

## О ПОЛОЖЕНИИ НАСЕКОМЫХ В КЛАДЕ PANCRUSTACEA

© 2009 г. В. В. Алёшин<sup>1\*</sup>, К. В. Михайлов<sup>2</sup>, А. В. Константинова<sup>1</sup>, М. А. Никитин<sup>1</sup>,  
Л. Ю. Русин<sup>3,4</sup>, Д. А. Буйнова<sup>5</sup>, О. С. Кедрова<sup>4</sup>, Н. Б. Петров<sup>1</sup>

<sup>1</sup>Научно-исследовательский институт физико-химической биологии им. А.Н. Белозерского Московского государственного университета им. М.В. Ломоносова, Москва, 119991

<sup>2</sup>Факультет биоинженерии и биоинформатики Московского государственного университета им. М.В. Ломоносова, Москва, 119991

<sup>3</sup>Институт проблем передачи информации им. А.А. Харкевича Российской академии наук, Москва, 127994

<sup>4</sup>Лаборатория геносистематики биологического факультета Московского государственного университета им. М.В. Ломоносова, Москва, 119991

Российский научный центр рентгенодиагностики Росмедтехнологий, Москва, 117997

Поступила в редакцию 04.02.2009 г.

Принята к печати 14.03.2009 г.

Современные взгляды на филогению членистоногих (Arthropoda) расходятся с традиционной системой, которая выделяла в их составе четыре независимых класса – хелицеровые, ракообразные, многоножки и насекомые. Имеются веские основания рассматривать насекомых как одну из филогенетических линий в составе единой клады вместе с ракообразными. Этой клade было присвоено наименование Pancrustacea, или Tetraconata. Однако вопрос, с какой именно группой ракообразных насекомые связаны наиболее тесным родством, остается открытым. В новейших филогенетических деревьях, построенных по последовательностям большого числа генов, насекомые объединяются с “нижними” жаброногими ракообразными Branchiopoda. До сих пор этому результату не придавали большого значения, полагая его артефактом притяжения длинных ветвей. Мы проанализировали совокупные данные по 77 рибосомным белкам, фактору элонгации 1A (EF1A), фактору инициации 5A (eIF5A), а также некоторым другим ядерным и митохондриальным белкам. Анализ ядерных генов подтверждает монофилию Hexapoda, т.е. таксона, объединяющего эктогнатных и эндогнатных насекомых. Гипотеза монофилии Hexapoda и Branchiopoda получает поддержку в большинстве сравнений. Maxillopoda, другой таксон “Entomostraca”, является сестринской группой относительно Hexapoda + Branchiopoda. “Высшие” ракообразные, т.е. Malacostraca, в большинстве из сконструированных деревьев оказываются более ранней ветвью Pancrustacea. Обнаружены молекулярные синапоморфии Hexapoda + Branchiopoda + Maxillopoda в сайтах, отличающихся низким уровнем гомоплазий, и синапоморфии Malacostraca (включая Phyllocarida), в отношении которых Hexapoda и “Entomostraca” плезиоморфны. Таким образом, монофилия Hexapoda и Branchiopoda и их принадлежность к Entomostraca кажутся, скорее, отражением реальных филогенетических отношений, чем вычислительным артефактом.

**Ключевые слова:** филогения, молекулярная эволюция, кладистика, белки EF1A, eIF5A, RpS28e, Arthropoda, Crustacea, Insecta.

ON THE PHYLOGENETIC POSITION OF HEXAPODA WITHIN THE PANCRUSTACEA, by V. V. Aleoshin<sup>1\*</sup>, K. V. Mikhailov<sup>2</sup>, A. V. Konstantinova<sup>1</sup>, M. A. Nikitin<sup>2</sup>, L. Yu. Rusin<sup>3,4</sup>, D. A. Buianova, O.S. Kedrova<sup>4</sup>, N. B. Petrov<sup>1</sup> (<sup>1</sup>Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, 119991 Russia, \*e-mail: Aleshin@genebee.msu.su; <sup>2</sup>Department of Bioengineering and Bioinformatics, Moscow State University, Moscow, 119991 Russia; <sup>3</sup>Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, 127994 Russia; <sup>4</sup>Department of Biology, Moscow State University, Moscow, 119991 Russia; <sup>5</sup>Russian Scientific Center of Roentgenoradiology, 117837 Russia). Contemporary views on the phylogeny of arthropods are at odds with the traditional system, which recognizes four independent arthropod classes: Chelicerata, Crustacea, Myriapoda and Insecta. There is compelling evidence that insects in fact comprise a monophyletic lineage with Crustacea within a larger clade of Pancrustacea (=Tetraconata). Which crustacean group is the closest living relative of insects remains an open question. Recent phylogenetic analyses based on multiple genes suggest their sistership with “lower” crustaceans, the Bran-

\* Эл. почта: Aleshin@genebee.msu.su

chiopoda. This relationship was often impeached to be caused by the long branch attraction artifact. We analyzed concatenated data on 77 ribosomal proteins, elongation factor 1 alpha (EF1A), initiation factor 5 alpha (eIF5A) and other selected nuclear and mitochondrial proteins. Nuclear protein data supports the monophyly of Hexapoda, the clade uniting entognath and ectognath insects. Hexapoda and Branchiopoda comprise a monophyletic lineage in most analyses. Maxillopoda occupies the sister position to the Hexapoda + Branchiopoda. "Higher" crustaceans, the Malacostraca, in most reconstructions comprise a more basal lineage within the Pancrustacea. Molecular synapomorphies in low homoplastic regions are found for the clades Hexapoda + Branchiopoda + Maxillopoda and the monophyletic Malacostraca containing the Phyllocarida. Therefore, the sistership of Hexapoda and Branchiopoda and their position within Entomostraca may in fact represent bona fide phylogenetic relationships.

