

ПРИРОДА

117049 Москва, Маро́новский пер., д. 26
тел. : [095] 238-24-56; e-mail: byalko@landau.ac.ru

№ 6 - 1999 г.

А.С. Антонов

Геномика и геносистематика

© Природа

*Использование или распространение этого материала
в коммерческих целях
возможно лишь с разрешения редакции*



Образовательный сетевой выпуск
VIVOS VOCO! - ЗОВУ ЖИВЫХ!
<http://www.techno.ru/vivovoco>

Геномика и геносистематика

А. С. Антонов



Андрей Сергеевич Антонов, доктор биологических наук, профессор, заведующий отделом эволюционной биохимии Научно-исследовательского института физико-химической биологии им. А.Н.Белозерского Московского государственного университета им. М.В.Ломоносова. Основные научные интересы связаны с геносистематикой растений. Неоднократно публиковался в «Природе».

ГЕНОМИКА в основе своей — та же геносистематика. Отличие заключается лишь в подходе к изучению геномов организмов. Исторически сложилось так, что сегодня геносистематики в основном изучают нуклеотидные последовательности фрагментов ДНК (например, генов) и на этой основе судят о родстве организмов. Геномика же исследует и сопоставляет целые геномы ядер или органелл клеток, в том числе и для целей филогенетики.

Задача эта и по сию пору, особенно в случае высших эвкарриот, очень сложная: ведь геномы наземных растений и животных построены из многих миллиардов нуклеотидов. Определить порядок их расположения в хромосомах можно только с применением новейших молекулярно-биологических методов, что требует больших затрат времени и труда. Однако, как говорится, овчинка стоит выделки: решение ее может пролить свет на целый ряд проблем филогенетики и систематики, не говоря уже о том, что геномика имеет и прикладное значение. Всем понятно, как важно, например, для лечения наследственных болезней человека знать строение полного генома здоровой и больной клетки. Как только выяснится характер повреждения «больного» гена, методы его «лечения» не заставят себя ждать, и сотни и тысячи ранее обреченных больных будут спасены.

Понятно, что одной медициной сфера применения геномики не ограничивается. Она будет полезна и в промышленности, и в сельском хозяйстве, и во многих других сферах деятельности человека, но об этом пусть напишут другие. Автор ставит перед собой иную задачу: показать, что геномика уже сегодня — весьма удобный и эффективный инструмент для решения проблем

классической биологии, связанных с изучением эволюции живого, и создания их систем.

Если геносистематика в целом — молодая наука, то геномика — вундеркинд в колыбели. Наибольших успехов она достигла при изучении микроорганизмов: размер геномов у них небольшой, и в настоящее время уже определены полные нуклеотидные последовательности хромосом нескольких десятков видов архе- и зубактерий. Такую работу обычно проводят многочисленные коллективы, включающие микробиологов, биохимиков, молекулярных биологов и геносистематиков. Например, в статье, где описывается строение генома сенной палочки, около 150 соавторов¹! Раньше, пожалуй, только физики-экспериментаторы создавали столь крупные исследовательские коллективы.

Из числа достижений геномики микробов отметим ее роль в выяснении распространенности латерального переноса генов у микроорганизмов. Интересные результаты получили при изучении генома одного из гипертермофильных микроорганизмов, *Aquifex aeolicus*². Оказалось, что его геном представляет собой сложную мозаику, состоящую как из «собственных» генов *A. aeolicus*, так и генов, «позаимствованных» им у других термофильных микробов. По мнению авторов этой работы, как минимум 10% генов появилось в геноме *A. aeolicus* в результате латерального переноса.

