

Молекулярная Зоология, весна 2015
лекция 4

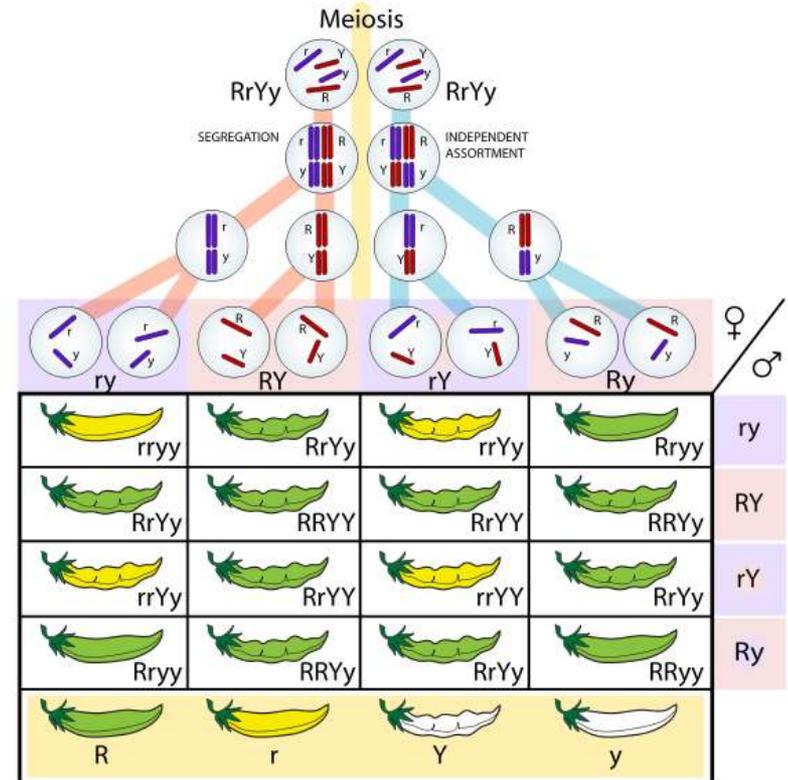
Введение в методы исследования ДНК, Популяционная генетика. Геногеография (филогеография) .

Вспоминаем школу. Законы Менделя.



Грегор Иоганн МЕНДЕЛЬ
Gregor Johann Mendel, 1822–84

- ▶ **1-й закон: Закон единообразия первого поколения гибридов.**
- ▶ **2- закон. Закон расщепления признаков (3:1) (закон чистоты гамет)**
- ▶ **3-й закон Закон независимого наследования признаков**



Условия выполнения законов Менделя

Условия выполнения закона расщепления при моногибридном скрещивании

- ▶ Расщепление 3 : 1 по фенотипу и 1 : 2 : 1 по генотипу выполняется приближенно и лишь при следующих условиях:
- ▶ Изучается большое число скрещиваний (большое число потомков).
- ▶ Гаметы, содержащие аллели А и а, образуются в равном числе (обладают равной жизнеспособностью).
- ▶ Нет избирательного оплодотворения: гаметы, содержащие любой аллель, сливаются друг с другом с равной вероятностью.
- ▶ Зиготы (зародыши) с разными генотипами одинаково жизнеспособны.

Условия выполнения закона независимого наследования

- ▶ Все условия, необходимые для выполнения закона расщепления.
- ▶ Расположение генов, отвечающих за изучаемые признаки, в разных парах хромосом (несцепленность).

Условия выполнения закона чистоты гамет

- ▶ Нормальный ход мейоза. В результате нерасхождения хромосом в одну гамету могут попасть обе гомологичные хромосомы из пары. В этом случае гамета будет нести по паре аллелей всех генов, которые содержатся в данной паре хромосом.



Харди—Вайнберг



Годфри Харолд ХАРДИ

Godfrey Harold Hardy, 1877–1947

Английский математик. Самую большую известность Харди принесли совместные работы с Джоном Идензором Литлвудом (John Edensor Littlewood, 1885–1977) и позднее с индийским математиком-самоучкой Сриниваса Рамануджаном (Srinivasa Aaiyengar Ramanujan, 1887–1920), который работал клерком в Мадрасе. В 1913 году Рамануджан послал Харди список доказанных им теорем. Признав гениальность юного клерка, Харди пригласил его в Кембридж, и в течение нескольких лет, предшествовавших безвременной смерти Рамануджана, они опубликовали серию блестящих совместных работ.

Вильгельм ВАЙНБЕРГ

Wilhelm Weinberg, 1862–1937

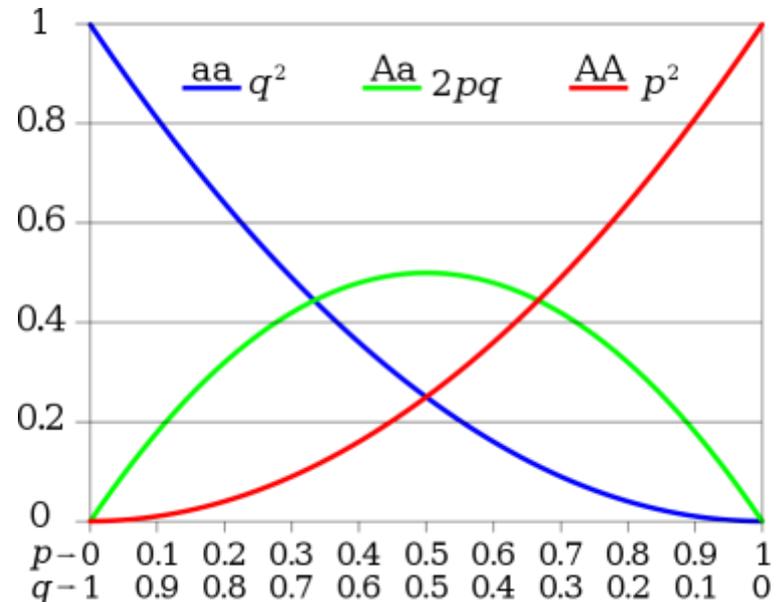
Немецкий врач, имевший большую частную практику в Штуттгарте. По воспоминаниям современников, помог появиться на свет 3500 младенцам, в том числе по крайней мере 120 парам близнецов. На основании собственных наблюдений над рождением близнецов и переоткрытых генетических законов Менделя пришел к выводу, что предрасположенность к рождению двуяйцевых (неидентичных) близнецов передается по наследству

Закон Харди-Вайнберга (1908)

- ▶ **Закон Харди-Вайнберга** — это закон популяционной генетики — в популяции бесконечно большого размера, в которой не действует отбор, не идет мутационный процесс, отсутствует обмен особями с другими популяциями, не происходит дрейф генов, все скрещивания случайны — **частоты генотипов по какому-либо гену** (в случае если в популяции есть два аллеля этого гена) **будут поддерживаться постоянными из поколения в поколение и соответствовать уравнению:**

$$p^2 + 2pq + q^2 = 1$$

Где p^2 — доля гомозигот по одному из аллелей; p — частота этого аллеля; q^2 — доля гомозигот по альтернативному аллелю; q — частота соответствующего аллеля; $2pq$ — доля гетерозигот.

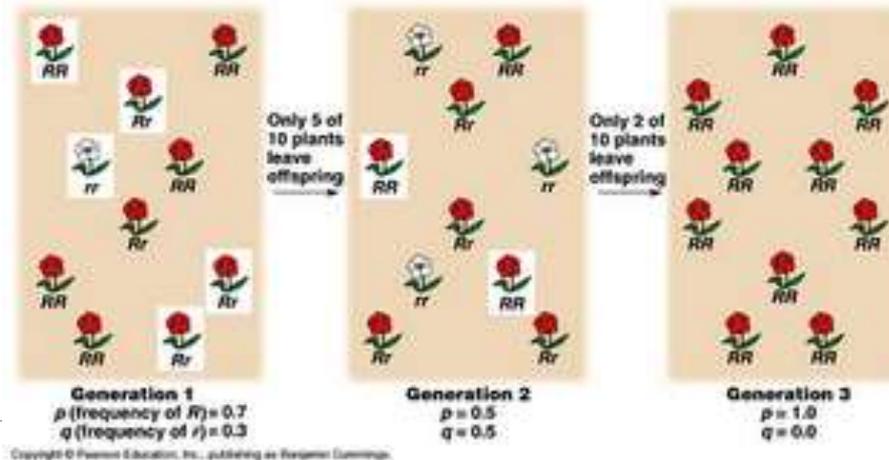


Обычно, нас интересует отклонение от

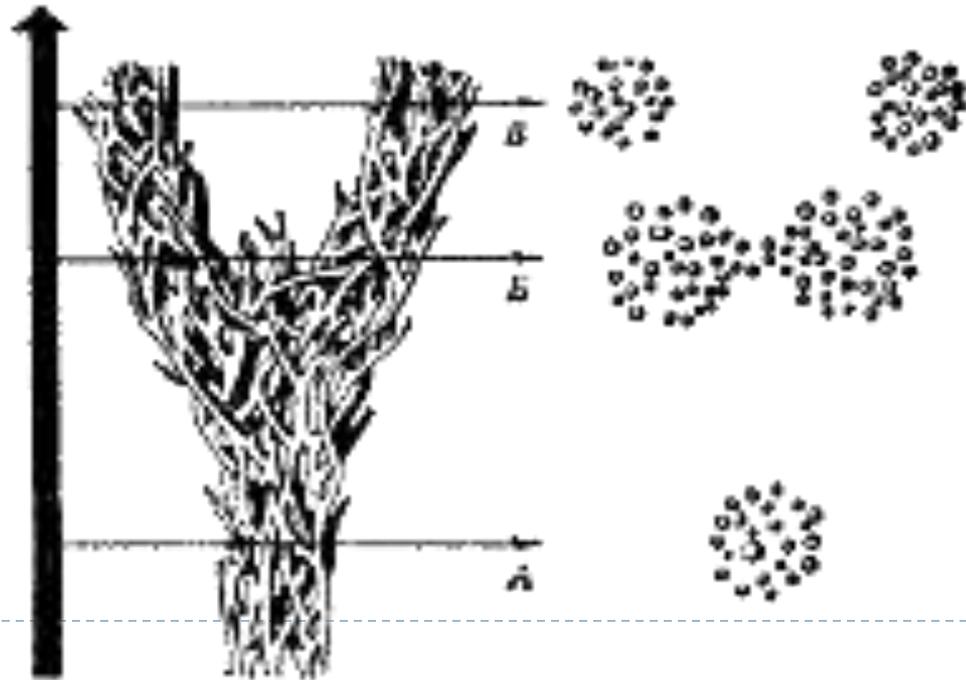
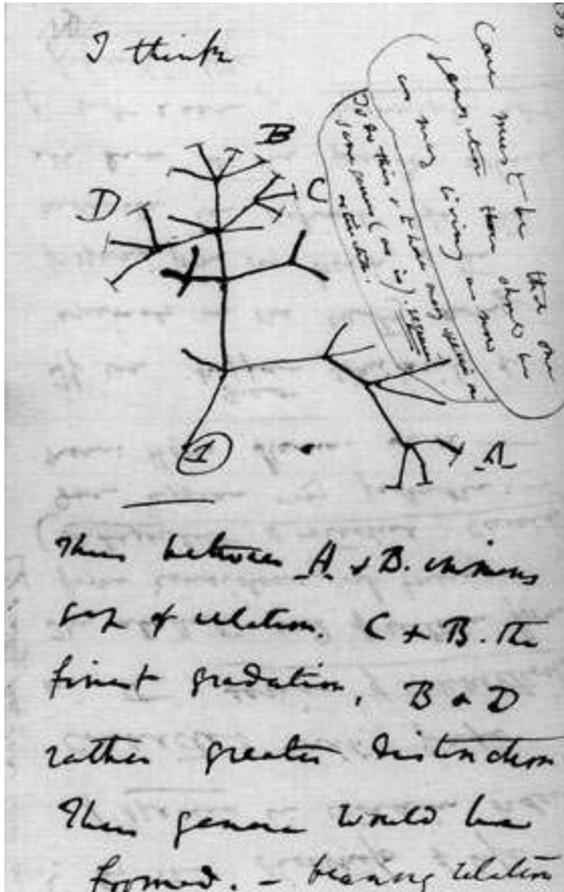
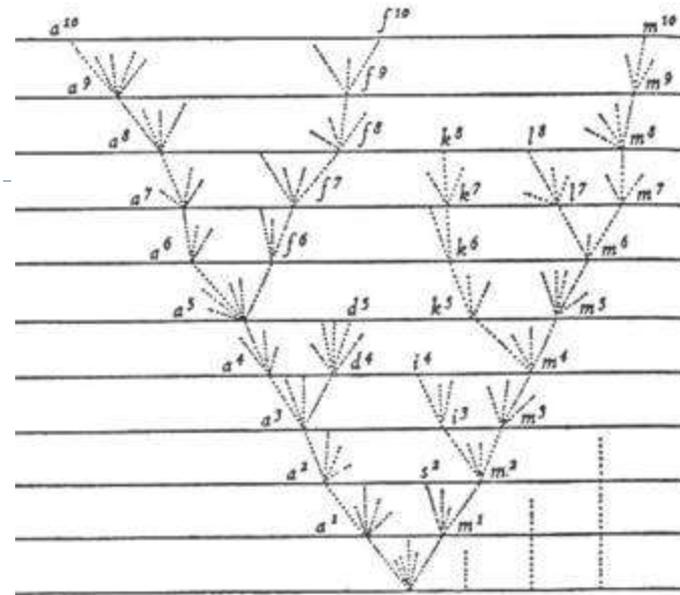
- ▶ Закона Харди-Вайнберга, проверяемое критерием χ^2

Генетический дрейф (дрейф генов, генетико-автоматические процессы)

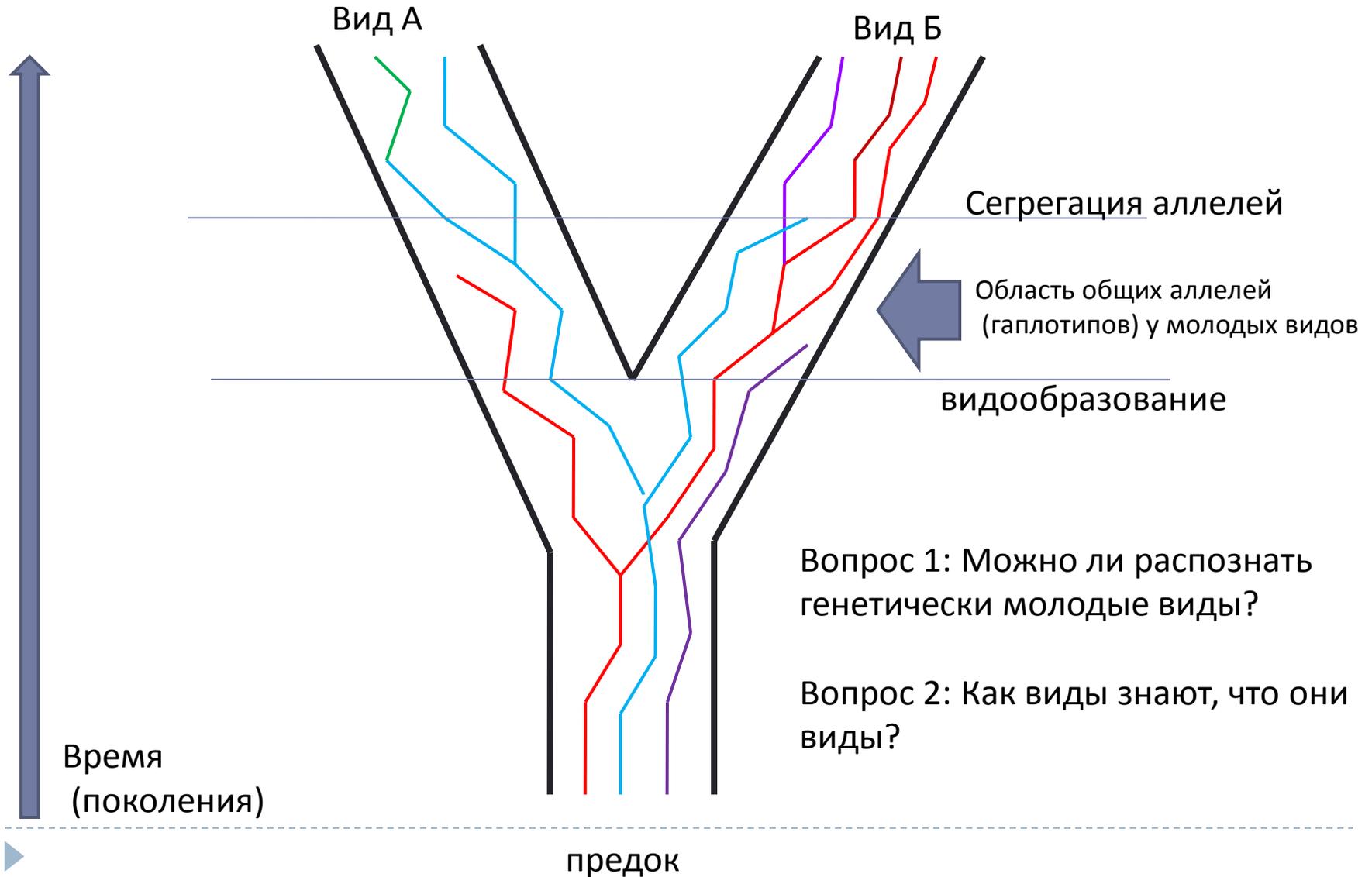
- ▶ Понятие «дрейф генов» (*genetic drift*) было введено в оборот Райтом (1931), а синонимичное понятие «генетико-автоматические процессы в популяциях» — Дубининым и Ромашовым (1932).
- ▶ С. Райт экспериментально доказал, что в маленьких популяциях частота мутантного аллеля меняется быстро и случайным образом. Его опыт был прост: в пробирке с кормом он посадил по две самки и два самца мух дрозофил, [гетерозиготных] по гену А (их генотип можно записать Аа). В этих искусственно созданных популяциях концентрация нормального (А) и мутационного (а) аллелей составила 50%. Через несколько поколений оказалось, что в некоторых популяциях все особи стали гомозиготными по мутантному аллелю (а), в других популяциях он был вовсе утрачен, и, наконец, часть популяций содержала как нормальный, так и мутантный аллель. Важно подчеркнуть, что, несмотря на снижение жизнеспособности мутантных особей и, следовательно, вопреки естественному отбору, в некоторых популяциях мутантный аллель полностью вытеснил нормальный. Это и есть результат случайного процесса - *дрейфа генов*.



Дерева- виды и гены



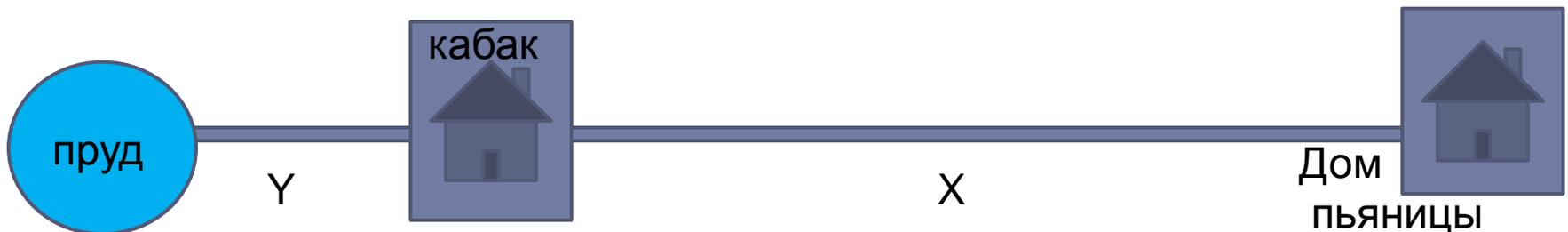
Генеалогия генов и видообразование



Дилемма пьяницы

В деревне Утконовка одна улица. С одной стороны улицы имеется пруд, с другой стороны – дом пьяницы, между ними на расстоянии X метров от дома и Y метров от пруда, расположен знатный кабак. Из кабака выходит мужик (в дугу пьяный) делает шаг в случайном направлении. После каждого шага мужик крепко задумывается, и затем снова делает шаг, опять в случайном направлении, не зависящем от направления предыдущего шага.