**Key words:** phylogeny, molecular evolution, cladistics, EF1A, eIF5A, RpS28e, Arthropoda, Crustacea, Insecta.

Насекомые – самая многочисленная группа живых существ на Земле. По этой причине шутка о том, что все животные – это, в первом приближении, насекомые [1], иногда кажется вполне серьезной. При этом филогенетические связи насекомых с другими классами членистоногих остаются неопределенными [2]. Традиционное сближение насекомых с многоножками в рамках одного подтипа под названием трахейных (Tracheata), или неполноусых (Atelocerata) [1–3], не получает поддержки при анализе последовательностей ДНК, по которому предпочтительнее выглядит гипотеза о более тесном родстве насекомых с ракообразными [4–14] в составе единой клады, названной Pancrustacea [15], или Tetracnata [16]. Новая филогенетическая гипотеза поддерживается некоторым числом признаков анатомии и развития нервной системы [7, 13–22], сходных у насекомых и ракообразных, но их синпоморфный статус остается спорным [23]. К настоящему времени гипотеза Pancrustacea еще не полностью победила традиционную точку зрения, которая пока не утратила симпатий зоологов [24, 25].

Если гипотеза Pancrustacea справедлива, то возникает новый вопрос [17], каким именно ракообразным родственнее насекомые. Если же насекомые ответвились от стволовой группы ракообразных, то возникает необходимость получить доказательства этого (другими словами, требуются доказательства монофилии современных ракообразных). До недавнего времени для основательного подхода к этому вопросу не было достаточно ни зоологических, ни молекулярных данных [24]. Масштабное секвенирование клонотек кДНК, которое стали применять в последние годы для изучения широкого набора организмов, а не только "модельных", быстро меняет ситуацию. Два года назад опубликовано первое филогенетическое дерево, основанное на действительно большом числе генов (133 предсказанных белка, более 30 тыс. аминокислотных остатков) [26], на котором две группы ракообразных – десятиногие и ветвистоусые – отходят от общего с насекомыми ствола двумя независимыми ветвями; другими словами, ракообразные оказываются парафилетичными относительно насекомых. Однако этот примечательный факт остался без обсуждения, поскольку

задачей упомянутой работы [26] было исследование систематических ошибок реконструкции филогении, возникающих даже при использовании столь обширных наборов молекулярных данных, а элемент дерева, правильность которого достоверно неизвестна (родственные отношения внутри Pancrustacea), не вызывал интереса в рамках избранной темы. С тех пор значительно увеличилась таксономическая выборка членистоногих, изученных по клонотекам кДНК. Это позволяет вновь рассмотреть вопрос о родстве насекомых и ракообразных и попытаться понять, действительно ли ракообразные парафилетичны относительно насекомых или это артефакт конструирования, подобный притяжению длинных ветвей.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

**Нуклеотидные последовательности**, в том числе кДНК, членистоногих и представителей внешней группы получены из базы данных GenBank, а также из необработанных хроматограмм (Trace Archive) NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) и ресурсов NEMBASE [27] (<http://www.nematodes.org>). Для поиска ортологов применяли программу BLAST [28]. Отобранные последовательности виртуально транслировали согласно таблице универсального генетического кода и выравнивали с помощью программы MUSCLE [29]. Выравнивание корректировали вручную в графическом редакторе BioEdit [30]. Нуклеотидные последовательности использовали в дальнейшем только в том случае, если возникала необходимость исправления рамки считывания в консервативных областях предсказанных белков. Ошибочно аннотированные последовательности фильтровали, анализируя дистанционные деревья, построенные для каждого семейства с помощью программ TREEFINDER и SEMPHY [31, 32]. Выравнивания по отдельным семействам конкатенировали с помощью программы ScaFoS [33], после чего удаляли гипервариабельные участки, однозначное выравнивание которых затруднено. Из близких видов, данные по каждому из которых неполны, были образованы химерные операциональные таксономические единицы: Peracarida (включающая последовательности

бокоплавов *Gammarus pulex*, *Parhyale hawaiiensis* и изоподы *Eurydice pulchra* и *Onychophora (Epipeperipatus sp. и Euperipatoides kanangrensis)*.

**На стадии отбора данных** мы преследовали две цели: заполнить матрицу с наименьшим числом пропусков и исключить паралоги. Для достижения обеих целей удобно использовать рибосомные белки. Они легко классифицируются на группы, паралоги в каждой группе не очень распространены и, как правило, легко выявляются; количество мРНК рибосомных белков в клетках очень велико, поэтому копии хорошо представлены в клонотеках кДНК, и по этой причине известны их последовательности у многих видов. Мы конкатенировали 77 рибосомных белков – почти полный набор “типичной” эукариотической рибосомы, за исключением короткого белка L41 и белков P0, P1 и P2. Ортологичность других, кодируемых ядром, белков проверяли с помощью однонаправленного поиска программой BLAST относительно полностью секвенированных геномов. Выравненные последовательности митохондриальных белков получали из базы данных NCBI ([www.ncbi.nlm.nih.gov/genomes/ORGANELLES/organelles.html](http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/organelles.html)) или OGRE (<http://drake.physics.mcmaster.ca/ogre/index.shtml>) [34]. Их выравнивание также корректировали вручную. Высокоизменчивые области удаляли с помощью программы Gblocks [35]. Для дальнейшего анализа были взяты все предсказанные белки митохондриального генома, за исключением переменных ATP6, ATP8 и NAD6.

**Для предварительного анализа** больших наборов последовательностей методом максимального правдоподобия использовали программу Phylml [36], для более тщательного анализа – программу MrBayes 3.1.2 [37]. Оптимальную матрицу аминокислотных замещений подбирали предварительно с помощью программы ModelGenerator [38] на платформе распределенных филогенетических вычислений [39] или же использовали смешанную модель в программе MrBayes 3.1.2. Все параметры, кроме топологии и длин ветвей, для всех белков в конкатенированном выравнивании вычисляли независимо (функция *partition*). Поиск потенциальных синапоморфий проводили в полуавтоматическом режиме. На первом этапе с использованием программы *protpars* пакета PHYLIP [40] выводили предсказанные состояния в узлах согласно заданному дереву, затем отбирали изменения, которые произошли в сайтах, характеризующихся низким уровнем гомопластичных замен.

