

НУЖНО ЛИ МНОГО ГЕНОВ ДЛЯ ФИЛОГЕНЕТИЧЕСКОГО ДЕРЕВА?

© 2007 г. В.В. Алешин^{1*}, А.В. Константинова¹,
К.В. Михайлов², М.А. Никитин², Н.Б. Петров¹

¹ НИИ физико-химической биологии им. А.Н. Белозерского
МГУ им. М.В. Ломоносова, 119992 Москва; факс: (495)939-3181,
электронная почта: Aleshin@genebee.msu.su

² Факультет биоинженерии и биоинформатики
МГУ им. М.В. Ломоносова, 119992 Москва

Поступила в редакцию 10.09.07

Методами филогеномики тестированы гипотезы о родственных отношениях многоклеточных животных (Metazoa) и других Opisthokonta, для чего отобрано 56 ядерных белоккодирующих генов из Taxonomically Broad EST Database и других баз данных. Результаты анализа подтверждают сестринские отношения Metazoa и воротничковых жгутиконосцев (Choanoflagellata), которые вместе с монофилетическим таксоном Ichthyosporae или Mesomycetozoea (включающим *Amoebidium parasiticum*, *Sphaeroforma arctica* и *Capsaspora owczarzakii*) объединяются как Holozoa. Эти филогенетические гипотезы получают высокую статистическую поддержку при учете как всего выравнивания, так и наборов длиной более 5000 сайтов, случайно выбранных из всего выравнивания. Согласно этим данным грибы подразделяются на Eumycota и кладу низших грибов. Низшие грибы образуют монофилетическую группу в составе Chytridiales + Spizellomycetales + Blastocladales (*Batrachochytrium*, *Spizellomyces*, *Allomyces*, *Blastocladia*), в противоположность более ранним сообщениям, основанным на анализе 18S рРНК и ограниченного набора белоккодирующих генов. Распределение на филогенетическом дереве гена, кодирующего убиквитин, слитый с рибосомным белком RpS30, показывает как минимум три независимых случая их слияния: у предков Holozoa, гетеротрофных Heterokonta (Oomycetes и *Blastocystis*) и предков Cryptophyta и Glaucophyta. Убиквитинподобные последовательности за пределами Holozoa не являются ортологами FUBI, и их эволюция в комплексе с белком RpS30 происходила независимо. Зарегистрировано повторное замещение FUBI на убиквитин в линии *C. owczarzakii* и, независимо, в монофилетической группе нематод Tylenchomorpha + Cephalobidae. У *Bursaphelenchus xylophilus* (Aphelenchoidoidea) сохраняется плезиоморфное, типичное для Metazoa состояние. Из имеющихся данных следует, что надежность филогенетических реконструкций определяется не столько числом сравниваемых генов или выраженным сходством в отдельных генах, но возможностью распознать артефакты как вычислительного, так и кладистического методов.

КЛЮЧЕВЫЕ СЛОВА: молекулярная эволюция, гомоплазии, происхождение многоклеточных, вторичная структура SSU рРНК, Tylenchida.

Настоящий выпуск посвящен 50-летию статьи [1], в которой впервые видовая специфичность нуклеиновых кислот была применена для решения конкретной таксономической задачи – ревизии систематического положения бактерии *Morganella* («*Proteus*») *morganii*. Редкая судьба оказалась у этой статьи: она открыла новое научное направление, по которому теперь в

университетах читают спецкурсы, сотни сотрудников работают в лабораториях во всем мире, издаются специализированные журналы, научные наработки которого применяют в экспресс-диагностике патогенов, селекционном деле, криминалистике, спорах по распределению квот на улов рыбы. За время активности одного научного поколения методики сравнительного анализа шагнули от определения нуклеотидного состава до сравнения полных геномов. Столь быстрые успехи высоко поднимают планку требований к методической стороне работ в этой области, оставляя без внимания методологию, списывая ее к принятой норме, моде, диктуемой методическим арсеналом. Сама постановка вопроса, вынесенного в заголовок, была немаловажна 50 лет назад: тогда не была известна структу-

Принятые сокращения: eEF1A – эвкариотический фактор элонгации 1A; EFL – паралог фактора элонгации 1A (elongation factor like); FUBI – убиквитинподобный пептид, слитый с рибосомным белком S30 животных и других Holozoa; Hsp – белок теплового шока, шаперон; RpL – рибосомный белок большой субчастицы; RpS – рибосомный белок малой субчастицы, *tef* – ген фактора элонгации 1A.

* Адресат для корреспонденции и запросов отпечатков.

ра ни одного гена. Сейчас известны первичные структуры всех генов примерно тысячи видов прокариотов и многих эукариотов, и филогенетические заключения «по одному гену» (например, 18S рРНК) теряют привлекательность в глазах широкой общественности, сообщения о них вызывают ставший банальным вопрос, подтверждаются ли эти выводы при анализе других генов. В основе такого вопроса лежит одновременно чуткость к живому течению современной науки (высоко поднимающей планку методических требований) и незнакомство с методологией систематики и филогенетики. Систематика призвана описать наиболее кратким способом все известное разнообразие [2]. Предмет ее изучения расширяется с каждым большим и малым открытием, он так же бесконечен, как познание. Геномные структуры для их краткого описания должны быть формализованы и систематизированы, как и структуры, видимые в микроскоп — световой и электронный, как типы дыхания и эмбрионального развития, способы регуляции активности генов и детали поведения, домики ручейников, гнезда муравьев и птиц — все, что уже открыто и будет открыто в будущем исследователями биоразнообразия; все это должно быть учтено в диагнозах таксонов, создаваемых систематикой, и повлиять на их состав. Задача филогенетики намного конкретнее, она конечна — установить родственные связи безотносительно к сходству и различию в целом [3–5]. Первая и простейшая теорема филогенетики на вопрос, подтверждаются ли филогенетические гипотезы анализом других признаков (генов), дает парадоксальный ответ: «не подтверждаются, но это, возможно, не имеет значения» (рис. 1).

Теоретически допустимо, что монофилетический таксон, возникший в ходе быстрой радиации или имеющий исключительно древний возраст, характеризуется одним единственным признаком. Даже если мы распознаем этот признак и каким-то образом сможем обосновать его надежность, это открытие нам мало что даст. Оправдание филогенетики состоит в надежде обнаружить у видов монофилетического таксона родственное сходство не по одному, а по многим признакам, в том числе не изученным до сих пор. Основания для такой надежды имеются: ведь эти виды родственнее друг другу, чем любым другим. Таким образом, от филогенетики ожидается прогностичность, выходящая, строго говоря, за ее рамки. И, что важнее, такому ожиданию противостоит более строгое предсказание, которое следует из самого факта эволюции: отдельные виды, эволюционируя от предка, неизбежно приобретают по некоторым признакам отличия, нарушающие их предковое

сходство. В результате виды, эволюционирующие медленнее хотя бы в отношении некоторых признаков, неизбежно по этим признакам будут более схожи между собой, чем с ближайшими родственниками. Это не имеет значения для филогенетики (но не может быть безразличным для систематики или практических приложений). Неопределенность биологической эволюции означает невозможность предсказать, по каким именно признакам предковое сходство утратится в результате эволюции, другими словами — в каких генах в отношении конкретного узла дерева обнаружится филогенетический

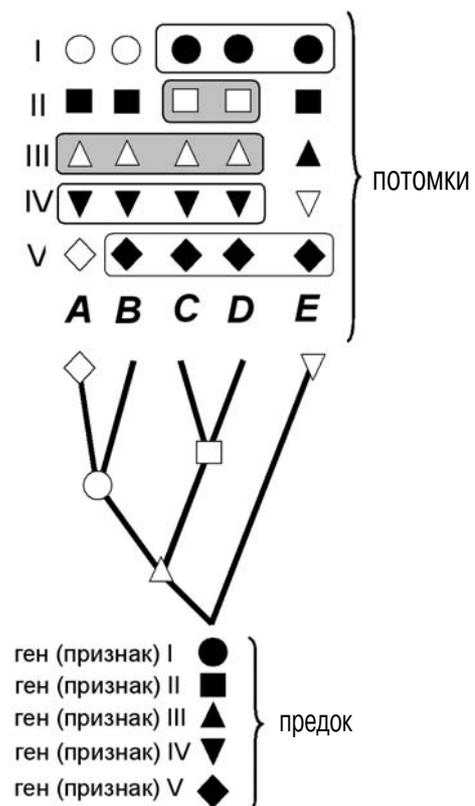


Рис. 1. Выделение монофилетических групп по совместным эволюционными приобретениями (синапоморфиям) и запрещенная в филогенетике процедура объединения по предковому (симплезиоморфному) сходству для известного дерева. A, B, C, D, E — современные виды (или расы, классы, типы — любые операциональные таксономические единицы), I–V — признаки или их группы (например, гены): предковое состояние — темные символы, апоморфное состояние — светлые символы; на дереве отмечены узлы, в которых приобретена соответствующая апоморфия, наследуемая видами-потомками; обведены группировки видов «по сходству» тех или иных признаков (некоторые варианты). Только объединения светлых фигур — монофилетические таксоны. Группировка по признаку IV выделена методически ошибочно, хотя и совпадает по объему с таксоном, выделенным по синапоморфии III; плезиоморфное сходство по признакам I и V не выделяет монофилетических групп

сигнал, а в каких — шум или «ложный» сигнал, например симплезиоморфия (рис. 1), и филогенетическая гипотеза «не подтвердится».

