

Лекция 14-ая

Часть 2

Компьютерное моделирование филогенетических связей видов полиморфных родов по методу SYNAP (на примере рода Молочай)

Лектор д.б.н. К.С.Байков.

Филогения ценит признаки с точки зрения их происхождения, она есть наука гомологий
Б.М. Козо-Полянский

Практика филогенетических построений имеет длительную историю, однако моделирование филогенеза как самостоятельное направление биологической систематики находится в стадии формирования. Традиционный подход основан на оценке общей примитивности и продвинутой семейств, родов и т.п. По небольшому числу специально отобранных признаков определяются основные направления эволюции, и затем строится филогенетическая схема, в большинстве случаев единственная (слайд 2). Единственность такой схемы связана с представлением об *уникальности конкретного исторического процесса*. Однако построение схемы, отражающей реальный филогенез, возможно лишь на основе *полных сведений о всех членах исследуемой группы*, как ныне живущих, так и вымерших, а также информации о конкретных исторических условиях, в которых проходило видообразование. Естественно полагать, что эти данные недоступны для исследователя в полной мере. Следовательно, *построение единственной схемы, соответствующей реальному ходу эволюции в данной группе, невозможно по независимой от исследователя, объективной причине*.

Возникает вопрос: что тогда можно предложить взамен? Не имея по объективной причине всех необходимых данных для воссоздания филогенеза, мы вынуждены отказаться от построения единственной эволюционной схемы, отражающей реальный исторический процесс во всех его подробностях. Мы можем рассматривать ее только как "идеал, к которому можно и нужно стремиться, но достичь которого заведомо невозможно" (Тихомиров, 1996: 141). Стремление к этому идеалу должно включать по крайней мере два момента: 1) *поиск и учет новых признаков*, важных для описания филогенеза исследуемой группы, и 2) *совершенствование нашего логического аппарата* в соответствии с процессом видообразования как природным феноменом с характерными для него чертами. Мы разделяем мнение А.К. Скворцова (1986) о том, что поиск и разработка гибких логических схем представляется наиболее эффективным средством филогенетического анализа. Обсуждению этого вопроса и посвящена данная лекция.

В основе филогенеза лежат процессы видообразования, поэтому монографию Ч.Дарвина "Происхождение видов путем естественного отбора" (Darwin, 1859; Дарвин, 1991) можно считать *первым логически завершенным руководством по филогенетическому анализу*. Подробно показанные на многочисленных примерах явления изменчивости, наследования признаков, дивергенции и конвергенции, гибридизации, адаптации, возврата к признакам предков легли в основу более поздних эволюционных построений. Многие из этих явлений обсуждались и до Дарвина, но именно он впервые соединил их в контексте процесса видообразования. К сожалению, здесь нет возможности подробно рассмотреть развитие представлений о видообразовании в рамках общей теории эволюции. В качестве наиболее полного и вместе с тем краткого обзора этих вопросов можно рекомендовать монографию А.С.Северцова "Основы теории эволюции" (1987).

В филогенетической систематике первой половины 20 века наблюдалась интеграция данных

из разных областей науки и усиление внимания к филогенетическим связям крупных групп (семейств, порядков и выше, см. слайд 3). При этом набор филогенетически значимых признаков был существенно увеличен, а А.Л. Тахтаджян обосновал важность их разделения на признаки уровня и признаки родства.

Для обсуждения конкретных процедур филогенетического анализа особое значение имеют методы, разработанные в последние 30-40 лет. Метод *В. Вагнера* (Wagner, 1961), основан на *дивергенции основного плана строения* (Ground-Plan-Divergence). Основной план строения по Вагнеру - это совокупность примитивных черт строения, свойственная исследуемой группе. Графическая схема дивергенции накладывается на последовательный ряд полуокружностей с общим центром (слайд 9). Виды располагаются на этих полуокружностях согласно числу уклонений от основного плана строения. При этом некоторые виды могут располагаться в точках ветвления схемы, другие - на концах ветвей. Следовательно, по Вагнеру современные виды могут быть связаны друг с другом отношением предок-потомок. Такая схема видообразования хорошо согласуется с представлением Дарвина о том, что после отделения дочернего вида материнский вид может длительное время продолжать свое существование (Дарвин, 1991).

Основоположник кладистики В. Хенниг (Hennig, 1950, 1966) предложил более строгую схему видообразования: отделение дочернего вида от материнского приводит к исчезновению материнского вида в его прежнем качестве (слайд 7). В силу достигнутой таким способом логической строгости метод Хеннига был положен в основу многих современных вычислительных алгоритмов, требующих максимальной формализации процедуры анализа. Однако нам представляется, что в случае воспроизводящихся систем, к которым относится биологический вид, потеря части генетической информации в результате отделения дочернего вида может восстанавливаться в процессе воспроизводства особей материнского вида. В схеме видообразования по Хеннигу это не допускается. Тем самым широкий спектр возможных путей образования новых видов сводится к простой дивергенции.

Логическая схема видообразования, подробно и в разных вариантах описанная Дарвином, сокращена в работе Хеннига до дихотомии и прекращения существования материнского вида в результате появления двух дочерних. Следовательно, упрощенное представление о процессах видообразования заключено уже в исходных установках его методологии. Это упрощение позволило Хеннигу создать строгую логическую модель, но ограничило возможности на ее основе корректно отразить реальные биологические процессы во всем их многообразии.