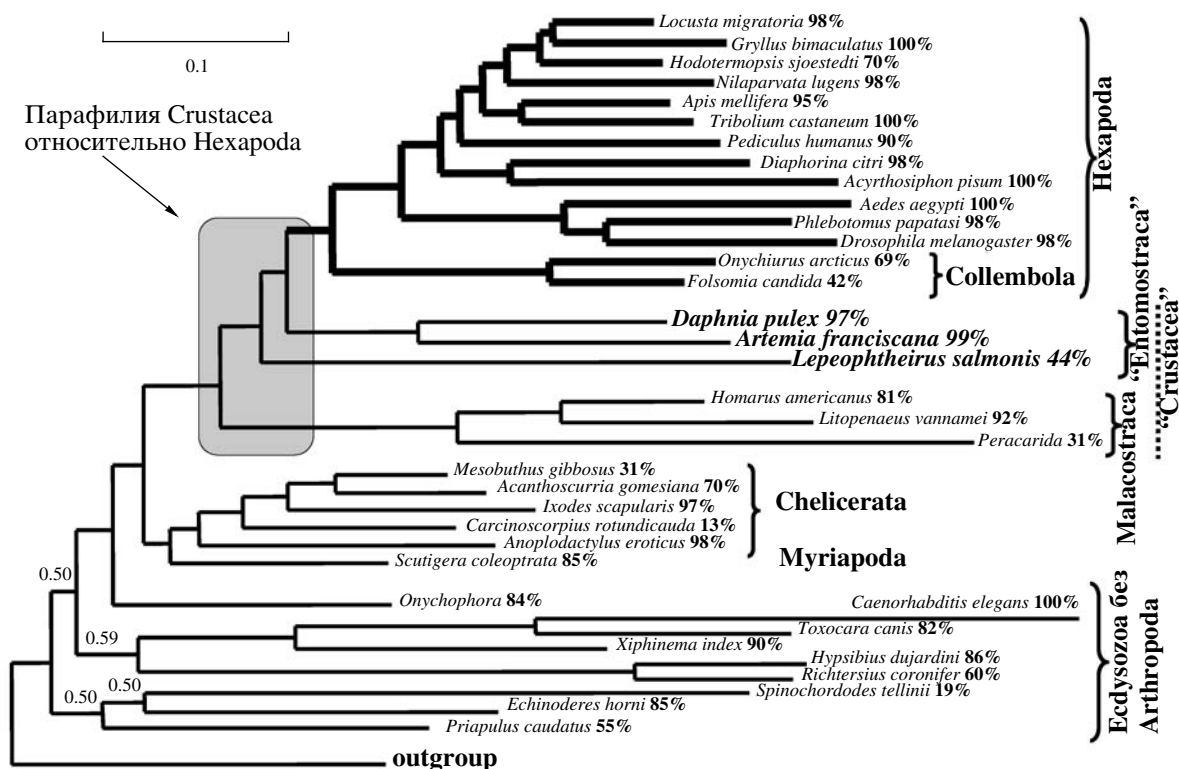
**Альтернативные топологии** (после их конструирования в редакторе TreeView [41]) сравнивали с помощью программы TREE-PUZZLE 5.2 [42] и статистического AU-теста [43] по программе CONSEL [44]. Для графического представления дерева митохондриальных белков использовали программу Treeconw [45].

## РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

### Полный набор рибосомных белков

Анализируемый набор включал выравненные и конкатенированные аминокислотные последовательности 77 рибосомных белков. После удаления переменных участков, выравнивание которых сомнительно, длина конкатената составила 11349 позиций. Степень полноты матрицы показана на рис. 1 как процент заполненных позиций выравнивания для каждой операциональной таксономической единицы. Отдельные рибосомные белки сильно отличаются по степени консервативности: расчетная доля инвариантных позиций варьирует от 0.02 (RpS12, RpL30, RpL18) до 0.19–0.21 (RpL13, RpL11, RpL10). Они отличаются также по разнице в скорости эволюции между сайтами:  $\alpha$ -параметр  $\Gamma$ -распределения варьирует в нашем наборе (содержащем представительную внешнюю группу) от 0.17–0.38 (RpS28, RpL40, RpS14, RpS23, RpS9, RpS5) до 1.31–1.49 (RpL28, RpS12, RpS19 RpL24, RpL24-like). Для большинства белков (54 из 77) характер аминокислотных замен наилучшим образом описывается моделью rtREV [46], для 16 белков – WAG [47], для 5 – JTT [48]; для двух белков модели rtREV и WAG примерно равноценны. Продолжительное вычисление с использованием процедуры Монте-Карло ( $n_{gen} = 10000000$ ), сопряженной в программе MrBayes с Марковскими цепями, не привело к сходимости в двух параллельных запусках. В каждом из них произошла стабилизация альтернативной топологии, что приводит к апостериорной вероятности 0.5 для некоторых узлов дерева. Однако эти отличия распространяются только на внешние относительно членистоногих группы (рис. 1). Насекомые на полученном дереве образуют с ногохвостками единую кладу, Hexaroda, которая, в свою очередь, объединяется с жаброногими ракообразными (Branchioroda). Такой результат не является совершенно новым. Гипотеза о монофилии насекомых и жаброногих ракообразных высказывалась и раньше [11, 14, 49, 50]. Несколько дальше от насекомых на дереве расположена ветвь веслоногих, единственной богато представленной в базе данных кДНК группы ракообразных, относимых к Maxillopoda. Наконец, от узла базальной дивергенции Pancrustacea отходит ветвь “высших раков”, Malacostraca. Все перечисленные группировки имеют апостериорную вероятность 1.0.