Таким образом, филогенетику должна интересовать степень надежности филогенетической гипотезы, а не то, как монофилия обосновывается: моногенным или мультигенным анализом. В любом случае найдутся признаки, если не противоречащие, то, по крайней мере, не подкрепляющие сделанный вывод. Другое дело, что при ограниченном числе признаков, например в последовательности одного гена 18S рРНК, их может вовсе не оказаться для установления монофилии конкретного таксона: ввиду слабого отличия его ближайшего общего предка от современных ему видов (малого времени существования стволковой группы) либо ввиду длительной независимой эволюции дочерних клад. Возможно, такие признаки будут обнаружены на большей выборке признаков (генов). Без выборки дополнительных признаков вряд ли можно распознать гомоплазии и точное обращение эволюции (нарушения закона Долло), а лабораторные ошибки загрязнения библиотек, хотя их можно выявить экспериментально, проще обнаруживаются по противоречиям с результатами анализа других генов. Далее мы рассмотрим на отдельных примерах филогенетические выводы, сделанные на наборах генов и на единичных генах.

ОДНОКЛЕТОЧНЫЕ РОДСТВЕННИКИ МНОГОКЛЕТОЧНЫХ ЖИВОТНЫХ

Молекулярные синапоморфии eEF1A. Поиск одноклеточных родственников животных — важная задача, решение которой может пролить свет на эволюцию клеточной коммуникации, закономерности исторического формирования онтогенеза многоклеточных, их роста и клеточной дифференцировки. Нынешние взгляды на этот вопрос берут начало от открытия в одном из белоккодирующих генов (фактора элонгации 1A, eEF1A — *tef*) животных и грибов специфической инсерции двенадцати триплетов [6]. Дело не просто в сходстве инсерций, отдельные черты сходства легко обнаружить у животных и с другими организмами. В случае eEF1A важно, что альтернативное состояние, наблюдаемое у растений, инфузорий, миксомицетов, дизентерийной амебы, эвглены, свойственно также EF-Tu, ортологичному белку прокариотов (архей и бактерий) — внешней группы по отношению к эвкариотам. Этот факт позволяет сделать выбор между двумя альтернативными эволюционными гипотезами: инсерции либо делеции 12

триплетов. По первой из них эволюционную историю EF1A можно представить очень просто — протекавшей от бактерий до эвкариотов без изменения длины. При этом достаточно допустить монофилию грибов и животных и однократное приобретение инсерции у их общего предка. Альтернативная гипотеза первоначально длинного гена потребует более одной независимой точной делеции указанного фрагмента в разных эволюционных линиях прокариотов и эвкариотов и, как менее экономная, отвергается. Таким образом, эволюцию eEF1A удается поляризовать, при этом грибы и животные, как имеющие не просто «общий признак», а эволюционно новый общий признак (синапоморфию), объявляются родственниками (рис. 1). Дополнительные аминокислотные остатки локализованы в неструктурированной области на поверхности белка, формируя пару антипараллельных β -листов. Остается неизвестным, связана ли их консервативность с какими-то функциональными особенностями eEF1A грибов и животных или только с внутри- или межмолекулярными ковариациями, замедляющими скорость молекулярной эволюции [7].

Результаты анализа других популярных маркеров: генов 18S и 28S рРНК, не противоречат гипотезе родства грибов и животных, дополняя группу одноклеточными (при умеренном уровне статистической поддержки): воротничковыми жгутиконосцами (*Choanoflagellata*), *Mesomycetozoea*, амебами нуклеариями и некоторыми другими протистами и «трихомицетами» [8]. Все эти группы были выделены в супертаксон *Opisthokonta*. Наряду с общим сходством последовательностей 18S рРНК, у всех их изученных представителей (более 3000) обнаружен мотив, входящий в консервативное основание спирали 49 (рис. 2) [9]. За пределами *Opisthokonta* этот мотив встречается как редкое исключение (ряд *Cercozoa*, по одному виду в родах *Goniomonas* и *Telonema*).

Хотя далеко не любой ген позволяет обнаружить монофилию *Opisthokonta*, в настоящее время она практически не ставится под сомнение [10, 11] (очень редко рассматриваются другие варианты [12, 13]), а использованная для ее обоснования кладистическая методология [3, 6] применяется и в решении других сложных вопросов, например, укоренения общего филогенетического дерева эвкариотов. Так, большинству эвкариотов свойственно объединение генов пиримидинового обмена дигидрофолатредуктазы и тимидилатсинтазы в один, кодирующий бифункциональный белок [14], тогда как у *Opisthokonta* — это два отдельных гена. Поскольку по этому признаку грибы и животные похожи на

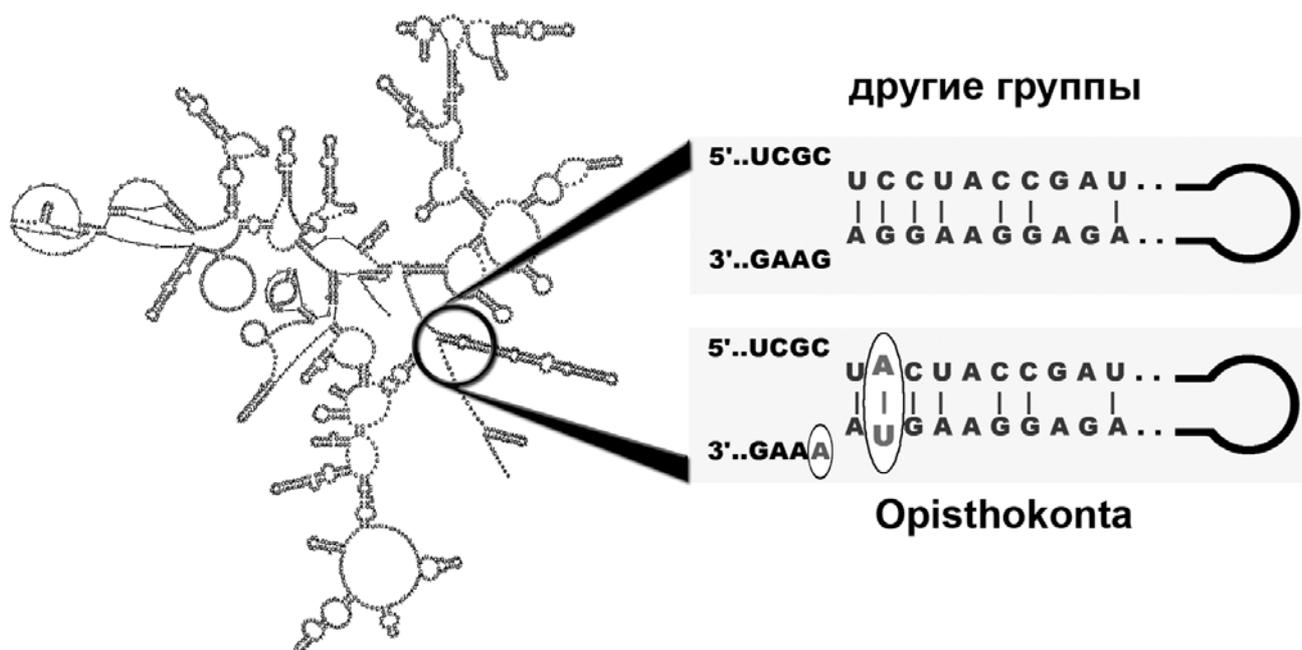


Рис. 2. Локализация в 18S рРНК нуклеотидного мотива, характерного для Opisthokonta

внешнюю группу – бактерий, то это сходство не может быть аргументом в пользу их родства [12], однако слияние дигидрофолатредуктазы и тимидилатсинтазы можно рассматривать как свидетельство родства (монофилии) большинства других эвкариотов, формирующих, таким образом, единую группу.