Из новых понятий, предложенных Хеннигом, именно синапоморфию следует считать основой его методологии. В чем особенность этого подхода и почему я считаю перспективным использовать его в филогенетических построениях? Говоря кратко, анализ синапоморфий дает возможность определить последовательность филогенетических событий и на этой основе проверить гомологичность признаков.

При сравнении разных методов филогенетической систематики были учтены работы А. Кронквиста (Cronquist, 1987), А. Тахтаджяна (1987), Г. Линдера (Linder, 1988), цикл статей российских зоологов (Емельянов, 1989, Несов, 1989, Песенко, 1989, Старобогатов, 1989) и др. Обзор многочисленных зарубежных публикаций по вопросам нумерической таксономии и кладистики составил М.Г. Пименов (1987). Несколько позднее А.И. Шаталкин (1988) опубликовал сравнительный анализ разных направлений в биологической систематике. Методам кладистики посвящена монография И.Я. Павлинова (1989), в которой дано краткое описание методологии и конкретных процедур анализа, а также его капитальный труд по теоретическим основам кладистики (Павлинов, 1990б). Обстоятельный анализ вопросов

методологии систематики выполнил Ю.А. Песенко (1989, 1991). Крайне полезными в методическом отношении были статьи из журнала "Taxon" (Jensen, 1981; Duncan, 1984, 1986; Churchill, Wiley, Hauser, 1984; Funk, 1986; Meacham, 1981; Smith, Phipps, 1984).

Установлено, что рассмотренные математические алгоритмы значительно упрощают представление о реальных биологических процессах. "Обычно у биологов сейчас преобладает стремление к математическим, т. е. абстрактным моделям, как имеющим более строгую форму. Но изоморфизм математической модели изучаемому явлению не всегда очевиден; нередко он весьма проблематичен - именно из-за абстрактного характера этой модели" (Скворцов, 1986). В связи с этим нами была поставлена задача разработать менее строгую, но более содержательную процедуру анализа, учитывающую основные особенности процесса видообразования.

Были внимательно изучены принципы работы компьютерных программ PAUP (Swofford, 1985), Hennig86 (Farris, 1988) и PHYLIP (Felsenstein, 1993). Эта работа выполнена при поддержке Международного научного фонда Дж. Сороса (МНФ). При разработке нового метода учтено моделирование филогении в родах *Anacyclus* (Churchill, Wiley, Hauser, 1984) и *Exacum* (Klackenberg, 1987).

Была опробована стандартная процедура анализа данных по методу MES (сравнение сначала самых примитивных таксонов, а затем - все более продвинутых) и обратная (начиная с наиболее продвинутых, а затем - все более и более примитивных), как это принято в некоторых кладистических процедурах (например, метод WISS: цит. по Павлинов, 1989). Установлено, что построение схемы сверху вниз не дает возможности отличить параллелизмы от других событий. Для достоверного различения уникальных и параллельных событий необходимо строить схему снизу вверх, начиная с наиболее примитивных таксонов в исследуемой группе. Только в этом случае можно распознать сходство по параллельно возникающим признакам. В результате было решено в рамках разрабатываемого метода проводить построение схемы только снизу вверх (Байков, 1996).

Новый метод получил название SYNAP по начальным буквам термина "synapomorphy". На его основе А.А. Зверев разработал первую версию компьютерной программы SYNAP017. После ее тестирования к концу 1994 г. были подготовлены версии SYNAP066 и SYNAP070, которые позволяли автоматически проводить оптимизацию исходных данных, создавать полный и краткий протоколы анализа.

Метод SYNAP в основном соответствует этапам деятельности систематиков-эволюционистов, но предъявляет более строгие требования к подбору и анализу исходных данных. Его вариант, изложенный в кладистической терминологии, был опубликован в "Журнале общей биологии" (Байков, 1996а). В такой форме он оказался трудным для восприятия многими отечественными систематиками, поэтому впоследствии было решено изложить его содержание в терминах традиционной филогенетики.

С учетом изложенных выше недостатков в вычислительную процедуру метода MES требовалось внести необходимые изменения и дополнения. Первым шагом в совершенствовании процедуры выбора лучшей медианы стала маркировка реверсий. Однако, одной маркировки оказалось недостаточно: нужно было снизить вероятность выбора таких медиан. Поэтому был введен новый показатель, одновременно учитывающий как индекс продвинутости медиан, так и число реверсий. Для этого число реверсий вычиталось из значения индекса продвинутости. Процедура вычитания была выбрана с тем, чтобы снизить вероятность выбора такой медианы. Новый показатель для выбора медианы сначала был назван индексом потенциальных синапоморфий (Байков, 1996а), а впоследствии - индексом филогенетической связи (Байков, 1997б). Ли выводит индекс продвинутости у медианы как показатель сходства по продвинутым признакам, которое он, вслед за Хеннигом, рассматривает как сходство по родству. Его расчет признаков медианы (Li, 1990) основан на совпадении кодов либо выборе наименьшего из двух кодов в случае их несовпадения.

