дополнительные методы для получения более детальной информации об особенностях этого сложного процесса.

Нельзя не упомянуть методы кластеризации, такие как *STRUCTURE* (Pritchard et al., 2000) и *ADMIXTURE* (Alexander et al., 2009), наиболее часто используемые для определения структуры популяций на основе мультилокусных данных, которые применяются также для точного определения особей со смешанными геномами, изучения гибридных зон, определения наличия интрогрессии и доли интрогрессированных локусов. Одним из пользующихся наибольшей популярностью методов исследования интрогрессии представляется *STRUCTURE*. Он базируется на байесовском подходе кластеризации, причисляющем особей на основе их генотипов к той или иной популяции при одновременной оценке частот популяционных аллелей и позволяющем оценить, какая часть генома от каждой особи принадлежит к каждой предковой популяции. Метод применяют для различных типов маркеров (например, микросателлиты, полиморфизмы длин рестриционных фрагментов – restriction fragment length polymorphism (RFLPs), однонуклеотидные полиморфизмы – single nucleotide polymorphism (SNP)). В более ранних исследованиях использовались микросателлиты, часто в комбинации с митохондриальными генами (например, Varilani et al., 2005). Быстрый прогресс в методах секвенирования привел к использованию в этом методе SNP и других полногеномных маркеров (Kraus et al., 2012; Gómez-Sánchez et al., 2018). К настоящему времени накопилось большое число исследований, успешно применяющих данный подход для исследования интрогрессии в популяциях и анализа гибридизации (Beaumont et al., 2001; Goossens et al., 2002; Thies et al., 2018; Ravagni et al., 2021). Однако, несмотря на популярность метода, моделирование в *STRUCTURE* сопряжено с ограничениями в отношении популяционно-генетических моделей, таких как соблюдение закона Харди–Вайнберга и равновесие сцепления (linkage equilibria) (Jombart et al., 2010). Способность *STRUCTURE* точно оценивать степень интрогрессии ставят под сомнение, поскольку есть вероятность недооценки количества особей со смешанными геномами (например, Sanchez-Donoso, 2014). В то же время в недавнем исследовании (Ravagni et al., 2021) было проведено моделирование интрогрессии и определен ряд подходов, которые могут повысить точность оценки степени интрогрессии при использовании *STRUCTURE*. Метод также реализован в удобной веб-программе *STRUCTURE HARVESTER* (Earl, vonHoldt, 2012), разработанной специально для быстрого анализа и обобщения выходных данных из *STRUCTURE*. Новая и обновленная версия *fastSTRUCTURE* позволяет анализировать большие наборы данных SNP (Raj et al., 2014; Elgvin et al., 2017). Для преодоления вышеописанных модельных ограничений были разработаны альтернативные методы, такие как

ADMIXTURE (Alexander et al., 2009) и дискриминантный анализ главных компонент или *DAPC* (Jombart et al., 2010). Для повышения надежности при проведении оценки интрогрессии лучше всего комбинировать разработанные для этой цели *STRUCTURE*, *ADMIXTURE* и *DAPC* (Frosch et al., 2014). Неправильное определение степени интрогрессии может повлиять на понимание эволюционной истории видов (Ravagni et al., 2021).

Многих исследователей интересует не только наличие или отсутствие интрогрессии, но и количественная оценка ее величины. В контексте филогеномных подходов под количественной оценкой понимают пропорцию заимствованного генома или, другими словами, степень интрогрессии. Для этой цели разработана серия популяционно-генетических методов.

Из множества используемых методов для определения и оценки степени интрогрессии наиболее часто применяется классическая *D*-статистика Петерсона, или так называемый АВВА-ВАВА тест. Данный тест, разработанный и впервые успешно примененный для оценки интрогрессии между древними людьми современного типа и неандертальцами (Green et al., 2010; Durand et al., 2011), позволяет различать паттерны, возникающие вследствие интрогрессии, от таковых, основанных на анцестральном полиморфизме и последующей неполной сортировке линий. В настоящее время накопилось большое количество исследований, успешно применяющих этот метод для обнаружения интрогрессии во многих других группах (например, Dasmahapatra et al., 2012; Fontaine et al., 2015; Martin et al., 2015; Elgvin et al., 2017; Kumar et al., 2017; Miao et al., 2017; Barlow et al., 2018; Mao et al., 2018; Andersen et al., 2021; Komarova et al., 2021; Meleshko et al., 2021). Данный подход применим для филогении трех таксонов и одной внешней группы (в литературе обозначаемой как four-taxon trees), отношения между которыми можно представить в виде $((P1, P2), P3), O)$, из которых три (за исключением внешней группы) могут быть вовлечены в процессы предполагаемой гибридизации (рис. 2в) (Durand et al., 2011). Если предположить подобный сценарий и обозначить родительский аллель как “А”, а дочерний аллель как “В”, то мы получим два филогенетических паттерна: “АВВА” включает сайты, где P2 и P3 несут дочерний аллель, в то время как P1 и внешняя группа (O) – родительский; паттерн “ВАВА” представляет сайты, где, наоборот, P1 и P3 включают общий родительский аллель, за исключением P2 и внешней группы (рис. 2в). При сценарии неполной сортировки линий без потока генов паттерны “АВВА” и “ВАВА” будут встречаться одинаково часто. Превышение “АВВА” и “ВАВА”, приводящее к значению *D*-статистики, значительно отличному от нуля, свидетельствует о потоке генов между двумя таксонами. Избыток пат-

тернов “АВВА” указывает на поток генов между таксонами P2 и P3, а избыток “ВАВА” указывает на поток генов между P1 и P3 (рис. 2в). Надо отметить, что все учитываемые сайты в выравниваниях последовательностей должны быть биваллельными (biallelic sites), при этом внешняя группа всегда определяет анцестральное (предковое) состояние (аллель А). Еще одно практическое ограничение метода заключается в том, что *D*-статистика для четырех таксонов не может обнаруживать интрогрессию между сестринскими таксонами P1 и P2 и позволяет определять не направление интрогрессии, а лишь ее наличие или отсутствие (неопределенное направление между $P2 \leftrightarrow P3$ или $P1 \leftrightarrow P3$; Durand et al., 2011). Кроме того, *D*-статистика была первоначально разработана для обнаружения интрогрессии в масштабе всего генома или всей хромосомы (Green et al., 2010). Вычисление этой статистики для небольших участков генома или конкретных локусов для количественной оценки интрогрессии может привести к ненадежным результатам, поскольку результаты *D*-статистики имеют тенденцию группироваться в регионах с пониженным генетическим разнообразием и низкой степенью рекомбинации (Martin et al., 2015). Примечательно исследование тиранных мухоловок (Tyranidae), выполненное на основе сильно ограниченного набора SNP локусов (2710 SNP) с использованием АВВА-ВАВА теста (Rheindt et al., 2014). В этом исследовании была выполнена реконструкция эволюционной истории фенотипически мозаичной популяции золотолобого карликового тирана (*Zimmerius viridiflavus*), особенности окраски оперения которого схожи с таковыми златолицего тирана (*Z. chrysops*). Картируя предполагаемые интрогрессированные локусы по результатам теста АВВА-ВАВА на геном зебровой амадины (*Taeniopygia guttata*), исследователи обнаружили тесную связь некоторых SNP-паттернов “АВВА” с генами, участвующими в производстве клеток и плазматических мембран. Учитывая связь между клеточными мембранами и окраской оперения, авторы предполагают, что между этими мозаичными популяциями и *Z. chrysops* могла происходить интрогрессия аллелей генов, определяющих окраску оперения. Однако, как отмечено выше, эти локусы не обязательно могли быть интрогрессированы, но могут представлять геномные области с низким генетическим разнообразием в результате действия вымещающего отбора. В целом классическая *D*-статистика видится надежным тестом для поиска следов интрогрессии в масштабе всего генома, однако необходимо использование дополнительных методов для получения более детальной информации об особенностях этого процесса.