Современное состояние. Пятнадцать лет, прошедшие с публикации [6], подтвердили эвристическую и определенную прогностическую силу филогенетической гипотезы, предложенной «по одному гену». С тех пор в eEF1A множества видов Opisthokonta, в том числе у микроспоридий, воротничковых жгутиконосцев и представителей описанного позднее «класса» Mesomy-cetozoa, обнаружена предсказанная инсерция [11], хотя ее последовательность и длина оказались не столь постоянны, как это казалось при анализе менее широкой выборки. Ни у одного вида за пределами Opisthokonta не найдено характерной инсерции в eEF1A, несмотря на значительное увеличение базы данных. Современная картина все же оказывается более сложной ввиду обнаружения паралога eEF1A – EFL [15], который столь же консервативен, но сильно отличается по первичной структуре от eEF1A и всегда содержит инсерцию сходного размера и локализации. У некоторых видов Opisthokonta и представителей других групп эвкариотов вместо eEF1A находится EFL. Так, в геноме *Monosiga brevicollis*, который к насто-

ящему времени полностью секвенирован, имеется только ген, кодирующий EFL [16], и отсутствует ген eEF1A. В библиотеке кДНК (<http://tbest-db.bcm.umontreal.ca/searches/login.php>) другого воротничкового жгутиконосца, *Monosiga ovata*, напротив, присутствуют 104 клон кДНК eEF1A и не обнаруживается кДНК EFL. EFL замещает eEF1A и у многих хитридиевых грибов. Чтобы выяснить распределение EFL среди разных групп эвкариотов, нужны дополнительные исследования. Если бы в 1993 г. в ограниченный набор eEF1A попали бы и их паралоги, то гипотеза родства грибов и животных на тот момент могла быть не сформулирована.

Интересно, что механизированное применение вычислительных процедур, лишь частично основанных на кладистических принципах, легко ведет к ошибочным результатам, независимо от количества сравниваемых генов. Так, дерево по объединенным (конкатенированным) последовательностям 780 генов, выбранных из восьми полностью секвенированных геномов, и супердерево по 780 индивидуальным деревьям дают ошибочный результат, отвергающий монофилию Opisthokonta на статистически достоверном уровне [12]. Очевидно, причина ошибки в этом случае – артефакт «притяжения длинных ветвей» [11, 17]. Причины усиления таких артефактов при механическом увеличении числа анализируемых генов обсуждались ранее [7, 17, 18].

Несмотря на многочисленные попытки установить истинные филогенетические связи внутри клады Opisthokonta по одному или нескольким генам, современные гипотезы на этот счет весьма противоречивы. Сравнение митохондриального протеома поддерживает традиционное мнение о воротничковых жгутиконосцах как ближайших родственниках многоклеточных животных [19], оно же поддерживается наличием в eEF1A воротничковых и многоклеточных специфического аминокислотного мотива STEPPYS [11]; с другой стороны, на деревьях, построенных по данным анализа 18S рРНК, воротничковые иногда оказываются сестринской группой Mesomycetozoa [20], а недавно по результатам комбинированного анализа 18S и 28S рРНК был сделан вывод, что их сестринской группой является симбиотическая амeba *Capsaspora owczarzaki* [8]. Положение этой амебы внутри Opisthokonta другие авторы считают неопределенным [21, 22], либо же относят этот организм к Mesomycetozoa [23, 24]. Новейшее компромиссное предложение состоит в том, чтобы признать неопределенность родственных отношений животных, воротничковых жгутиконосцев, Mesomycetozoa, *Ministeria*, *Corallochytrium* и *Capsaspora*, графически изображаемое на дереве политомией [25].

Мультигенный анализ. Мы предприняли попытку разрешить указанную неопределенность с помощью геномных данных. Недавно в различных лабораториях [25] получены представительные библиотеки кДНК нескольких видов одноклеточных Opisthokonta, а геном одного вида, *Monosiga brevicollis*, полностью секвенирован в Joint Genome Institute (<http://genome.jgi-psf.org/Monbr1/Monbr1.home.html>). Из имеющихся баз данных мы отобрали 56 ортологов, которые имели представителей во всех группах Opisthokonta; 35 из них кодируют рибосомные белки, среди остальных имеются гены, кодирующие транскрипционные факторы, шапероны, белки цитоскелета, аппарата трансляции и энергетического обмена. Их аминокислотные последовательности выравняли вручную, после чего конкатенировали и строили дерево методом максимальной экономии и байесовым методом. В последнем случае для вычисления максимального правдоподобия использовали оптимизированные для каждого гена параметры (модель аминокислотных замен, степень гетерогенности набора по скорости накопления замен (α -параметр Γ -распределения), долю инвариантных сайтов). Общая длина выравнивания составила 10 678 позиций. Консенсусное дерево (рис. 3) согласуется с гипотезой сестринских отношений многоклеточных животных и воротничко-

вых жгутиконосцев. Ее апостериорная вероятность – 100%. Такова же апостериорная вероятность гипотезы монофилии Mesomycetozoa, включая *Capsaspora owczarzaki*. Обе эти группировки получают также 100%-ную поддержку при бутстрэп-анализе. На максимально экономном и байесовом дереве все Opisthokonta распадается на две эволюционные ветви. Одна объединяет Metazoa + Choanoflagellata и кладу Mesomycetozoa в таксон, для которого ранее предложено наименование Holozoa [19], другая в нашем наборе представлена грибами. Среди грибов со 100%-ной апостериорной вероятностью выделяется монофилетическая группа «низших» – хитридиевых грибов, для которых характерны жгутиковые стадии в жизненном цикле. Интересно, что анализ последовательностей 18S рРНК [26] и шести белкокодирующих генов грибов [27] не выявлял сестринских отношений хитридиевых, отделяя от них виды порядка Blastocladales. Анализ конкатенированных последовательностей 56 белков свидетельствует в пользу монофилии хитридиевых (рис. 3).

Преимущество топологий, содержащих указанные группы, подтверждается КН- [28], SH-

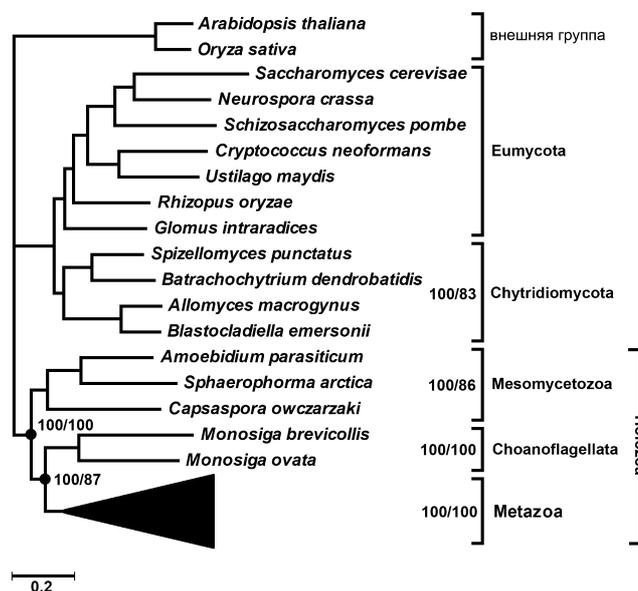


Рис. 3. Филогенетическое дерево Opisthokonta, построенное для 56 белков (10 678 позиций выравнивания) программой MrBayes 3.1.2 [58] с учетом параметров, оптимизированных для каждого белка, и программой protpars пакета PHYLIP [31] (100 реплик бутстрэпа). Общие параметры MrBayes 3.1.2: nchains = 4; nrns = 2; rates = invgamma; ngamcat = 8; ngen = 2 000 000; burnin = 1 000 000. Указаны значения апостериорной вероятности и бутстрэп-поддержки обсуждаемых в работе таксонов

[29] и AU- [30] тестами. Альтернативные топологии, полученные из дерева на рис. 3 перемещением отдельных ветвей, хуже консенсусной согласно использованным критериям (таблица). В соответствии с тестами КН и AU, во всех случаях, за исключением изменений, нарушающих монофилию хитридиевых, отличия статистически значимы. Более либеральный SH-тест отвергает с вероятностью 0,995 гипотезу более тесного родства животных с Mesomycetozoea, чем с воротничковыми, и гипотезу отсутствия монофилии Holozoa.

Для оценки влияния длины последовательностей на статистическую поддержку рассматриваемых филогенетических гипотез из конкатенированного выравнивания 56 белков генерировали наборы частичных последовательностей, используя процедуру случайного удаления сайтов (jackknife) [31]. Было генерировано по 100 реплик длиной 75 и 50% от первоначальной; для последующих длин в 25, 12,5, 6,3 и 3,1% было генерировано по 20 наборов, для каждого из которых получено 100 реплик, т.е. всего получено 10 200 частичных последовательностей. На их основе программой protpars из пакета PHYLIP [31] построены максимально экономные деревья и установлен уровень поддержки интересующих нас узлов, который зависел линейно от логарифма длины используемого выравнивания (рис. 4, а, б). Подобная процедура была предложена ранее [32] и применена в работе [33]. При длине выравнивания 5000 и более случайно выбранных сайтов, гипотеза сестринских отношений многоклеточных животных и воротничковых жгутиконосцев получает высокую статистическую поддержку, так же как и гипотеза монофилии Mesomycetozoea, включая *Capsaspora*.