Следовательно, в методе MES появление продвинутого признака не связывалось с конкретным филогенетическим событием, а выявление последовательности филогенетических событий не рассматривалось как основная задача исследования.

Согласно определению Хеннига, принцип синапоморфии допускает исключение из анализа уникальных продвинутых признаков. Действительно, такие признаки не могут указать на сходство с любым другим таксоном, но показывают отличие их обладателя от остальных. Поэтому при совершенствовании процедуры анализа нами было предложено временное исключение уникальных продвинутых признаков для сокращения расчетов. Эта операция впоследствии была названа оптимизацией данных.

Исключение уникальных продвинутых признаков вынуждает нас отказаться от такого показателя как индекс продвинутости таксонов. Построение филогенетической схемы необходимо производить снизу вверх, начиная с наиболее примитивных членов исследуемой группы. В этом случае необходимо провести сортировку таксонов с учетом нового показателя, который равен индексу продвинутости за вычетом уникальных продвинутых признаков. Изначально он был назван индексом потенциальных синапоморфий (Байков, 1996а), впоследствии - индексом потенциальной связи (сокращенно ИПС). Следовательно, порядок сравнения таксонов теперь определяется по возрастанию ИПС (слайд 33). Итак, первые отличия метода SYNAP от метода MES состояли в следующем: исключение уникальных продвинутых признаков, объединение ставших одинаковыми таксонов; замена индекса продвинутости таксонов на ИПС; маркировка реверсий в протоколе; вычитание реверсий при расчете ИПС медианы; сортировка медиан в порядке убывания ИПС на каждом шаге построения; выбор лучшей медианы по максимальному значению ИПС (слайд 33); оформление графической схемы в соответствии со шкалой значений ИПС (слайды 34 и 35).

Для обработки реверсии следует сформулировать логические правила ее выявления. В общем случае реверсия есть возврат продвинутого признака в исходное состояние, поэтому условиями реверсии нужно считать: 1) предварительное возникновение продвинутого признака; 2) последующее его исчезновение. В контексте описанной выше процедуры это будет соединение таксона без продвинутого признака с медианой или связанным таксоном, у которых этот признак уже возник (слайд 25).

Выполнение первого условия предполагает возможность логического различения ситуации, когда для продвинутого признака показан момент его возникновения и когда такой момент еще не выявлен. Для этого в методе SYNAP предложена перекодировка продвинутых признаков, которая позволяет по измененному коду судить о том, что данный продвинутый признак уже возник. Измененный код также будет показывать, что таксон уже присоединен к схеме. В результате перекодировки продвинутый признак, который уже возник, получает новый логический статус, по которому можно будет отличать его от других продвинутых признаков (слайды 23 и 34). В процедуре MES сравниваются только два таксона - несвязанный и верхний (связанный), а предок последнего не учитывается. По методу SYNAP также сравниваются пары (таксон с таксоном, таксон с медианой, иногда медиана с медианой), но перекодировка дает возможность отразить в признаках связанных таксонов и медиан признаки их ближайших предков (слайды 23 и 24).

В методе SYNAP впервые разделены момент возникновения продвинутого признака (код *N*) и его передача в ряду потомков (код *S*). Код *N* может встретиться только у медианы (слайды 23 и 24). Если несвязанный таксон без продвинутого признака сравнивается с такой медианой, то код *N* будет указывать на то, что у ближайшей медианы (по цепочке в сторону основания схемы) этого признака еще нет, поэтому соединение таксона с этой медианой возможно без реверсии (слайд 27). Когда таксон без продвинутого признака связывается с медианой, имеющей код *S* по данному признаку, то необходимо вводить реверсию на

соответствующий фрагмент схемы (слайд 25). Появление и развитие логической процедуры SYNAP следует рассматривать как стремление соединить основы эволюционной теории Дарвина с процедурой анализа синапоморфий по Хеннигу. При этом новый метод много ближе к эволюционной теории, поскольку он разработан как логическое обоснование традиционной практики филогенетических построений. Кладистическая форма его изложения, в которой он был впервые обнародован (Байков, 1996а), скрыла некладистический характер самой процедуры. В связи с этим возникла потребность описать процедуру филогенетического анализа по методу SYNAP в терминах и понятиях традиционной филогенетики.

Новые термины и понятия. Синапоморфия определяется через апоморфию, поэтому прежде всего следовало найти адекватную замену понятия апоморфии. Учитывая относительный характер введенных Хеннигом представлений об апоморфии и плезиоморфии, следовало найти понятия, характеризующие прежде всего порядок возникновения признаков в филогенезе. Некоторое время пришлось пользоваться не совсем удачным переводом терминов Хеннига на русский язык: относительно продвинутое (вторичное, производное) состояние признака - для апоморфии, относительно примитивное (первичное, исходное) - для плезиоморфии. Такое именование плезиоморфий и апоморфий нередко приводило к путанице: апоморфия в одном случае именовалась продвинутым состоянием признака, в другом - вторичным, в третьем - производным.

<В качестве эквивалента апоморфии было введено понятие о новом признаке (Байков, 1997б). Термин "новый признак" позволяет подчеркнуть его возникновение в пределах исследуемой группы. Новый признак - это любой признак, возникший в пределах исследуемой группы на основе исходного для него признака (слайд 10).