Использование лишь *D*-статистики не дает объективной оценки доли интрогрессированных

локусов в геноме (Martin et al., 2015; Hamlin et al., 2020). Одно из предложенных расширений D -статистики, так называемое D_p , вносит одно простое дополнение – добавление в знаменатель формулы подсчета сайтов “ВВАА”, что значительно улучшает оценку пропорции интрогрессированных локусов по сравнению с классической D -статистикой (Hamlin et al., 2020). Однако данный способ имеет тенденцию недооценивать реальную пропорцию заимствований, и его точность повышается при более низкой степени интрогрессии.

Другой альтернативный подход, статистика отношений F_4 (F_4 -ratio estimation) или a -статистика, ранее называемая f_4 -оценкой происхождения (f_4 -ancestry estimation; Reich et al., 2009), специализируется на оценке доли заимствованных генов в смешанной популяции на основе SNP всего генома (Patterson et al., 2012). Тест проводит сравнение наблюдаемого значения интрогрессии с ожидаемым в сценарии, при котором интрогрессирован весь геном. Статистика применима для данных от пяти таксонов, включая внешнюю группу. Метод разработан и впервые применен для исследования истории популяций людей из Индии (Reich et al., 2009) и в дальнейшем использовался в основном для оценки примесей в различных популяциях человека (Gallego Llorente et al., 2015; Isshiki et al., 2020; Mineta et al., 2021). Важным преимуществом F - и D -статистик будет то, что они зависят только от частот популяционных аллелей, и для оценки не требуется информация о гаплотипах (Gautier et al., 2021). *HyDe*, реализованный как пакет на языке программирования Python, аналогичным образом оценивает степень интрогрессии в сценарии гибридного видообразования, используя линейные комбинации производных паттернов сайтов (Blischak et al., 2018; Kubatko, Chifman, 2019). *HyDe* автоматизирует обнаружение гибридизации для большого числа видов и может проводить проверку гипотез как на индивидуальном уровне, так и на уровне популяций. Метод дает очень точную оценку степени интрогрессии, когда выполняется его предположение о гибридизации, и приемлемую оценку, когда допущения нарушаются (Kong, Kubatko, 2021).

Статистика fd (Martin et al., 2015) может быть применима к данным по небольшому участку генома. Подобно D_p , этот метод также чувствителен к направлению интрогрессии, поскольку оценивает долю генома, полученную от популяции донора. В некоторой степени статистика fd решает эту проблему, предполагая, что популяция с более высокой производной частотой аллелей будет донором на каждом сайте. Тем не менее fd имеет тенденцию недооценивать долю интрогрессированных локусов в том случае, когда таксон P2 является донором.

Чтобы оценить направление интрогрессии для филогении трех или четырех таксонов, недостаточно располагать информацией, содержащейся в частотах паттернов “АВВА” и “ВАВА”. Тем не менее, если доступна информация о пятом таксоне (Pease, Hahn, 2015) или о полиморфизме для четырех данных таксонов (Martin, Amos, 2021), возможно определение направления интрогрессии. Так называемая “*partitioned D-statistics*” стала первой попыткой получить сведения о направлении интрогрессии для филогении с пятью таксонами (Eaton, Ree, 2013). Однако избыточность в подсчете паттернов сайтов делает результаты этого теста не интерпретируемыми. Для решения этой проблемы была разработана *DFOIL* статистика (Pease, Hahn, 2015) для филогении из пяти таксонов (укорененное дерево с четырьмя таксонами и одной внешней группой), отношения между которыми можно представить в виде $((P_1, P_2), (P_3, P_4)), O$, из которых четыре (за исключением внешней группы) могут быть вовлечены в процессы предполагаемой гибридизации (рис. 2а). *DFOIL* создает систему из четырех D -статистик (применимость к каждому из четырех таксонов), тестируя каждое из 16 возможных событий и направлений интрогрессии. Помимо определения таксонов, вовлеченных в интрогрессию, статистика *DFOIL* также предоставляет информацию о направлении межгрупповых интрогрессий, в частности идентифицируя таксоны-доноры и реципиенты. В совокупности все четыре рассматриваемые статистики содержат больше информации, чем сумма отдельных D -тестов. Метод *DFOIL* исходит из предположения, что филогения четырех таксонов симметрична, с двумя парами сестринских видов (рис. 2а). При такой конфигурации дерева становится возможным определять направление интрогрессии (т.е. идентифицировать виды-доноры и реципиенты), поскольку направление влияет на отношения между видом-донором, видом-реципиентом и его сестринским таксоном. К сожалению, тест не работает, если дерево видов асимметрично (рис. 2б).

Мартин и Эмос (Martin, Amos, 2021) показали, что информация о степени, направлении и времени интрогрессии в филогении пяти таксонов будет доступной, если включать в анализ несколько особей от одной линии. Подход, названный частотным спектром D , или Dfs , оценивает D -статистику в каждой ячейке общего полученного частотного спектра аллелей, построенного для двух сестринских таксонов. Ожидается, что форма спектра Dfs будет зависеть от направления интрогрессии. Если один из сестринских таксонов реципиент – спектр смещается влево. Напротив, если сестринская линия донор, а несестринская – реципиент, следует ожидать, что спектр будет плоским, поскольку частотный спектр несестринской линии не используется для построе-

ния Dfs . Степень асимметрии влево зависит от времени интрогрессии, в то время как доля интрогрессии влияет на величину D -статистики в ячейках. На форму спектра Dfs также могут влиять демографическая история и смещения при более сложных сценариях интрогрессии, поэтому необходимо предварительно выполнять моделирование, чтобы протестировать различные сценарии интрогрессии (Martin, Amos, 2021).