Вопрос: с какой вероятностью мужик попадет домой и с какой – свалится в пруд? (Ответ $P_{\text{будет спать дома}} = Y/(X+Y)$)



Биологический смысл: В популяции из N особей нейтральный аллель A , имеющий частоту P через $2N$ поколений зафиксируется с вероятностью P либо исчезнет с вероятностью $1-P$.

Следствия дилеммы пьяницы:

- ▶ Новые мутации редко закрепляются, чаще исчезают из-за дрейфа генов, даже в больших популяциях (табл.).
- ▶ Чем меньше популяция, тем ниже гетерозиготность.

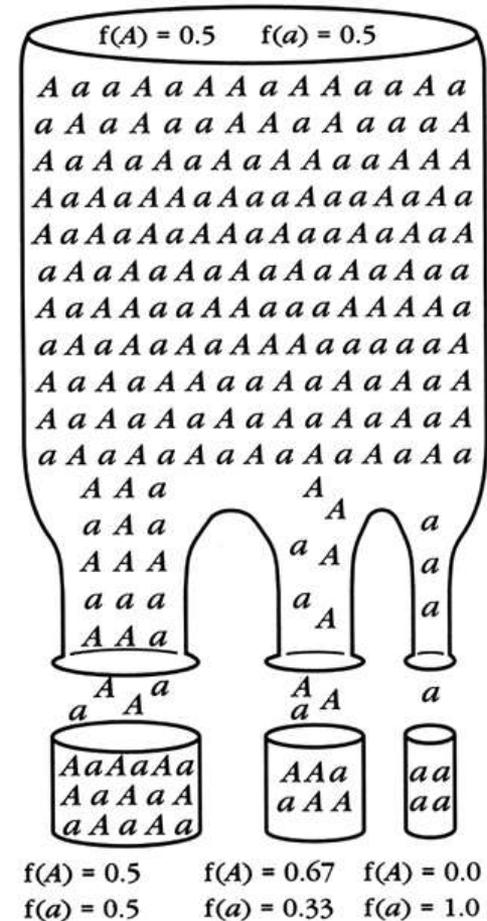
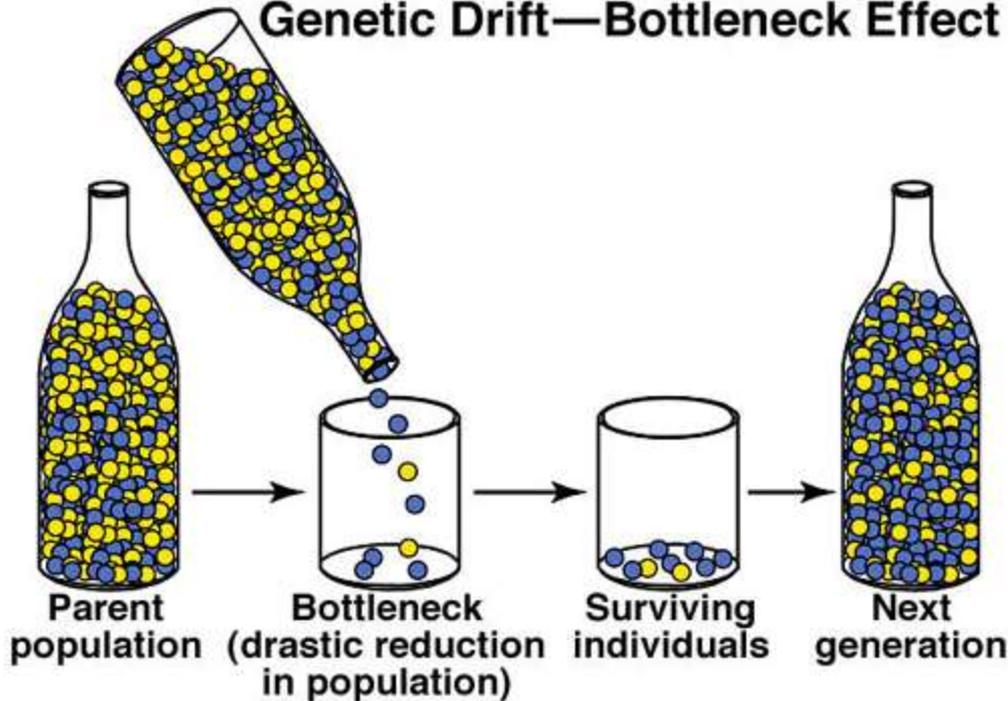
| Вероятность исчезновения мутации, возникшей у одной особи (Fisher, 1930, 1958-) | | |
|---|--------------------------------------|-----------------------------------|
| Число поколений | Вероятность исчезновения при условии | |
| | Нейтральности | 1%-ного селективного преимущества |
| 1 | 0.368 | 0.364 |
| 3 | 0.626 | 0.620 |
| 7 | 0.790 | 0.782 |
| 15 | 0.887 | 0.878 |
| 31 | 0.941 | 0.931 |
| 63 | 0.970 | 0.959 |
| 127 | 0.985 | 0.973 |



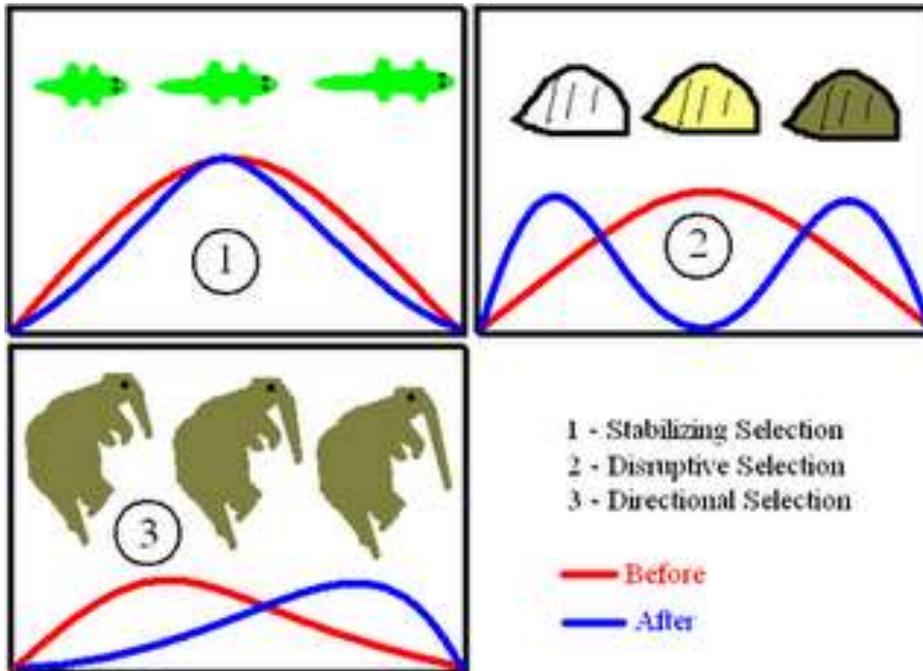
Эффект «бутылочного горлышка», он же «Эффект основателя»

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

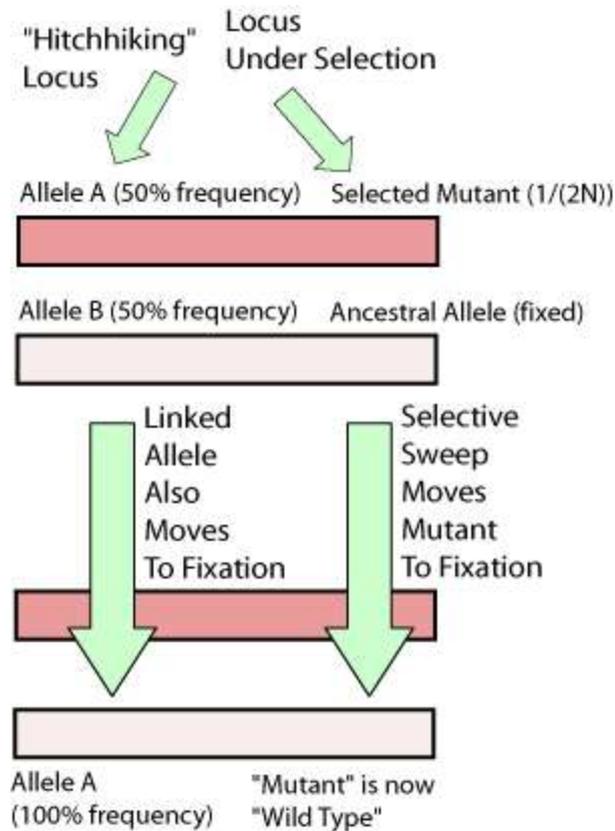
Genetic Drift—Bottleneck Effect



Три разных формы естественного отбора

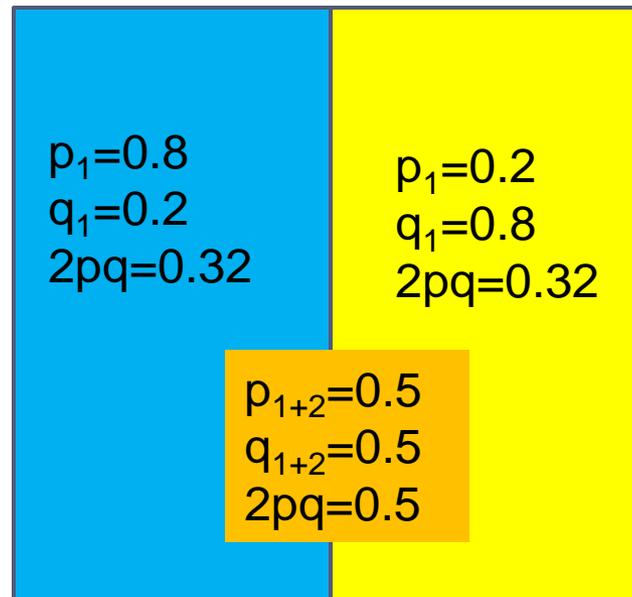


Hitchhiking, или феномен «попутчика»



Эффект Валунда

which is always smaller than $2p(1 - p)$ ($= 2pq$) unless $p_1 = p_2$



$$H = \frac{2p_1q_1 + 2p_2q_2}{2}$$

- ▶ Недостаток гетерозигот часто свидетельствует о присутствии в выборке «смеси» двух популяций (Wahlund effect) (к Воланду и другим темным силам отношения не имеет –шведский генетик Sten Wahlund, 1928г.)

F-статистика Фишера

F-статистика (индексы фиксации)
основаны на измерении F –
индекса имбридинга:

$p^2(1 - F) + pF$ for AA; $2pq(1 - F)$ for Aa; and $q^2(1 - F) + qF$ for aa.

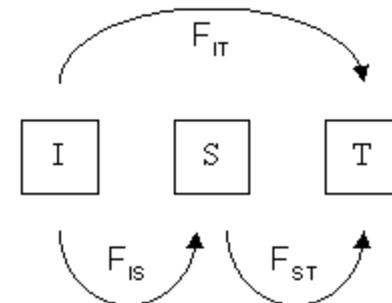
Решение – через нахождение числа
наблюдаемых и ожидаемых гетерозигот

$$F = 1 - \frac{O(f(\mathbf{Aa}))}{E(f(\mathbf{Aa}))} = 1 - \frac{\text{ObservedNumber}(\mathbf{Aa})}{nE(f(\mathbf{Aa}))},$$

here the expected value at Hardy–Weinberg equilibrium is given by

$$E(f(\mathbf{Aa})) = 2pq,$$

Наиболее важно – коэффициент инбридинга
по отношению I (индивидуум), S
(субпопуляция) и T (общая популяция)



$$1 - F_{IT} = (1 - F_{IS})(1 - F_{ST}).$$

$$F_{ST} = \frac{\text{var}(\mathbf{p})}{p(1 - p)}$$

Способ оценить наличие бутылочного горлышка и возраст популяции

▶ *Garza-Williamson index (G-W)*

$$G-W = \frac{k}{R+1}$$

- ▶ K – число наблюдаемых аллелей,
- ▶ R – диапазон возможных аллелей

▶ Попарные расстояния между гаплотипами:

8.1.2.1 Mean number of pairwise differences (π)

Mean number of differences between all pairs of haplotypes in the sample. It is given by

$$\hat{\pi} = \frac{n}{n-1} \sum_{i=1}^k \sum_{j=1}^k p_i p_j \hat{d}_{ij},$$

where \hat{d}_{ij} is an estimate of the number of mutations having occurred since the divergence of haplotypes i and j , k is the number of haplotypes, p_i is the frequency of haplotype i , and n is the sample size. The total variance (over the stochastic and the sampling process), assuming no recombination between sites and selective neutrality, is obtained as

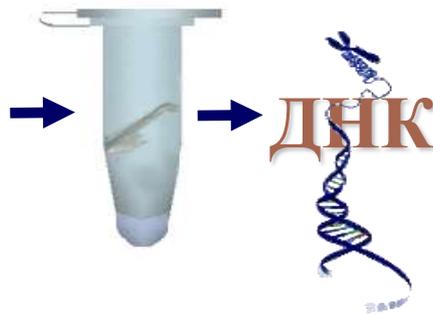
$$V(\hat{\pi}) = \frac{3n(n+1)\hat{\pi} + 2(n^2 + n + 3)\hat{\pi}^2}{11(n^2 - 7n + 6)}. \quad (\text{Tajima, 1993})$$



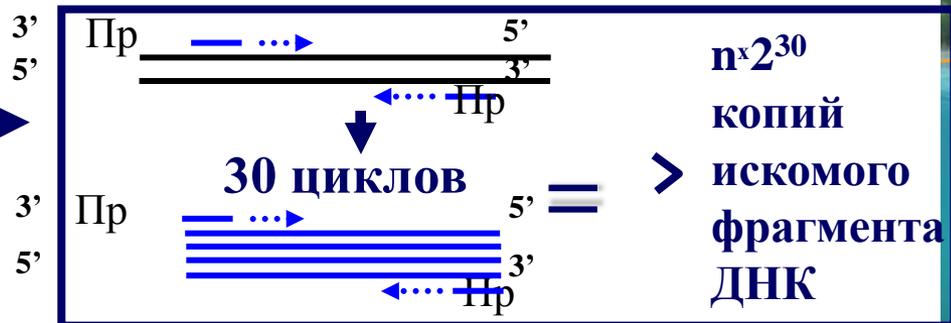
Общая схема изучения нуклеотидных последовательностей