<На основе понятия "новый признак" можно сформулировать новое определение синапоморфии в узком смысле, более строго, чем у Хеннига и Фэрриса. Синапоморфия - это сходство таксонов по новому признаку, возникшему в результате одного и того же события в филогенезе. Такое определение синапоморфии предполагает разработку динамической модели филогенеза.

<В основе филогенеза лежат процессы видообразования, поэтому в логической процедуре SYNAP важно было корректно отобразить такие особенности видообразования как дивергенция, адаптивная радиация, параллельная эволюция, реверсия и, что особенно важно, гибридизация. Поскольку моделируется процесс, то модель должна быть динамической, контекстно зависимой и эволюционно осмысленной. Суть ее состоит в том, что начальные условия анализа определяют только общие черты итоговой схемы, тогда как конкретные особенности ее определяются в динамике, как контекстно зависимые от текущего этапа построения. В этом заключается ее отличие от динамической модели, кратко описанной Павлиновым при обсуждении первого пункта "программы Хеннига" (Павлинов, 1990б: 16), и положенной в основу вычислительных алгоритмов, объединяемых под названием "кладистические".

<Методы, основанные на расчете фенетических дистанций между таксонами и построении "сетей Вагнера" с последующим их "укоренением" через внешнюю группу (Farris, 1970; Jensen, 1981), представляются мне некорректными для отражения филогенеза. Динамическая модель отсутствует и при анализе совместимости признаков (Estabrook, 1972; LeQuesne, 1982). Структурная кладистика, устанавливающая иерархию признаков на основе их появления в онтогенезе (Nelson, 1979, Platnick, 1979), также не реализует динамическую модель, оставляя без должного внимания вопросы появления новых признаков в контексте текущей признаковой среды.

<Понятие признаковой среды обсуждается в литературе нечасто. По своему содержанию оно приближается к представлениям о филетических корреляциях или координациях (Северцов, 1939; Шмальгаузен, 1939). В методе SYNAP оно использовано для оценки возможности возникновения нового признака в текущих условиях признаковой среды как приспособления к определенным условиям среды обитания. Другими словами, в контексте признаковой среды мы можем задать биологически осмысленную вероятность возникновения нового признака, то есть разрабатывать вероятностную модель филогенеза, не прибегая к чисто количественным оценкам (Павлинов, 1989). Сама признаковая среда постоянно изменяется и развивается в пределах существующих в ней разрешений и ограничений (Северцов, 1987). Ее развитие можно проследить по последовательности возникновения новых признаков, выделяя тем самым основные направления этих изменений (слайд 11).

Чтобы следить за изменениями признаковой среды в процессе моделирования филогенеза, нам представляется важным оставаться в "мире реальных признаков и объектов", не уходя в "мир фенетических дистанций" или в "мир совместимости признаков". Именно поэтому при разработке метода SYNAP пришлось отказаться от многих процедур кладистического анализа и обратить внимание на малоизвестные методы.

Возвращаясь к расчету признаков медианы в методе MES (Li, 1990), можно предположить, что, выбирая наименьший из двух признаков в случае их несовпадения, Ли хотел подчеркнуть цельность состояний признака, их качественный характер. Такая процедура дает возможность рассматривать медианы как таксоны, только в большинстве своем вымершие. Если медиана обладает свойствами реальных (современных) видов, то она должна иметь или исходный признак, или продвинутый признак (в методе MES кодируются только эти варианты), но не среднее арифметическое значение (1/2). Вопрос о переходных (промежуточных) признаках Ли не разбирает.

Это ограничение метода MES, на первый взгляд вполне разумное, излишне формализует модель филогенеза. Оно оставляет в ней дискретные качества, но не отражает непрерывности процесса, поскольку исключает переходные варианты. Такой подход вызван следующими причинами: для определения направлений филогенеза анализируемые признаки должны быть представлены более чем одним состоянием, эти состояния должны быть отличными и не пересекаться у одного таксона. В методе SYNAP это ограничение частично снимается, поскольку введена обработка полиморфных признаков (одновременное присутствие исходного и продвинутого признаков), что особенно важно, когда объектами анализа выступают группы сложного состава (например, роды или полиморфные видовые комплексы). Разрабатывается также логическая процедура обработки варианта "промежуточный признак" (уже не исходный, но еще не новый).

Концепция элементарного эволюционного вектора. Связывание признаков только в пары, а не в более крупные группы, иногда вызывает критику. Нам представляется, что анализ пар имеет более высокую разрешающую способность. Он позволяет рассматривать переход от исходного признака к новому как отдельное филогенетическое событие, относительно независимое от других событий. Порядок возникновения новых признаков будет определяться их комбинациями у членов исследуемой группы.