Алгоритмы классической D -статистики и ее модификации реализованы в доступных пакетах в различных языках программирования (Python, R) (Patterson et al., 2012; Blischak et al., 2018; Petr et al., 2019; Mussmann et al., 2020; Gautier et al., 2021; Malinsky et al., 2021; Martin, Amos, 2021). Стоит отметить комплексную программу *COMP-D*, позволяющую объединять результаты классической D и *DFOIL* статистик. Она эффективна в выполнении всех необходимых вычислений из общих филогенетических и популяционно-генетических форматов программных файлов (Mussmann et al., 2020). Особое внимание стоит обратить на простую в использовании программу *Dsuite*, в которой реализованы возможности вычисления нескольких статистик одновременно (D и f_4). *Dsuite* значительно более эффективна с точки зрения вычислений, чем сопоставимые программы, и предоставляет удобный набор инструментов, облегчающий оценку наличия потока генов, особенно в больших геномах (Malinsky et al., 2021). Для анализа данных, полученных методами полногеномного секвенирования целого пула образцов (Pool-seq) и подходами F -статистик, разработан специальный программный пакет в R – *poolfstat* (Gautier et al., 2021). И, наконец, *admixr* – программный пакет в R, обеспечивающий удобный интерфейс для выполнения комплексного популяционно-генетического анализа (f_3 , D , f_4 -ratio), ранее реализованного в программном пакете *ADMIXTOOLS* с использованием командной строки (Petr et al., 2019). Объединяя все этапы рабочего процесса в единой среде R, этот пакет позволяет автоматизировать анализ, значительно улучшая воспроизводимость популяционно-генетических исследований.

ФИЛОГЕНЕТИЧЕСКИЕ СЕТИ: МЕТОДЫ ПРЕДСТАВЛЕНИЯ ЭВОЛЮЦИОННЫХ ИСТОРИЙ В ПРИСУТСТВИИ РЕТИКУЛЯРНЫХ СОБЫТИЙ

Широкий набор методов в филогенетике основан на алгоритмах, представляющих эволюцию исследуемых таксонов в виде так называемых филогенетических сетей (phylogenetic networks). Переход от филогенетических деревьев к филогенетическим сетям – еще один способ справиться с разграничением процессов интрогрессии и неполной сортировки линий. В литературе можно

встретить различные применения этого термина и поэтому дать общее определение весьма сложно. Под филогенетическими сетями можно понимать любой направленный граф, на котором таксоны представлены в виде узлов, а их эволюционные отношения – в виде ребер (на филогенетических деревьях ребра соответствуют ветвям) (Huson, Bryant, 2006). Филогенетические сети – альтернативный вариант представления филогенетических взаимоотношений между таксонами, а также основной инструмент для иллюстрации сложных эволюционных историй, сопровождающихся ретикулярными событиями (гибридизация, интрогрессия, горизонтальный перенос генов (horizontal gene transfer, HGT) и другие), которые не могут быть представлены в виде дерева (Morrison, 2014). Даже если эволюция исследуемых таксонов происходила согласно древовидной модели, сети позволяют более наглядно отображать неоднозначные филогенетические сигналы в наборах данных, которые могли быть вызваны небиологическими причинами, такими как, например, некорректное выравнивание последовательностей, слишком короткие отдельные последовательности, неправильно подобранные модели нуклеотидной эволюции и другие (Huson, Bryant, 2006; Degnan, 2018). Существуют два больших класса филогенетических сетей: неукорененные абстрактные филогенетические сети (implicit phylogenetic networks) и укорененные явные сети (explicit phylogenetic networks) (Klopper, Huson, 2008). Первый класс сетей строят для отображения и анализа конфликтующих филогенетических сигналов в имеющемся наборе данных; для учета этих неразрешенных сигналов часто необходимо наличие узлов на сети, не представляющих предковый вид (Huson, Bryant, 2006). Поэтому такие сети представляют “невные” эволюционные сценарии (Klopper, Huson, 2008). К этому классу принадлежат так называемые сети расщепления (split networks), где каждое ребро делит набор таксонов на две непересекающиеся группы (расщепления), расположенные на разных сторонах ветвей (Huson, Bryant, 2006). Существует целый ряд признанных методов для вычисления неукорененных сетей (*median networks* – Bandelt et al., 2000; *split decomposition* – Bandelt, Dress, 1992; *neighbor-net* – Bryant, Moulton, 2004; *consensus networks* – Holland et al., 2004; *supernetworks* – Hassanzadeh et al., 2012). Более подробную характеристику данного типа сетей и алгоритмов их вычисления можно найти в обзорной работе Хусона и соавт. (Huson et al., 2010). Второй класс сетей вычисляют для представления “явных” историй исследуемых таксонов, которые эволюционируют в рамках модели, включающей как события видообразования, так и ретикулярные процессы. Поэтому такие сети называют также ретикулярными сетями (рекомбинантными сетями – Huson, Klopper, 2005; се-

тиями гибридизации – Huson et al., 2005). Методы для построения укорененных ретикулярных сетей представлены ограниченным набором алгоритмов, отображающих эволюционные сценарии в виде обобщенного филогенетического дерева из набора укорененных реконструированных генетических деревьев (Morrison, 2014). Внутренние узлы в такой сети представляют предковые таксоны, а узлы с двумя родительскими таксонами (ретикулярные узлы) соответствуют ретикулярным событиям, таким как гибридизация и ее последствия (интрогрессия), рекомбинация или горизонтальный перенос генов (Huson, Kloerper, 2005). Подобного типа сети позволяют наглядно иллюстрировать наличие ретикулярных событий и способствуют пониманию того, в какой степени эти процессы повлияли на эволюцию. Стоит подчеркнуть, что ни один из существующих методов построения эволюционных сетей (другое название для “явных” сетей) не пользуется популярностью в качестве рутинного инструмента реконструкции эволюционной истории исследуемых таксонов в присутствии тех или иных ретикулярных событий (Huson et al., 2010). Однако возможности существующих на данный момент инструментов отображения эволюционных сетей могут стать многообещающей стратегией для обнаружения событий интрогрессии в эволюции видов.

На построение укорененных ретикулярных сетей гибридизации направлен один из наиболее известных комбинаторных подходов, называемый *HYBRIDNUMBER* и основанный на подсчете минимального числа событий гибридизации для комбинаций из двух укорененных бифуркационных деревьев (T_1 и T_2 , рис. 2д) (Bordewich et al., 2007). Он отличается большей точностью и отсутствием модельных ограничений в отличие от его предыдущих аналогов (Hallett, Lagergren, 2001; Huson et al., 2005). Его задача заключается в том, что модель, учитывая набор из двух укорененных филогенетических деревьев (например, одно на основе ядерных генов, другое – митохондриальных) от X таксонов, вычисляет укорененную филогенетическую сеть N (рис. 2д) (предполагая, что гибридизация – причина неконгруэнтности двух деревьев), содержащую обобщенное дерево и имеющую минимальное количество ретикулярных узлов, необходимых для объяснения эволюции рассматриваемых таксонов. Нижняя граница числа событий гибридизации будет показателем того, в какой степени гибридизация повлияла на эволюционную историю исследуемых таксонов. Однако, используя данный подход, нельзя провести анализ эволюционных историй при наличии других биологических событий, таких как ILS. Подход был разработан и успешно применен для исследований процессов гибридизации в семействе мятликовых Poaceae (Barker et al., 2001; Bordewich et al., 2007). Этот алгоритм реализован

как пакет в языке программирования Perl и находится в открытом доступе. Предыдущие аналоги подхода для построения сетей гибридизации реализованы в универсальном программном пакете *SplitsTree4* (Huson, Bryant, 2006).