1. Сбор материала



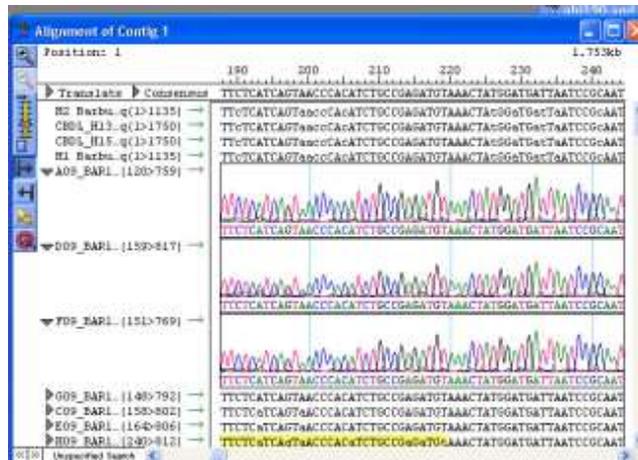
2. Выделение ДНК



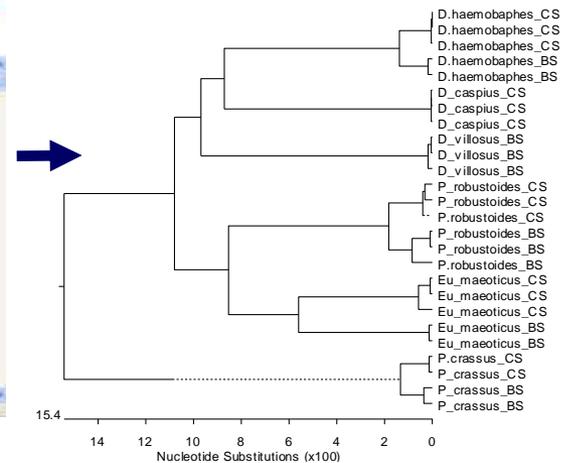
3. ПЦР-амплификация фрагмента мтДНК



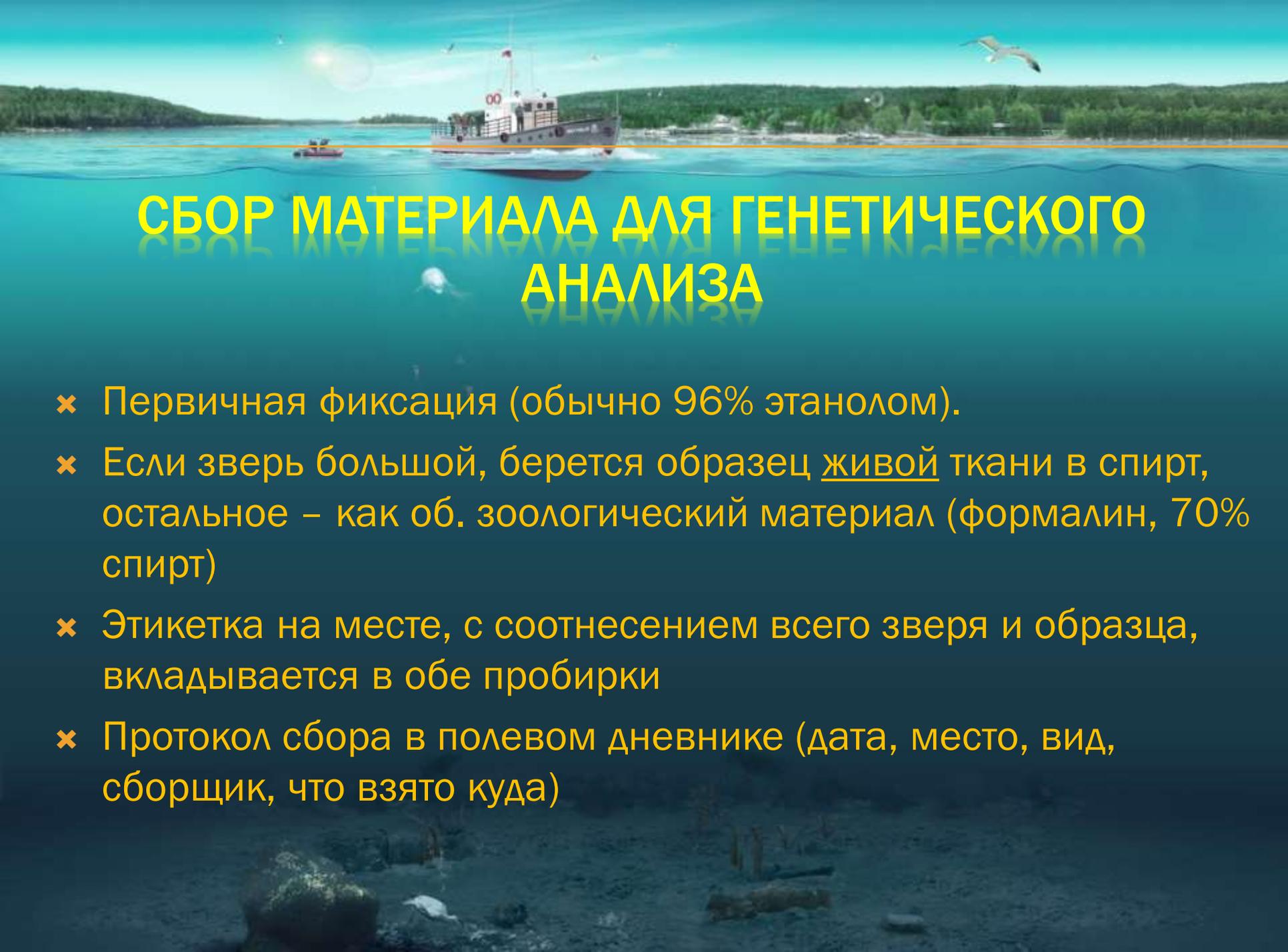
4. Сиквенс пцр-фрагмента



5. Стыковка и множественное выравнивание последовательностей

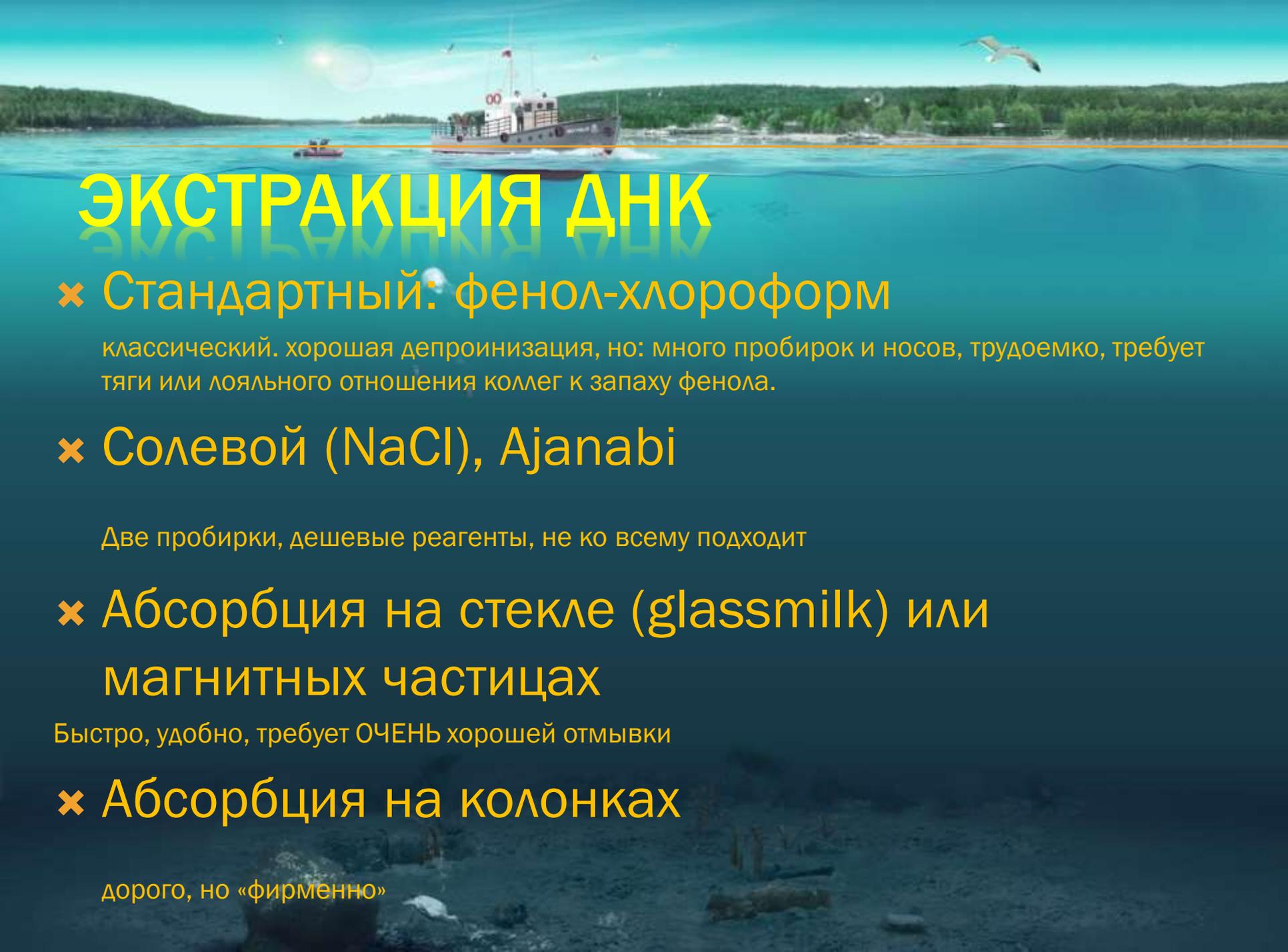


6. Филогенетическая реконструкция

A scenic view of a lake with a boat and a seagull. The background shows a bright blue sky with a sun, a green forested shoreline, and a seagull flying in the air. In the foreground, a boat is on the water, and a jellyfish is visible underwater. The text is overlaid on the image.

СБОР МАТЕРИАЛА ДЛЯ ГЕНЕТИЧЕСКОГО АНАЛИЗА

- ✘ Первичная фиксация (обычно 96% этанолом).
- ✘ Если зверь большой, берется образец живой ткани в спирт, остальное – как об. зоологический материал (формалин, 70% спирт)
- ✘ Этикетка на месте, с соотнесением всего зверя и образца, вкладывается в обе пробирки
- ✘ Протокол сбора в полевом дневнике (дата, место, вид, сборщик, что взято куда)



ЭКСТРАКЦИЯ ДНК

✘ Стандартный: фенол-хлороформ

классический. хорошая депронизация, но: много пробирок и носов, трудоемко, требует тяги или лояльного отношения коллег к запаху фенола.

✘ Солевой (NaCl), Ajanabi

Две пробирки, дешевые реагенты, не ко всему подходит

✘ Абсорбция на стекле (glassmilk) или магнитных частицах

Быстро, удобно, требует ОЧЕНЬ хорошей отмывки

✘ Абсорбция на колонках

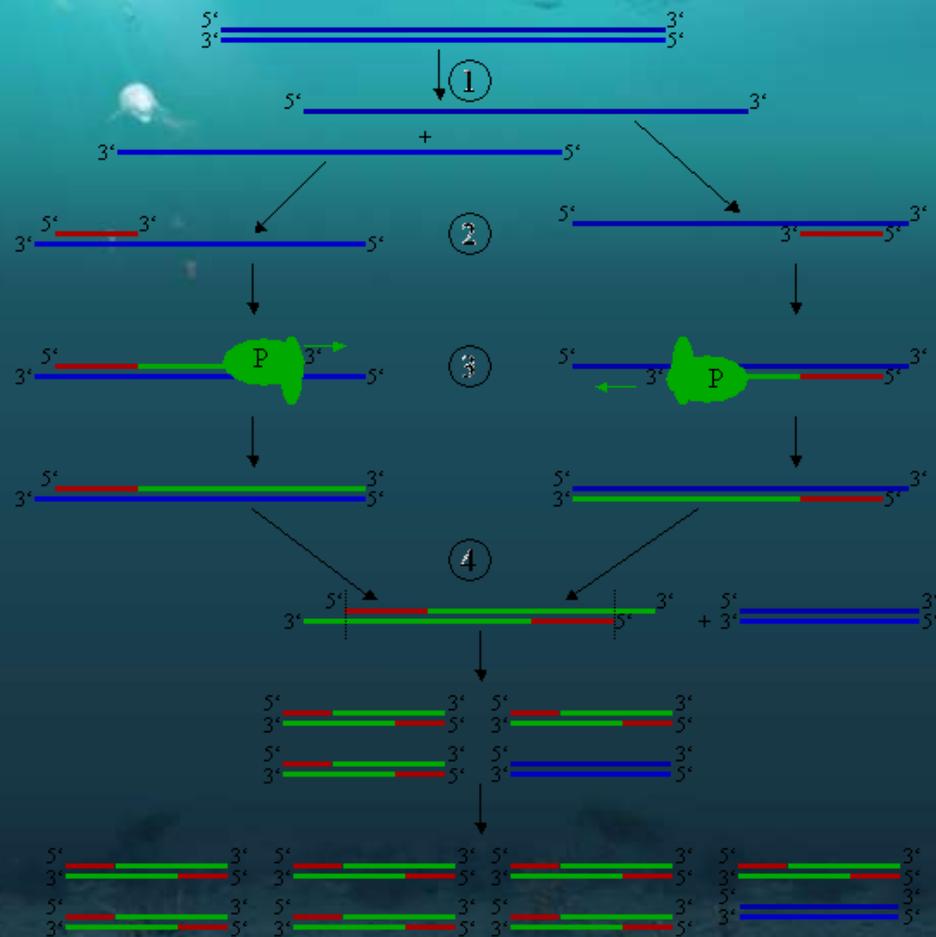
дорого, но «фирменно»

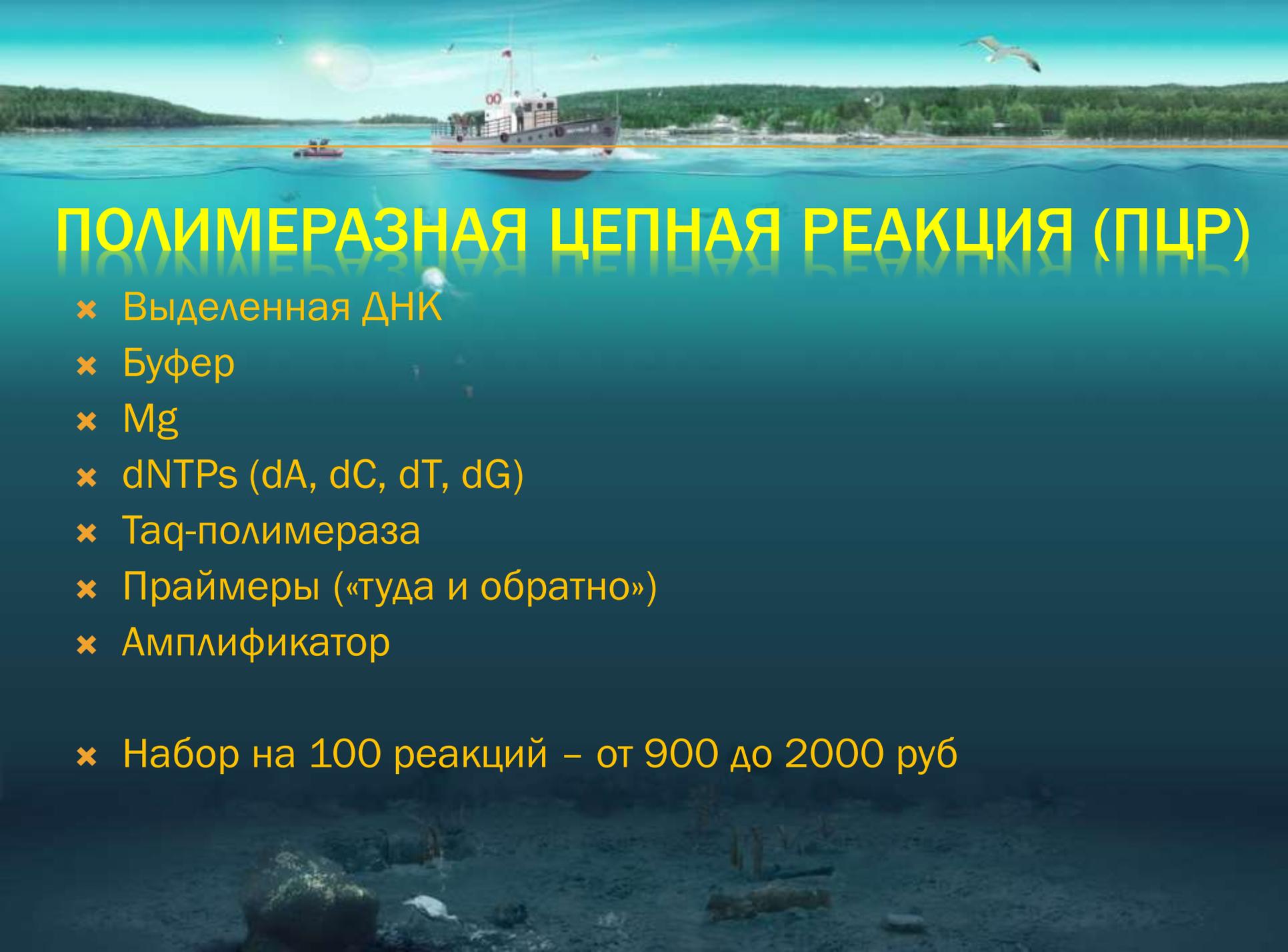
A scenic view of a lake with a boat and a seagull. The background shows a bright blue sky with a sun, a green forested shoreline, and a seagull flying in the upper right. In the middle ground, a boat with a red cabin is on the water, and a smaller boat is visible to the left. The foreground shows the dark, rocky bottom of the lake.

ИСТОРИЯ ПЦР

- ✘ Первоначально искусственный синтез ДНК с использованием олигонуклеотидов был предложен в 1971 (!) г. но не нашел применения (Kleppe et al., 1971)
- ✘ Предложен в 1983 г Кэрри Маллисом (Kary Mullis, (публикация 1985г).
- ✘ Нобелевская премия по химии 1993 г.
- ✘ Патент Cetus Corporation, затем продан за 300000 Hoffmann-La Roche. Значительно тормозил развитие методов, основанных на ПЦР, неоднократно оспаривался в судах (Du Pont, Promega)
- ✘ Патент истек в 2005 г.!

СХЕМА ПЦР



A scenic view of a lake with a boat and a seagull. The background shows a bright blue sky with a sun, a green forested shoreline, and a seagull flying in the air. In the foreground, a boat is on the water, and a small boat is visible in the distance. The water is clear, and the overall scene is peaceful and natural.

ПОЛИМЕРАЗНАЯ ЦЕПНАЯ РЕАКЦИЯ (ПЦР)

- × Выделенная ДНК
 - × Буфер
 - × Mg
 - × dNTPs (dA, dC, dT, dG)
 - × Taq-полимераза
 - × Праймеры («туда и обратно»)
 - × Амплификатор
-
- × Набор на 100 реакций – от 900 до 2000 руб

Что дают микросателлиты, SNP, AFLP, RAPD и т.п.?

- ▶ Наличие генетической дифференциации популяций и выявление популяционной структуры вида
- ▶ Криптические симпатрические виды на ранних этапах видообразования (с оговорками)
- ▶ Поток генов (миграция) между популяциями

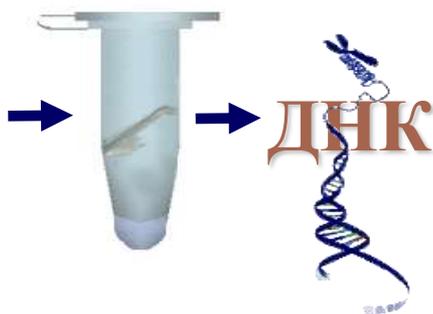
НО: генетические маркеры надо искать для каждой группы!



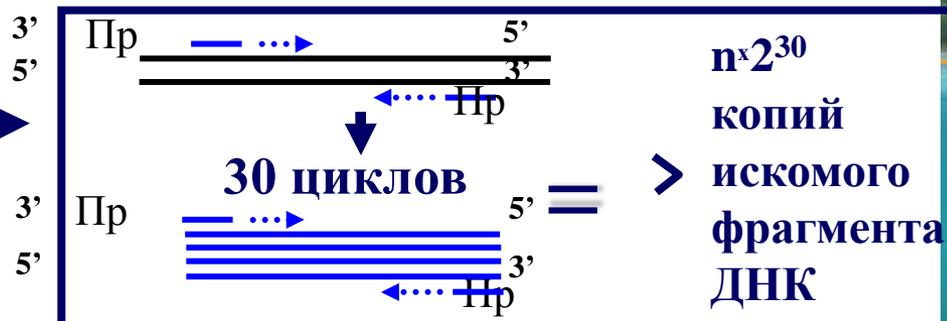
Общая схема изучения нуклеотидных последовательностей



1. Сбор материала



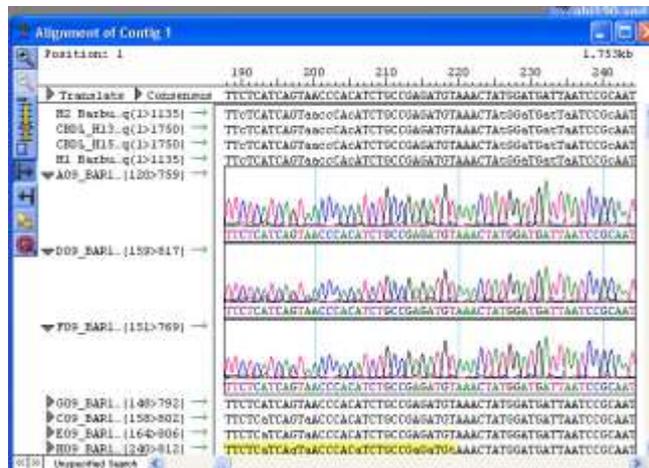
2. Выделение ДНК



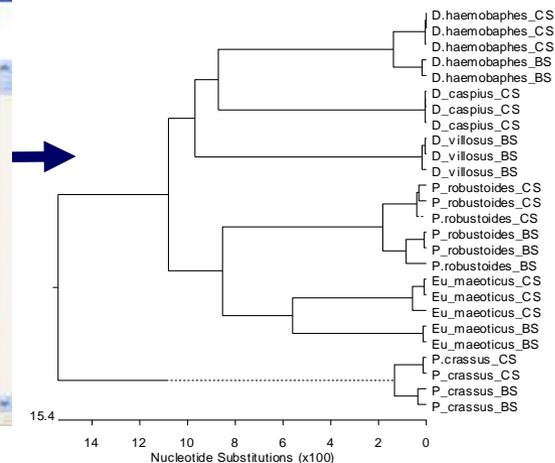
3. ПЦР-амплификация фрагмента мтДНК



4. Сиквенс пцр-фрагмента

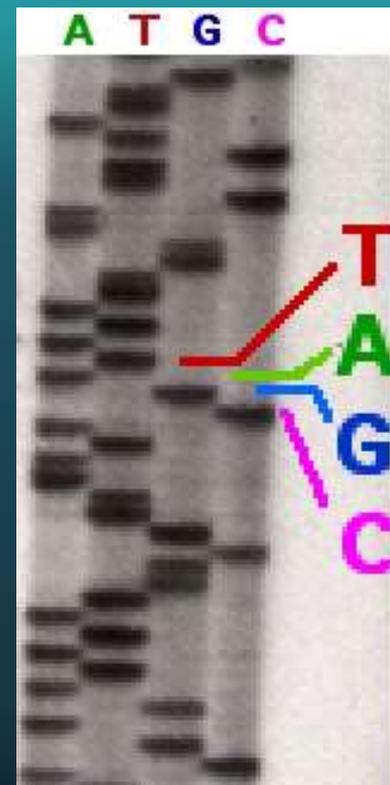
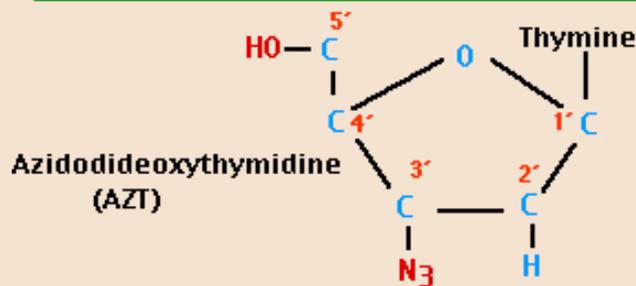
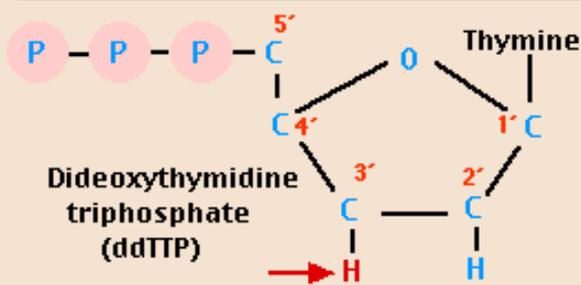
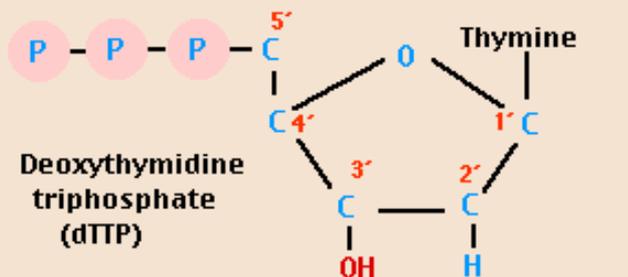