Перенос генов удалось зафиксировать и количественно оценить не только между термофильными архебактериями, но и между ними и зубактериями, адаптировавшимися к высоким температурам, т.е. между разными надцарствами живого. Не менее важные результаты получены при изучении обмена генами между патогенными зубактериями. Недавно показано, что к такому обмену способны бактерии из родов *Haemophilus* и *Neisseria*, в частности,

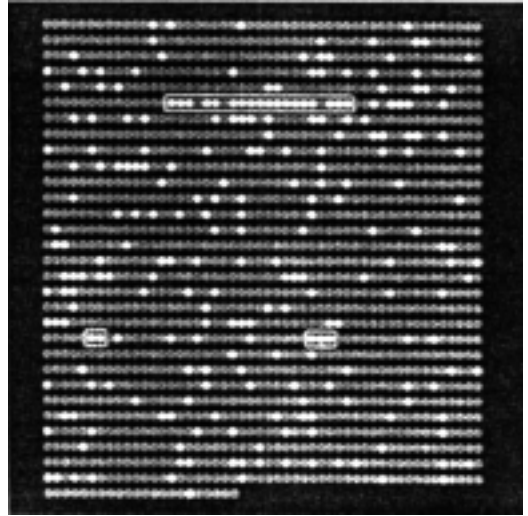


Схема строения генома гипертермофила *Aquifex aeolicus*. Светлыми кружочками показаны гены, перешедшие из геномов архебактерий. Блоки таких генов обведены рамками.

генами вирулентности. Это тревожный сигнал для медиков³!

В последнее время латеральный перенос фрагментов ДНК наблюдался уже во многих случаях: так, например, элемент «mariner» генома дрозофилы обнаружен в геноме простейшего, лейшмании, а целые блоки генов двудольных растений — в геномах однодольных.

Существование микроорганизмов с мозаичными геномами ставит перед систематиками сложную задачу. Совсем недавно, отказавшись построить филогенетическое дерево микробов по результатам изучения их фенотипов, микробиологи решили, что разумнее строить систему по молекулярным признакам, по строению генов рибосомной РНК (рРНК). При этом исходили из вполне разумного предположения, что чем более похожи эти гены, тем в более близком родстве находятся сравниваемые виды микробов. Иными словами, решили строить систему по результатам сопоставления всего одного гена, оставив в стороне все остальные. Хотя

¹ Kunst F., Ogasawara N., Moszer I. et al. // Nature. 1997. V.390. P.249-256.

² Aravind L., Tatusov R.I., Wolf Yu.I., Walker D.R., Koonin E.V. // TIG. 1998. V.14. P.442-444.

³ Kroll J.S., Wilks K.E., Farrant J.L., Langford P.R. // Proc. Natl. Acad. Sci. USA. 1998. V.95. P.12381-12385.

этот принцип и был положен в основу последнего определителя микробов Берджи, вряд ли он был самый правильный. Ведь число генов даже у наиболее просто устроенных микроорганизмов измеряется сотнями, и создавать систему по результатам изучения всего одного из них, пусть даже филогенетически весьма информативного, по меньшей мере легкомысленно.

Для других систематиков, изучающих более сложно устроенные организмы, такой подход к построению систем — давно пройденный и отвергнутый этап. Представьте себе, например, что мы строим систему семенных растений, исходя только из строения цветка. Образование этого важнейшего органа кодируется множеством генов и дает нам несравненно больше информации о родстве растений, чем сравнительное изучение генов их РНК. Однако такой путь не привел ботаников к успеху. Через некоторое время они убедились, что этот подход не позволяет построить филогенетическую систему растений, в которой группы видов (роды, семейства, порядки и пр.) объединялись бы, исходя из степени их эволюционного родства. Только комплексное исследование всех признаков ныне живущих и вымерших растений позволило им приблизиться к решению этой проблемы и выдвинуть ряд правдоподобных гипотез об эволюционной истории растений.

Иными словами, микробиологам предстоит выбрать еще какой-нибудь ген, а лучше несколько, изучить их строение не менее тщательно, чем нуклеотидные последовательности гена рРНК, и сопоставить полученные результаты. Такая работа уже начата, и четкие схемы, построенные с помощью одного гена рРНК, постепенно размываются. Микробиологам придется поломать голову над тем, что делать дальше.