Разбиение сложных комплексов признаков на пары оправдано и с точки зрения более точного установления гомологичности признаков. Как было показано Т.В. Кузнецовой, наиболее универсальным критерием гомологии у растений является не критерий положения или критерий одинакового строения, а критерий переходных форм. "Сравнительно-морфологический ряд строится по линии минимального изменения признаков, связанных с первыми двумя критериями, и позволяет уловить истинное морфологическое соответствие структур" (Кузнецова, 1986: 31). В связи с этим в методе SYNAP предложено понятие элементарного эволюционного вектора, который выражается в филогенезе как появление

нового признака на основе исходного для него в масштабе "минимальных изменений морфологического ряда" (слайд 12).

Элементарный эволюционный вектор имеет начало (исходный признак) и конец (новый признак). Если их соединение представляет собой общую эволюционную цепочку, то конец первого такого вектора будет одновременно началом следующего, и т. д. (слайд 13). Любые эволюционные преобразования морфо-функциональных систем можно разложить на последовательности таких векторов. Теперь для построения филогенетической схемы нам понадобится определить такую последовательность этих переходов, которая будет наиболее точно соответствовать распределению новых признаков в пределах исследуемой группы. Она может быть линейной (слайд 14), дивергентной (слайд 15), конвергентной (слайд 16), представлять собой псевдоцикл (слайд 17).

Процедура анализа. Определение филогенетических связей по методу SYNAP начинается с подробного сравнительно-морфологического изучения исследуемой группы. "Никакой, даже самый изощренный формальный метод не может дать интересного и достоверного результата, если в анализ вводятся плохо разработанные морфологические данные (Любарский, 1997: 29). Для последующего анализа выделяются признаки таксонов, с учетом существующих гипотез определяются наиболее вероятные направления их изменений. На этой основе выделяются новые признаки, которые, по мнению исследователя, возникли в пределах данной группы. Некоторые признаки могут быть связаны в длинные эволюционные цепочки, поэтому для выявления каждого конкретного перехода, который рассматривается как отдельное филогенетическое событие, они разбиваются на пары - элементарные эволюционные векторы. При расчете филогенетической связи каждый такой вектор будет учитываться отдельно. Направление векторов в методе SYNAP согласуется с традиционной практикой филогенетических построений и основано на данных по сравнительной морфологии, особенностям онтогенеза, палеонтологических материалах, а также на оценке адаптивного значения нового признака. Каждый вектор рассматривается как рабочая гипотеза, требующая проверки и согласования с другими рабочими гипотезами.

Уже стало традицией исходный признак кодировать нулем, а производный - единицей. Нам представляется более подходящим кодировать их буквами. Во-первых, здесь мы имеем дело с логическими вариантами, а не цифрами, и их сравнение друг с другом основано на логических операциях, но не на чисто математических процедурах. Во-вторых, это дает возможность реализовать в методе SYNAP процедуру априорного взвешивания векторов в баллах. Поэтому в качестве стандартного сокращения далее предлагается использовать обозначения: P (от англ. "primitive character") - для исходного признака и A (от англ. "advanced character") - для нового признака (слайд 22).

Оптимизация данных. Для сокращения общего числа расчетов в методе SYNAP предложена новая процедура - оптимизация исходных данных. Она состоит во временном удалении той информации, которая не влияет на расчет связи по новым признакам.

После их удаления некоторые таксоны становятся одинаковыми по оставшимся признакам, поэтому их можно объединить, что позволит дополнительно сократить объем расчетов. Так, оптимизация данных по гипотетической матрице Dendrogrammaceae позволила снизить объем расчетов на 60% (Байков, 1996а).

Сортировка объектов. Для проведения анализа необходимо определить общий порядок сравнения таксонов. Поскольку ранее была показана необходимость построения схемы снизу вверх, то логично выбрать для этого индекс потенциальной связи (ИПС), который равен сумме новых признаков после исключения уникальных векторов. В случае равного индекса

таксоны располагаются в алфавитном порядке их названий. Расположив таксоны в порядке возрастания ИПС, мы получим последовательность, примерно соответствующую расхождению таксонов по ветвям филогенетической схемы. Сортировка по возрастанию ИПС (а не индекса продвинутости) точнее соответствует порядку расхождения по ветвям схемы, поскольку учитывает только новые признаки, возникшие совместно (слайд 33).

Если в исследуемой группе встречаются другие логические варианты признаков (например, полиморфный признак, неизвестный признак, отсутствие вектора), то сортировка таксонов по методу SYNAP проводится с учетом этих параметров (слайд 18).

Расчет медианы для первой пары таксонов. Первыми сравниваются два таксона с минимальным значением ИПС. Целью такого сравнения будет обоснование возможности появления у них нового признака в результате одного и того же события в филогенезе. Возможные варианты такого сравнения показаны на (слайды 19-23).

В общем случае, если таксоны X и Y обладают исходным признаком или только один из них обладает новым признаком, то их медиана имеет исходный признак (слайд 20). Если оба они имеют новый признак, можно считать, что их медиана также имеет этот новый признак (слайд 23). Момент возникновения нового признака фиксируется кодом N , а его передача в ряду потомков обозначается кодом S . Так последовательно проводится сравнение таксонов X и Y по каждому элементарному эволюционному вектору. Затем рассчитывается индекс связи медианы. Его значение зависит от числа (в случае априорного взвешивания векторов также и "веса") новых признаков, а также числа реверсий. Если значение индекса связи меньше или равно нулю, то такая медиана не нумеруется. Если ИПС имеет положительное значение, то медиане присваивается номер (например, $m01$). На втором шаге построения будет уже три медианы (слайд 34). Из них выбирается лучшая (с максимальным значением ИПС) и получает следующий номер (например, $m02$).