В диапазоне длин выравнивания 668–2670 сайтов статистическая поддержка этих гипотез умеренная, у нижней границы этого диапазона рассматриваемые группы во многих наборах не попадают в консенсусное мажоритарное дерево. Меньшая длина случайно выбранных сайтов анализируемых белков не позволяет методом максимальной экономии выявить рассматриваемые группы.

Проведенная грубая оценка свидетельствует, что длина большинства отдельных белков недостаточна, чтобы наудачу выбранный белок позволил выявить рассматриваемые группы. Эта оценка хорошо согласуется с результатами анализа четырех белков: eEF1A, Hsp70, актина и β-тубулина, ни один из которых по отдельности не обеспечивает высокой поддержки объединению животных и воротничковых, тогда как четыре белка вместе обеспечивают умеренную поддержку в сконструированном дереве такой группе (включающей также одноклеточный организм *Ministeria vibrans*) (данные табл. 2 работы [11]). Из нашего набора только четыре белка из 56 поддерживают монофилию животных и воротничковых жгутиконосцев с уровнем более 50% бутстрэпа при построении максимально экономных деревьев по отдельным выравниваниям и всего два белка поддерживают при таких условиях монофилию Mesomycetozoea.

Конечно, 5000 сайтов выравнивания нельзя рассматривать как универсальный размер выборки, пригодный для статистической оценки филогенетических гипотез. Например, монофилия Holozoa получает поддержку в более чем 85% псевдовыборок при длине анализируемого выравнивания всего в 333 случайных сайта и более 95% – при 668 сайтах (рис. 3). По-видимому,

Статистическая оценка альтернативных гипотез филогении Opisthokonta

Перемещены относительно лучшей топологии, рис. 3	КН	SH	AU
Blastocladales к Eumycota, [26, 27]	0,139	0,823	0,278
Blastocladales в корень Fungi	0,096	0,786	0,193
Mesomycetozoea к Choanoflagellata [20]	0,011	0,396	0,040
<i>Capsaspora owczarzaki</i> в корень Metazoa + Choanoflagellata	0,042	0,286	7e–005
<i>Capsaspora owczarzaki</i> в корень Holozoa	0,019	0,168	0,015
<i>Capsaspora owczarzaki</i> к Choanoflagellata [8]	0,014	0,087	0,004
Choanoflagellata в корень Holozoa	2e–004	0,003	5e–087
Mesomycetozoea к Fungi	0	0,004	8e–006

Примечание: Тестирование осуществляли с помощью программ TREE-PUZZLE 5.2 [57] и CONSEL [30]; правдоподобие для сайтов вычисляли согласно модели WAG + Γ (6 категорий).

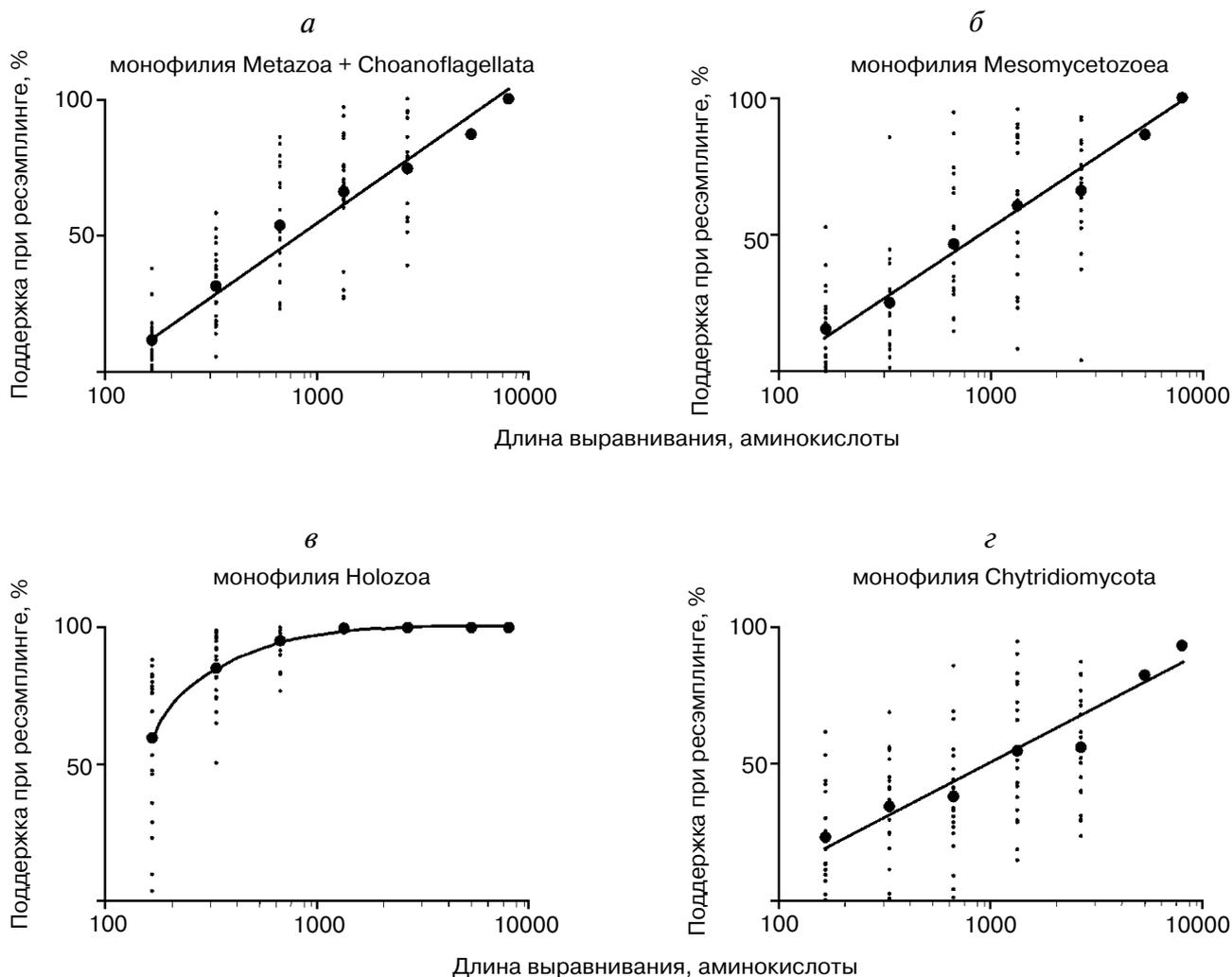


Рис. 4. Поддержка монофилии некоторых групп Opisthokonta в максимально экономных деревьях в зависимости от длины выравнивания: *a* – животные + воротничковые, *б* – Mesomycetozoea, *в* – Holozoa, *г* – хитридиевые (Chytridiales + Spizellomycetales + Blastocladales)

стволовая группа Holozoa существовала значительно дольше и накопила больше отличий от сестринской, ведущей к грибам, по сравнению со ствольными группами дочерних клад Choanoflagellata + Metazoa и Mesomycetozoea, рано разделившихся на, соответственно, линии воротничковых жгутиконосцев и животных и линии *Capsaspora* и *Amoebidium* + *Sphaeroforma*. По этой причине, вероятно, при одной и той же длине выравнивания монофилия Holozoa получает высокую статистическую поддержку, а монофилия Choanoflagellata + Metazoa или Mesomycetozoea – умеренную.