Приняв за основу дерево, сконструированное программой MrBayes (рис. 1), мы протестировали 63 альтернативные топологии, которые отличаются положением насекомых (включая ногохвосток) и многоножек (представленных в нашем наборе единственным видом). Разбиение выравнивания на отдельные белки было отменено, и конкатенированное выравнивание рассматривали как единую последовательность. Статистическую достоверность различий топологий вычисляли для



**Рис. 1.** Байесово дерево конкатенированных последовательностей 77 рибосомных белков. Общая длина выравнивания после удаления участков, трудных для выравнивания, – 11349 аминокислотных остатков. Процент заполненных позиций в выравнивании указан после названий операциональных таксономических единиц. Приведена апостериорная вероятность, отличная от 1.0, по результатам двух независимых запусков. Параметры MrBayes 3.1.2: nruns = 2, nchains = 4, rates = invgamma, ngammacat = 8, aamodelpr = mixed, ngen = 10000000, burnin = 5000000, partition = by\_gene, partition by\_gene = 77, unlink statefreq = (all) shape = (all) pinvar = (all) aamodel = (all). Используемые химеры описаны в тексте. Жирными линиями показаны ребра, относящиеся к насекомым и ногохвосткам. Полу жирным шрифтом набраны видовые названия "Entomostraca" (Branchiopoda и Maxillopoda).

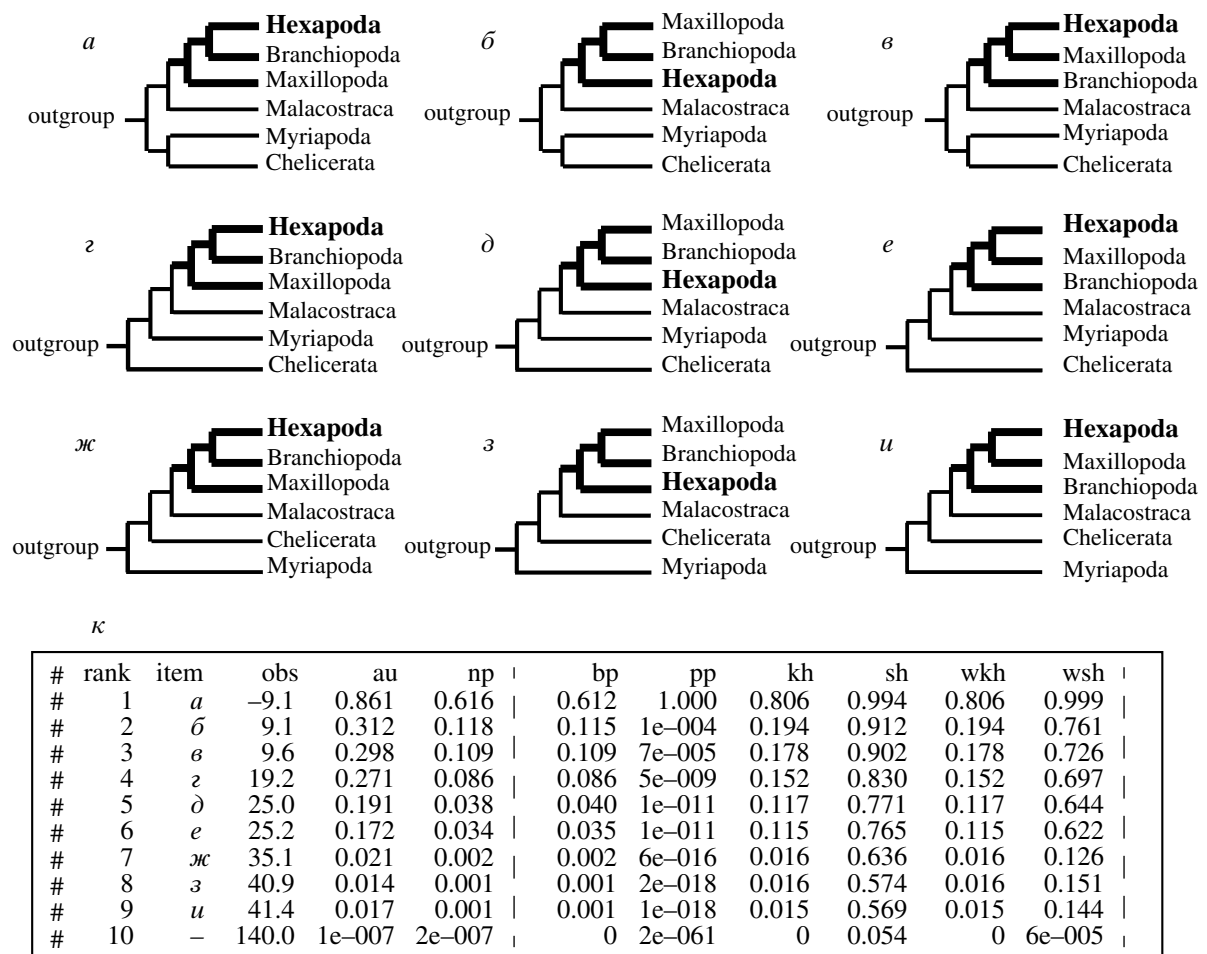
Г-распределения, аппроксимируемого восемью категориями и инвариантными позициями, при WAG-модели аминокислотных замещений. Наибольшее значение правдоподобия получила топология, признанная лучшей по результатам байесова анализа. Всего шесть из 63 альтернативных топологий преодолели 5%-ный порог значимости, согласно AU-тесту [43] и тесту Кишины-Хасегавы [51], и девять топологий – согласно тесту Шимодаиры-Хасегавы [51].