Геномика также представляет собой способ комплексного изучения микроорганизмов, но вместо фенотипов она исследует их генотипы. Если оказывается, что микроорганизмы легко обмениваются генами, то создается любопытная ситуация: микробы становятся химера-

ми, несущими черты сразу нескольких видов, и, по мере все более подробного изучения их генотипов, все сложнее отнести каждый конкретный вид к какой-то определенной ветви их эволюционного развития.

Для решения этой нелегкой задачи нам, возможно, придется вспомнить и «пространство логических возможностей» Г.А.Заварзина, и другие аналогичные предложения систематиков⁴. Действительно, к какому отряду млекопитающих пришлось бы отнести кентавров? К рыбам или млекопитающим принадлежат русалки?

Успехи геномики других групп организмов несравненно скромнее. Продолжаются исследования генома человека, крестоцветного растения — арабидопсиса — и других сложно устроенных эукариот, количество ДНК в которых на порядки выше, чем в геномах микроорганизмов. В настоящее время известны полные первичные структуры всего у двух видов эукариот: дрожжей и возбудителя малярии. Однако работа над геномами других видов идет все быстрее и близится к завершению. Технический арсенал геномики все время обогащается методами, которые за меньшее время позволяют получать больше информации о строении ДНК. Именно поэтому некоторые специалисты считают, например, что всего через 5–7 лет будут определены полные первичные структуры геномов у представителей всех отрядов млекопитающих⁵. Получив в свои руки такую информацию, исследователи впервые смогут построить полное молекулярное дерево, отражающее родственные связи между этими животными и сопоставить эти связи со схемами, полученными при помощи классических методов филогенетики и палеонтологии.

Кажется, уже близка давнишняя мечта известного генетика Ф.Добжанского, который 30 лет назад писал: «Со временем нам, возможно, станет известной последовательность генетических

⁴ Заварзин Г.А. Пространство логических возможностей в многообразии бактерий и филогения // Природа. 1979. № 6. С.9–19.

⁵ O'Brien S.J., Weinderg J., Lyons L.A. // TIG. 1997. V.13. P.393–399.