Особое внимание стоит обратить на комплексный программный пакет *PhyloNet* на языке программирования Java (Than et al., 2008; Wen et al., 2018), разработанный в основном для реконструкции и интерпретации эволюционных сетей. *PhyloNet* при помощи алгоритмов максимального правдоподобия (*InferNetwork_ML*) выводит ретикулярную сеть на основе набора генетических деревьев и проводит дополнительный анализ, используя алгоритмы максимального псевдоподобия (*InferNetwork_MPL*). Кроме того, в своей расширенной версии (Wen et al., 2018) *PhyloNet* включает дополнительный набор алгоритмов для построения эволюционных сетей с учетом ILS. Метод проявил себя как надежный инструмент при оценке наличия интрогрессии между сестринскими видами. Так, анализ геномных данных всех ныне живущих видов медведей с использованием *PhyloNet* продемонстрировал, что интрогрессия между парами близкородственных видов гораздо более обычное событие, чем это предполагалось ранее (Kumar et al., 2017). Недавнее исследование эволюции серогрудой муравьеловки (*Myrmoborus lugubris*) хорошо иллюстрирует необходимость применения методов построения филогенетических сетей для оценки филогенетических взаимоотношений в присутствии потока генов (Thom et al., 2018). Так, использование этого метода позволило показать, что два из четырех подвидов этого вида (*M. lugubris femininus* и *M. lugubris lugubris*) вовлечены в процессы интрогрессивной гибридизации и не являются сестринскими кладами, как считали раньше на основе построения видовых деревьев, не учитывающих на стадии филогенетического вывода возможный поток генов (Thom et al., 2018). Филогенетические сети, построенные с помощью *PhyloNet*, визуализируют с помощью другой интерактивной программы для построения широкого спектра неукорененных филогенетических сетей – *Dendroscope 3* (Huson, Scornavacca, 2012).

SNaQ (Solís-Lemus, Ané, 2016), пожалуй, единственный метод моделирования событий гибридизации с помощью неукорененных сетей. Так же, как и *PhyloNet*, он позволяет проводить анализ ретикулярных событий с учетом ILS, однако для точной оценки событий гибридизации может потребоваться большое количество локусов. Основываясь на алгоритмах максимального псевдоподобия (maximum pseudolikelihood), *SNaQ* выполняет оценку филогенетических сетей на основе мультилокусных данных, используя коэффициенты согласованности (concordance factors, CF) (Ané et al., 2007; Baum, 2007). Отли-

чительная особенность *SNaQ* – ускорение вероятностных вычислений, поскольку вместо вычисления вероятности для всех таксонов одновременно, метод рассматривает подмножества из четырех таксонов. Кроме того, предполагаемая сеть не подвержена потенциальным ошибкам укоренения, поскольку входной набор данных представлен неукорененными деревьями. Данный метод хорошо зарекомендовал себя на первых этапах тестирования при исследовании событий гибридизации и интрогрессии, сопровождающих эволюционную историю тропических рыб р. *Xiphophorus*. К настоящему моменту накопилось большое количество исследований, успешно применяющих этот метод для обнаружения событий интрогрессии и в других группах организмов (Roda et al., 2017; Blair et al., 2019). Метод реализован в доступном программном комплексе *PhyloNetworks* с удобным программным интерфейсом (Solís-Lemus et al., 2017).

В целом *PhyloNet* и *SNaQ* являются надежными инструментами для представления сложных эволюционных историй в виде филогенетических сетей и успешно применяются на практике для обнаружения событий интрогрессии. Например, моделирование филогенетических сетей с использованием этих двух методов позволило получить первые доказательства геномной интрогрессии у кораллов р. *Acropora* (Mao et al., 2018). Однако эти методы предъявляют особые требования к формату входных данных и наличию набора заранее построенных генетических деревьев. Кроме того, *PhyloNet* применим для исследования ограниченных эволюционных сценариев, поскольку скорость вычисления вероятностей сильно снижается при исследовании более чем 10 таксонов (Mao et al., 2020). Недавно разработанный универсальный программный комплекс *TREEasy* (Mao et al., 2020) воплотил в себе возможности обоих инструментов, *PhyloNet* и *SNaQ*, для вывода укорененных ретикулярных сетей. Кроме того, интерфейс не предъявляет особых требований к формату (только наличие формата FASTA и некоторых входных параметров), самостоятельно выполняет построение генетических деревьев, необходимых для построения сетей, и может выводить филогенетические сети одновременно двумя методами на основе не выровненных последовательностей.

Широкий набор алгоритмов, направленных на реконструкцию, интерпретацию и визуализацию сетей минимальной гибридизации, реализован в удобном и простом программном пакете *Dendroscope* (Huson, Scornavacca, 2012). Представление сетей минимальной гибридизации основано на подсчете минимального числа событий гибридизации для комбинаций из двух укорененных филогенетических деревьев. Этот общий принцип работы алгоритмов был представлен выше при описании

HYBRIDNUMBER (Bordewich et al., 2007). Однако математическая база предыдущих алгоритмов по вычислению сетей гибридизации и большинства тех, что реализованы в *Dendroscope*, предполагает, что входные деревья были бифуркационными, правильно укорененными и содержат одни и те же таксоны. Стоит заметить, что в реальных биологических исследованиях такое встречается крайне редко. Особое внимание из представленных в *Dendroscope* подходов заслуживает относительно новый алгоритм *Autumn* (Huson, Linz, 2016). В отличие от предыдущих аналогов, он направлен на вычисление сетей минимальной гибридизации для пары “реалистичных” укорененных филогенетических деревьев. Важной особенностью метода является способность обнаруживать существование предполагаемых гибридных видов. Например, недавнее исследование, основанное на филогеномном анализе гусеобразных, показало, что краснозобая казарка (*Branta ruficollis*) может иметь гибридное происхождение (Ottenburghs et al., 2017).

Сравнительно недавно возможности реконструкции эволюционной истории в присутствии ретикулярных событий с помощью филогенетических сетей были реализованы в пакете *SpeciesNetwork* программного комплекса BEAST2 (Zhang et al., 2018). *SpeciesNetwork* выводит ретикулярную сеть на основе мультилокусных данных при помощи алгоритмов байесовской статистики. *SpeciesNetwork* позволяет проводить анализ ретикулярных событий с учетом ILS, однако, как и *PhyloNet*, сопряжен с высокой вычислительной нагрузкой при исследовании большого количества таксонов.

ЗАКЛЮЧЕНИЕ

В данном обзоре мы обобщили большинство накопленных к настоящему времени сведений о существующих методах обнаружения интрогрессии, основанных на использовании как отдельных генетических маркеров, так и данных по всему геному. Рассмотренные ограничения и недостатки этих методов позволяют сделать вывод о том, что остаются не решенными несколько методологических проблем, таких как надежность обнаружения интрогрессии в присутствии неполной сортировки линий при анализе небольшого числа локусов; более точные оценки доли интрогрессированных локусов, времени и направления интрогрессии; обнаружение интрогрессии между сестринскими таксонами; количественная оценка событий интрогрессии в отдельной кладе; более точная автоматизированная визуализация филогенетических сетей. Несмотря на эти проблемы, доступные в настоящее время подходы позволяют эффективно обнаруживать следы интрогрессии и характеризовать ее особенности, однако следует помнить, что не существует универ-

сального метода, позволяющего представить надежные доказательства интрогрессии, ее количественную характеристику, а также наглядно проиллюстрировать сложные эволюционные истории в присутствии ретикулярных событий. Поэтому необходимо применять комбинации нескольких методов, что позволит получить более детальную информацию об особенностях этого сложного процесса.