Для выявления позиций выравнивания, согласующихся или не согласующихся с деревом на рис. 3, программой *protpars* пакета PHYLIP [31]

были сконструированы аминокислотные последовательности для узлов дерева по рис. 3 или же деревьев альтернативной топологии. Предположительными синапоморфиями Mesomycetozoea являются аминокислотные замены в белке 2 потенциалзависимого анионного канала (порина 2 внешней мембраны митохондрий, VDAC-2) (G94 → N), белке теплового шока Hsp86 (HS90A_HUMAN) (S165 → T), рибосомных белках RpS2 (P208 → G, V209 → T), RpS25 (G13 → T, K17 → E), RpL2 (V62 → H, R184 → A), RpL6 (V121 → T), RpL10 (D55 → L) и другие (здесь и далее нумерация по белкам человека). Потенциальными синапоморфиями хитридиевых являются аминокислотные замены в метионинаденилтрансферазе II (A68 → L, H122 → L, D129 → K,

V195 → P, L261 → V), факторе процессивности ДНК-полимеразы δ (PCNA) (E93 → G, Q184 → P), один–три С-концевых кислотных остатка убиквитинлигазы E2 (UBC9) и другие. Эти признаки локализованы в эволюционно консервативных областях макромолекул. Тем не менее почти все они подвержены гомоплазиям. Например, у всех позвоночных, от лучеперых рыб до млекопитающих, имеется два паралога: метионинаденилтрансфераза II с «обычным», свойственным множеству беспозвоночных, грибов и растений остатком L261, и S-аденозилметионинсинтаза I с остатком V261, как у хитридиевых. Судя по родословному дереву этих белков (данные не представлены), дупликация произошла у общего предка позвоночных и сопровождалась заменой L261 → V в одном из паралогов. Признав два независимых случая замены L261 → V: у родоначальника S-аденозилметионинсинтазы I позвоночных и предков хитридиевых, нельзя отрицать возможности и большего числа аналогичных событий, т.е. независимого возникновения признака V261 в различных линиях хитридиевых. Таким образом, этот, как и другие найденные признаки, в отдельности кажутся недостаточными для доказательства монофилии рассматриваемых групп. Среди потенциально значимых филогенетических маркеров выделяется рибосомный белок RpS30, который у всех Holozoa отличается от соответствующего ему белка грибов наличием на N-конце специфического пептида из нескольких десятков аминокислот.

ИСТОРИЯ ГЕНА РИБОСОМНОГО БЕЛКА RPS30

Opisthokonta. Потенциальным маркером, пригодным для распознавания ближайших родственных многоклеточных животных, является рибосомный белок RpS30. У Metazoa он кодируется единственным геном *fau*, направляющим синтез слитного белка, N-конец которого FUBI напоминает убиквитин [34, 35]. FUBI человека отличается от убиквитина заменами 47 аминокислот из 74, и кроме того, делецией двух аминокислот. Филогенетическое значение FUBI определяется тем, что у других Opisthokonta – грибов, RpS30 не слит с другими белками [36], так же как во внешней группе – у растений и большинства протистов. Другими словами, слияние RpS30 – это эволюционное приобретение (апоморфия) предков Metazoa. Чтобы установить, на каком уровне филогенетического дерева произошло это событие, мы провели поиск гомологов RpS30 в доступных базах данных раз-

личных Opisthokonta. Обнаружилось, что в библиотеках кДНК всех одноклеточных Holozoa (воротничковых жгутиконосцев *Monosiga ovata* и *M. brevicollis* и Mesomycetozoa: *Amoebidium parasiticum* и *Sphaeroforma arctica*) белок RpS30 кодируется слитно с убиквитинподобной последовательностью, за исключением одного вида, *Capsaspora owczarzaki*, где он слит с убиквитином (рис. 4). Такое распределение признаков среди современных видов согласуется с гипотезой о слиянии этих генов у общего предка Holozoa, подтверждая, таким образом, их монофилию. FUBI многоклеточных и одноклеточных Holozoa, несмотря на вариабельность, имеет характерные особенности, подтверждающие его общее происхождение (делецию двух аминокислотных остатков вблизи N-конца, замены I13 → H, Q39 → P, G53 → E/D (преимущественно), H70 → E/D (преимущественно), преимущественную утрату функционального остатка K48 – сайта убиквитинилирования убиквитина).

По наиболее экономному эволюционному сценарию на первом этапе произошло слияние RpS30 и убиквитина, причем в линии *C. owczarzaki* это предковое состояние и сохранилось, а в линии остальных Holozoa в убиквитиновой половине накопились изменения, превратившие ее в FUBI. В пользу этого говорит факт отсутствия ортологов FUBI за пределами Holozoa. Многочисленные убиквитинподобные белки других эвкариотов и паралоги в геномах многоклеточных животных, отличаясь от типичного убиквитина, не обнаруживают специфического сходства с FUBI. Значит, FUBI как специфическая модификация убиквитина возник именно у Holozoa и именно в связи с RpS30. Другой слитый с убиквитином рибосомный белок, RpS27a, показывает высокую консервативность убиквитиновой части на протяжении всей эволюции эвкариотов. Например, в этой области белка человека и *Arabidopsis thaliana* имеется три аминокислотные различия. Однако последовательности человека и нематоды *Caenorhabditis elegans* отличаются в этой области 39 аминокислотными остатками и тремя делециями [37]. Очевидно, у далекого предка *C. elegans* произошло некое изменение, сделавшее необязательным сохранение консервативной структуры убиквитина, слитого с RpS27a, и у различных видов Rhabditidae (*C. briggsae*, *Nippostrongylus brasiliensis*, *Pristionchus pacificus* и др.) он превратился в вариабельный убиквитинподобный белок (отличающийся от FUBI). У нематод *Ascaris*, *Xiphinema* и других, не принадлежащих к кладе Rhabditidae, сохраняется консервативное состояние убиквитина, слитого с RpS27a. По-видимому, убиквитин, слившийся у предка Holozoa с

RpS30, в общих чертах повторил более наглядную историю убиквитина нематод, слитого с RpS27a.

Несмотря на кажущуюся логичность, описанный максимально экономный сценарий эволюции RpS30 Holozoa вызывает серьезные сомнения. Ему противоречат приведенные выше данные мультигенного анализа в пользу монофилии Mesomycetozoea. Значит, эволюция RpS30 отклонялась от наиболее экономного сценария. Принимая выраженное сходство FUBI *Amoebidium parasiticum* и *Sphaeroforma arctica* с последовательностями Metazoa и Choanoflagellata, остается признать, что после формирования слитого белка FUBI-RpS30 у ближайшего общего предка Holozoa, в линии *C. owczarzaki* произошло повторное замещение FUBI на убиквитин, своеобразное нарушение закона Долло.

Попытка найти аналогичный пример выявления замещения FUBI на убиквитин у нематод Tylenchomorpha (*Meloidogyne chitwoodi*, *M. incognita*, *M. javanica*, *Globodera rostochiensis*, *Heterodera glycines*, *Pratylenchus vulnus*) и Cephalobidae (*Zeldia punctata*). Типичное для Metazoa состояние, наряду с другими видами, сохраняет *Bursaphelenchus xylophilus*, традиционно рассматриваемый в составе семейства Aphelenchoidoidea отряда Tylenchida. Анализ последовательностей 18S рРНК согласуется с монофилией Cephalobidae и Tylenchomorpha и самостоятельностью ветви Aphelenchoidoidea [38], хотя резко различающаяся длина ветвей масштабированного дерева заставляла, учитывая опасность артефактов «притяжения длинных ветвей», проявлять осторожность и воздерживаться от таксономических выводов. Замещение FUBI на убиквитин представляет ясную синапоморфию Cephalobidae и Tylenchomorpha, свидетельствующую об их монофилии и невозможности объединения Tylenchomorpha с Aphelenchoidoidea в филогенетической системе. Таким образом, находят естественное объяснение существенные различия раннего эмбрионального развития *Aphelenchoides* и *Aphelenchus* [39] и другие их отличия. Технически замещение FUBI на убиквитин могло проходить по механизму перетасовки экзонов, поскольку последовательности, кодирующие FUBI и RpS30, разделены интроном [36].

Наличие сложных преобразований слитных белков в эволюции Holozoa поднимает вопрос о функциональной роли соединения убиквитина или FUBI с рибосомными белками. Ранее было экспериментально установлено, что убиквитиновые хвосты RpS27a и RpL40, еще одного слитого с убиквитином рибосомного белка, не являются жизненно необходимыми для дрожжей, но их отсутствие замедляет рост штаммов и мо-

жет быть компенсировано увеличением копииности этих генов [40]. Возможно, убиквитин способствуют фолдингу RpS27a и RpL40 и биогенезу рибосом [40]. Учитывая, что назначение FUBI в настоящее время неизвестно [41], этих предположений, конечно, недостаточно для раскрытия физиологических оснований обнаруживаемой конвергенции.

Другие эвкарियोты. Слияние генов убиквитина и RpS30 представляет наглядный пример гомоплазии — независимого возникновения строго гомологичных структур у предков Holozoa и гетеротрофных разножгутиковых (*Blastocystis* и Oomycetes) — видов, разделенных значительным расстоянием на дереве эвкарियोтов (рис. 5). В самом деле, ген, кодирующий RpS30 — уникальный, а в семействе убиквитинов не составляет труда распознать собственно убиквитин, отличающийся крайней консервативностью, от убиквитинподобных паралофов. Одинаковое соединение ортологов приводит к независимому возникновению полностью гомологичных слитных белков у *C. owczarzaki*, Tylenchomorpha и Heterokonta.