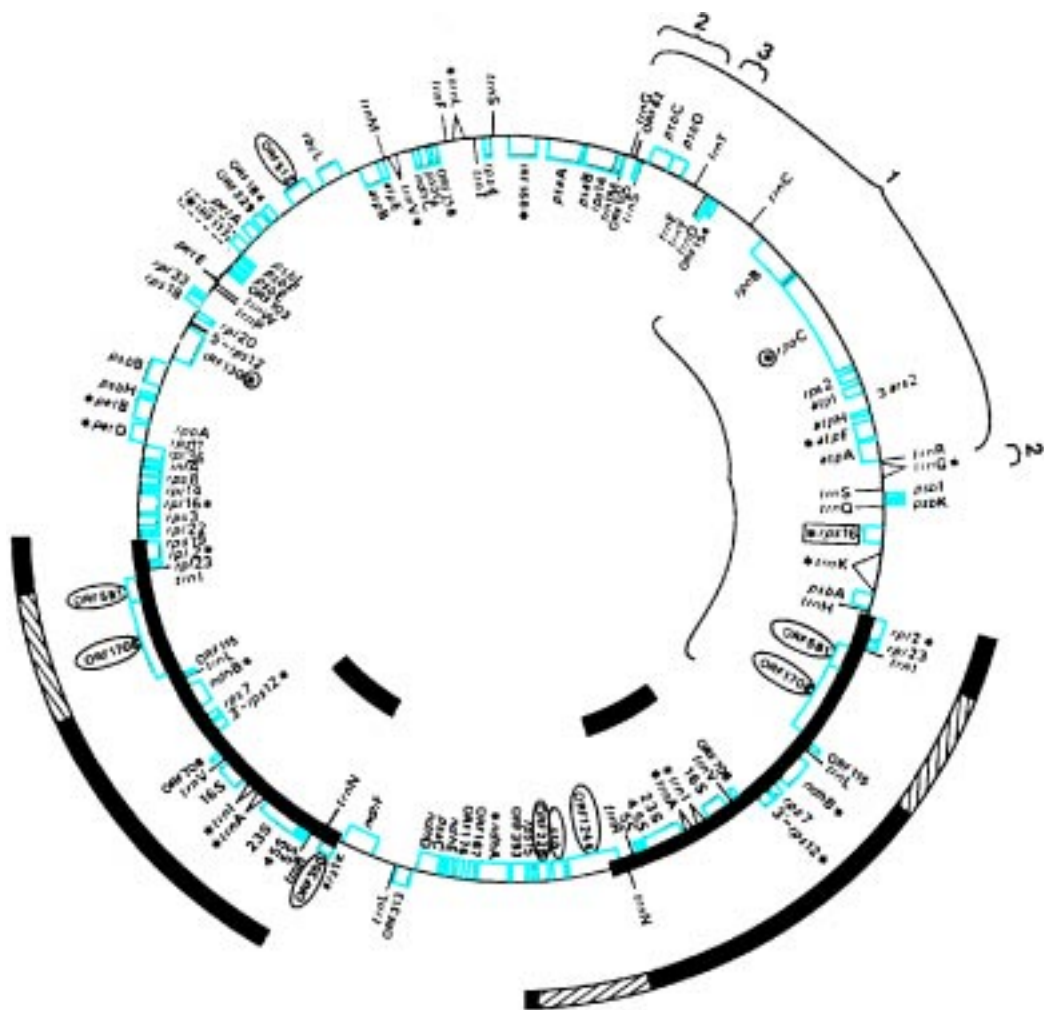


Схема строения хлДНК трех видов высших растений: маршанции, табака и риса. В основу положена хлДНК табака (156 тыс. пар нуклеотидов). Показано расположение отдельных генов в этой хромосоме. Снаружи кольца – гены, которые транскрибируются против часовой стрелки, а изнутри – по часовой стрелке. Кругочками обведены гены, присутствующие в хлДНК табака и маршанции, но отсутствующие в хлДНК риса. Цветом показаны гены, общие для табака и риса, но отсутствующие у маршанции. Ген, имеющийся только в хлДНК табака, выделен черным прямоугольником. Звездочками обозначены гены, имеющие во всех трех хлДНК одинаковые интроны. Цветные звездочки указывают на гены, интроны в которых есть только у маршанции и табака. Утолщенные части кольцевой хромосомы – обращенные повторы в геноме табака. Такие же утолщенные линии внутри и снаружи центральной хромосомы соответствуют обращенным повторам в хлДНК маршанции и риса. Участки, deletированные из обращенных повторов риса (при сопоставлении с хлДНК табака), на этих линиях показаны диагональной штриховкой. Скобкой внутри выделен участок, инвертированный в хлДНК маршанции. Пронумерованные скобки снаружи кольца указывают на инверсии в хлДНК риса. Гены обозначены общепотребительными символами. ORF – открытые рамки считывания (т.е. участки геномов, которые могут кодировать белки). (Palmer J.D. // *TIG*. 1990. V.6. P.115–120.)

«букв», нуклеотидов, во всех генах у всех организмов. Тогда окажется возможным количественно оценить сходства и различия между ними и создать на этой основе классификации живых существ, основанные на точной информации об их генетической составляющей...»⁶ Правда, пока неясно, будут ли схемы, построенные на основании изучения фенотипов и генотипов, топологически подобны. Ведь в процессах молекулярной и организменной эволюции есть ряд специфических особенностей, которые не позволяют считать, что эволюция фенотипов всегда полностью повторяет эволюцию генотипов.