Очевидно, что дальнейшее развитие и использование рассмотренных выше методов будет способствовать выявлению природы и распространения феномена интрогрессии у разных групп живых организмов.

Работа выполнена при финансовой поддержке РФФИ (проект № 20-34-90032).

СПИСОК ЛИТЕРАТУРЫ

- Абрамсон Н.И., 2007. Филогеография: итоги, проблемы, перспективы // Информ. вестн. ВОГиС. Т. 11. № 2. С. 307–331.
- Абрамсон Н.И., 2009. Молекулярные маркеры, филогеография и поиск критерия разграничения видов // Тр. Зоол. ин-та РАН. Прил. 1. С. 185–189.
- Боркин Л.Я., Литвинчук Р.Н., 2013. Гибридизация, видообразование и систематика животных // Тр. Зоол. ин-та РАН. Прил. 2. С. 83–139.
- Майр Э., 1968. Зоологический вид и эволюция. М.: Мир. 597 с.
- Alexander D.H., Novembre J., Lange K., 2009. Fast model-based estimation of ancestry in unrelated individuals // Genome Res. V. 19. № 9. P. 1655–1664.
- Anderson E., 1949. Introgressive Hybridization. N.-Y.: J. Wiley. 109 p.
- Anderson E., Hubricht L., 1938. Hybridization in *Tradescantia*. III. The evidence for introgressive hybridization // Am. J. Bot. V. 25. № 6. P. 396–402.
- Andersen M.J., McCullough J.M., Gyllenhaal E.F., Mapel X.M., Haryoko T. et al., 2021. Complex histories of gene flow and a mitochondrial capture event in a nonsister pair of birds // Mol. Ecol. V. 30. № 9. P. 2087–2103.
- Ané C., Larget B., Baum D.A., Smith S.D., Rokas A., 2007. Bayesian estimation of concordance among gene trees // Mol. Biol. Evol. V. 24. № 2. P. 412–426.
- Arnold M.L., 2006. Evolution through Genetic Exchange. Oxford: Oxford Univ. Press. 272 p.
- Arnold M.L., Kunte K., 2017. Adaptive genetic exchange: A tangled history of admixture and evolutionary innovation // Trends Ecol. Evol. V. 32. № 8. P. 601–611.
- Arnold M.L., Martin N.H., 2009. Adaptation by introgression // J. Biol. V. 8. № 9. Art. 82.
- Arnold M.L., Sapir Y., Martin N.H., 2008. Genetic exchange and the origin of adaptations: prokaryotes to primates // Philos. Trans. R. Soc. Lond. B. Biol. Sci. V. 363. № 1505. P. 2813–2820.
- Avise J.C., 2000. Phylogeography: The History and Formation of Species. Cambridge: Harvard Univ. Press. 286 p.
- Bandelt H.J., Dress A.W., 1992. Split decomposition: A new and useful approach to phylogenetic analysis of distance data // Mol. Phylogenet. Evol. V. 1. № 3. P. 242–252.
- Bandelt H.J., Macaulay V., Richards M., 2000. Median networks: Speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA // Mol. Phylogenet. Evol. V. 16. № 1. P. 8–28.
- Barilani M., Derégnaucourt S., Gallego S., Galli L., Mucci N. et al., 2005. Detecting hybridization in wild (*Coturnix c. coturnix*) and domesticated (*Coturnix c. japonica*) quail populations // Biol. Conserv. V. 126. № 4. P. 445–455.
- Barker N.P., Clark L.G., Davis J.I., Duvall M.R., Guala G.F. et al., 2001. Phylogeny and subfamilial classification of the grasses (Poaceae) // Ann. Missouri Bot. Gard. V. 88. № 3. P. 373–457.
- Barlow A., Cahill J.A., Hartmann S., Theunert C., Xenikoudakis G. et al., 2018. Partial genomic survival of cave bears in living brown bears // Nat. Ecol. Evol. V. 2. № 10. P. 1563–1570.
- Baum D.A., 2007. Concordance trees, concordance factors, and the exploration of reticulate genealogy // Taxon. V. 56. № 2. P. 417–426.
- Beaumont M., Barratt E.M., Gottelli D., Kitchener A.C., Daniels M.J. et al., 2001. Genetic diversity and introgression in the Scottish wildcat // Mol. Ecol. V. 10. № 2. P. 319–336.
- Blair C., Bryson R.W., Jr., Linkem C.W., Lazcano D., Klicka J., McCormack J.E., 2019. Cryptic diversity in the Mexican highlands: Thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times of montane rattlesnakes (Viperidae: *Crotalus*) // Mol. Ecol. Res. V. 19. № 2. P. 349–365.
- Blischak P.D., Chifman J., Wolfe A.D., Kubatko L.S., 2018. HyDe: A Python package for genome-scale hybridization detection // Syst. Biol. V. 67. № 5. P. 821–829.
- Bordewich M., Linz S., John K.S., Semple C., 2007. A reduction algorithm for computing the hybridization number of two trees // Evol. Bioinform. V. 3. P. 86–98.
- Bryant D., Moulton V., 2004. Neighbor-net: An agglomerative method for the construction of phylogenetic networks // Mol. Biol. Evol. V. 21. № 2. P. 255–265.
- Burgarella C., Barnaud A., Kane N.A., Jankowski F., Scarcelli N. et al., 2019. Adaptive introgression: An untapped evolutionary mechanism for crop adaptation // Front. Plant Sci. V. 10. Art. 4.
- Chan K.O., Hutter C.R., Wood P.L., Jr., Grismer L.L., Das I., Brown R.M., 2020. Gene flow creates a mirage of cryptic species in a Southeast Asian spotted stream frog complex // Mol. Ecol. V. 29. № 20. P. 3970–3987.
- Choleva L., Musilova Z., Kohoutova-Sediva A., Paces J., Rab P., Janko K., 2014. Distinguishing between incomplete lineage sorting and genomic introgressions: Complete fixation of allospecific mitochondrial DNA in a sexually reproducing fish (*Cobitis*; Teleostei), despite clonal reproduction of hybrids // PLoS One. V. 9. № 6. Art. e80641.
- Dasmahapatra K.K., Walters J.R., Briscoe A.D., Davey J.W., Whibley A., Nadeau N.J., 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species // Nature. V. 487. № 7405. P. 94–98.
- Degnan J.H., 2018. Modeling hybridization under the network multispecies coalescent // Syst. Biol. V. 67. № 5. P. 786–799.