Интересно, что у разножгутиковых рода *Phytophthora* также возникает тенденция к видоизменению убиквитиновой части, слитой с RpS30. При этом не копируются изменения, происходившие в эволюции RpS30 Holozoa, а происходит делеция большей части убиквитина (*P. infestans*, *P. sojae*, *P. ramorum*, *P. brassicae*, *P. parasitica*). В оставшемся фрагменте убиквитина, содержащем 21 аминокислотный остаток, несмотря на произошедшие существенные изменения, сохраняется консервативный С-концевой мотив. Неизвестно, связано ли его сохранение с небольшим временем существования делеционной формы или с ковариациями в RpS30, препятствующими полной утрате фрагмента убиквитина. У *Phytophthora citrophthora* и других оомицетов (*Pythium ultimum*, *Saprolegnia parasitica*, *Aphanomyces cochlioides*) полная убиквитиновая последовательность, слитая с RpS30, сохраняет высокую консервативность.

Роль слияния этих двух белков как потенциального филогенетического маркера определяется тем, что у близких к разножгутиковым Alveolata [42, 43] и Excavata, так же как у автотрофных разножгутиковых (диатомовых, бурых водорослей), RpS30 не слит с убиквитином. Таким образом, слитный белок оомицетов и *Blastocystis* — очевидная синапоморфия. Положение *Blastocystis*, безжгутикового паразита человека и животных, в системе разножгутиковых остается спорным. Дерево последовательностей 18S рРНК Heterokonta [44] состоит из кроны с короткими ветвями, ведущими к современным

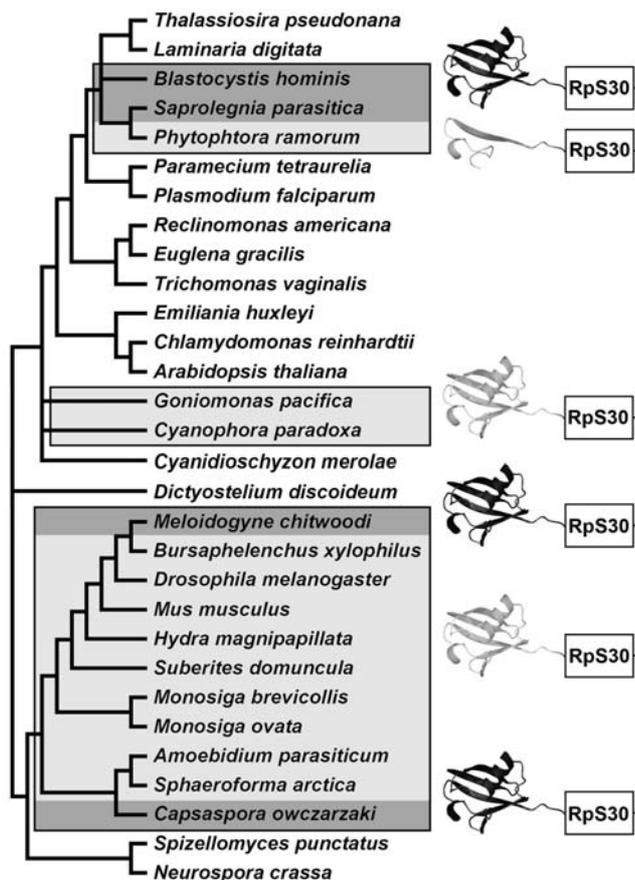


Рис. 5. Распределение на филогенетическом дереве убиквитина, слитого с рибосомным белком S30. Темная пиктограмма – убиквитин, светлая пиктограмма – FUBI или иной убиквитинподобный пептид, который в роде *Phytophthora* укорочен с N-конца; светлый прямоугольник – RpS30

фототрофным таксонам, и пучка многочисленных длинных ветвей, независимо отходящих от основания дерева к гетеротрофным таксонам, причем порядок их отделения не получает значимой статистической поддержки (рис. 3 в работе [44]). Было бы соблазнительно объединить некоторые из них на основании синапоморфий в независимых генах, например Oomycetes и *Blastocystis* – по наличию слитого с убиквитином RpS30. Однако, учитывая неоднократность таких слияний в эволюции эвкариотов, этот признак вряд ли можно признать решающим. Для подтверждения или опровержения гипотезы монофилии Oomycetes и *Blastocystis* желательны признаки, менее подверженные гомоплазии.

К настоящему времени обнаружено еще два таксона с RpS30, слитым с убиквитинподобным пептидом. Это одноклеточные глаукофитовые и криптофитовые водоросли – небольшие по чис-

лу видов изолированное группы эвкариотов. Глаукофитовые известны тем, что в качестве фотосинтетических органелл содержат в клетках цианеллы – облигатные симбиотические цианобактерии с геномом, близким по размеру к хлоропластному, а пластида фототрофных видов криптофитовых – это эвкариотический эндосимбионт с рудиментом ядра и хромосомами. Можно предполагать, что криптофитовые и глаукофитовые занимают центральное место в истории возникновения фотосинтезирующих эвкариотов, между тем родственные отношения «хозяйского» компонента их составных клеток остаются до сих пор мало изученными. На деревьях рРНК ядерные гены криптофитовых и глаукофитовых объединяются в единый кластер [45–47], однако истинность этой группировки обычно ставится под сомнение ввиду ее недостаточной статистической поддержки. Слияние RpS30 с убиквитинподобным пептидом (рис. 4) – дополнительный, хотя и недостаточно весомый аргумент в пользу монофилии криптофитовых и глаукофитовых. В отличие от FUBI животных, воротничковых и Mesomycetozoa, между убиквитинподобными последовательностями криптофитовых и глаукофитовых не наблюдается очевидного сходства в первичной структуре, возможно из-за недостаточной таксономической выборки. Это не позволяет исключить независимого слияния и независимого превращения убиквитина в слитном белке у их предков. В этом случае признак слияния обесценивается как филогенетический маркер, но делает более интригующим вопрос о причинах потери консервативности убиквитином независимо в разных таксономических группах именно при его слиянии с RpS30.

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Работы по мультигенной реконструкции филогении, появившиеся в последние годы благодаря геномным проектам и анализу целых библиотек кДНК, дают массовое подтверждение гипотезам, прежде сформулированным или получившим известность в результате анализа отдельных генов, в первую очередь рРНК малой субчастицы. Вряд ли здесь следует останавливаться на таксонах, фигурирующих в учебниках с XIX в. Родство составляющих их видов не вызывает удивления. Интереснее подтверждение обоснованности ранее выделенных по скудным молекулярным данным групп, монофилия которых неожиданна и в свете морфологической эволюции и не так легко может быть обоснована фактическим материалом из традиционного,

неполного багажа зоологии и ботаники. К ним относится объединение животных, грибов и некоторых одноклеточных в Opisthokonta [43]; иглокожих, полухордовых и *Xenoturbella* – в Ambulacraria [48]; круглых червей, членистоногих и других животных – в Ecdysozoa; кольчатых червей, моллюсков и плоских червей – в Lophotrochozoa [49]; сосудистых растений – с антоцеротовыми мхами [50, 51]; перемещение нимфейных, магнолиевых и *Amborella* в корень цветковых прежде выделения клад двудольных и однодольных [33]; включение «протистов» микроспоридий в состав многоклеточных животных [24] и другие нетривиальные предложения. Никто также не подвергает сомнению высокий таксономический ранг Archaea, установленный первоначально сравнением олигонуклеотидных карт рРНК. Как правило, установленные «по одному гену» группы сохраняются при условии значимой статистической поддержки, достаточной таксономической выборки, бесспорной ортологичности сравниваемых генов и отсутствия признаков притяжения длинных ветвей при построении дерева. Авторам не известно примера групп, установленных на основании последовательностей генов 18S рРНК, удовлетворяющих указанным условиям, которые были отвергнуты более поздними данными филогеномики.

И все же эпоха «великих филогенетических открытий по одному гену» уходит: привлечение внушительного массива данных становится нормой современных работ. Во многом это обстоятельство является следствием стремления поставить выдвигаемые гипотезы на прочный статистический базис. Кроме того, предполагается, что анализом множества генов можно усилить филогенетический сигнал до его превышения над шумом и добиться правильного разрешения конфликтов между различными генами. В настоящей работе мы обнаружили, что гипотезы монофилии Choanoflagellata + Metazoa, Mesomycetozoea (включая *Capsaspora owczarzaki*), Chytridiales + Blastocladales получают высокую статистическую поддержку при использовании более 5000 случайно выбранных позиций аминокислотного выравнивания, что превышает размеры большинства отдельных белков. При анализе меньшего числа признаков поддержка этих гипотез существенно понижается. Надо иметь в виду, однако, что число в 5000 позиций выравнивания является пороговым именно для указанных таксонов и не может претендовать на универсальность (в качестве примера можно привести работу [52]). Например, для выявления Holozoa на сконструированном дереве достаточно на порядок меньшего числа сайтов. По-видимому, это различие связано с быстрой ради-

ацией в одних случаях и длительной филетической эволюцией стволовой группы, позволившей накопить большое количество синапоморфий в различных генах, в последнем случае. Таким образом, уровень поддержки зависит не только от длины выравнивания, но и других факторов.