Так обстоят дела с животными. Что же касается растений, то заниматься геномикой этой группы высших эукариот несравненно сложнее: размер генома у них, как правило, гораздо больше, чем у животных. Геносистематики не случайно выбрали для своих опытов именно арабидопсис: у него один из самых маленьких геномов среди покрытосеменных растений (у некоторых лилий, например, он чуть ли не в 1000 раз больше).

Но, как говорится, голь на выдумку хитра. Поняв, что ядерный геном большинства растений сегодня геномике не по зубам, они обратили свое внимание на геномы органелл — хлоропластов и митохондрий. Я расскажу лишь о результатах, полученных при изучении геномов хлоропластов.

Основы для геномики пластид растений были заложены давно. Еще в самом начале применения методов определения нуклеотидных последовательностей ДНК японские ученые попытались прочитать полные первичные структуры хлоропластных ДНК (хлДНК) двух далеко отстоящих друг от друга видов высших растений: примитивного бриофита — маршанции и одного из продвинутых видов покрытосеменных — табака. Вскоре эта задача была успешно решена⁷. К

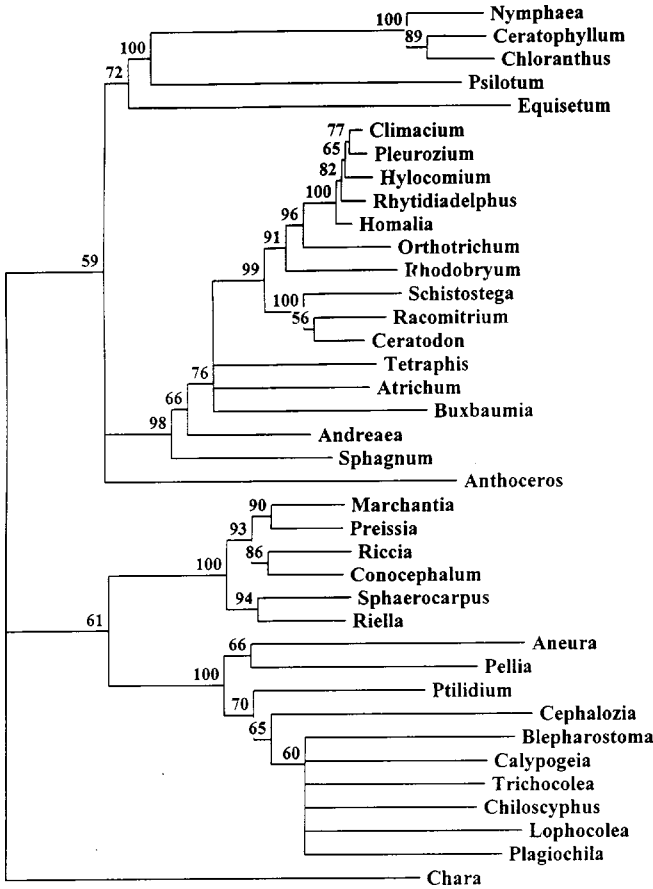
удивлению исследователей, хлДНК этих двух удаленных видов оказались весьма схожими, но не настолько, чтобы не заметить произошедшие в них в ходе эволюции изменения. Однако, чтобы составить представление об эволюции всех высших растений, необходимо знать структуры хлДНК по крайней мере у одного представителя из каждого крупного таксона. Но такие исследования развивались хаотически: одни считали, что прежде всего нужно определить строение ДНК у хозяйственно важных видов (риса, кукурузы и т.п.), а другие, увлеченные эволюционными проблемами, исследовали хлДНК сосны, папоротникообразного растения псилота и других видов. К началу 1999 г. достоверно было известно строение хлДНК около десяти видов высших растений.