- Degnan J.H., Rosenberg N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent // Trends Ecol. Evol. V. 24. № 6. P. 332–340.
- Després L., 2019. One, two or more species? Mitonuclear discordance and species delimitation // Mol. Ecol. V. 28. № 17. P. 3845–3847.
- Di Candia M.R., Routman E.J., 2007. Cytonuclear discordance across a leopard frog contact zone // Mol. Phylogenet. Evol. V. 45. № 2. P. 564–575.
- Durand E.Y., Patterson N., Reich D., Slatkin M., 2011. Testing for ancient admixture between closely related populations // Mol. Biol. Evol. V. 28. № 8. P. 2239–2252.
- Earl D.A., vonHoldt B.M., 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method // Conserv. Genet. Resour. V. 4. № 2. P. 359–361.
- Eaton D.A.R., Ree R.H., 2013. Inferring phylogeny and introgression using RADseq data: An example from flowering plants (*Pedicularis*: Orobanchaceae) // Syst. Biol. V. 62. № 5. P. 689–706.
- Edelman N.B., Frandsen P.B., Miyagi M., Clavijo B., Davey J. et al., 2019. Genomic architecture and introgression shape a butterfly radiation // Science. V. 366. № 6465. P. 594–599.
- Elgvin T.O., Trier C.N., Tørresen O.K., Hagen I.J., Lien S. et al., 2017. The genomic mosaicism of hybrid speciation // Sci. Adv. V. 3. № 6. Art. e1602996.
- Fontaine M.C., Pease J.B., Steele A., Waterhouse R.M., Neafsey D.E. et al., 2015. Extensive introgression in a malaria vector species complex revealed by phylogenomics // Science. V. 347. № 6217. Art. 1258524.
- Frosch C., Kraus R.H., Angst C., Allgöwer R., Michaux J. et al., 2014. The genetic legacy of multiple beaver reintroductions in Central Europe // PloS One. V. 9. № 5. Art. e97619.
- Gallego Llorente M., Jones E., Eriksson A., Siska V., Arthur K.W. et al., 2015. Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent // Science. V. 350. № 6262. P. 820–822.
- Gautier M., Vitalis R., Flori L., Estoup A., 2021. *f*-statistics estimation and admixture graph construction with Pool-Seq or allele count data using the R package *poolstat* // bioRxiv. <https://doi.org/10.1101/2021.05.28.445945>
- Gómez-Sánchez D., Olalde I., Sastre N., Enseñat C., Carrasco R. et al., 2018. On the path to extinction: Inbreeding and admixture in a declining grey wolf population // Mol. Ecol. V. 27. № 18. P. 3599–3612.
- Gompert Z., Forister M.L., Fordyce J.A., Nice C.C., 2008. Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides* // Mol. Ecol. V. 17. № 24. P. 5231–5244.
- Goossens B., Funk S.M., Vidal C., Latour S., Jamart A. et al., 2002. Measuring genetic diversity in translocation programmes: Principles and application to a chimpanzee release project // Anim. Conservat. V. 5. № 3. P. 225–236.
- Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U. et al., 2010. A draft sequence of the Neandertal genome // Science. V. 328. № 5979. P. 710–722.
- Hallett M.T., Lagergren J., 2001. Efficient algorithms for lateral gene transfer problems // Proc. 5th Annual International Conference on Computational Biology. N.Y.: ACM Press. P. 149–156.
- Hamlin J.A., Hibbins M.S., Moyle L.C., 2020. Assessing biological factors affecting postspeciation introgression // Evol. Lett. V. 4. № 2. P. 137–154.
- Harrison R.G., Larson E.L., 2014. Hybridization, introgression, and the nature of species boundaries // J. Hered. V. 105. № S1. P. 795–809.
- Hassanzadeh R., Eslahchi C., Sung W.K., 2012. Constructing phylogenetic supernetworks based on simulated annealing // Mol. Phylogenet. Evol. V. 63. № 3. P. 738–744.
- Hedrick P.W., 2013. Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation // Mol. Ecol. V. 22. № 18. P. 4606–4618.
- Heiser C.B., 1973. Introgression re-examined // Bot. Rev. V. 39. № 4. P. 347–366.
- Holder M.T., Anderson J.A., Holloway A.K., 2001. Difficulties in detecting hybridization // Syst. Biol. V. 50. № 6. P. 978–982.
- Holland B.R., Huber K.T., Moulton V., Lockhart P.J., 2004. Using consensus networks to visualize contradictory evidence for species phylogeny // Mol. Biol. Evol. V. 21. № 7. P. 1459–1461.
- Hudson R.R., Turelli M., 2003. Stochasticity overrules the “three-times rule”: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA // Evolution. V. 57. № 1. P. 182–190.
- Huerta-Sánchez E., Jin X., Bianba Z., Peter B.M., Vinckenbosch N. et al., 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA // Nature. V. 512. № 7513. P. 194–197.
- Hufford M.B., Lubinsky P., Pyhäjärvi T., Devengenzo M.T., Ellstrand N.C., Ross-Ibarra J., 2013. The genomic signature of crop-wild introgression in maize // PLoS Genet. V. 9. № 5. Art. e1003477.
- Huson D.H., Bryant D., 2006. Application of phylogenetic networks in evolutionary studies // Mol. Biol. Evol. V. 23. № 2. P. 254–267.
- Huson D.H., Klopper T.H., 2005. Computing recombination networks from binary sequences // Bioinformatics. V. 21. № S2. P. ii159–ii165.
- Huson D.H., Linz S., 2016. Autumn algorithm—computation of hybridization networks for realistic phylogenetic trees // IEEE/ACM Trans. Comput. Biol. Bioinform. V. 15. № 2. P. 398–410.
- Huson D.H., Scornavacca C., 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks // Syst. Biol. V. 61. № 6. P. 1061–1067.
- Huson D.H., Rupp R., Scornavacca C., 2010. Phylogenetic Networks: Concepts, Algorithms and Applications. Cambridge: Cambridge Univ. Press. 362 p.
- Huson D.H., Klöpper T., Lockhart P.J., Steel M.A., 2005. Reconstruction of reticulate networks from gene trees // Annual International Conference on Research in Computational Molecular Biology. Berlin; Heidelberg: Springer. P. 233–249.
- Irisarri I., Singh P., Koblmüller S., Torres-Dowdall J., Henning F. et al., 2018. Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes // Nat. Commun. V. 9. Art. 3159.