Более того, простое увеличение числа генов отнюдь не гарантирует реконструкции истинного дерева. Пример неверного дерева, не содержащего монофилетической группы Opisthokonta [12], упомянутый выше, не единственный. Так, результаты анализа 146 белкокодирующих генов (35 346 аминокислотных остатков) поддерживают, в зависимости от состава внешней группы (только дрожжи или дрожжи и воротничковые), альтернативные гипотезы монофилии либо целомических животных, либо Ecdysozoa [53]. Одна (по меньшей мере) из этих топологий ложная. Возможно та, которая содержит группу целомат и согласуется с деревом из другой работы [54], сконструированным по матрице наличия/отсутствия интронов в 3000 генов у 11 видов эвкариотов (всего 25 676 признаков). Такое же дерево можно получить по одному гену 18S рРНК этого же набора видов, если использовать заведомо неадекватные модели эволюции нуклеотидных последовательностей. На дереве 23-х видов, построенном по конкатенированному выравниванию 133 белков, плоские черви группируются с круглым червем *Caenorhabditis elegans*, однако эта группировка распадается, если *C. elegans* заменить другим видом круглых червей, *Xiphinema index*, образующим на дереве менее длинную ветвь [49]. Дерево полных хлоропластных геномов (более 150 т.п.н.) и дерево 18S рРНК (1,7 т.п.н.) тех же видов содержат одинаковые ошибки, возникающие, как было показано [55], от недостаточной таксономической выборки. Все это говорит о том, что нынешние методы реконструкции филогении по большому числу генов не избавляют от артефактов, известных для единичных генов. Здесь точно так же могут сказываться допущения моделей эволюции, разница в скорости эволюции видов, ошибки выравнивания и выбора ортологов, недостаточная репрезентативность таксономической выборки. Для устранения артефактов мультигенного филогенетического анализа предлагается селекция данных [17], что, конечно, делает его не столь формальным. Таким образом, практика современной филогеномики показывает, что статистическая поддержка реконструкций филогении повышается с увеличением числа сравниваемых генов, однако высокий уровень статистической поддержки дерева в целом или его отдельных узлов не может служить показателем правильности филогенетической реконструкции.

В то время как причины возможных ошибок при использовании численных методов построения филогенетического дерева по нуклеотидным или аминокислотным последовательностям весьма многочисленны и недостаточно исследованы, существует только три причины ошибок кладистического анализа: неверная «поляризация» эволюционного перехода (при этом за синапоморфию ошибочно принимается симплезиоморфия); неверная гомологизация (например, сравнение паралофов); гомоплазии и точные обращения эволюции. Две первых несут субъективный, а третья — объективный характер. Очевидно, что для обоснования филогенетических гипотез годятся только редкие события [7, 56], так как количество различных состояний молекулярных признаков конечно, что делает их уязвимыми к гомоплазиям. Устранение ошибок третьего рода возможно только путем сопоставления с независимыми маркерами и итеративного уточнения базовых представлений о закономерностях молекулярной эволюции: какие события могут повторяться, а какие следует полагать уникальными. Взятый отдельно факт замещения FUBI убиквитином требует объединения амебы *Capsaspora owczarzaki* с нематодами *Serphalobidae* и *Tylenchomorpha*, чему противоречат многие другие признаки. Но если такое замещение происходило в эволюции неоднократно, мы уже не можем рассматривать его как бесспорное доказательство монофилии этих нематод, а всего лишь как один из аргументов (пока все же весомый) в ряду других. Под сомнением оказываются и вычислительные, и кладистические реконструкции. Лучшим выходом кажется их совмещение, то есть проверка

правильности дерева одного или множества генов молекулярными синапоморфиями отдельных реконструированных узлов.

Как же найти для проверки ген или нуклеотид, достойный безграничного доверия? Чем меньший геологический срок существовала стволовая группа, тем меньше вероятность, что выбранный наудачу ген будет нести синапоморфию, притом не подверженную гомоплазиям и реверсиям. Чтобы добыть наверняка выигрышный билет в лотерее, есть способ — скупить весь тираж. Учитывая скорость развития технологии секвенирования и компьютерной обработки, применительно к геномам это может через несколько лет показаться не такой глупой идеей. С другой стороны, если родственное сходство у видов большое, то оно обнаружится во многих генах из числа выбранных наудачу и даже, вероятно, в одном достаточно протяженном гене, как 18S или 28S рРНК. Это родственное сходство будет проявляться если не в морфологии, то во многих других признаках организма, и установление факта монофилии группы вооружит нас множеством оправдывающихся предсказаний. Возникает вопрос, находящийся, однако, полностью за рамками филогенетики: сколько мы готовы заплатить за синапоморфные индикаторы родства, если они окажутся единственными чертами родственного сходства сравниваемых видов.

Авторы благодарны Л.Ю. Русину и П.В. Трошину за помощь в проведении компьютерных вычислений. Работа выполнена при финансовой поддержке РФФИ (гранты 05–04–49705 и 06–04–49288).

СПИСОК ЛИТЕРАТУРЫ

1. Спиринов А.С., Белозерский А.Н., Шугаева Н.В., Ванюшин Б.Ф. (1957) *Биохимия*, **22**, 744–754.
2. Беклемишев В.Н. (1994) *Методология систематики*, Изд-во КМК, Москва.
3. Hennig, W. (1966) *Phylogenetic systematics*, Illinois Univ. Press, Urbana.
4. Шаталкин А.И. (1988) *Биологическая систематика*, Изд-во МГУ, Москва.
5. Павлинов И.Я. (2005) *Введение в современную филогенетику (кладогенетический анализ)*, Изд-во КМК, Москва.
6. Baldauf, S.L., and Palmer, J.D. (1991) *Proc. Natl. Acad. Sci. USA*, **90**, 11558–11562.
7. Петров Н.Б., Алешин В.В. (2002) *Генетика*, **38**, 1043–1062.
8. Moreira, D., von der Heyden, S., Bass, D., Lopez-Garcia, P., Chao, E., and Cavalier-Smith, T. (2007) *Mol. Phylogenet. Evol.*, **44**, 255–266.
9. Cavalier-Smith, T., and Chao, E.E.-Y. (2003) *Protist*, **154**, 341–358.
10. Шаталкин А.И. (2005) *Журн. общ. биологии*, **66**, 389–415.
11. Steenkamp, E.T., Wright, J., and Baldauf, S.L. (2006) *Mol. Biol. Evol.*, **23**, 93–106.
12. Philip, G.K., Creevey, C.J., and McInerney, J.O. (2005) *Mol. Biol. Evol.*, **22**, 1175–1184.
13. Серавин Л.Н., Гудков А.В. (2005) *Журн. общ. биологии*, **66**, 212–223.
14. Stechmann, A., and Cavalier-Smith, T. (2002) *Science*, **297**, 89–91.
15. Keeling, P.J., and Inagaki, Y. (2004) *Proc. Natl. Acad. Sci. USA*, **101**, 15380–15385.
16. King, N., and Carroll, S.B. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 15032–15037.
17. Jeffroy, O., Brinkmann, H., Delsuc, F., and Philippe, H. (2006) *Trends Genet.*, **22**, 225–231.