Как было отмечено выше, новый признак, по которому таксон был включен в схему филогенетических связей, следует логически отличить от остальных. Для этого в методе SYNAP предложена перекодировка: у соответствующих таксонов код A (имеется новый признак) заменяется на код S (новый признак возник в результате установленного филогенетического события). Тем самым задается логическое различие между такими признаками (слайд 23).

По сравнению с другими процедурами филогенетического анализа, в методе SYNAP разработан самый полный набор вариантов для кодирования признаков. Различие между двумя признаками, определяющими элементарный эволюционный вектор, может быть хорошо выраженным. В этом случае имеются только два логических варианта: исходный признак (код P) и новый признак (код A). Однако в реальной ситуации нередко возникают и другие варианты. Например, требуется обработать вариант полиморфизма, когда одновременно имеется и новый признак, и исходный для него. Этот логический вариант отмечается кодом B (от англ. "both" - оба). Вариант "отсутствие элементарного эволюционного вектора" (нет ни исходного, ни нового признака) обозначается латинской буквой M (от англ. "missing" - отсутствующий). Наличие такого варианта дает возможность сравнивать между собой группы таксонов с разными наборами векторов: те векторы, которые имеются в одной группе, но отсутствуют в другой, следует отмечать кодом M . В практической работе часто приходится сталкиваться с ситуацией, когда не для всех векторов можно указать конкретный признак. Без привлечения дополнительных данных невозможно установить, какой из признаков данного вектора характерен для таксона - новый или исходный для него (например, требуется дополнительное изучение материала, который в данный момент недоступен). Для этого случая введен логический вариант "неизвестный признак", который кодируется латинской буквой U (от англ. "unknown" - неизвестный). В результате была разработана новая логическая схема построения признаков медианы с

учетом дополнительных вариантов кодирования признаков (слайд 18).

Для выбора лучшей медианы из нескольких, имеющих одинаково высокий индекс связи, в методе SYNAP разработана специальная процедура, с помощью которой можно задать ту или иную эволюционную модель. В случае выполнения анализа вручную такой выбор можно провести субъективно, с учетом всей доступной информации. В этом заключается основное преимущество ручной обработки: можно контекстно изменять "вес" новых признаков, вводить поправки на морфо-функциональные корреляции, учитывать климатические и географические параметры формирования видов и т. д.

Для компьютерной обработки необходимы формальные критерии выбора, связанные только с особенностями исходных данных и процедуры анализа. Когда несколько медиан имеют одинаково высокий индекс связи, в методе MES выбирается та, которая расположена ближе к основанию схемы. В методе SYNAP дополнительно предлагается правило политомии, или слияния одинаковых медиан, которое позволяет исключить из построения медианы-дубликаты и тем самым сократить итоговую схему. При этом сохранена возможность построения схемы и по строго дихотомическому принципу, что бывает необходимо в случае сравнения результатов с другими методами, основанными на строгой дихотомии.

Другими критериями выбора лучшей медианы среди равных по значению ИПС могут быть: сумма новых признаков, возникших на данном шаге построения; сумма ранее возникших новых признаков; сумма реверсий. Если в исходных данных есть другие логические варианты, то критериями выбора медиан могут быть также: сумма полиморфных признаков, сумма неизвестных признаков, сумма отсутствующих векторов. Помимо этого предлагаются еще два критерия - возникновение парного нового признака, характерного только для двух таксонов в исследуемой группе и соединение с "висящим" (ранее введенным, но несвязанным) таксоном.

В ручной процедуре есть возможность гибко применять эти правила с учетом *конкретной ситуации*), т. е. в контексте признаковой среды. В компьютерной программе эти правила работают без учета конкретной ситуации, поэтому построение эволюционной схемы будет более формальным. Однако можно получить для каждого набора критериев индивидуальные протоколы, сравнить их между собой и сделать заключение по подбору критериев (слайд 36).

При обработке данных вручную применение того или иного критерия и его приоритет можно определять субъективно. В компьютерной программе требуется до начала процедуры указать не только набор, но и приоритет критериев, т. е. указать, в каком порядке они будут применяться. Некоторые из них имеют противоположное действие. Поскольку выбор делается среди медиан с одинаково высоким ИПС, который определяется как сумма кодов S и N за вычетом реверсий, то увеличение доли кодов S обязательно сопровождается снижением доли кодов N .)

Для определения относительного положения таксонов на схеме введена шкала филогенетической связи, которая одновременно будет и шкалой продвинутости таксонов. Проекция таксонов и медиан на эту шкалу не всегда будут точно совпадать со значением их ИПС, что обусловлено особенностями графического отображения связей и маркировкой признаков.

Форма графического изображения филогенетической схемы имеет важное значение для восприятия результатов анализа. Как справедливо отметила А.В. Положий, "сложнейшей задачей построения филогенетической схемы является разработка метода графического изображения ее. ... Задача состоит в том, как показать в системе, что речь идет о генетических связях не между ныне существующими таксонами, а их наиболее близкими предками" (Положий, 1978: 79). Цепочки связанных медиан наиболее точно показывают

такие связи (слайды 41, 43).