- Isshiki M., Naka I., Watanabe Y., Nishida N., Kimura R. et al.*, 2020. Admixture and natural selection shaped genomes of an Austronesian-speaking population in the Solomon Islands // *Sci. Rep.* V. 10. Art. 6872.
- Ivanov V., Lee K.M., Mutanen M.*, 2018. Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression // *Mol. Ecol.* V. 27. № 7. P. 1681–1695.
- Joly S., Bruneau A.*, 2006. Incorporating allelic variation for reconstructing the evolutionary history of organisms from multiple genes: An example from *Rosa* in North America // *Syst. Biol.* V. 55. № 4. P. 623–636.
- Joly S., McLenachan P.A., Lockhart P.J.*, 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting // *Am. Nat.* V. 174. № 2. P. E54–E70.
- Jombart T., Devillard S., Balloux F.*, 2010. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations // *BMC Genet.* V. 11. Art. 94.
- Klopper T.H., Huson D.H.*, 2008. Drawing explicit phylogenetic networks and their integration into SplitsTree // *BMC Evol. Biol.* V. 8. Art. 22.
- Komarova V.A., Kostin D.S., Bryja J., Mikula O., Bryjová A. et al.*, 2021. Complex reticulate evolution of speckled brush-furred rats (*Lophuromys*) in the Ethiopian centre of endemism // *Mol. Ecol.* V. 30. № 10. P. 2349–2365.
- Kong S., Kubatko L.*, 2021. Comparative performance of popular methods for hybrid detection using genomic data // *Syst. Biol.* V. 70. № 5. P. 891–907.
- Kraus R.H., Kerstens H.H., Hooft P., van, Megens H.J., Elmerberg J. et al.*, 2012. Widespread horizontal genomic exchange does not erode species barriers among sympatric ducks // *BMC Evol. Biol.* V. 12. Art. 45.
- Kubatko L.S., Chifman J.*, 2019. An invariants-based method for efficient identification of hybrid species from large-scale genomic data // *BMC Evol. Biol.* V. 19. № 1. Art. 112.
- Kumar V., Lammers F., Bidon T., Pfenninger M., Kolter L. et al.*, 2017. The evolutionary history of bears is characterized by gene flow across species // *Sci. Rep.* V. 7. Art. 46487.
- Lohse K., Frantz L.A.*, 2014. Neandertal admixture in Eurasia confirmed by maximum-likelihood analysis of three genomes // *Genetics.* V. 196. № 4. P. 1241–1251.
- Lynd A., Weetman D., Barbosa S., Egyir Yawson A., Mitchell S. et al.*, 2010. Field, genetic, and modeling approaches show strong positive selection acting upon an insecticide resistance mutation in *Anopheles gambiae* s.s. // *Mol. Biol. Evol.* V. 27. № 5. P. 1117–1125.
- Maddison W.P.*, 1997. Gene trees in species trees // *Syst. Biol.* V. 46. № 3. P. 523–536.
- Mallet J.*, 2005. Hybridization as an invasion of the genome // *Trends Ecol. Evol.* V. 20. № 5. P. 229–237.
- Mallet J., Besansky N., Hahn M.W.*, 2016. How reticulated are species? // *BioEssays.* V. 38. № 2. P. 140–149.
- Malinsky M., Matschiner M., Svardal H.*, 2021. Dsuite – Fast D-statistics and related admixture evidence from VCF files // *Mol. Ecol. Resour.* V. 21. № 2. P. 584–595.
- Mao Y., Economo E.P., Satoh N.*, 2018. The roles of introgression and climate change in the rise to dominance of *Acropora corals* // *Curr. Biol.* V. 28. № 21. P. 3373–3382.
- Mao Y., Hou S., Shi J., Economo E.P.*, 2020. TREEasy: An automated workflow to infer gene trees, species trees, and phylogenetic networks from multilocus data // *Mol. Ecol. Resour.* V. 20. P. 832–840.
- Martin S.H., Amos W.*, 2021. Signatures of introgression across the allele frequency spectrum // *Mol. Biol. Evol.* V. 38. № 2. P. 716–726.
- Martin S.H., Davey J.W., Jiggins C.D.*, 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci // *Mol. Biol. Evol.* V. 32. № 1. P. 244–257.
- Meier J.I., Marques D.A., Mwaiko S., Wagner C.E., Excoffier L., Seehausen O.*, 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations // *Nat. Commun.* V. 8. Art. 14363.
- Meleshko O., Martin M.D., Korneliusen T.S., Schröck C., Lamkowski P. et al.*, 2021. Extensive genome-wide phylogenetic discordance is due to incomplete lineage sorting and not ongoing introgression in a rapidly radiated bryophyte genus // *Mol. Biol. Evol.* V. 38. № 7. P. 2750–2766.
- Miao B., Wang Z., Li Y.*, 2017. Genomic analysis reveals hypoxia adaptation in the Tibetan Mastiff by introgression of the gray wolf from the Tibetan Plateau // *Mol. Biol. Evol.* V. 34. № 3. P. 734–743.
- Mineta K., Goto K., Gojobori T., Alkuraya F.S.*, 2021. Population structure of indigenous inhabitants of Arabia // *PLoS Genet.* V. 17. № 1. Art. e1009210.
- Morrison D.A.*, 2014. Phylogenetic networks: A review of methods to display evolutionary history // *Annu. Res. Rev. Biol.* V. 4. № 10. P. 1518–1543.
- Musmann S.M., Douglas M.R., Bangs M.R., Douglas M.E.*, 2020. Comp-D: A program for comprehensive computation of D-statistics and population summaries of reticulated evolution // *Conserv. Genet. Resour.* V. 12. № 2. P. 263–267.
- Nevado B., Koblmüller S., Sturmbauer C., Snoeks J., Usano-Alemayn J., Verheyen E.*, 2009. Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish // *Mol. Ecol.* V. 18. № 20. P. 4240–4255.
- Nichols R.*, 2001. Gene trees and species trees are not the same // *Trends Ecol. Evol.* V. 16. № 7. P. 358–364.
- Osborne O.G., Chapman M.A., Nevado B., Filatov D.A.*, 2016. Maintenance of species boundaries despite ongoing gene flow in ragworts // *Genome Biol. Evol.* V. 8. P. 1038–1047.
- Ottensburghs J., Kraus R.H., Hooft P., van, Wieren S.E., van, Ydenberg R.C., Prins H.H.*, 2017. Avian introgression in the genomic era // *Avian Res.* V. 8. № 1. Art. 30.
- Palkopoulou E., Lipson M., Mallick S., Nielsen S., Rohland N. et al.*, 2018. Comprehensive genomic history of extinct and living elephants // *Proc. Natl. Acad. Sci.* V. 115. № 11. P. E2566–E2574.
- Patterson N., Moorjani P., Luo Y., Mallick S., Rohland N. et al.*, 2012. Ancient admixture in human history // *Genetics.* V. 192. № 3. P. 1065–1093.
- Pease J.B., Hahn M.W.*, 2015. Detection and polarization of introgression in a five-taxon phylogeny // *Systematic Biol.* V. 64. № 4. P. 651–662.
- Pease J.B., Haak D.C., Hahn M.W., Moyle L.C.*, 2016. Phylogenomics reveals three sources of adaptive variation during a rapid radiation // *PLoS Biol.* V. 14. № 2. Art. e1002379.