5. Стыковка и множественное выравнивание последовательностей

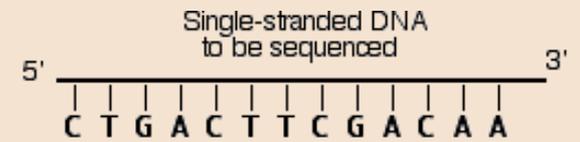


6. Филогенетическая реконструкция

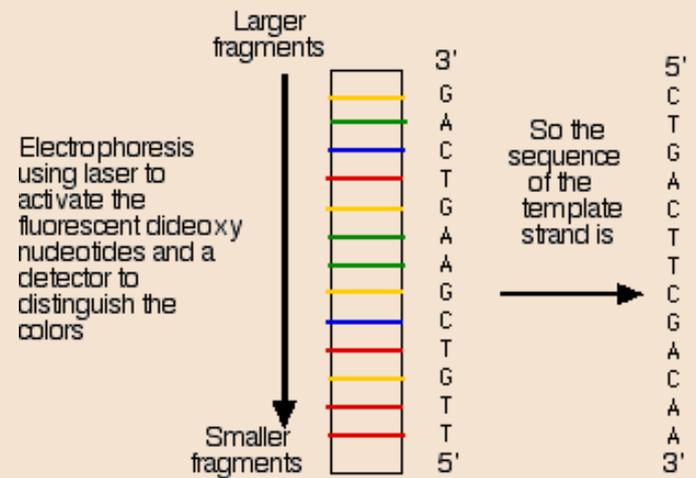
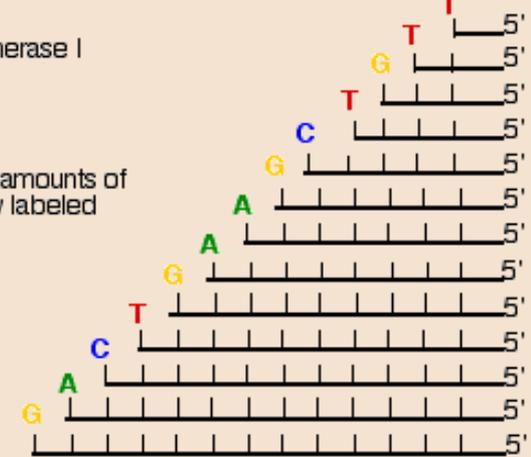
СЕКВЕНИРОВАНИЕ ПО САНГЕРУ (SANGER)



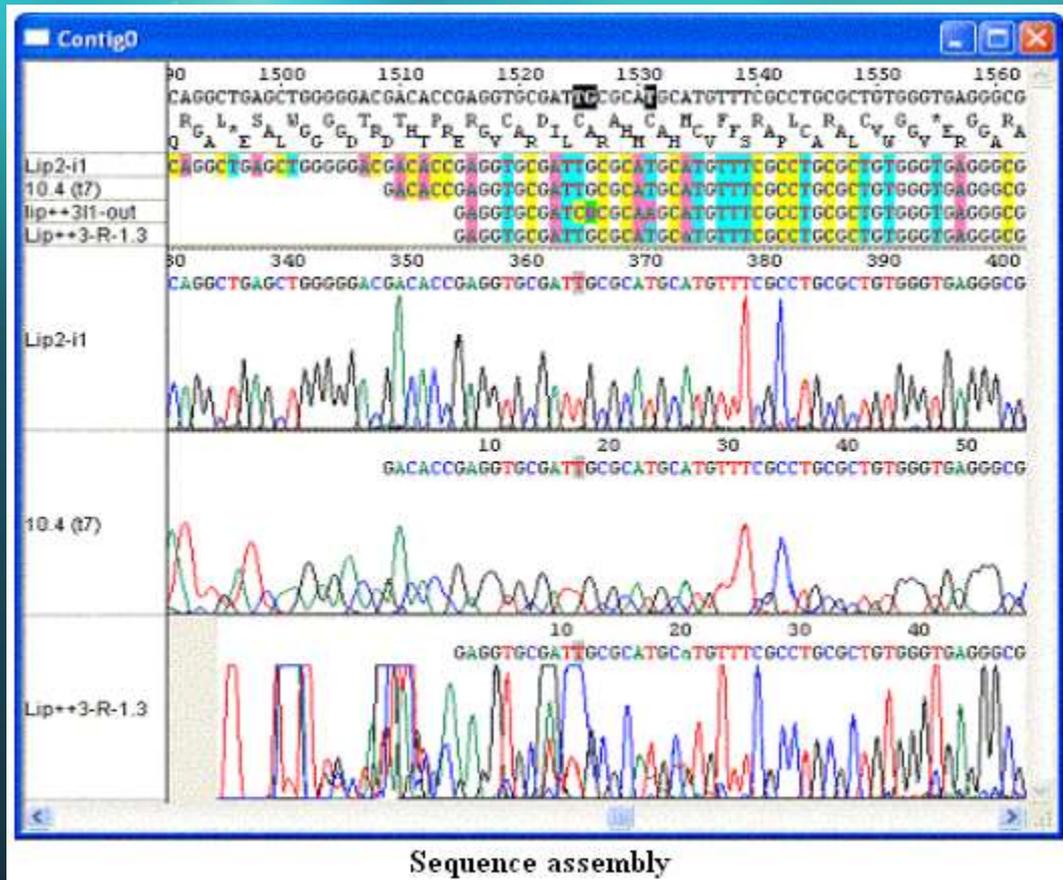
✘ Секвенирование с флуоресцентными ddNTPs



- Add:**
 DNA polymerase I
 dATP
 dGTP
 dCTP
 dTTP
 plus limiting amounts of fluorescently labeled
ddATP
ddGTP
ddCTP
ddTTP

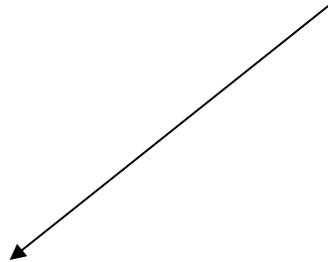


ПРОГРАММЫ ДЛЯ ПЕРВИЧНОЙ ОБРАБОТКИ ПОСЛЕДОВАТЕЛЬНОСТЕЙ CHROMAS И CHROMAS PRO.

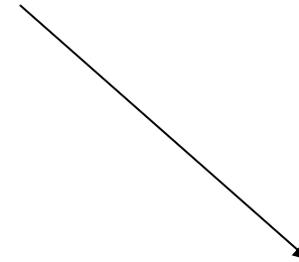


<http://www.technelysium.com.au/ChromasPro.html>

Изучение нуклеотидных последовательностей

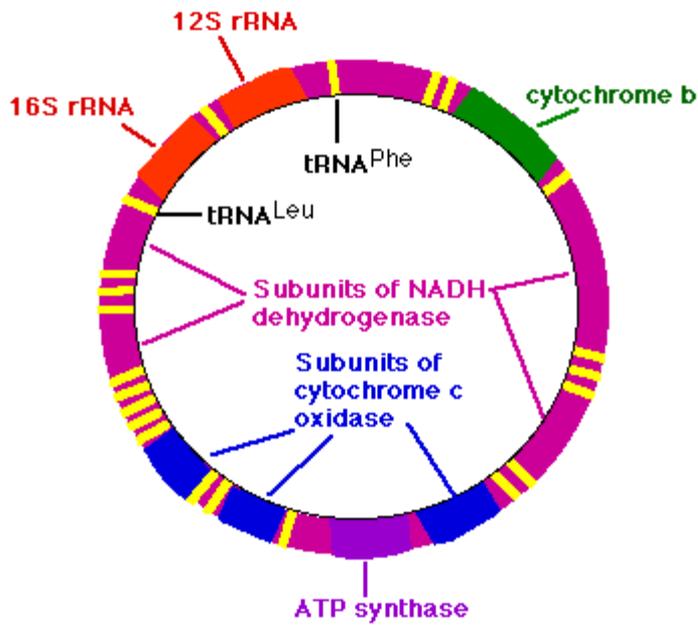


ядерные гены –
18S рРНК, 28S рРНК, 5.8S рРНК, 5S
рРНК; тРНК; мРНК, мяРНК, гены
белков (опсин и др.) и т.д.



митохондриальные гены –
>cyt b (Vertebrata), >COI
(Invertebrata)

Митохондриальная ДНК



- не участвует в рекомбинационном процессе
- обладает высокой скоростью эволюции и значительным внутривидовым полиморфизмом
- нейтральная природа мутаций
- большое количество копий на клетку
- Универсальные праймеры – баркодинг

- изучение последовательностей мтДНК позволяет делать внутривидовые филогенетические построения
- в эволюционно молодых группах расхождение линий происходит быстрее, чем в ядерных генах
- филогеография

ограничения: метод не может быть применим для групп, далеко отстоящих друг от друга по времени образования

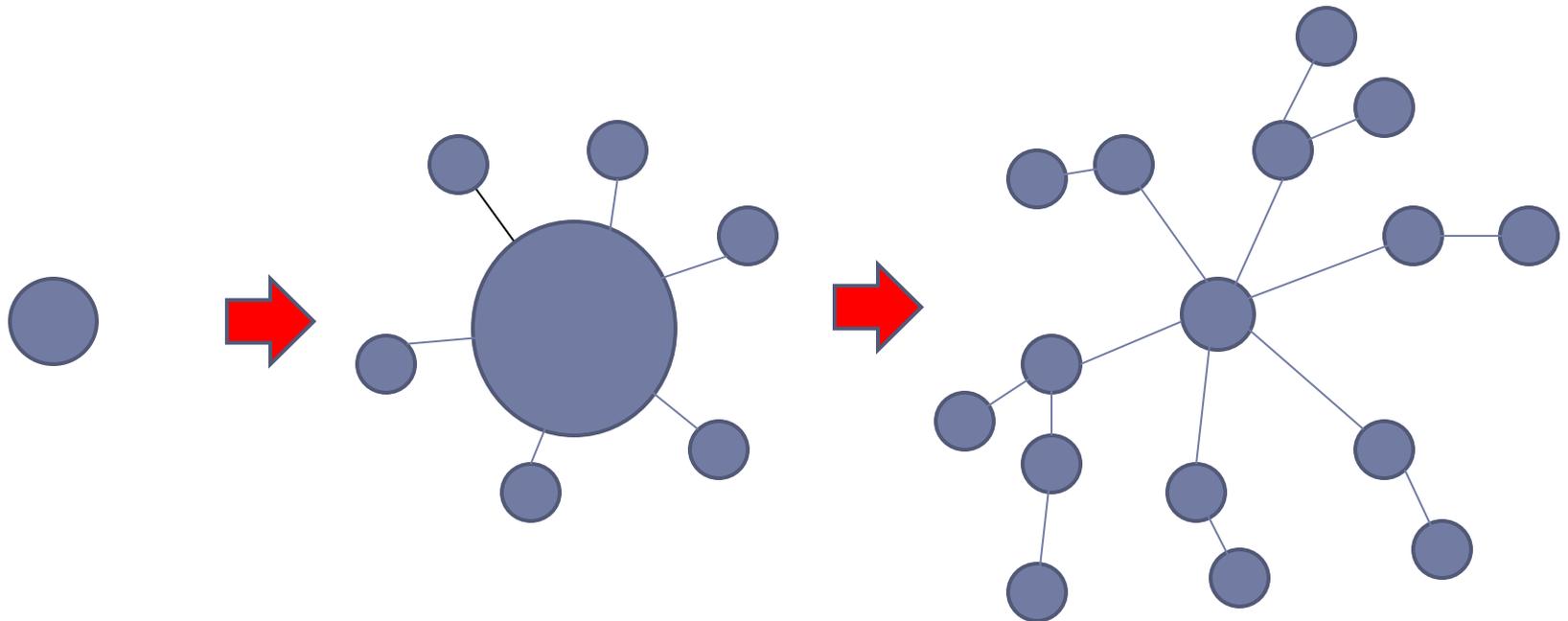


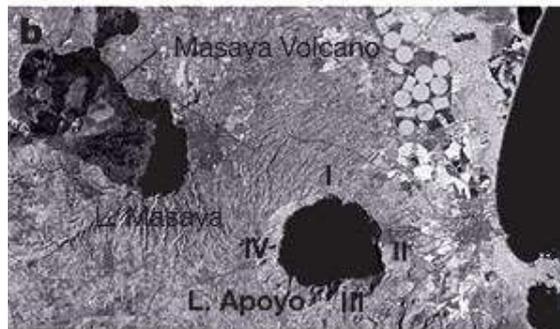
Ядерная ДНК

- большое количество генов, по которым можно проводить филогенетические построения
 - имеет консервативное строение
 - нейтральная природа мутаций
-
- изучение последовательностей ядДНК позволяет делать межвидовые и более высокого ранга филогенетические построения

ограничения: метод слабо применим для внутривидовых филогенетических построений, требует предварительной модификации для работы с новым организмом

Возраст и демографическая история по данным ДНК

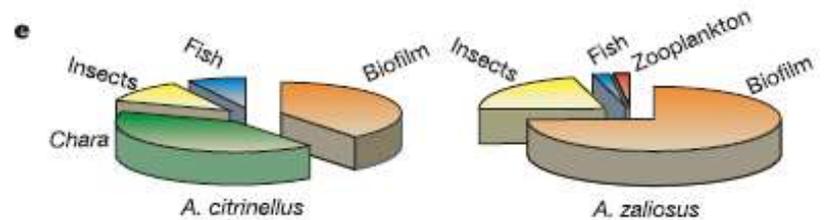
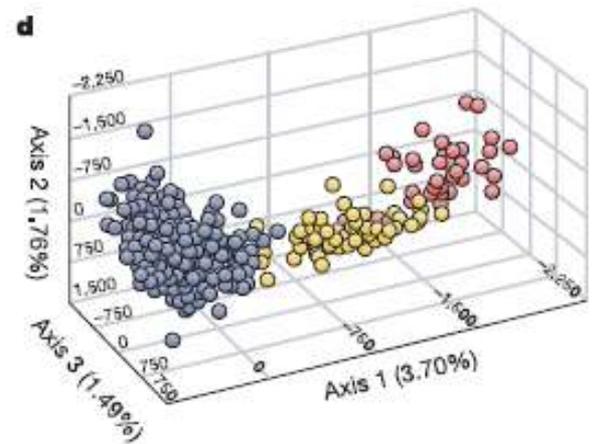
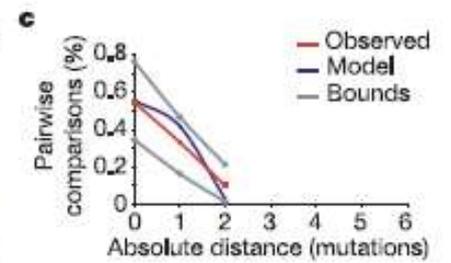
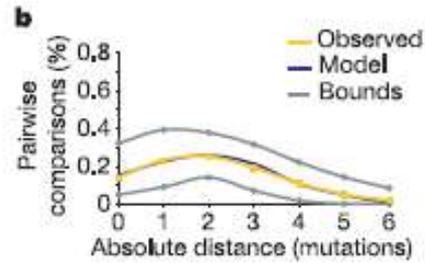
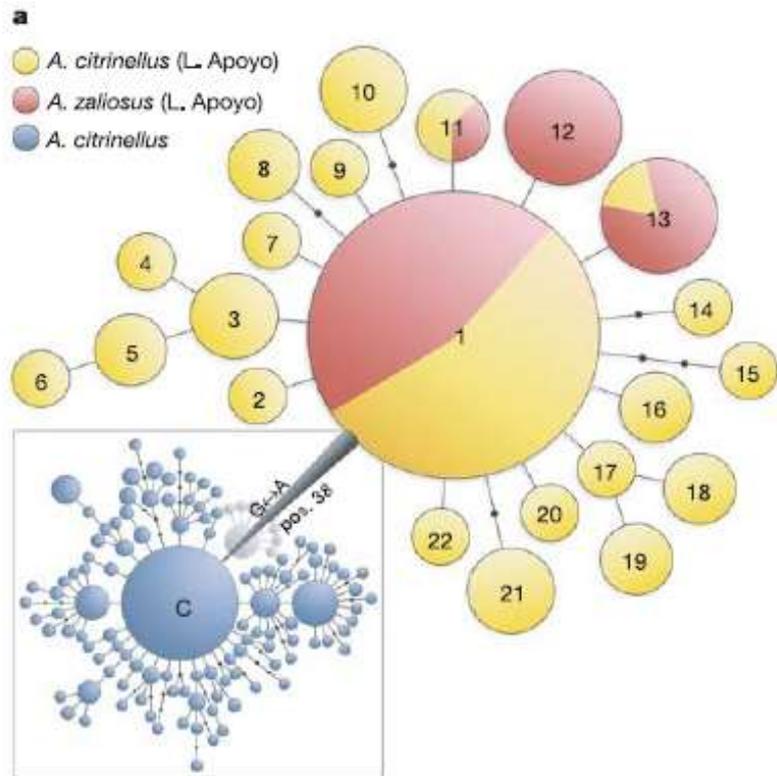