Конечно, прочитать первичную структуру хлДНК несравненно легче, чем, скажем, ядерного генома арабидопсиса, но все же требует не меньших усилий, чем изучение строения генома микроба. И тут, что называется, геносистематиков растений бес попутал. Вместо этой тяжелой, но нужной работы, они занялись сравнением отдельных ядерных, хлоропластных и митохондриальных генов, надеясь, что таким, более легким, путем они быстрее разгадают тайну эволюции высших растений. При этом они, как правило, не утруждали себя поисками наиболее подходящего для этих целей гена или другого участка хлДНК, а пошли по принципу «поиска кошелька под фонарем».

В качестве основного объекта исследований первоначально взяли два гена: один из них кодировал структуру 18S рРНК, а другой — одну из субъединиц ключевого фермента фотосинтеза, ген *rbcL*. К началу 1999 г. структуру гена 18S рРНК определили у более чем 400 видов высших растений, а число изученных генов *rbcL* перевалило за 2,5 тыс. Однако уже в 1996–1997 гг. стало ясно, что работа эта ведется в значительной мере впустую: как показал специальный анализ, проведенный учеными Гарвардского университета, построить на основании этих результатов достаточно устой-

⁶ Dobshansky Th. Genetics of the evolutionary process. Columbia Univ. Press. N.Y.; L., 1970. P.414.

⁷ Shinozaki K., Ohme M., Tanaka M. et al. // EMBO J. 1986. V.5. P.2043–2049; Ohyama K., Fukuzawa H., Kohchi T. et al. // Nature. 1986. V.322. P.572–574.



Филогенетическое дерево, отражающее взаимосвязи важнейших таксонов наземных растений. Построено NJ-методом по результатам изучения эволюции спейсерных последовательностей ITS 2-4.

живые филогенетические деревья высших растений в принципе невозможно⁸.

Ведущий исследователь гена *rbcL* М.Чейз, группа которого исследовала строение этого гена у нескольких сотен видов растений, в конце концов признал, что наибольшее количество филогенетически значимой информации можно получить, изучая некодирующие последовательности ДНК (затем идут гены белков, а на последнем месте находятся гены рРНК)⁹.

Но Чейз запоздал с таким признанием. Еще в 1994 г. появилась статья, в которой прямо указывалось на преиму-

щества исследования некодирующих участков ДНК¹⁰. Эта работа лишь убедила в правильности выбора фрагмента некодирующей ДНК для изучения эволюции растений группой под руководством А.В.Троицкого в МГУ. Они исследовали эволюционные изменения (так называемый участок ITS 2-4) в участке хлоропластной ДНК, располагающемся между генами 23S рРНК и геном тРНК_{Карг}. Исследования велись в содружестве с группой У.Мартина (Генетический центр университета г.Брауншвейга, ФРГ), которая взяла на себя основную

⁸ Kellog E.A., Juliano N.D. // Amer. L. Bot. 1997. V.84(3). P.413-428.