18. Felsenstein, J. (1978) *Syst. Zool.*, **27**, 401–410.
19. Lang, B.F., O'Kelly, C., Nerad, T., Gray, M.W., and Burger, G. (2002) *Curr. Biol.*, **12**, 1773–1778.
20. Medina, M., Collins, A.G., Taylor, J.W., Valentine, J.W., Lipps, J.H., Amaral-Zettler, L., and Sogin, M.L. (2003) *Int. J. Astrobiol.*, **2**, 203–211.
21. Ruiz-Trillo, I., Inagaki, Y., Davis, L.A., Sperstad, S., Landfald, B., and Roger, A.J. (2004) *Curr. Biol.*, **14**, R946–947.
22. Ruiz-Trillo, I., Lane, C.E., Archibald, J.M., and Roger, A.J. (2006) *J. Eukaryot. Microbiol.*, **53**, 379–384.
23. Hertel, L.A., Bayne, C.J., and Loker, E.S. (2003) *Int. J. Parasitol.*, **32**, 1183–1191.
24. Jimenez-Guri, E., Philippe, H., Okamura, B., and Holland, P.W. (2007) *Science*, **317**, 116–118.
25. Ruiz-Trillo, I., Burger, G., Holland, P.W., King, N., Lang, B.F., Roger, A.J., and Gray, M.W. (2007) *Trends Genet.*, **23**, 113–118.
26. Lutzoni, F., Kauff, F., Cox, C.J., McLaughlin, D., et al. (2004) *Am. J. Bot.*, **91**, 1446–1480.
27. James, T.Y., Kauff, F., Schoch, C., Matheny, P.B., et al. (2006) *Nature*, **443**, 818–822.
28. Kishino, H., and Hasegawa, M. (1989) *J. Mol. Evol.*, **29**, 170–179.
29. Shimodaira, H., and Hasegawa, M. (1999) *Mol. Biol. Evol.*, **16**, 1114–1116.
30. Shimodaira, H., and Hasegawa, M. (2001) *Bioinformatics*, **17**, 1246–1247.
31. Felsenstein, J. (1993) *PHYLIP (Phylogeny Inference Package) version 3.5c*, Department of Genetics, University of Washington, Seattle.
32. Philippe, H. (1993) *Nucleic Acids Res.*, **21**, 5264–5272.
33. Логачева М.Д., Пенин А.А., Самигуллин Т.Х., Вальехо-Роман К.М., Антонов А.С. (2007) *Биохимия*, **72**, 1639–1646.
34. Kas, K., Michiels, L., and Merregaert, J. (1992) *Biochem. Biophys. Res. Commun.*, **187**, 927–933.
35. Perina, D., Cetkovic, H., Harcet, M., Premzl, M., Lukic-Bilela, L., Muller, W.E., and Gamulin, V. (2006) *Gene*, **366**, 275–284.
36. Baker, R.T., Williamson, N.A., and Wettenhall, R.E. (1996) *J. Biol. Chem.*, **271**, 13549–13555.
37. Jones, D., and Candido, E.P. (1993) *J. Biol. Chem.*, **268**, 19545–19551.
38. Meldal, B.H., Debenham, N.J., De Ley, P., De Ley, I.T., Vanfleteren, J.R., Vierstraete, A.R., Bert, W., Borgonie, G., Moens, T., Tyler, P.A., Austen, M.C., Blaxter, M.L., Rogers, A.D., and Lamshead, P.J. (2007) *Mol. Phylogenet. Evol.*, **42**, 622–636.
39. Дроздовский Э.М. (1968) *Докл. АН СССР*, **180**, 750–753.
40. Finley, D., Bartel, B., and Varshavsky, A. (1989) *Nature*, **338**, 394–401.
41. Rossman, T.G., Visalli, M.A., and Komissarova, E.V. (2003) *Oncogene*, **22**, 1817–1821.
42. Burki, F., Shalchian-Tabrizi, K., Minge, M., Skjaeveland, A., Nikolaev, S.I., Jakobsen, K.S., and Pawlowski, J. (2007) *PLoS ONE*, **2**, e790.
43. Rodriguez-Ezpeleta, N., Brinkmann, H., Burger, G., Roger, A.J., Gray, M.W., Philippe, H., and Lang, B.F. (2007) *Curr. Biol.*, **17**, 1420–1425.
44. Cavalier-Smith, T., and Chao, E.E.-Y. (2006) *J. Mol. Evol.*, **62**, 388–420.
45. Bhattacharya, D., Helmchen, T., Bibeau, C., and Melkonian, M. (1995) *Mol. Biol. Evol.*, **12**, 415–420.
46. Hoef-Emden, K., Marin, B., and Melkonian, M. (2002) *J. Mol. Evol.*, **55**, 161–179.
47. Deane, J.A., Strachan, I.M., Saunders, G.W., Hill, D.R.A., and McFadden, G.I. (2002) *J. Phycol.*, **38**, 1529–8817.
48. Bourlat, S.J., Juliusdottir, T., Lowe, C.J., Freeman, R., Aronowicz, J., Kirschner, M., Lander, E.S., Thorndyke, M., Nakano, H., Kohn, A.B., Heyland, A., Moroz, L.L., Copley, R.R., and Telford, M.J. (2006) *Nature*, **444**, 85–88.
49. Baurain, D., Brinkmann, H., and Philippe, H. (2006) *Mol. Biol. Evol.*, **24**, 6–9.
50. Samigullin, T.K., Yacentyuk, S.P., Degtyaryeva, G.V., Valieho-Roman, K.M., Bobrova, V.K., Capesius, I., Martin, W.F., Troitsky, A.V., Filin, V.R., and Antonov, A.S. (2002) *Arctoa*, **11**, 31–43.
51. Qiu, Y.L., Li, L., Wang, B., Chen, Z., Knoop, V., Groth-Malonek, M., Dombrovska, O., Lee, J., Kent, L., Rest, J., Estabrook, G.F., Hendry, T.A., Taylor, D.W., Testa, C.M., Ambros, M., Crandall-Stotler, B., Duff, R.J., Stech, M., Frey, W., Quandt, D., and Davis, C.C. (2006) *Proc. Natl. Acad. Sci. USA*, **103**, 15511–15516.
52. Rokas, A., Williams, B.L., King, N., and Carroll, S.B. (2003) *Nature*, **425**, 798–804.
53. Philippe, H., Lartillot, N., and Brinkmann, H. (2005) *Mol. Biol. Evol.*, **22**, 1246–1253.
54. Dopazo, H., Santoyo, J., and Dopazo, J. (2004) *Bioinformatics*, **20** (suppl. 1), i116–i121.
55. Degtyareva, G.V., Samigullin, T.Y., Sokoloff, D.D., and Vallejo-Roman, C.M. (2004) *Botanic J.*, **89**, 896–907.
56. Rokas, A., and Holland, P.W.H. (2000) *Trends Ecol. Evol.*, **15**, 454–459.
57. Schmidt, H.A., Strimmer, K., Vingron, M., and von Haeseler, A. (2002) *Bioinformatics*, **18**, 502–504.
58. Huelsenbeck, J.P., and Ronquist, F. (2001) *Bioinformatics*, **17**, 754–755.

**DO WE NEED MANY GENES
FOR PHYLOGENETIC INFERENCE?****V. V. Aleshin¹, A. V. Konstantinova¹,
K. V. Mikhailov², M. A. Nikitin², N. B. Petrov¹**¹ *A. N. Belozersky Institute of Physico-Chemical Biology,
M. V. Lomonosov Moscow State University, Moscow 119991,
Russia; fax: (495) 939-3181, E-mail: Aleshin@genebee.msu.su*² *Department of Bioengineering and Bioinformatics,
M. V. Lomonosov Moscow State University, Moscow 119991, Russia*

Received September 10, 2007

Fifty-six nuclear protein coding genes from Taxonomically Broad EST Database and other databases were selected for phylogenomic-based examination of alternative phylogenetic hypotheses concerning intergroup relationship between multicellular animals (Metazoa) and other representatives of Opisthokonta. The results of this work support sister group relationship between Metazoa and Choanoflagellata. Both of these groups form the taxon Holozoa along with the monophyletic Ichthyosporea or Mesomycetozoa (a group that includes *Amoebidium parasiticum*, *Sphaeroforma arctica*, and *Capsaspora owczarzaki*). These phylogenetic hypotheses receive high statistical support both when utilizing whole alignment and when only 5000 randomly selected alignment positions are used. The presented results suggest a division of Fungi into Eumycota and lower fungi. The latter form a monophyletic group that comprises Chytridiales + Spizellomycetales + Blastocladales (*Batrachochytrium*, *Spizellomyces*, *Allomyces*, *Blastocladia*), contrary to the earlier reports based on the analysis of 18S rRNA and a few protein coding genes. The phylogenetic distribution of a gene coding for a ubiquitin-fused ribosomal protein S30 implies at least three independent cases of gene fusion: in the ancestors of Holozoa, in heterotrophic Heterokonta (Oomycetes and Blastocystis) and in the ancestors of Cryptophyta and Glaucophyta. Ubiquitin-like sequences fused with ribosomal protein S30 outside of Holozoa are not FUBI orthologs. Two independent events of FUBI replacement by a ubiquitin sequence were detected in the lineage of *C. owczarzaki* and in the monophyletic group of nematode worms Tylenchomorpha + Cephalobidae. *Bursaphelenchus xylophilus* (Aphelenchoidoidea) retain a state typical of the rest of the Metazoa. The data demonstrates the fact that the reliability of phylogenetic reconstructions depends on the number of analyzed genes to a lesser extent than on our ability to detect reconstruction artifacts.

Key words: molecular evolution, homoplasy, origin of metazoans, secondary structure of SSU rRNA, Tylenchida