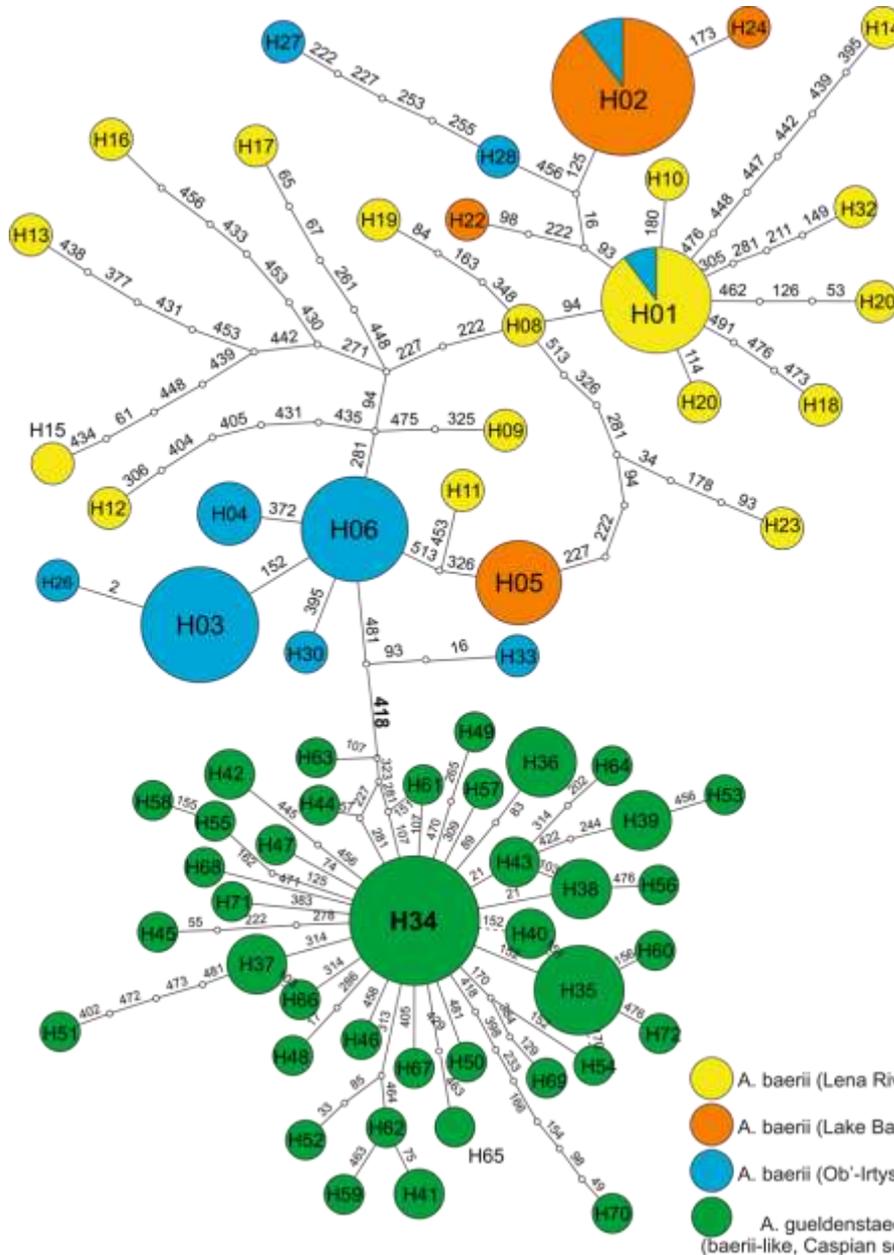


● *Amphilophus citrinellus*

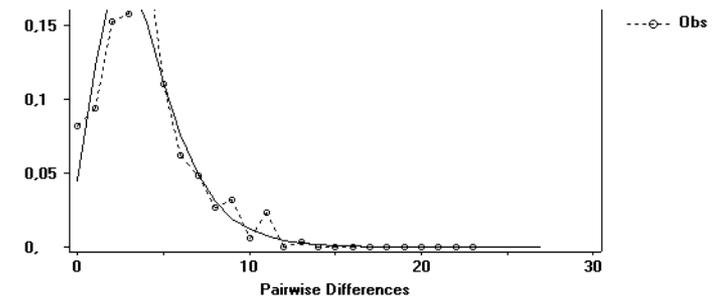
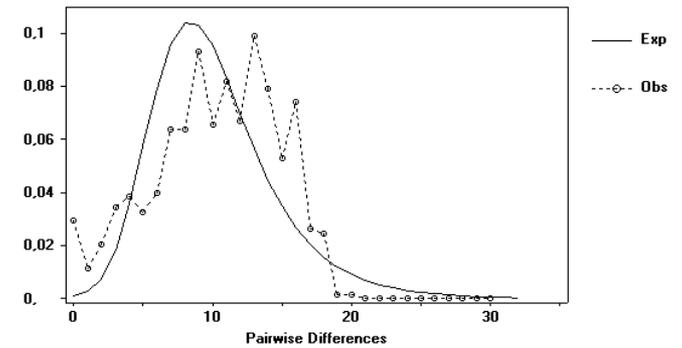


■ *Amphilophus zaliosus*





baerii and baerii-like haplotype MST

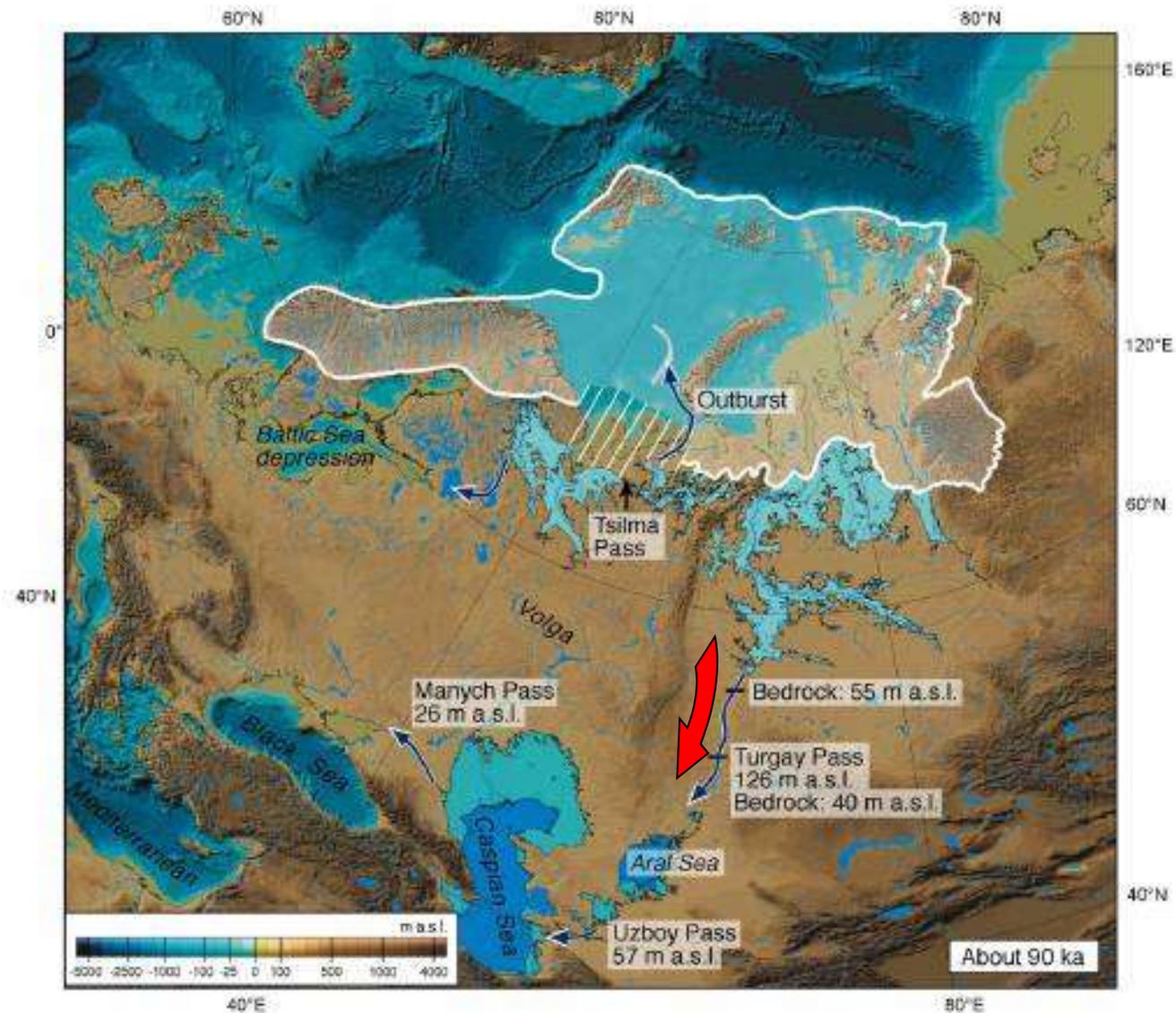


Возможные пути миграции *A. baerii* в Каспийское море

1. через Тургайский сброс в бассейн Арала, 90 000 у.а.

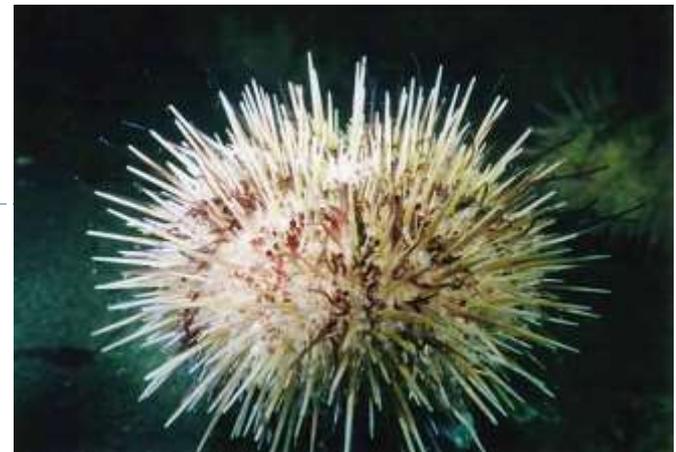
J. Mangerud et al / Quaternary Science Reviews 23 (2004) 1313-1332

1315



From Mangerud
et al., 2005

Ежики в Белом море



538

J. A. ADDISON AND M. W. HART

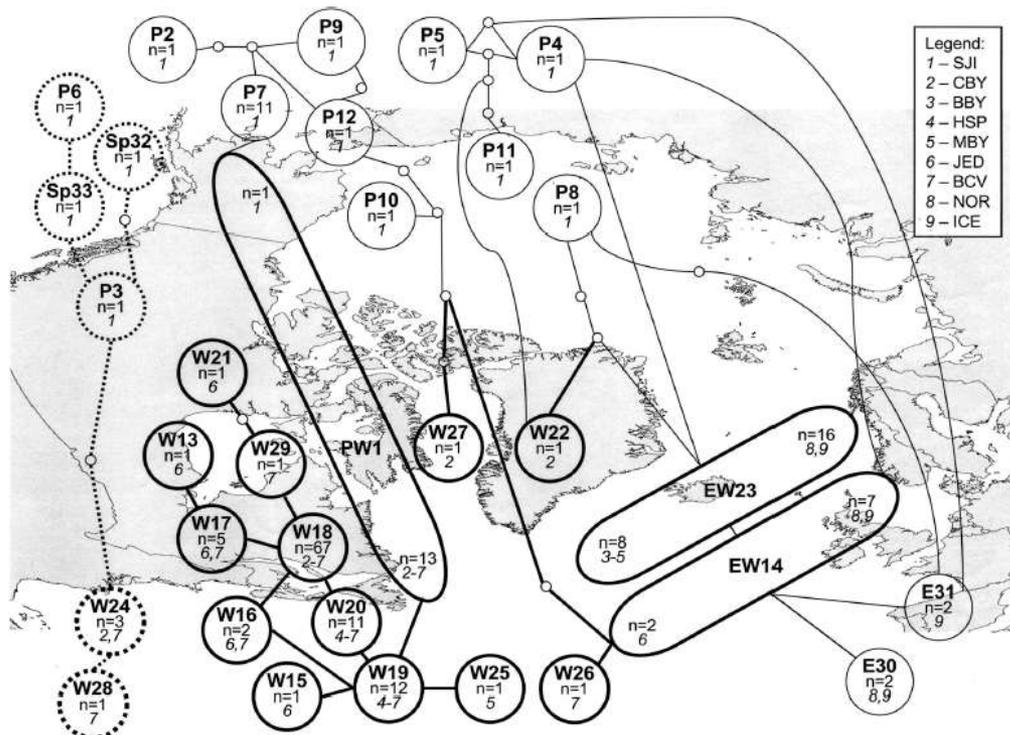


FIG. 3. 95% plausible networks for *Strongylocentrotus* COI haplotypes. Bold lines surround and join haplotypes found in the NW Atlantic; those shared among oceanic regions (PW1, EW14, EW23) are represented by ovals. Haplotype identifiers (as in Table 1 and Fig. 2) are in bold; n is the haplotype sample size; numbers in italics indicate the population in which the haplotypes were detected (SJI, Pacific; NOR, ICE, NE Atlantic; for detailed resolution of NW Atlantic populations see fig. 1 in Addison and Hart 2004). Lines connect haplotypes that differ by one mutation; small open circles indicate missing intermediate haplotypes. Samples of *S. pallidus* (Sp32, Sp33) and four haplotypes from *S. droebachiensis* (P3, P6, W24, W28) form a separate network in the Pacific and NW Atlantic that is indicated by heavy dotted lines.

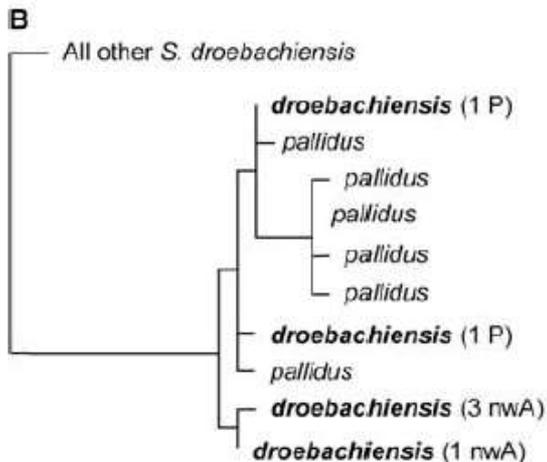
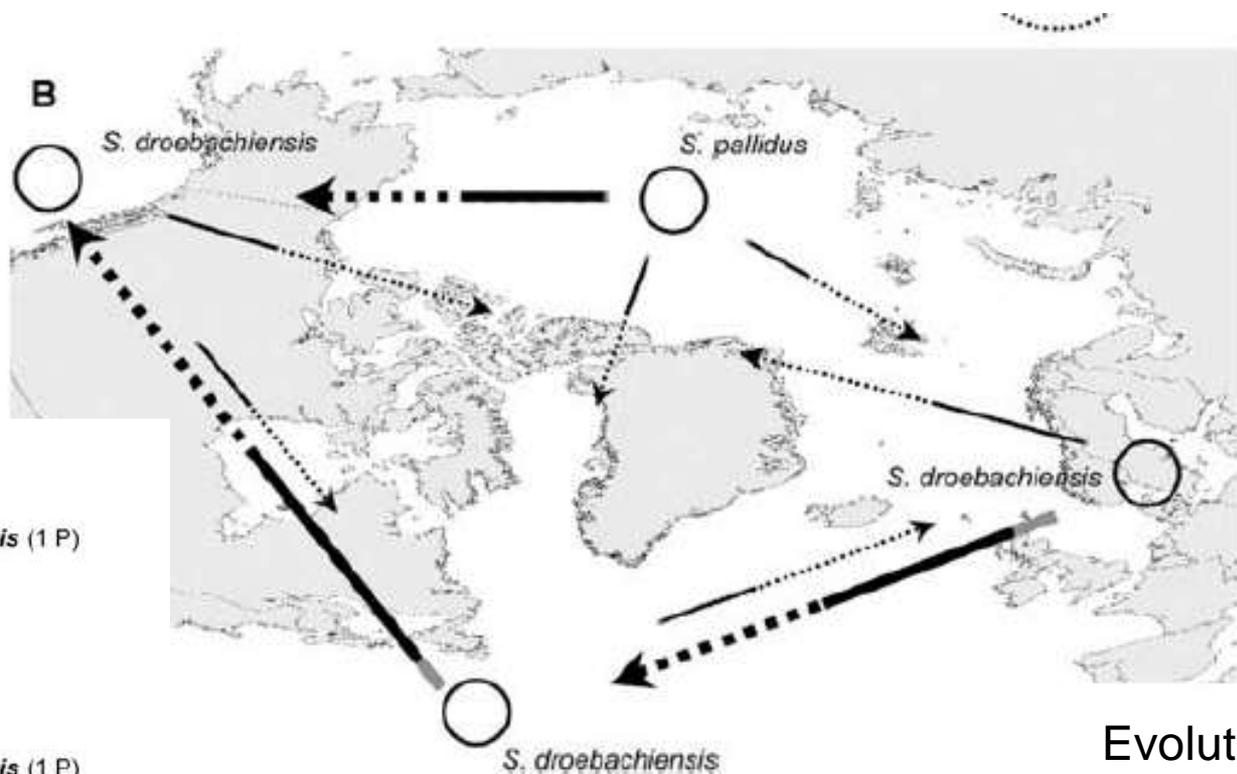
Evolution 2005

031, 2005, pp. 510-541

COLONIZATION, DISPERSAL, AND HYBRIDIZATION INFLUENCE
PHYLOGEOGRAPHY OF NORTH ATLANTIC SEA URCHINS
(*STRONGYLOCENTROTUS DROEBACHIENSIS*)

JASON A. ADDISON¹ AND MICHAEL W. HART²
Department of Biology, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada

Пример интрогрессии мтДНК



Evolution 2007

**INTROGRESSION VERSUS IMMIGRATION
IN HYBRIDIZING HIGH-DISPERSAL
ECHINODERMS**

- unknown
- cement
- byssal attach
- free-living
- recess
- glide
- nestle

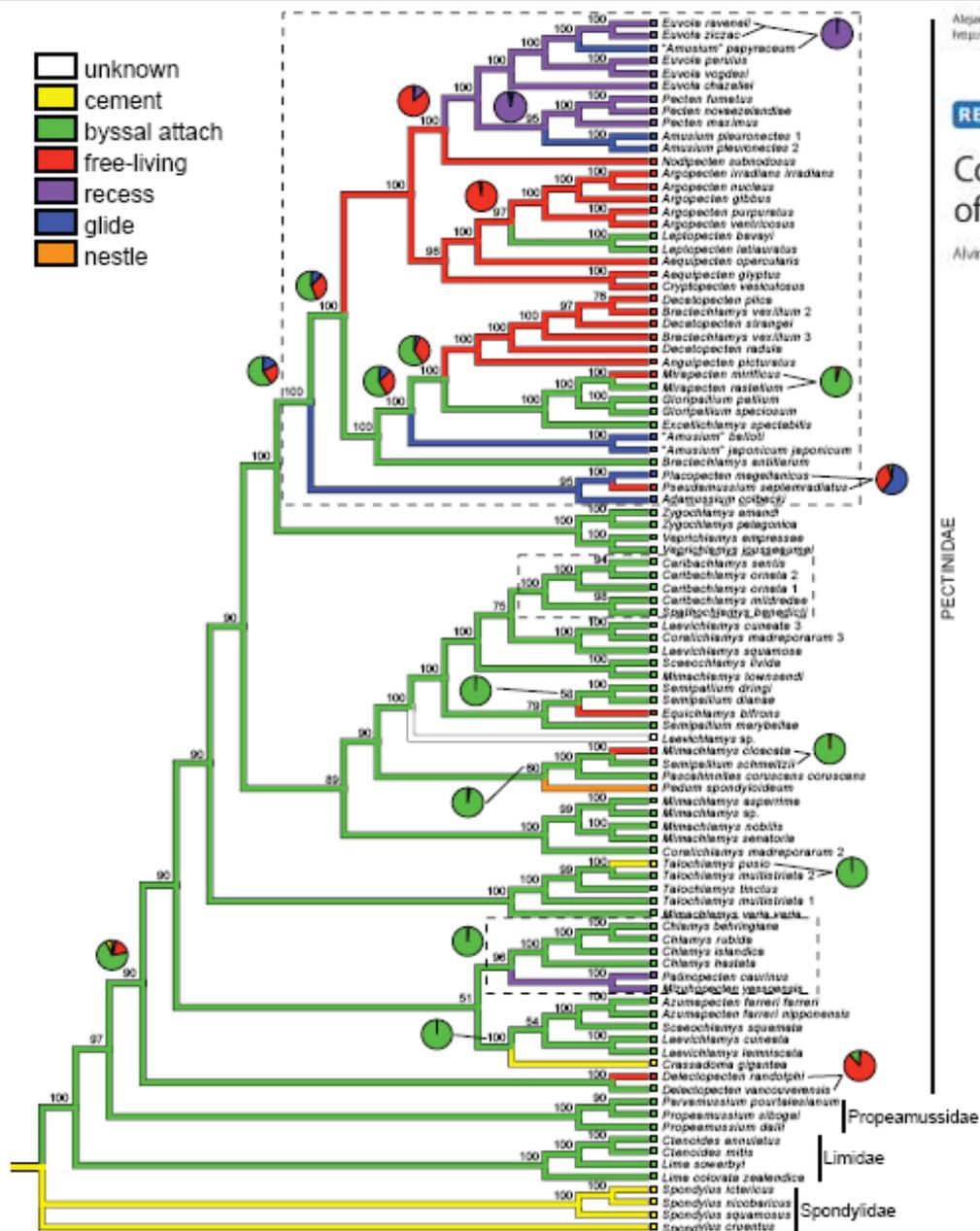


Figure 1 Bayesian inference majority-rule consensus topology. Posterior probability support values (> 50) above respective nodes. Branch colors represent MP reconstruction of life habit and pie charts represent their relative probabilities from ML reconstructions. If probability of ML reconstruction equals 10, no pie chart is given. ML ancestral state reconstructions are used to illustrate the 17 life habit transitions described in the text. Dashed boxed represent densest taxonomic sampling.