Эти филограммы представлены на рис. 2. Все они обладают одной особенностью: в них насекомые объединены тем или иным образом с "низшими" ракообразными "Entomostraca" и удалены от многоножек. Многоножки расположены на дереве вне Pancrustacea, т.е. таксон "трахейные" ("Tracheata"), или "неполноусые" ("Atelocerata"), оказывается полифилетическим. На "лучшем" дереве многоножки объединяются с хелицеровыми, тем не менее их филогенетическое положение остается не вполне определенным, поскольку имеются два альтернативных варианта, которые незначительно отличаются от "лучшей" топологии, согласно использованным

статистическим тестам: если многоножки включены в состав Mandibulata (рис. 2d–e) и если помещены в корень членистоногих (рис. 2ж–и). Это согласуется с другими работами, в которых положение многоножек признается не установленным [10, 12, 52].

### Синапоморфии в отдельных генах

**1. Белок RpS28.** Очевидно, что за объединением насекомых не просто с ракообразными, а, более точно, с "Entomostraca" (рис. 1 и рис. 2), стоят вполне определенные черты сходства их рибосомных белков. Однако результаты предыдущего раздела не позволяют ответить на вопрос, унаследованы ли эти сходные состояния признаков от удаленного общего предка, т.е. являются симплезиоморфными, или происходят от ближайшего общего предка и являются истинными синапоморфными индикаторами родства, или возникли независимо вследствие гомопластического изменения генетического материала. Между тем в логике филогенетического анализа симплезиоморфным и синапоморфным сходствам придается совершенно разное значение (для обзора см. [53, 54]). Чтобы провести оценку общих призна-



**Рис. 2.** Альтернативные топологии для конкатенированного набора 77 рибосомных белков, незначимо отличающиеся согласно статистическим тестам AU, KH и/или SH. Деревья для анализа сконструированы на основе байесова дерева (рис. 1) путем перемещения насекомых и многоножек относительно всех выделенных операциональных таксономических единиц членистоногих. Правдоподобие по сайтам вычислено программой TREE-PUZZLE 5.2 для модели WAG + I + Г для 8 категорий. *a–и* – 9 лучших филограм, согласно ранжированию по результатам AU-теста; *к* – первые 10 строк выходного файла CONSEL с оценкой 10 лучших топологий из 63 проверенных. Жирными линиями показаны ребра, объединяющие насекомых и “Entomostraca”.

ков насекомых и “Entomostraca” с позиций кладистики, мы вывели, используя программу *protpars*, конкретные признаки, поддерживающие эту группу. В первую очередь нас интересовали сходства в консервативных позициях, замены в которых вообще редки, а также замены в сайтах с переменной скоростью эволюции.

RpS28 – небольшой, относительно консервативный рибосомный белок. Вычисленная доля инвариантных сайтов для нашего набора RpS28 равна 0.089; это значение не сильно отличается от среднего значения для рибосомных белков. Однако RpS28 выделяется среди остальных рибосомных белков самым высоким уровнем гетерогенности сайтов по скорости эволюции: значение  $\alpha$ -параметра Г-распределения для него минимально (во всем наборе) и составляет 0.169. Иными словами, в составе RpS28

имеются и очень консервативные сайты (но не инвариантные, число которых не сильно превышает “среднее”, см. выше), и очень вариабельные сайты. Это сочетание весьма выгодно для последовательностей, выбираемых для филогенетического анализа.

На рис. 3 приведен небольшой (по числу видов) фрагмент выравнивания белка RpS28, где выделены три синапоморфные замены у Hexapoda и “Entomostraca”. Выборка видов “Entomostraca” с известными последовательностями RpS28 в настоящий момент исчерпывается приведенными (два вида Branchiopoda и один вид Maxillopoda), но число видов насекомых, “высших” раков Malacostraca и представителей внешней группы (хелицероных и других беспозвоночных, не являющихся членистоногими), у которых последовательность этого белка определена,