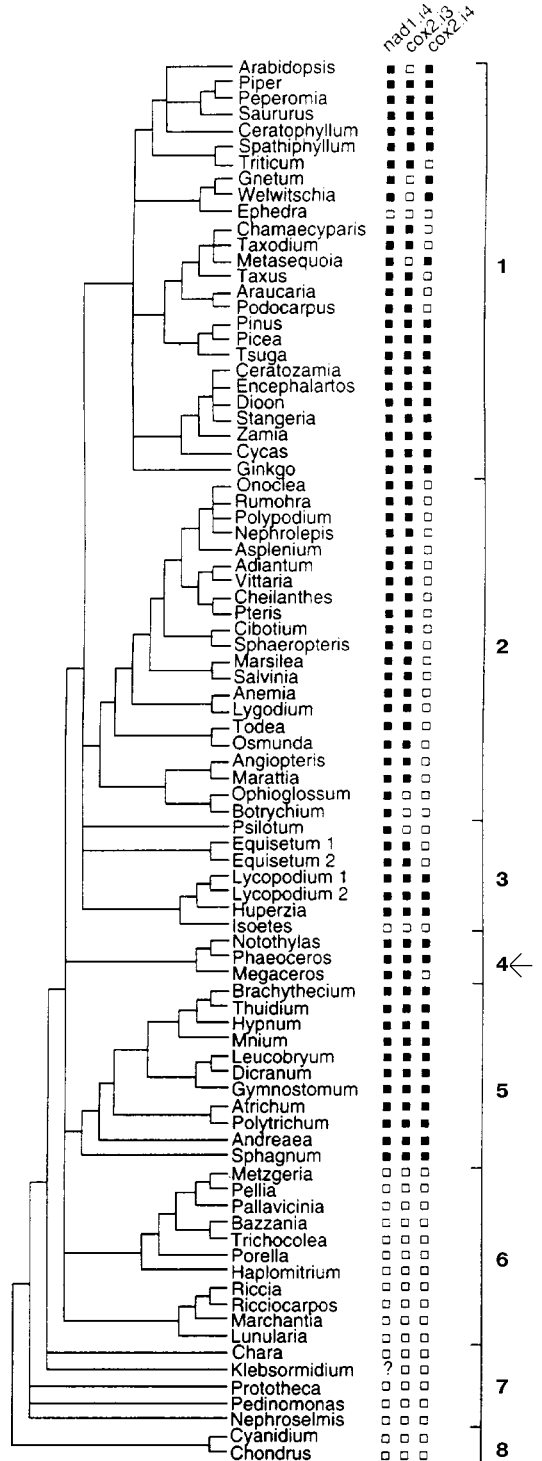
⁹ Chase M.H., Cox A.V., Soltis D.E. // Ibidem. V.84(6). P.181.

¹⁰ Bohle U.R., Hilger H.H., Cerff R., Martin W. Non-coding chloroplast DNA for plant molecular systematics at the infrageneric level // Molecular ecology and evolution: approaches and applications / Eds B.Schierwater et al. Basel, 1994. P.391-403.

филогенетическое дерево высших растений и водорослей, построенное по результатам сопоставления гена *rbcL*. Черные квадратики — наличие интронов, белые — их отсутствие в некоторых генах (указаны сверху) митохондриальной ДНК. 1 — семенные растения, 2 — папоротники, 3 — другие сосудистые растения, 4 — антоцеротовые мхи, 5 — мхи, 6 — печеночники, 7 — зеленые водоросли, 8 — красные водоросли. Стрелкой показано необычное положение антоцеротовых мхов. (Qiu Y-L. et al., 1998; Holzman D., 1998.)

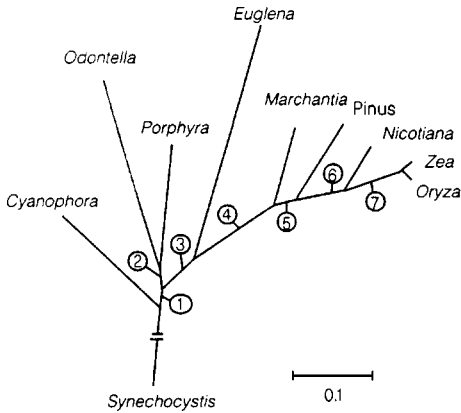
нагрузку по изучению ITS 2-3 папоротников, а группа А.В.Троицкого — ITS 2-4 у остальных высших растений. Конечным результатом работы стало филогенетическое дерево высших растений. Оно включает в себя далеко не все изученные виды, и построено исключительно для того, чтобы показать основные узлы дивергенции наиболее крупных групп растений¹¹.