RESEARCH ARTICLE

Open Access

Convergent and parallel evolution in life habit of the scallops (Bivalvia: Pectinidae)

Alvin Alejandrino¹, Louise Puslednik^{1,2} and Jeanne M. Serb^{1*}



PECTINIDAE



Propeamussidae

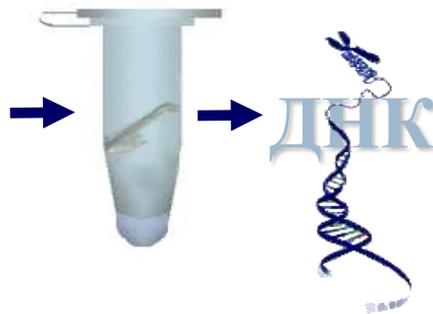
Limidae

Spondyliidae

Общая схема изучения нуклеотидных последовательностей



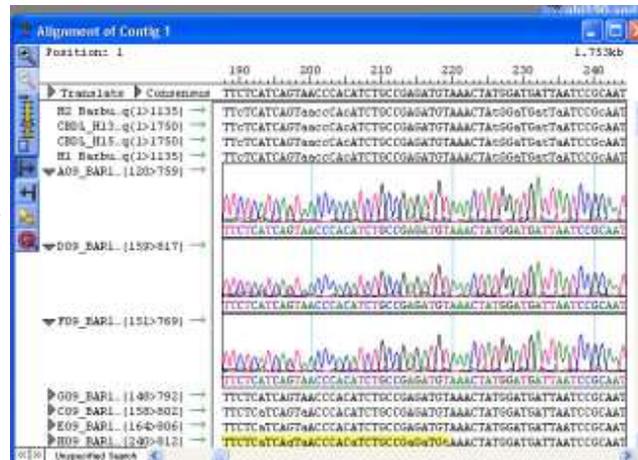
1. Сбор материала



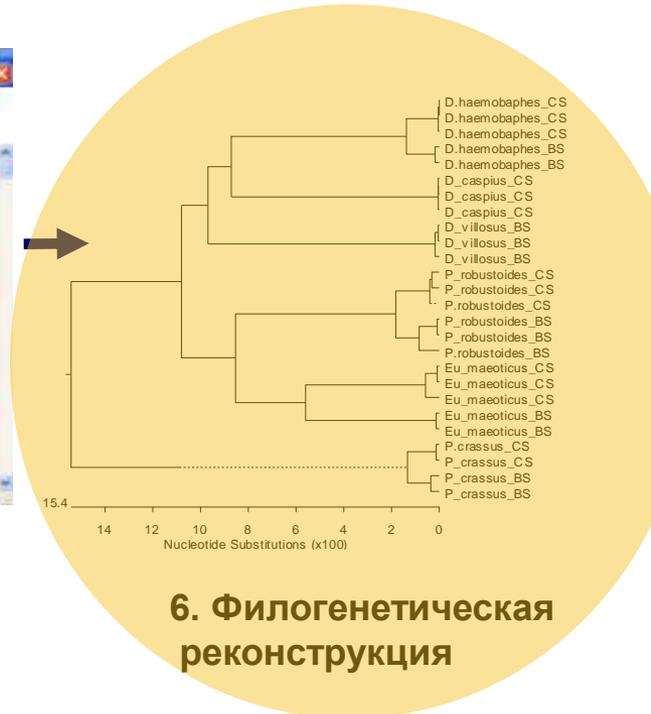
2. Выделение ДНК



4. Сиквенс пцр-фрагмента



5. Стыковка и множественное выравнивание последовательностей



Парсимония

- Ведет начало от кладистики Хеннинга (Henning, 1966)
- Еск и Dayhoff (1966) – первый анализ МР для белков
- Fitch (1971) и Hartigan (1973) – анализ нуклеотидных последовательностей

Программы для анализа парсимонии:

Phylip <http://evolution.genetics.washington.edu/phylip.html>

Paup <http://paup.csit.fsu.edu/>

Mega <http://www.megasoftware.net/>

parsimony сущ. (multitran)

общ. бережливость; экономия; скупость; скряжничество

ЭКОН. ЭКОНОМНОСТЬ



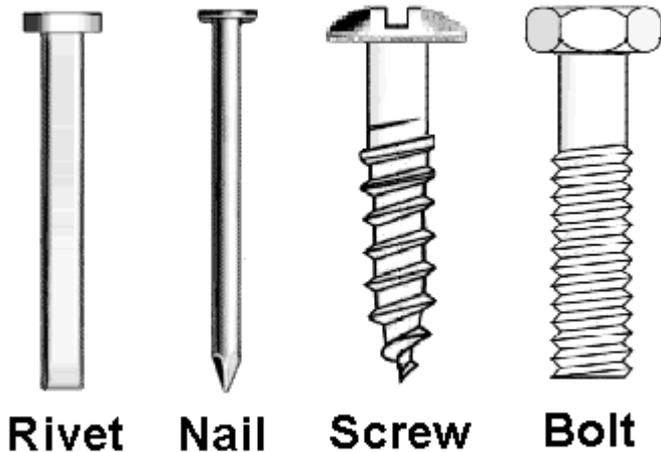
Принципы кладистики: Хенниг (1966)

«Филогенетическая систематика»

- ▶ 1. Кладограммы (филогенетические схемы) строятся по дихотомическому принципу.
- ▶ 2. Таксоны выделяются только по вертикальному принципу.
- ▶ 3. Ранг таксонов определяется последовательностью их ответвления на кладограмме, понижаясь от основания кладограммы к вершине; таким образом, степень родства таксонов соответствует времени их разделения.
- ▶ 4. Все признаки, характеризующие таксон, подразделяются на **плезиоморфные** (унаследованные, примитивные) и **апоморфные** (производные, прогрессивные).
- ▶ 5. Таксоны выделяются только по апоморфным признакам.
- ▶ 6. Критерием родства является синапоморфия; соответственно последовательность обособления различных таксонов на кладограмме определяется путем сопоставления их апоморфных признаков.
- ▶ 7. Пары таксонов, исходящие на кладограмме из одной точки, образуют «сестринские группы», связанные друг с другом максимальным родством и характеризующиеся наиболее полной синапоморфией.
- ▶ 8. Из пары сестринских групп одна обычно сохраняет значительно большее сходство с предковым таксоном, чем другая (правило девиации); обоим сестринским таксонам придается тем не менее одинаковый ранг.
- ▶ 9. Предковый таксой, давая начало двум сестринским, исчезает, что определяется требованиями дихотомического принципа построения кладограмм.



Разные подходы к сравнению организмов (фенетика, кладистика)



| признак | Заклепка | гвоздь | шуруп | винт |
|--------------------|----------|--------|-------|------|
| шлиц | 0 | 0 | 1 | 0 |
| закругл. гол-ка | 0 | 0 | 1 | 0 |
| шестигр гол-ка | 0 | 0 | 0 | 1 |
| резьба | 0 | 0 | 1 | 1 |
| сужающ вниз | 0 | 0 | 1 | 0 |
| острый конец | 0 | 1 | 1 | 0 |
| большой диаметр | 0 | 0 | 1 | 1 |

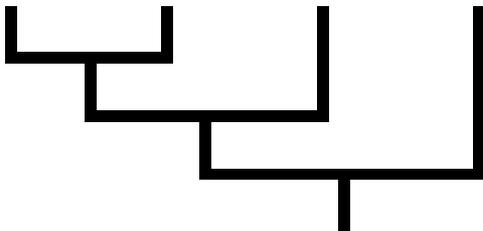
Plesiomorph / Plesiomorphic / Plesiomorphy (плезиоморфный, плезиоморфия) - признак, который присутствует в наследственной форме и также сохранен в потомке или потомках.

Apomorph / Apomorphy / Apomorphic (апоморф, апоморфный, апоморфия) - Признак в организме, который получен из другого, но уже больше не являющийся тем же самым как предыдущий признак.

Фенетический анализ
(сумма общих признаков)

| | Заклепка | гвоздь | шуруп | винт |
|----------|----------|--------|-------|------|
| Заклепка | - | 6 | 1 | 4 |
| гвоздь | | - | 2 | 3 |
| шуруп | | | - | 2 |
| винт | | | | - |

Rivet Nail Bolt Screw

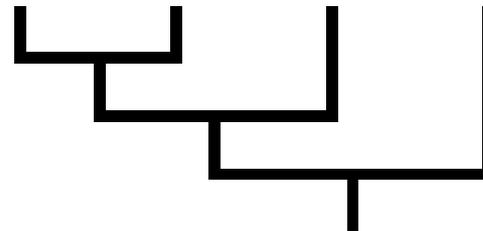


Phenogram

Кладистический анализ
(сумма производных признаков)

| | Заклепка | гвоздь | шуруп | винт |
|----------|----------|--------|-------|------|
| Заклепка | - | 0 | 0 | 0 |
| гвоздь | | - | 1 | 0 |
| шуруп | | | - | 2 |
| винт | | | | - |

Bolt Screw Nail Rivet



Cladogram

Три основных метода реконструкции филогении:



▶ Парсимония (Parsimony) (PAUP, MEGA, Phylip)



▶ Максимального правдоподобия (maximum likelihood) - (PAUP, Phylip)



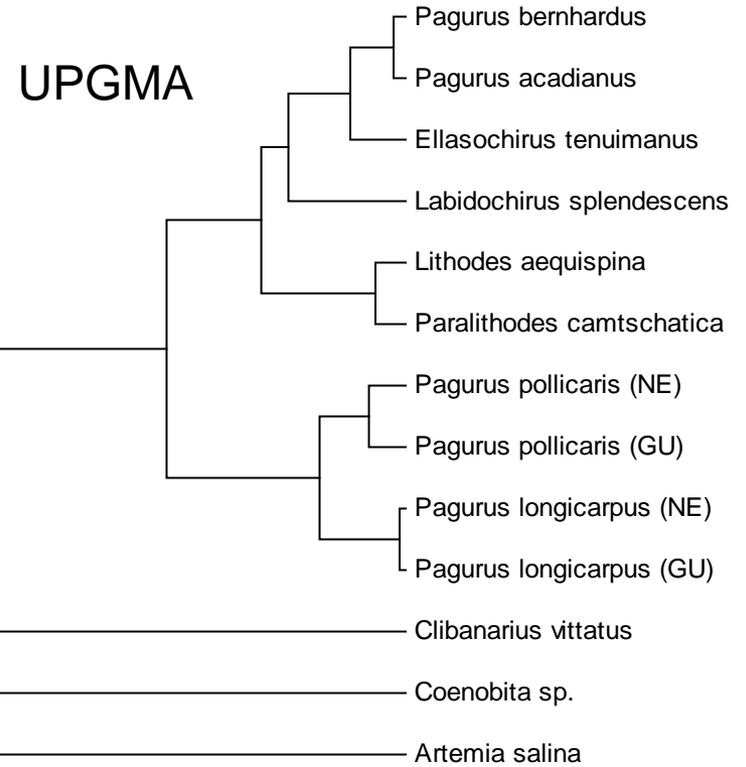
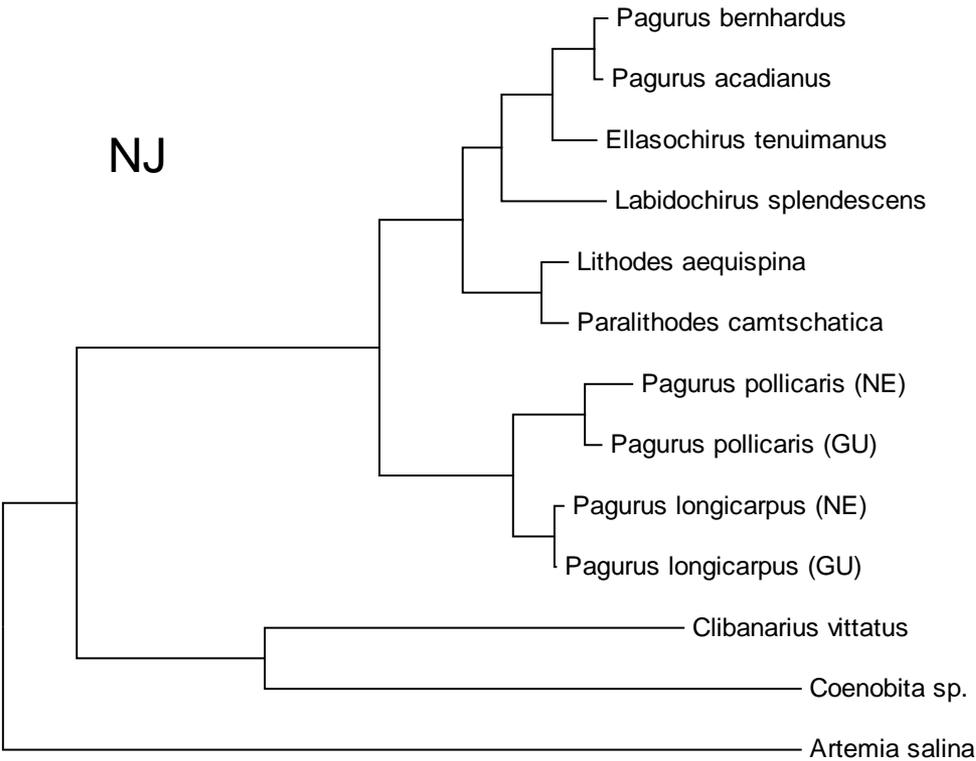
▶ Обратной вероятности, байезиан (bayesian) – (MrBayes)



NJ

MEGA 4.0
Crab_RNA

UPGMA



0.05

0.15

0.10

0.05

0.00



Операторы теории множеств

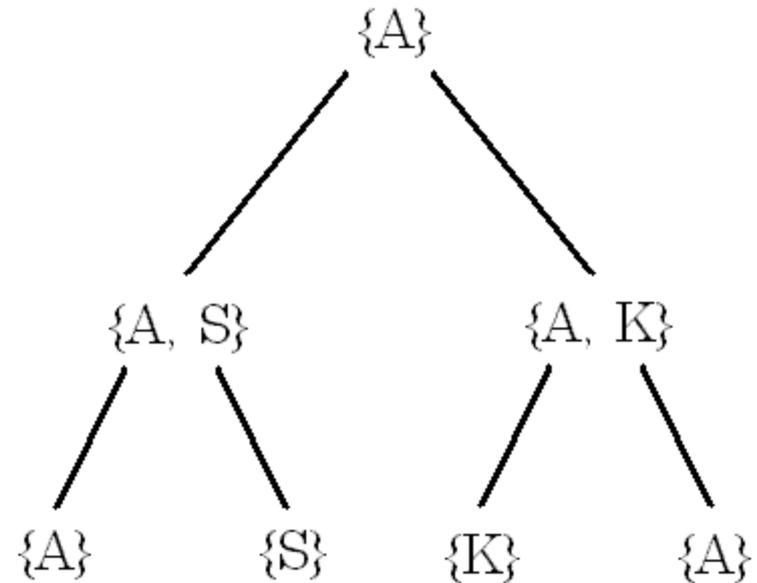
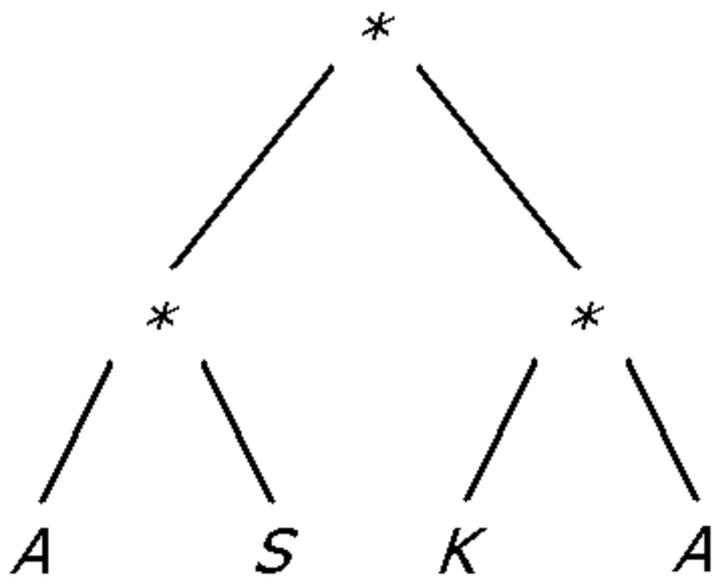
U объединение $[X, Y] \cup [X, Z] = [X, Y, Z]$

∩ пересечение $[X, Y] \cap [X, Z] = [X]$

<> Экономия $[X, Y] \langle \rangle [X, Z] = [X]$

(парсимония) $[X] \langle \rangle [Z] = [X, Z]$





Ожидаемый результат – найти топологию дерева , с наименьшим количеством замен (не всегда это будет одно дерево)



Правдоподобие (Likelihood)

➤ В модели, использующейся для анализа нуклеотидных последовательностей методом правдоподобия, определяется вероятность перехода за определенное время от одной последовательности к другой в результате мутаций.

Программы для анализа парсимонии:

| | |
|--------|---|
| Phylip | http://evolution.genetics.washington.edu/phylip.html |
| PAUP | http://paup.csit.fsu.edu/ |
| PhyML | http://atgc.lirmm.fr/phyml/ |



▶ **JC69 model (Jukes and Cantor, 1969)**

- ▶ JC69 is the simplest substitution model. There are several assumptions. It assumes equal base frequencies (

$$\pi_T = \pi_C = \pi_A = \pi_G = \frac{1}{4}$$

- ▶) and equal mutation rates. The only parameter of this model is therefore μ , the overall substitution rate.