- Peters J.L., Zhuravlev Y., Fefelov I., Logie A., Omland K.E., 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.) // *Evolution*. V. 61. № 8. P. 1992–2006.
- Petit R.J., Excoffier L., 2009. Gene flow and species delimitation // *Trends Ecol. Evol.* V. 24. № 7. P. 386–393.
- Petr M., Vernot B., Kelso J., 2019. admixr – R package for reproducible analyses using ADMIXTOOLS // *Bioinformatics*. V. 35. № 17. P. 3194–3195.
- Poroshina A.A., Sherbakov D.Y., Peretolchina T.E., 2020. Diagnosis of the mechanisms of different types of discordances between phylogenies inferred from nuclear and mitochondrial markers // *Vavilov J. Genet. Breed.* V. 24. № 4. P. 420–426.
- Pritchard J.K., Stephens M., Donnelly P., 2000. Inference of population structure using multilocus genotype data // *Genetics*. V. 155. № 2. P. 945–959.
- Racimo F., Sankararaman S., Nielsen R., Huerta-Sánchez E., 2015. Evidence for archaic adaptive introgression in humans // *Nat. Rev. Genet.* V. 16. № 6. P. 359–371.
- Raj A., Stephens M., Pritchard J.K., 2014. fastSTRUCTURE: Variational inference of population structure in large SNP data sets // *Genetics*. V. 197. № 2. P. 573–589.
- Ravagni S., Sanchez-Donoso I., Vilà C., 2021. Biased assessment of ongoing admixture using STRUCTURE in the absence of reference samples // *Mol. Ecol. Resour.* V. 21. № 3. P. 677–689.
- Reich D., Thangaraj K., Patterson N., Price A.L., Singh L., 2009. Reconstructing Indian population history // *Nature*. V. 461. № 7263. P. 489–494.
- Rheindt F.E., Fujita M.K., Wilton P.R., Edwards S.V., 2014. Introgression and phenotypic assimilation in *Zimmerius* flycatchers (Tyrannidae): Population genetic and phylogenetic inferences from genome-wide SNPs // *Syst. Biol.* V. 63. № 2. P. 134–152.
- Rieseberg L.H., Wendel J.F., 1993. Introgression and its consequences in plants // *Hybrid Zones and the Evolutionary Process* / Ed. Harrison R. N.Y.: Oxford Univ. Press. P. 70–109.
- Rochus C.M., Tortereau F., Plisson-Petit F., Restoux G., Moreno-Romieux C. et al., 2018. Revealing the selection history of adaptive loci using genome-wide scans for selection: An example from domestic sheep // *BMC Genomics*. V. 19. № 1. Art. 71.
- Roda F., Mendes F.K., Hahn M.W., Hopkins R., 2017. Genomic evidence of gene flow during reinforcement in Texas *Phlox* // *Mol. Ecol.* V. 26. № 8. P. 2317–2330.
- Rosenberg N.A., 2002. The probability of topological concordance of gene trees and species trees // *Theor. Popul. Biol.* V. 61. № 2. P. 225–247.
- Sanchez-Donoso I., Huisman J., Echeagaray J., Puigcerver M., Rodríguez-Teijeiro J.D. et al., 2014. Detecting slow introgression of invasive alleles in an extensively restocked game bird // *Front. Ecol. Evol.* V. 2. Art. 15.
- Sang T., Zhong Y., 2000. Testing hybridization hypotheses based on incongruent gene trees // *Syst. Biol.* V. 49. № 3. P. 422–434.
- Schaefer J., Duvernell D., Campbell D.C., 2016. Hybridization and introgression in two ecologically dissimilar *Fundulus* hybrid zones // *Evolution*. V. 70. № 5. P. 1051–1063.
- Schrider D.R., Ayroles J., Matute D.R., Kern A.D., 2018. Supervised machine learning reveals introgressed loci in the genomes of *Drosophila simulans* and *D. sechellia* // *PLoS Genet.* V. 14. № 4. Art. e1007341.
- Schwenk K., Spaak P., 1995. Evolutionary and ecological consequences of interspecific hybridization in cladocerans // *Experientia*. V. 51. № 5. P. 465–481.
- Shurtliff Q.R., 2013. Mammalian hybrid zones: A review // *Mammal Rev.* V. 43. № 1. P. 1–21.
- Solis-Lemus C., Ané C., 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting // *PLoS Genet.* V. 12. № 3. Art. e1005896.
- Solis-Lemus C., Bastide P., Ané C., 2017. PhyloNetworks: A package for phylogenetic networks // *Mol. Biol. Evol.* V. 34. № 12. P. 3292–3298.
- Suarez-Gonzalez A., Lexer C., Cronk Q.C., 2018. Adaptive introgression: A plant perspective // *Biol. Lett.* V. 14. № 3. Art. 20170688.
- Taylor S.A., Larson E.L., 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature // *Nat. Ecol. Evol.* V. 3. № 2. P. 170–177.
- Than C., Ruths D., Nakhleh L., 2008. PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships // *BMC Bioinform.* V. 9. № 1. Art. 322.
- Thies L., Tomkovich P., Remedios N., dos, Lislevand T., Pinchuk P. et al., 2018. Population and subspecies differentiation in a high latitude breeding wader, the Common Ringed Plover *Charadrius hiaticula* // *Ardea*. V. 106. № 2. P. 163–176.
- Thom G., Amaral F.R.D., Hickerson M.J., Aleixo A., Araujo-Silva L.E. et al., 2018. Phenotypic and genetic structure support gene flow generating gene tree discordances in an Amazonian floodplain endemic species // *Syst. Biol.* V. 67. № 4. P. 700–718.
- Toews D.P., Brelsford A., 2012. The biogeography of mitochondrial and nuclear discordance in animals // *Mol. Ecol.* V. 21. № 16. P. 3907–3930.
- Twyford A.D., Ennos R.A., 2012. Next-generation hybridization and introgression // *Heredity*. V. 108. № 3. P. 179–189.
- Weigand H., Weiss M., Cai H., Li Y., Yu L. et al., 2017. Deciphering the origin of mito-nuclear discordance in two sibling caddisfly species // *Mol. Ecol.* V. 26. № 20. P. 5705–5715.
- Wen D., Yu Y., Zhu J., Nakhleh L., 2018. Inferring phylogenetic networks using PhyloNet // *Syst. Biol.* V. 67. № 4. P. 735–740.
- Whitney K.D., Randell R.A., Rieseberg L.H., 2010. Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus* // *New Phytol.* V. 187. № 1. P. 230–239.
- Zhang C., Ogilvie H.A., Drummond A.J., Stadler T., 2018. Bayesian inference of species networks from multilocus sequence data // *Mol. Biol. Evol.* V. 35. № 2. P. 504–517.

Approaches to detecting hybridization events and genetic introgression under phylogenetic incongruence

V. A. Komarova^{a, *}, L. A. Lavrenchenko^{a, **}

*^aSevertsov Institute of Ecology and Evolution, RAS
Leninsky pr., 33, Moscow, 119071 Russia*

**e-mail: k_yakka@yahoo.com*

***e-mail: llavrenchenko@gmail.com*

Hybridization and consequent exchange of genetic material (introgression) are widespread phenomena in the evolution of plants and animals. Interspecies gene flow can become obvious in phylogenetic analyses of different loci resulting in gene tree discordance. However, other evolutionary processes, such as incomplete lineage sorting, can also result in the phylogenetic incongruence. The rate and evolutionary significance of hybridization and introgression are difficult to evaluate due to the hardship in distinguishing hybridization from incomplete lineage sorting. The review provides a brief characterization of the main processes (hybridization and introgression, incomplete lineage sorting) that are responsible for the phylogenetic incongruence between gene trees. The review also summarizes classical phylogenetic and analytical methods contributing to detection and distinguishing hybridization and introgression from incomplete lineage sorting under comparing a phylogenetic trees inferred from separate genetic markers. The approaches developed not only for detection of introgression patterns in the presence of incomplete lineage sorting, but also for characterizing the revealed introgression (proportion of admixture, direction, time) are examined; the advantages and methodological limitations of each approach are briefly discussed. The methods of phylogenetic networks are also discussed, it is demonstrated that they are an integral tool for reconstructing complex evolutionary histories under reticulate events.