С помощью программы SYNAP выполнен анализ основных направлений в эволюции плодов молочаев из подрода *Esula* Pers., дана оценка филогенетических связей в группе сибирских молочаев (Байков, 1997а) и однолетних молочаев из Средней Азии, выполнено исследование филогенетических связей сибирских и среднеазиатских молочаев (около 90 видов) и др.

Секция *Esula* рода *Euphorbia* представлена в Северной Азии тремя подсекциями. Наиболее близки к исходному типу два вида из подсекции *Sieboldianae*, а также *E. potaninii* из монотипной подсекции *Potaniniae*. Более эволюционно продвинутые виды принадлежат подсекции *Esulae*. Анализ филогенетических связей в этой подсекции выполнен с использованием 26 вектор-признаков (слайды 30, 31, 32). Набор признаков отражает основные эволюционные преобразования в этой секции. Характеристика видов по указанным векторам показана на слайде 33, где виды упорядочены по возрастанию индекса потенциальной филогенетической связи (точкой обозначен исходный признак, единицей - новый).

С учетом активного видообразования в этой секции выбрана эволюционная модель, в которой приоритет отдан возникновению новых признаков в сочетании с политомией. Полученная схема показана на слайдах 41 и 43. Рассмотрим этапы построения филогенетической схемы, отраженные в пошаговом протоколе анализа (слайды 35, 36, 38-40, 42). На первом и втором шагах построения филогенетической схемы сравниваются два наиболее примитивных по заданному набору вектор-признаков вида - *E. lucida* (ИПС=4) и *E. mandshurica* (ИПС=7). Их ближайший общий гипотетический предок (медиана) *m01* характеризуется четырьмя новыми признаками **1**, **2** (нектарники двурогие), **3** (длина наружных брактеев меньше их ширины) и **5** (листья плотные, до 15 мм шир.) (слайд 34). Далее с ними сравнивается *E. virgata*. Наилучший показатель связи имеет медиана *m02*, для которой характерны, помимо ранее возникших, новые признаки **4** (рожки отходят от края нектарников) и **6** (столбики 2-3 мм дл.).

Другие медианы, соединяющие *E. virgata* с *E. lucida* и *m01*, не отличаются новыми признаками от *m01*.)

На четвертом шаге *E. latifolia* (ИПС=8) посредством медианы *m03* соединяется с *E. mandshurica*. Эта медиана характеризуется новым признаком **7** (рожки длинные, плотные), но предполагает реверсию по вектору **5**, поскольку листья *E. latifolia* шире 15 мм. Эта реверсия возможна в контексте признаковой среды, так как виды этого родства характеризуются относительно широкими листьями, а экологические условия обитания *E. latifolia* (приречные галечники) способствовали сохранению этого признака.

Далее, *E. borealis* (ИПС=8) соединяется с *E. latifolia*, при этом задействованы все новые признаки *E. borealis*. Медиана *m04* характеризуется двумя новыми признаками **8** (рожки нектарников короткие, плотные) и **10** (плоды шероховатые), при этом отмечена возможность закрепления реверсии по признаку **5** у *E. latifolia*, если *E. borealis* будет присоединен непосредственно к *E. latifolia*. Высокий индекс имеют также варианты соединения *E. borealis* с *E. mandshurica* и медианой *m03* (приложение 4.1).

На следующем этапе *E. uralensis* (ИПС=9) обнаруживает максимальное сходство по новым признакам с *E. virgata* (слайд 35). Медиана *m05* характеризуется новым признаком **9** (наружные брактеев шире срединных листьев). *E. dahurica* (ИПС=10) через медиану *m06* соединяется с ближайшим общим предком для *E. borealis* и *E. latifolia*, с которыми имеет значительное сходство. Эта медиана характеризуется новым признаком **8** (рожки нектарников короткие, плотные), то есть момент возникновения этого признака

перемещается на более ранний этап (ранее он был выявлен для медианы *m04*, производной от *m06*). Высокий индекс связи имеет вариант присоединения к медиане *m05*, но признак 7 здесь отсутствует.

Для *E. chankoana* (ИПС=11) выявлены два равноценных варианта связывания. Первый характеризуется новыми признаками **9** и **14** (лучи соцветия простые или однажды двураздельные), но предполагает реверсию по ранее возникшему признаку **8**. Признак **9** возникает в этой линии параллельно (ранее он был выявлен для пары *E. virgata* и *E. uralensis*)).

Второй вариант предполагает соединение *E. chankoana* с *E. uralensis* по новому признаку 21, но лишен ранее возникшего признака 7, который есть в первом случае. Предпочтение отдано первому, поскольку выбранная эволюционная модель имеет приоритет вновь возникших признаков (здесь их два).

E. rossica (ИПС=13), в результате сравнения со всеми ранее введенными видами и выбранными медианами, соединен с *E. chankoana* (слайд 37). Медиана *m08* характеризуется возникновением двух новых признаков: **11** (растения обычно не превышают 40 см выс.) и **21** (столбики 1-2 мм дл.) и предполагает реверсию по вектору **8**. Высокий показатель филогенетической связи *E. rossica* имеет с *E. uralensis* и *E. dahurica*).