▶ **K80 model (Kimura, 1980)**

- ▶ The K80 model distinguishes between transitions (A <-> G, i.e. from purine to purine, or C <-> T, i.e. from pyrimidine to pyrimidine) and transversions (from purine to pyrimidine or vice versa) (α/β).
- ▶ It also assumes equal base frequencies

$$\pi_T = \pi_C = \pi_A = \pi_G = \frac{1}{4}$$



F81 model (Felsenstein 1981)

Unequal base frequencies ($\pi_T \neq \pi_C \neq \pi_A \neq \pi_G$)

$$\text{Rate matrix } Q = \begin{pmatrix} * & \pi_T & \pi_T & \pi_T \\ \pi_C & * & \pi_C & \pi_C \\ \pi_A & \pi_A & * & \pi_A \\ \pi_G & \pi_G & \pi_G & * \end{pmatrix}$$

HKY85 model (Hasegawa, Kishino and Yano 1985)

The HKY85 model distinguishes between [transitions](#) and [transversions](#) (α/β).

It allows unequal base frequencies ($\pi_T \neq \pi_C \neq \pi_A \neq \pi_G$).

$$\text{Rate matrix } Q = \begin{pmatrix} * & \kappa\pi_T & \pi_T & \pi_T \\ \kappa\pi_C & * & \pi_C & \pi_C \\ \pi_A & \pi_A & * & \kappa\pi_A \\ \pi_G & \pi_G & \kappa\pi_G & * \end{pmatrix}$$

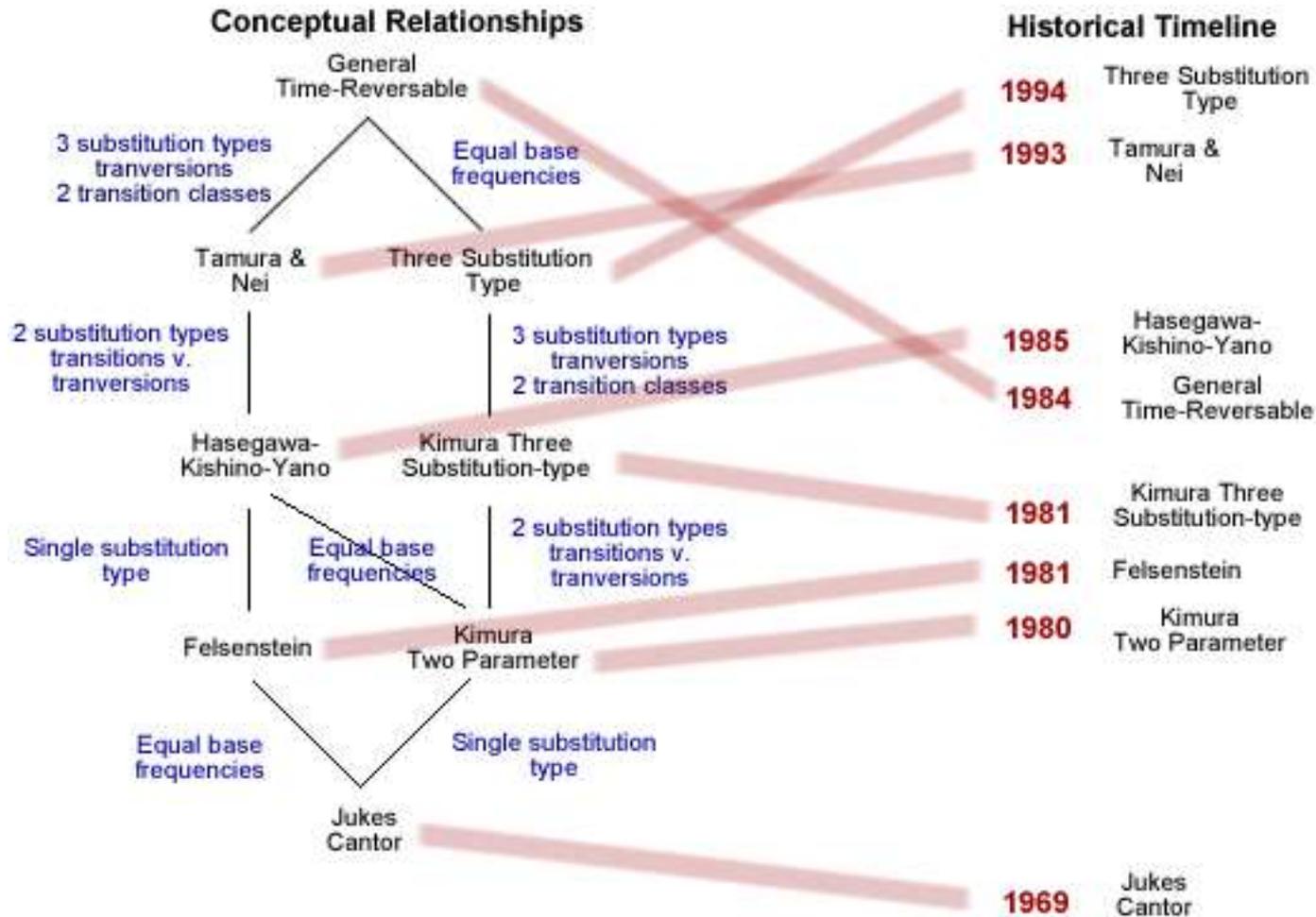


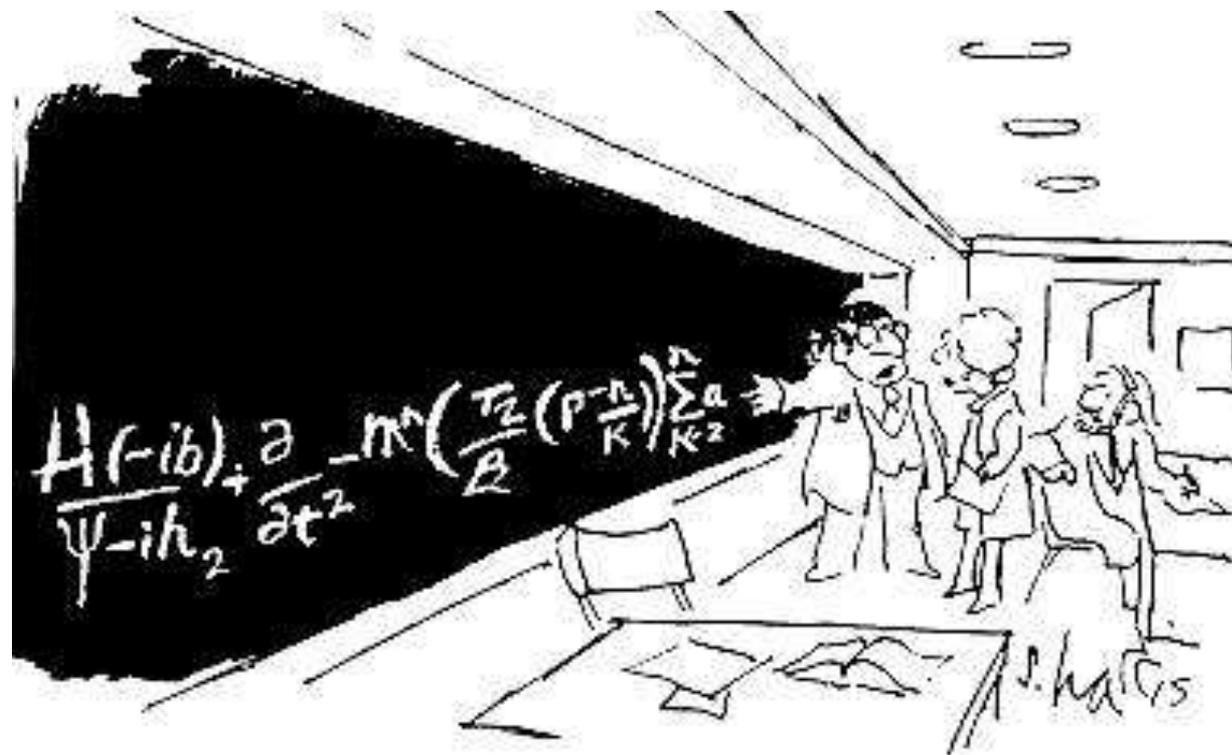
GTR: Generalised time reversible

$$Q = \begin{pmatrix} -\left(\frac{\pi_1 x_1}{\pi_2} + \frac{\pi_1 x_2}{\pi_3} + \frac{\pi_1 x_3}{\pi_4}\right) & \frac{\pi_1 x_1}{\pi_2} & \frac{\pi_1 x_2}{\pi_3} & \frac{\pi_1 x_3}{\pi_4} \\ x_1 & -\left(x_1 + \frac{\pi_2 x_4}{\pi_3} + \frac{\pi_2 x_5}{\pi_4}\right) & \frac{\pi_2 x_4}{\pi_3} & \frac{\pi_2 x_5}{\pi_4} \\ x_2 & x_4 & -\left(x_2 + x_4 + \frac{\pi_3 x_6}{\pi_4}\right) & \frac{\pi_3 x_6}{\pi_4} \\ x_3 & x_5 & x_6 & -(x_3 + x_5 + x_6) \end{pmatrix}$$



Эволюция моделей эволюции ДНК





“But this is the simplified version for the general public.”

Томас Байес 1702-1761



Беркли утверждает, что Логика и Метафизика откроют математикам глаза и выведут их из всех затруднений... Но если склоки среди профессоров любой науки позорят саму науку, а Логика и Метафизика намного более склочны, нежели математика, то почему же, раз я наполовину слеп, я должен выбирать себе в проводники того, кто вообще ничего не видит?

Преподобный Томас Байес, "В защиту математиков..."



Академик (член королевского общества), не опубликовавший ни одной работы по математике

- ▶ Единственная работа отца Байеса, опубликованная им под своим именем (в 1731 году):

"Благость господня, или попытка доказать, что конечной целью божественного провидения и направления является счастье его созданий".



-
- ▶ Фундаментальное исследование Байеса в области теории вероятностей было изложено им в

"Эссе о решении проблем в теории случайных событий".

Эту работу математика лишь после его смерти обнаружил друг Ричард Прайс, который и переслал статью в академию. В 1764 году это "Эссе" было опубликовано в "Трудах Лондонского Королевского общества", откуда и берет начало его мировая слава.



Теорема преподобного Байеса

Формула Байеса позволяет «переставить причину и следствие»: по известному факту события вычислить вероятность того, что оно было вызвано данной причиной.

События, отражающие действие «причин», в данном случае обычно называют **гипотезами**, так как они — *предполагаемые* события, повлекшие данное. Безусловную вероятность справедливости гипотезы называют **априорной** (насколько вероятна причина вообще), а условную - с учетом факта произошедшего события — **апостериорной** (насколько вероятна причина *оказалась с учетом данных о событии*).

Формула Байеса:

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

где

$P(A)$ — априорная вероятность гипотезы A (смысл такой терминологии см. ниже);

$P(A|B)$ — вероятность гипотезы A при наступлении события B (апостериорная вероятность);

$P(B|A)$ — вероятность наступления события B при истинности гипотезы A ;

$P(B)$ — вероятность наступления события B .



Формула полной вероятности

Важным следствием формулы Байеса является формула полной вероятности события, зависящего от нескольких несовместных гипотез (и только от них!).

$$P(B) = \sum_{i=1}^N P(A_i)P(B|A_i)$$

вероятность наступления события B , зависящего от ряда гипотез A_i , если известны степени достоверности этих гипотез (например, измерены экспериментально)



<http://mrbayes.csit.fsu.edu/>



MrBayes: Bayesian Inference of Phylogeny

Search results for 'mrbayes: (Нет результатов)'

[Home](#)

[Download](#)

[Manual](#)

[Online Help](#)

[Bug Report](#)

[Authors](#)

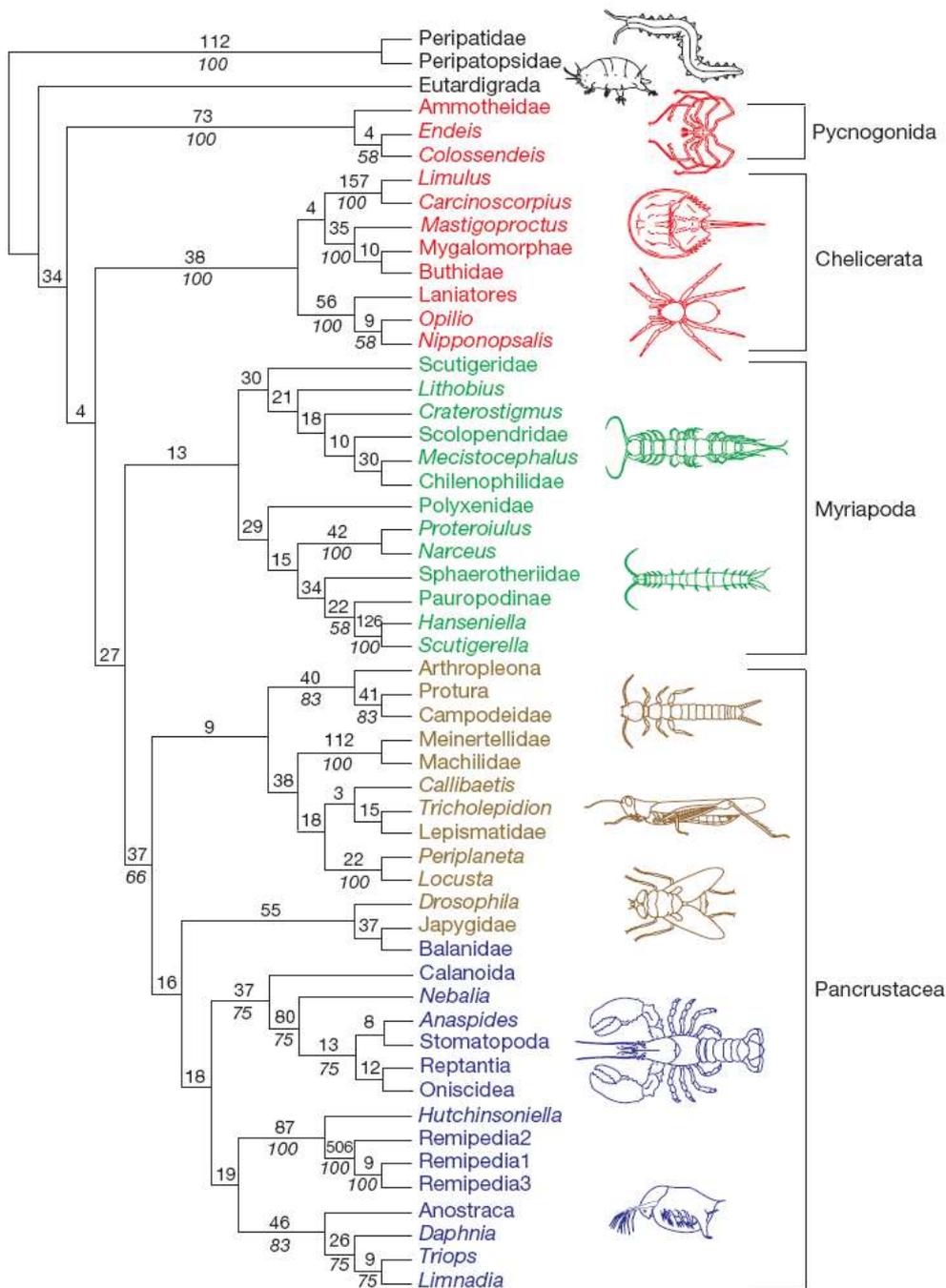
[Wiki](#)

[Links](#)

MrBayes is a program for the Bayesian estimation of phylogeny. Bayesian inference of phylogeny is based upon a quantity called the posterior probability distribution of trees, which is the probability of a tree conditioned on the observations. The conditioning is accomplished using Bayes's theorem. The posterior probability distribution of trees is impossible to calculate analytically; instead, MrBayes uses a simulation technique called Markov chain Monte Carlo (or MCMC) to approximate the posterior probabilities of trees.

The program takes as input a character matrix in a NEXUS file format. The output is several files with the parameters that were sampled by the MCMC algorithm. MrBayes can summarize the information in these files for the user. The program features include:





Arthropod phylogeny based on eight molecular loci and morphology

Ignacio Giribet*, Gregory D. Edgecombe† & Ward C. Wheeler‡

*Department of Organismic and Evolutionary Biology, Harvard University, Divinity Avenue, Cambridge, Massachusetts 02138, USA
 †Australian Museum, 6 College Street, Sydney, New South Wales 2010, Australia
 ‡Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024, USA

Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences

Jerome C. Regier¹, Jeffrey W. Shultz^{1,2,3}, Andreas Zwick¹, April Hussey¹, Bernard Ball¹, Regina Wetzler², Joel W. Martin² & Clifford W. Cunningham¹

NATURE

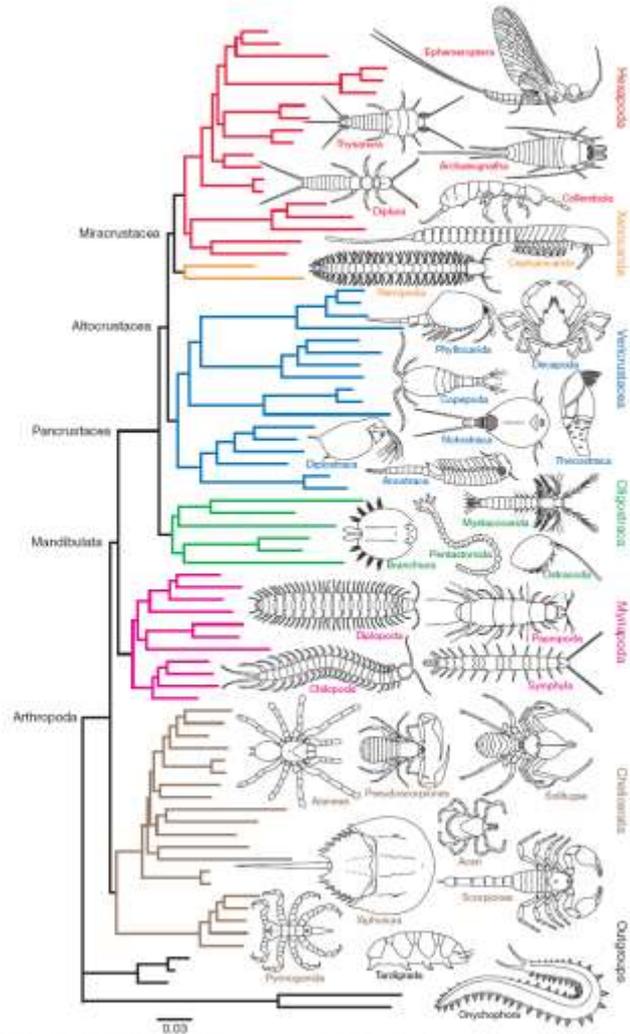


Figure 2 | Phylogram of relationships for 75 arthropod and five outgroup species. Based on likelihood analyses of 62 nuclear protein-coding genes.