Эта схема не уступает любому из построенных, предложенных ранее палеоботаниками или филогенетиками, исследовавшими фенотип растений. Есть у нее и небольшое отличие: необычное положение группы антоцеротовых мхов. Часть палеоботаников склонна считать, что эти мхи — эволюционно очень древняя группа, но, может быть, они принимают за предков антоцеротовых какие-то другие растения? Ведь в результате изучения митохондриальных интронов Ю.-Л.Кю с соавторами также построили дерево¹², на котором антоцеротовые мхи занимают то же положение, что и на дереве, построенном нами по результатам изучения эволюции ITS 2-4. Если бы мы знали полную первичную структуру хпДНК одного из видов антоцеротовых мхов, то окончательный ответ на этот вопрос был бы найден. К сожалению, такие исследования только планируются.



¹¹ См.: Антонов А.С. Происхождение основных групп наземных растений // Природа. 1997. № 10. С. 55-63.

¹² Qiu Y-L., Cho Y., Cox C., Palmer J.D. // Nature. 1998. V.394. P.671-674; Holzman D. // ASM News. 1998. V.64. P.679-680.



Филогенетическое дерево пластид водорослей и растений и их предполагаемого эволюционного предка, цианобактерии *Synechocystis*, построенное по результатам сопоставления 45 ортологичных белков NJ-методом. Цифрами показаны ветви со 100% статистической достоверностью. (Martin W. et al., 1998.)

В основном все схемы, построенные на основе изучения эволюции отдельных участков ДНК, почти одинаково описывают эволюцию высших растений. Однако есть обстоятельства, которые не позволяют считать их окончательным решением проблемы: во-первых, большинство таких деревьев топологически малостабильно, во-вторых, достоверность узлов ветвления обычно невелика, и, наконец, они не дают однозначных ответов на ряд ключевых вопросов филогенетики. Так, уже много лет продолжается дискуссия, есть ли среди современных голосеменных близкие родственники покрытосеменных растений (на эту роль обычно претендуют гнетовые).

Эти задачи и взялась решать геномика. Все началось с того, что бывший аспирант нашего отдела В.В.Горемыкин, будучи в командировке в Германии, в лаборатории У.Мартина, решил посмотреть, какие изученные к тому времени гены хлДНК несут максимальную филогенетическую информацию во всем спектре видов, от цианобактерий

до покрытосеменных¹³. Таких генов оказалось 45 (кстати, гена *rbcl* среди них не было). Стало ясно, что если изучать первичные структуры хлДНК, то исследователи получают несравненно более богатый материал для сопоставлений, чем при анализе отдельных фрагментов ДНК. Тогда на небольшом материале Горемыкину с соавторами удалось выяснить, что дивергенция голо- и покрытосеменных растений произошла приблизительно 350 млн лет назад, а двудольных и однодольных растений — 160 млн лет назад¹⁴. Надежность построенного ими простейшего филогенетического дерева была во много раз выше всех предыдущих деревьев. Ведь в его основу были положены результаты анализа эволюции не одного-двух, а сразу нескольких десятков генов. Очевидно, такую работу необходимо продолжать, и сегодня ряд сильных лабораторий идет по этому пути.

Надо сказать, что другие авторы предложили иное решение. Они стали соединять воедино данные, полученные для разных генов, анализировать такие искусственно сцепленные последовательности и в конечном счете отметили, что надежность выводов при этом повышается. Жаль только, что в число таких генов частенько попадали все те же гены *rbcl*, 18S рРНК и др., далеко не лучшие кандидаты на роль поводырей в темном прошлом наземных растений. Если Ч.Дарвин даже происхождение покрытосеменных растений называл «проклятой загадкой», то как можно назвать события, происходившие многие сотни миллионов лет назад, когда на поверхности Земли появились высшие растения? Надежду, что со временем ученые решат эту задачу, дает геномика.

Работа выполнена при поддержке РФФИ. Грант 96-15-97970.

¹³ Martin W., Stoebe B., Goremykin V. et al. // Nature. 1998. V.393. P.162-165.

¹⁴ Goremykin V., Hausman S., Martin W. // Syst. Evol. 1997. V.206. P.337-351.