CHAOHOFLAGELLATA ( <i>Monosiga brevicollis</i> )	MDA-----PPKLAQVTKVGLGRTGSGGQAIQVRVEFI-DDT-QRSILRNKVGPPVRVGDILVLMSESEARRMR*
PORIFERA ( <i>Amphimedon queenslandica</i> )	MDQ-K---P-RYGVKIKLGLRTGSGGQCTQVRVEFL-DDS--RSIIRNVKGPVREGDVLCLLESEARRRLL*
PORIFERA ( <i>Suberites domuncula</i> )	MD--K---P-RYAKVIKILGRTGSGGQATQVRVEFL-DDG--RAIIRNVKGPVREGDVLCLLESEARRRLL*
CNIDARIA ( <i>Acropora millepora</i> )	MD--K---PVLKAVTKVGLGRTGSGGQCTQVRVEFL-DDT-NRSIIRNVKGPVREGDILTLLESEARRRLL*
CHORDATA ( <i>Homo sapiens</i> )	MDTSR-VQPIKLAQVTKVGLGRTGSGGQCTQVRVEFL-DDT-SRSIIRNVKGPVREGDVLTLLESEARRRLL*
ECHINODERMATA ( <i>Patria pectinifera</i> )	MD--K---PVLKAVTKVGLGRTGSGGQCTQVRVEFL-DDS-NRSIIRNVKGPVREGDILTLLESEARRRLL*
HEMICHORDATA ( <i>Saccoglossus kowalevskii</i> )	MD--K---PIKLAQVTKVGLGRTGSGGQCTQVRVEFL-DDS-NRSIIRNVKGPVREGDVLTLLESEARRRLL*
ANNELIDA ( <i>Hirudo medicinalis</i> )	MDA-K-AAPIKLQVTKVGLGRTGSGGQCTQVRVEFI-DET-NRSIIRNVKGPVREGDILTLLESEARRRLL*
ANNELIDA ( <i>Lumbricus rubellus</i> )	MDA-K-AVPIKLAQVTKVGLGRTGSGGQCTQVRVEFL-DES-SRSIIRNVKGPVREGDILTLLESEARRRLL*
ANNELIDA ( <i>Arenicola marina</i> )	MEA-K-IAPIKLAQVTKVGLGRTGSGGQCTQVRVEFL-DDS-NRLIIRNVKGPVREGDILTLLESEARRRLL*
NEMERTINI ( <i>Cerebratulus lacteus</i> )	MDA-K-LQPLKLAQVIRILGRTGSGGQCTQVRVEFL-DDT-NRSIIRNVKGPVREGDILTLLESEARRRLL*
BRACHIOPODA ( <i>Terebratalia transversa</i> )	MD--K-LAPVLAQVTKVGLGRTGSGGQCTQVRVEFL-DDS-NRSIIRNVKGPVREGDILTLLESEARRRLL*
MOLLUSCA ( <i>Crassostrea virginica</i> )	MD--K-LQPVLAQVTKVGLGRTGSGGQCTQVRVEFL-DDS-NRSIIRNVKGPVREGDILTLLESEARRRLL*
MOLLUSCA ( <i>Aplysia californica</i> )	MD--K-LQPVLAQVTKVGLGRTGSGGQATQVRVEFL-DDS-NRSIIRNVKGPVREGDILTLLESEARRRLL*
PLATYHELMINTHES ( <i>Schistosoma japonicum</i> )	MD--KAVVQTKLAQVAVVGRGSGGQCTQVRVEFL-DDT-NRTIIRNVKGPVREGDILTLLETEREARRRLL*
ROTIFERA ( <i>Philodina roseola</i> )	MD--KATAGGQVAVVAVVVKILGRTGSGGQCTQVRVEFI-DSSSKRSIIRNVKGPVREGDVLTLLESEARRRLL*
PRIAPULIDA ( <i>Priapulius caudatus</i> )	MEA-K-TAPIRLARVNVKGLGRTGSGGQCTQVRVEFL-DET-TRSIIRNVKGPVREGDILTLLESEARRRLL*
KINORHYNCA ( <i>Echinoderes horni</i> )	MD--K-SQPIT-AKVTKVGLGRTGSGGQCTQVRVELL-DSS-SRSIIRNVKGPVREGDILTLLESEARRRLL*
NEMATOMORPHA ( <i>Spiniochordodes tellinii</i> )	MD--K---PTLVRVEKLLGRTGSGGQCTQVRVEFL-NRSLIIRNVKGPVREGDMLTLLETEREARRRLL*
NEMATODA ( <i>Caenorhabditis elegans</i> )	MD--K---LTLARVTKVGLGRTGSGGQCTQVRVEFI-NDQNNRSIIRNVKGPVREGDILTLLESEARRRLL*
NEMATODA ( <i>Xiphinema index</i> )	MD--K---PVQKLAQVTKVGLGRTGSGGQCTQVRVEFL-DDP-KRSIIRNVKGPVREGDILTLLESEARRRLL*
TARDIGRADA ( <i>Hypsibius dujardini</i> )	MD--K---IRLAKVTKVGLGRTGSGGQATQVRVEFL-DES-NRSIIRNVKGPVREGDILTLMESEARRRLL*
TARDIGRADA ( <i>Richtersius coronifer</i> )	MD--K---AKQAVTKVGLGRTGSGGQATQVRVELI-DEQ-SRSIIRNVKGPVREGDILTLMESEARRRLL*
ONYCHOPHORA ( <i>Euperipatoides kanangrensis</i> )	MD--K---VKLARVTKVGLGRTGSGGQCTQVRVEFL-DET-SRSIIRNVKGPVREGDILTLLESEARRRLL*
<b>ARTHROPODA</b>	
PANTOPODA ( <i>Anoplodactylus eroticus</i> )	MD--K---PRLARVVNLLGRTGSGGQCTQVRVEFL-DES-NRTIIRNVKGPVREGDILTLLESEARRRLL*
CHELICERATA ( <i>Mesobuthus gibbosus</i> )	MD--K---VKLARVTKVGLGRTGSGGQCTQVRVEFL-DES-NRSIIRNVKGPVREGDILTLLESEARRRLL*
CHELICERATA ( <i>Acanthoscurria gomesiana</i> )	MDT-Q---VKLAVVVKVGLGRTGSGGQCTQVRVFL-DDQ-NRLIMRNKGPVREGDILTLLESEARRRLL*
CHELICERATA ( <i>Ixodes scapularis</i> )	ME--K---VKLARVTKVGLGRTGSGGQCTQVRVEFL-DET-NRSIIRNVKGPVREGDILTLLESEARRRLL*
CHELICERATA ( <i>Suidasia medanensis</i> )	MD--K---IRLAKVSKVGLGRTGSGGQCTQVRVEFL-DES-YRSIIRNVKGPVREGDILTLLESEARRRLL*
MYRIAPODA ( <i>Scutigera coleoptrata</i> )	MD--K---AKLARVTKVGLGRTGSGGQCTQVRVEFL-DET-NRTIIRNVKGPVREGDILTLLESEARRRLL*