In Fig. 1, line drawings of representatives of the *tr* show the morphological disparity across Arthropoda

degen1 BP noLRall1+nt2 BP
codon BP amino-acid BP

x: x% = BP and node recovered in MLA
-: BP < 50% and node not recovered in MLA
[53]: BP = 53% and node not recovered in MLA

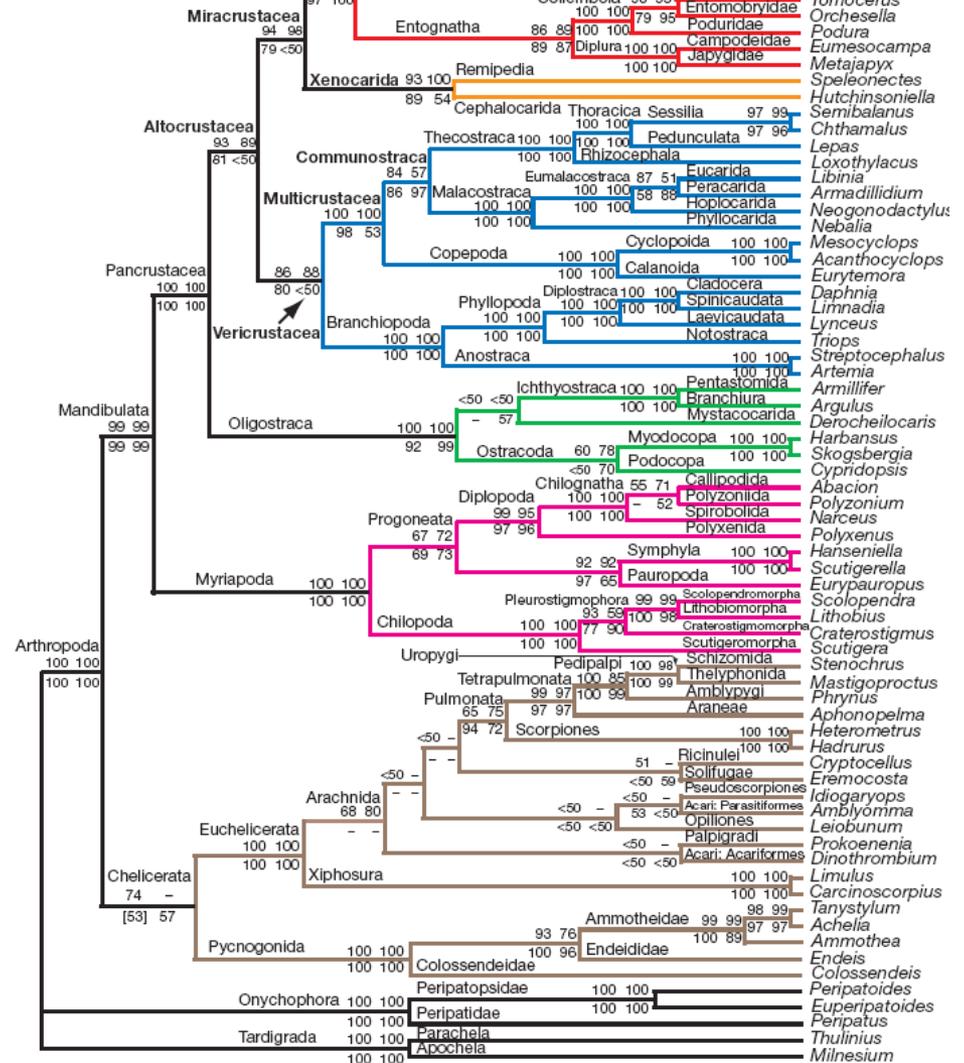
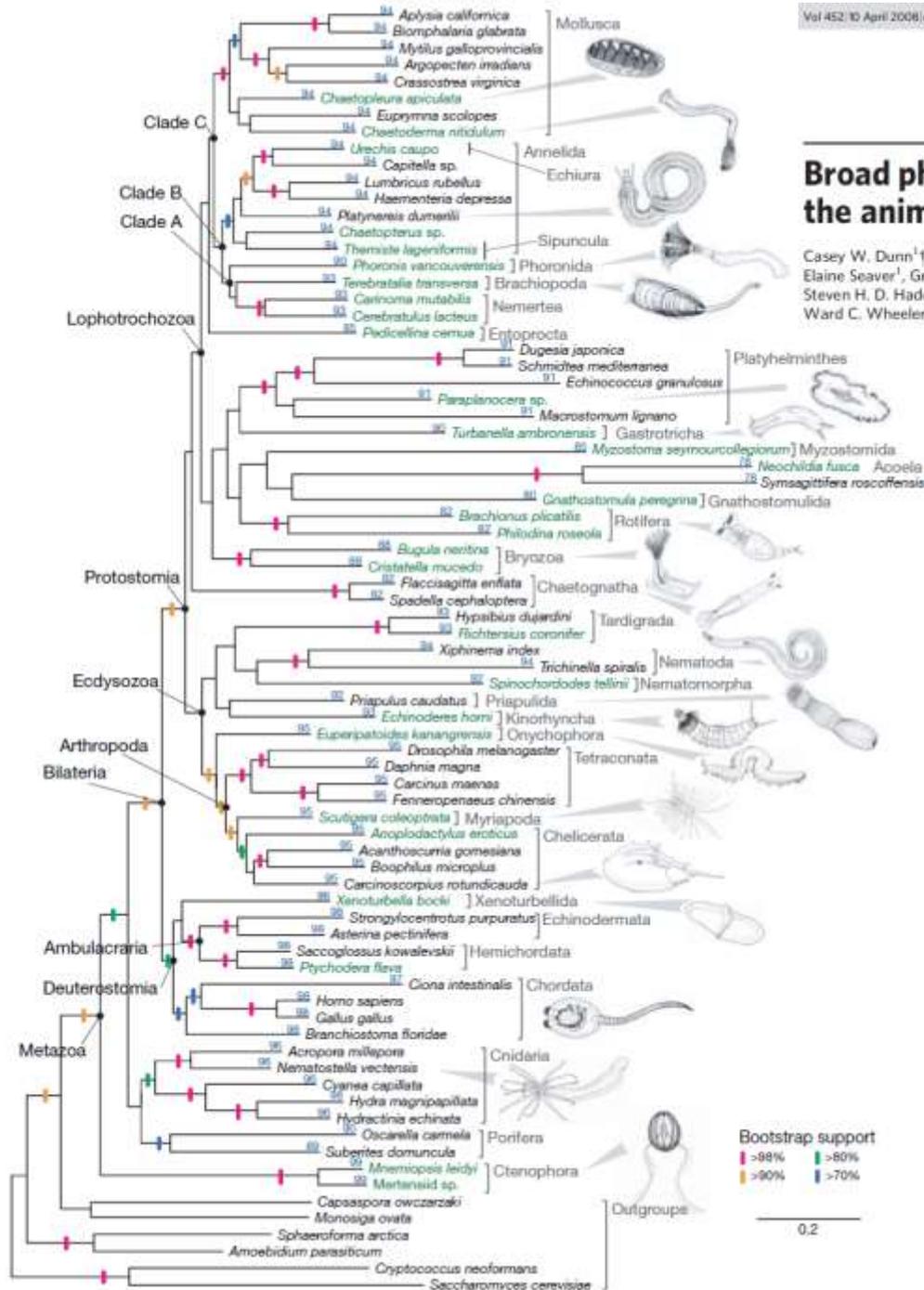


Figure 1 | Phylogenetic relationships of 75 arthropod and five outgroup species. Aligned sequences for 62 nuclear protein-coding genes were

leucine or arginine codons¹⁶; codon model, by nucleotides are analysed directly under a model

Broad phylogenomic sampling improves resolution of the animal tree of life

Casey W. Dunn^{1†}, Andreas Hejnol¹, David Q. Matus¹, Kevin Pang¹, William E. Browne¹, Stephen A. Smith², Elaine Seaver¹, Greg W. Rouse², Matthias Obst¹, Gregory D. Edgecombe³, Martin V. Sørensen⁴, Steven H. D. Haddock⁵, Andreas Schmidt-Rhaesa⁶, Akiko Okusu⁶, Reinhardt Møbjerg Kristensen¹⁰, Ward C. Wheeler¹¹, Mark Q. Martindale¹ & Gonzalo Giribet^{12,13}



THE NEW VIEW OF ANIMAL PHYLOGENY

Kenneth M. Halanych

Department of Biological Sciences, Auburn University, Auburn,

Annu. Rev. Ecol. Evol. Syst. 2004. 35:229–56

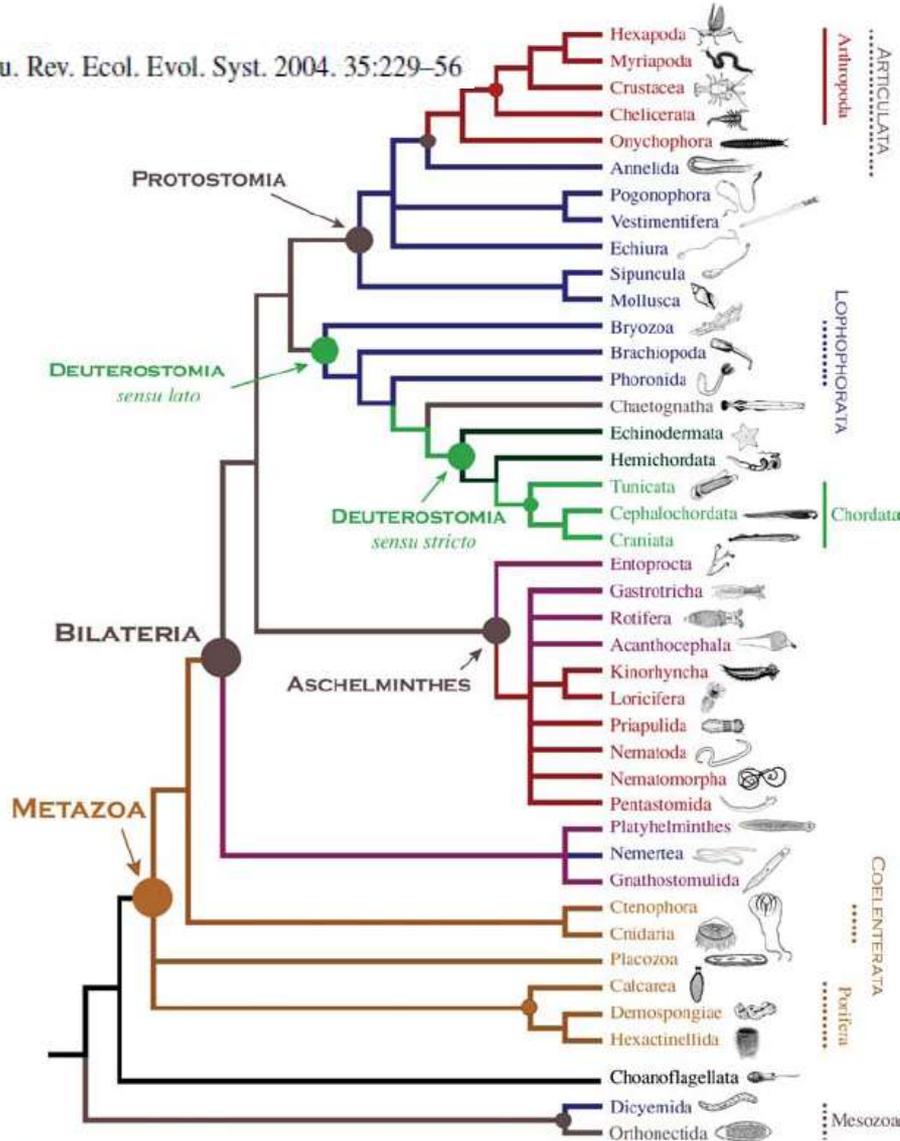


Figure 1 Traditional synthesis. The understanding of metazoan evolution prior to Field et al. (1988). This tree is drawn to illustrate major concepts. The tree is color

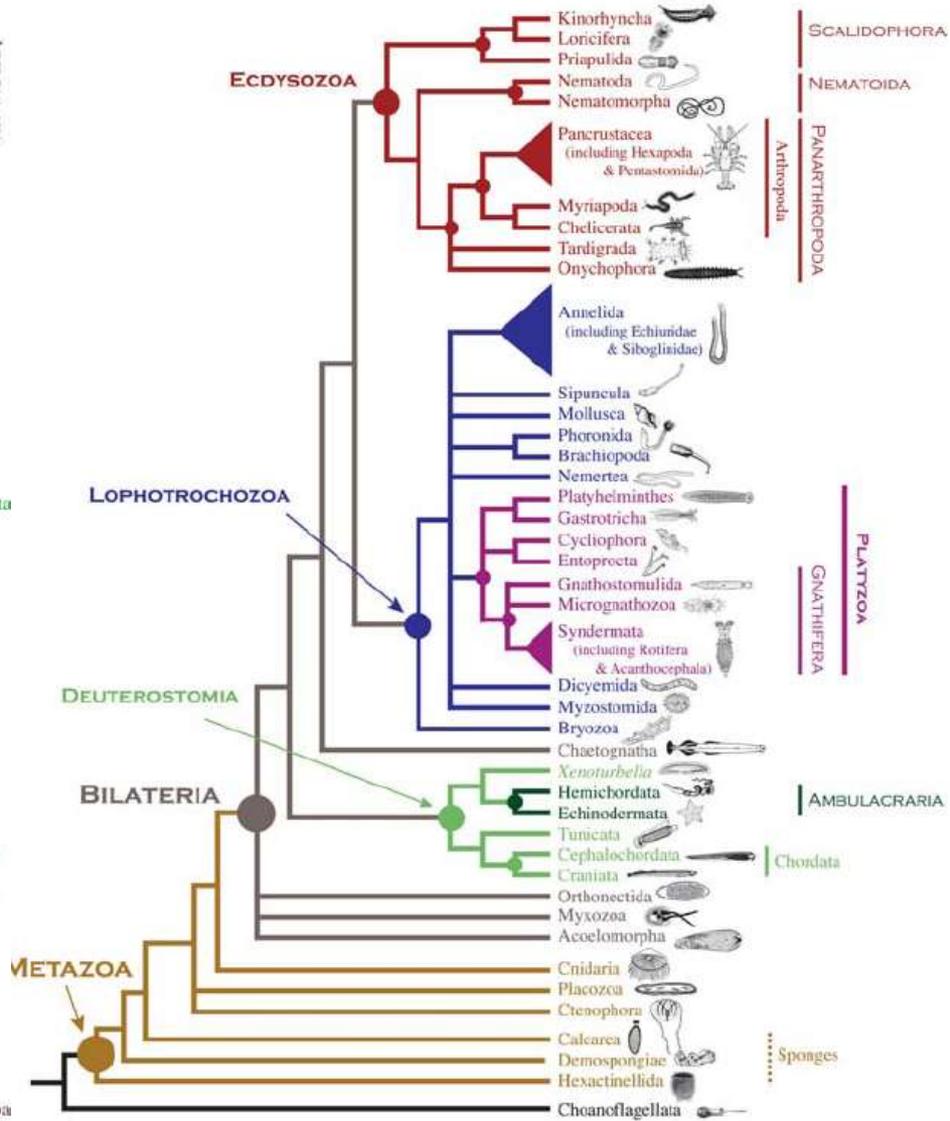


Figure 2 Modern synthesis. The new view of animal phylogeny based largely on molecular data. Details and support for various clades are discussed in the text. This

LETTERS

Non-random decay of chordate characters causes bias in fossil interpretation

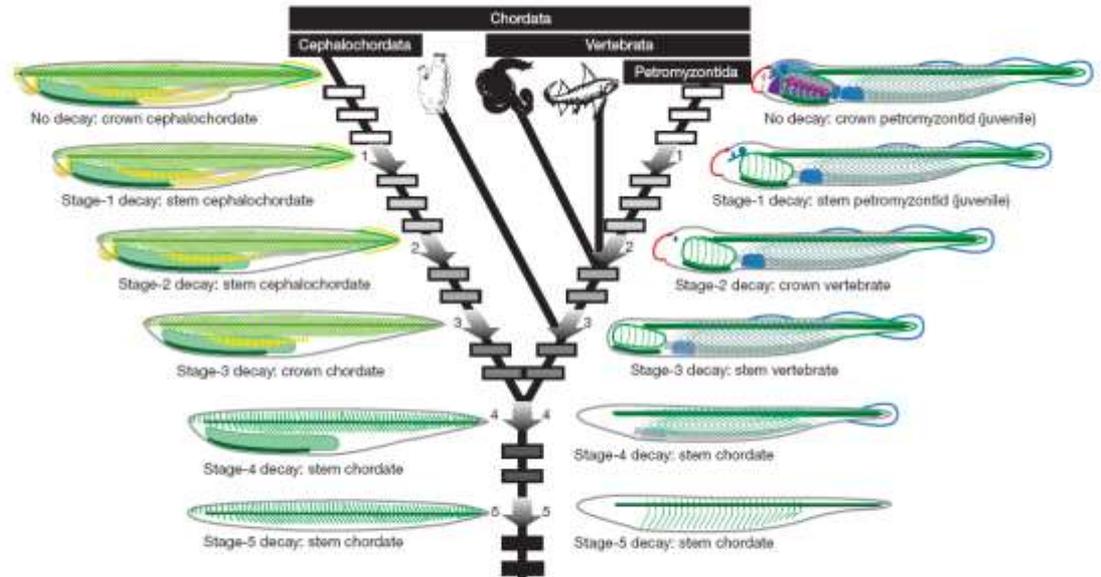
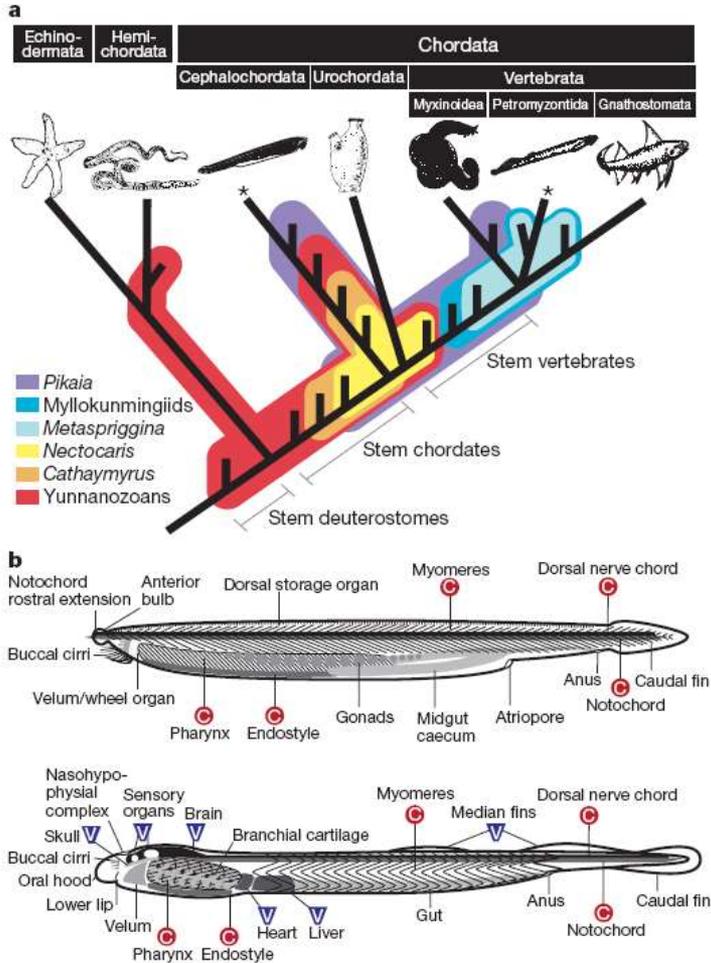


Figure 3 | Morphological decay stages of *Branchiostoma* (left) and larval *Lampetra* (right) and the phylogenetic position of each stage if interpreted as a fossil. Rectangles on branches of the phylogeny are morphological

taphonomic bias exists in identification of fossil chordates. Characters are colour coded according to the hierarchical level for which they are informative (green, chordate; yellow, cephalochordate; blue, vertebrate;