Далее, *E. caesia* соединяется с *E. rossica* (слайд 38). Новая медиана *m09* характеризуется возникновением новых признаков **13** (верхушечный зонтик 4-8-лучевой, с малым числом боковых лучей) и **15** (листья узкие), но предполагает реверсию по признаку **3**, поскольку его брактеи длиннее своей ширины (примитивный признак, не характерный для большинства видов из секции *Esula*)).

Далее, *E. kirimzjulica* (ИПС=13) присоединяется к медиане *m09* - ближайшему общему предку *E. rossica* и *E. caesia*, с которыми он имеет максимальный показатель филогенетической связи. В данном случае правило политомии, или слияния одинаковых медиан, обусловило трихотомию в этом фрагменте схемы. К этой же группе видов присоединяется *E. esula*), однако отсутствие у него нового признака **14** требует соединения не с медианой *m09*, а введения нового предка *m10*, соединяющего его с *E. kirimzjulica* (слайд 39). Медиана *m10* встраивается на схеме под медианой *m09*, т. е. представляет ее предка (отличается отсутствием признака 14).

E. tshuiensis (ИПС=13) также встраивается в схему ниже медианы *m09*. Новая медиана *m11* указывает на возникновение нового признака 15 ранее, чем у медианы *m09* (слайд 40). Согласно полученной схеме, этот вид одновременно близок к *E. rossica*, *E. caesia*, *E. chankoana* и *E. esula*.)

E. subcordata (ИПС=14) по правилу политомии присоединяется к ранее установленной медиане *m09* и обнаруживает значительное родство с *E. rossica* и *E. caesia*. Очень близкую позицию к нему получает следующий вид - *E. lunulata*.)

К ним примыкает *E. chankoana*, который следует сблизать с видами, расположенными ближе к основанию филогенетической схемы. Выше них на схеме расположены узколистные виды (признак **15**). К паре близких видов *E. cyparissias* и *E. leoncroizatii* примыкает *E. tshuiensis*. Далее, *E. lunulata* и *E. subcordata* близки к *E. rossica* и *E. caesia*, вместе образуют естественную группу видов степных и полупустынных местообитаний (ряд *Subcordatae*). Группа видов лугово-лесных и приречно-галечниковых местообитаний занимает промежуточное положение между предыдущими видами и узколистными представителями ряда *Leptocaula* (*E. microcarpa* и *E. karoi*): это - *E. esula*, *E. lenensis*, *E. borealis* и *E. maackii* (ряд *Esulae* s. str.)

. Ближе к основанию схемы расположены *E. virgata* и *E. uralensis* (ряд *Virgatae*) и наиболее

широколистные виды: *E. latifolia* и *E. boro-dinii* (ряд *Latifolia* e), *E. mandshurica* и *E. lucidae* (ряд *Lucidae*). Полученная схема в целом отражает последовательность филогенетических событий в подсекции *Esula*, а выявленные группы близкородственных видов позволяют уточнить границы видовых рядов и описать новые. Наличие параллелизмов по многим признакам объясняет сложность принятия таксономических решений в этой подсекции, а существование тесных филогенетических связей между видами из разных рядов, подтверждает относительно молодой возраст этой группы. Некоторые экологические и географические основания этой схемы опубликованы (Байков, 2001б).

По механизмам видообразования секция *Esula* значительно отличается от остальных, что, на наш взгляд, связано с ее распространением на равнинных территориях (преимущественно степь и лесостепь), где отсутствуют физико-географические барьеры для изоляции видов. В этих условиях в качестве изолирующего фактора выступает лишь значительная удаленность и различия в микрорельефе (для симпатрических видов), вызывающие смещение сроков цветения и, соответственно, относительную репродуктивную изоляцию. Многие современные виды из секции *Esula* представляют собой недавние образования, генетически слабо обособленные друг от друга, что при определенных условиях вызывает их гибридизацию. Этот процесс сначала приводит к возникновению нового гибридного образования в зоне контакта родительских видов. Постепенно в этот процесс могут быть вовлечены соседние участки ареалов, что в конце концов приведет к формированию нового вида с расширенным диапазоном изменчивости многих признаков. Такие варианты встречаются и в Восточной Европе, и в Северной Азии (например, комплексы *E. subcordata*, *E. caesia*, *E. rossica* и *E. esula*, *E. borealis*, *E. lenensis*, *E. maackii*).

В заключение хочется отметить, что не следует подходить формально к моделированию филогенетических отношений в группах, имеющих большой объем и значительный возраст. Особенно осторожно следует отражать полученные результаты в таксономии рода, учитывая прежде всего содержание понятий подрода, секции, подсекции и ряда в данной группе. Данные филогенетического анализа, если он проведен с максимально полным учетом морфологических признаков и эколого-географического обоснования эволюционных векторов, в большинстве случаев соответствует хорошо проработанной системе рода, созданной традиционным способом. Его преимущество заключается в том, что имеется возможность различить строгие и нестрогие гомологии и дать более точную оценку диагностического значения признаков в разных филогенетических линиях.