<b>PANCRUSTACEA</b>	
DECAPODA ( <i>Homarus americanus</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-DES-HRSIIRNVKGPVREGDILTLLESEARRRLL*
DECAPODA ( <i>Panulirus japonicus</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-DES-HRSIIRNVKGPVREGDILTLLESEARRRLL*
DECAPODA ( <i>Litopenaeus vannamei</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-DES-HRSIIRNVKGPVREGDILTLLESEARRRLL*
DECAPODA ( <i>Penaeus monodon</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-DES-HRSIIRNVKGPVREGDILTLLESEARRRLL*
EUPHAUSIIDA ( <i>Euphausia superba</i> )	-----PVLKARVXKVLGRTGSGGQCTQVRVEFL-DET-QRSIIRNVKGPVREGDILTLLESEARRRLL*
<b>"ENTOMOSTRACA" + HEXAPODA</b>	
MAXILLOPODA ( <i>Lepeophtheirus salmonis</i> )	MD--K---QVTVARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
BRANCHIOPODA ( <i>Artemia franciscana</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVDFI-GEQ-NRSIIRNVKGPVREGDILTLLESEARRRLL*
BRANCHIOPODA ( <i>Daphnia pulex</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRSIIRNVKGPVREGDILTLLESEARRRLL*
COLLEMBOLA ( <i>Folsomia candida</i> )	ME--R---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
COLLEMBOLA ( <i>Onychiurus arcticus</i> )	ME--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Locusta migratoria</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Gryllus bimaculatus</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Diaphorina citri</i> )	MD--K---QVTVARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRTIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Nilaparvata lugens</i> )	MDN-K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Acyrtosiphon pisum</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Macroneuraphis hirsutus</i> )	ME--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Myzus persicae</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Aedes aegypti</i> )	MD--K---SVLVARVTKVGLGRTGSGGQCTQVRVEFI-SEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Culex pipiens</i> )	MD--K---SVLVARVTKVGLGRTGSGGQCTQVRVEFI-SEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Anopheles gambiae</i> )	MD--K---TFVLAQVTKVGLGRTGSGGQCTQVRVEFL-NEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Chironomus tentans</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-NDP-TRAIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Phlebotomus papatasi</i> )	ME--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEV-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Drosophila melanogaster</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Diabrotica virgifera</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Diaprepes abbreviatus</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Tribolium castaneum</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Apis mellifera</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Solenopsis invicta</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Nasonia vitripennis</i> )	ME--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFL-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Ctenocephalides felis</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GEQ-NRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Bombyx mori</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GET-SRQIIRNVKGPVREGDILTLLESEARRRLL*
INSECTA ( <i>Spodoptera frugiperda</i> )	MD--K---PVLKARVTKVGLGRTGSGGQCTQVRVEFI-GET-SRQIIRNVKGPVREGDILTLLESEARRRLL*

Рис. 3. Фрагмент (по видам) выравнивания рибосомного белка S28. Серым фоном выделены синапоморфии Hexapoda + "Entomostaca". В линии двукрылых инфраотряда Culicomorpha прослеживается последовательное замещение Gly32 → Asn32 → Ser32 (нумерация остатков – по белку RpS28 *Drosophila melanogaster*).

весьма велико. Закономерность изменчивости выделенных сайтов (их апоморфного либо плезиоморфного состояния согласно ожиданиям для группы) сохраняется также и на существенно большей, проверенной нами, выборке. Таким образом, направление эволюции от плезиоморфных Lys6, Agr27 и Asp32 к апоморфным состояниям Val6, Lys27 и Gly32, (нумерация остатков по белку RpS28 *Droso-*

*phila melanogaster*) в последовательностях RpS28 членистоногих можно считать установленным.

**2. Белок eIF5A.** Фактор инициации 5A (eIF5A) – высококонсервативный и жизненно важный белок эукариот и архей с неясными функциями [55]. Подавление экспрессии гена eIF5A с помощью специфических микроРНК снижает общий уровень инициации трансляции более чем на четверть, однако, по-видимому, существенна также регуляторная

CHAOHOFAGELLATA ( <i>Monosiga</i> )	LLVTTIQKAVGEEAAVAVKQLP--K*
PORIFERA ( <i>Amphimedon</i> )	FLVTVVTAMGTEAVMGTKNMK--E*
CNIDARIA ( <i>Nematostella</i> )	FMVTVLKAMGEEETVGVKVMSDT-K*
CNIDARIA ( <i>Hydra</i> )	FLVTVLSA <sup>C</sup> NEEMVVGTKVLT--G*
PLACHOZOA ( <i>Trichoplax</i> )	SLVTTILSAMGEEAAIAVKQMA--K*
CHORDATA ( <i>Homo</i> )	ILITVLSAMTEEA <sup>A</sup> VAIKAMA--K*
HEMICHORDATA ( <i>Saccoglossus</i> )	LLVTVLSAMGEE <sup>C</sup> AI <sup>S</sup> VKNCP--K*
ANNELIDA ( <i>Lumbricus</i> )	VLATVWKAMDEEVVVDLKTAKA--N*
BRACHIOPODA ( <i>Terebratalia</i> )	TLVTVQKAVGEEIPI <sup>S</sup> VKTSK--K*
MOLLUSCA ( <i>Agropecten</i> )	FNVTVLKAMGEE <sup>T</sup> IIAVKTLIN--K*
MOLLUSCA ( <i>Crassostrea</i> )	IAVTVLNSMGEEQITGVKNLS*--
MOLLUSCA ( <i>Venerupis</i> )	LLATVLSAMGEEVIVSVK <sup>T</sup> NE--K*
MOLLUSCA ( <i>Haliotis</i> )	FMVTVQKAMDEEIVVALKSNMK--ND*
MOLLUSCA ( <i>Aplysia</i> )	IMVTVLKAMGEE <sup>M</sup> AVGKIKNAK--D*
BRYOZOA ( <i>Bugula</i> )	LMVTVLKS <sup>M</sup> GT <sup>E</sup> EMISSKPSK*--
PLATYHELMINTHES ( <i>Schistosoma</i> )	VIVTVVTMD <sup>E</sup> EQAHAVRTSS--DK*
ROTIFERA ( <i>Brachionus</i> )	LLVTVLKS <sup>M</sup> GEEAIIQFKVENDSKK*--
GNATHOSTOMULIDA ( <i>Gnathostomula</i> )	ILVTVLAA <sup>C</sup> G <sup>O</sup> EAAIASKPMTK--DK*
NEMATODA ( <i>Caenorhabditis</i> )	ILVTVVSAIGEEAII <sup>L</sup> GWKVSTK--E*
NEMATODA ( <i>Meloidogyne</i> )	ILVTVVSAIGEEVVLG <sup>W</sup> KKNM <sup>P</sup> NRD*--
NEMATODA ( <i>Ascaris</i> )	ILVTVVSAIGEEQILG <sup>F</sup> KKNM <sup>P</sup> N--K*
NEMATODA ( <i>Trichinella</i> )	ILITVVSAMGEEAII <sup>M</sup> AYKNM <sup>F</sup> K--D*
NEMATODA ( <i>Xiphinema</i> )	ILVTVVTAMGEE <sup>M</sup> MAFK <sup>N</sup> M <sup>F</sup> K--D*
TARDIGRADA ( <i>Richtersius</i> )	RLVTVGLKALPDE <sup>C</sup> ATAWKPNTAAEK*
ONYCHOPHORA ( <i>Euperipatoides</i> )	LLR <sup>T</sup> TVLRA <sup>C</sup> GEEAVIAKQNTQSDK*
NEMATOMORPHA ( <i>Spiniochordodes</i> )	LLIVIQKAMGLETAIA <sup>F</sup> KTQ <sup>E</sup> SKK*
ARTHROPODA	.....
PANTOPODA ( <i>Anoploaetylus</i> )	YSITVLTAMGIEQAI <sup>A</sup> IKQSS*--
CHELICERATA ( <i>Acanthoscurria</i> )	VIVTVLSAVGEEAIV <sup>A</sup> CKN <sup>S</sup> *--
CHELICERATA ( <i>Ixodes</i> )	VIVTVMSAVGTE <sup>C</sup> LI <sup>T</sup> HKNAN*--
CHELICERATA ( <i>Ornithoconus</i> )	VIVTVLSAVGEEAIV <sup>A</sup> CKN <sup>S</sup> *--
DECAPODA ( <i>Callinectes</i> )	LLLTVLAAMGHEM <sup>V</sup> VATKPNM--K*
DECAPODA ( <i>Celca</i> )	LLLTVLAAMGHEM <sup>V</sup> VATKPNM--K*
DECAPODA ( <i>Penaeus</i> )	LLLTVLAAMEEM <sup>V</sup> VATKPNM--K*
AMPHIPODA ( <i>Gammarus</i> )	LLCTVLCAMGSEMV <sup>I</sup> ATKPN <sup>S</sup> --NK*
“ENTOMOSTRACA” + HEXAPODA	.....C.....C.....C.....TA <sup>V</sup> /L <sup>D</sup> /E <sup>K</sup> .....
MAXILLOPODA ( <i>Calanus</i> )	LLVTVLGA <sup>C</sup> GEECVIA <sup>T</sup> KANTAVDK*--
MAXILLOPODA ( <i>Lepeophtheirus</i> )	ILCNVLSACGEEAVI <sup>A</sup> TKVNSAVDK*--
MAXILLOPODA ( <i>Caligus</i> )	ILCTVLSACGEXAVI <sup>A</sup> TKINTAVDK*--
BRANCHIOPODA ( <i>Artemia</i> )	LLCTVLKACGEECVIAIK <sup>T</sup> STIGDK*--
BRANCHIOPODA ( <i>Daphnia</i> )	LLCTVLKACGEECVIAV <sup>A</sup> KANTAAEK*--
COLLEMBOLA ( <i>Folsomia</i> )	ILCTVLKSCGEEV <sup>V</sup> IAVKNNTALEKTN*
COLLEMBOLA ( <i>Onychiurus</i> )	LLCTVLKACGEEV <sup>V</sup> IAINNNNTALEK*
INSECTA ( <i>Locusta</i> )	LLCTVLKSCGEECVIAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Gryllus</i> )	LLCTVLKSCGEECVIAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Pediculus</i> )	LLCTVLKSCGEECVIAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Diaphorina</i> )	LLCTVLKACGEECVIAIK <sup>T</sup> NTA <sup>A</sup> DK*
INSECTA ( <i>Nilaparvata</i> )	LLCTVLKSCGEECVIAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Acyrtosiphon</i> )	LLCTVLKSCGEECVIAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Macronellicoccus</i> )	LLCTVLKSCGEECVIAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Diabrotica</i> )	ILCTVLKSCGEEV <sup>V</sup> IAVKNNTALDK*
INSECTA ( <i>Tribolium</i> )	ILCTVLKSCGEEV <sup>V</sup> IAVKNNTALDK*
INSECTA ( <i>Lysiphlebus</i> )	LLCTVLKACGEEV <sup>V</sup> IAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Solenopsis</i> )	LLCTVLKACGEEV <sup>V</sup> IAVKNNTALDK*
INSECTA ( <i>Drosophila</i> )	LLCTVLKACGEECVIAIK <sup>T</sup> NTALDK*
INSECTA ( <i>Chironomus</i> )	LLCTVLKSCGEECVIAIK <sup>T</sup> NTAAEK*--
INSECTA ( <i>Aedes</i> )	LVCTVLKSCGEE <sup>T</sup> VIAIK <sup>N</sup> NTA <sup>A</sup> DKN*
INSECTA ( <i>Culex</i> )	LLCTVLKSCGEE <sup>T</sup> VIAIK <sup>N</sup> NTA <sup>A</sup> DKN*
INSECTA ( <i>Bombyx</i> )	LLCTVLKSCGEECVIAV <sup>A</sup> KANTALDK*--
INSECTA ( <i>Danaus</i> )	LLCTVLKSCGEECVIAV <sup>A</sup> KANTALDK*--
INSECTA ( <i>Spodoptera</i> )	LLCTVLKSCGEECVIAV <sup>A</sup> KANTALDK*--

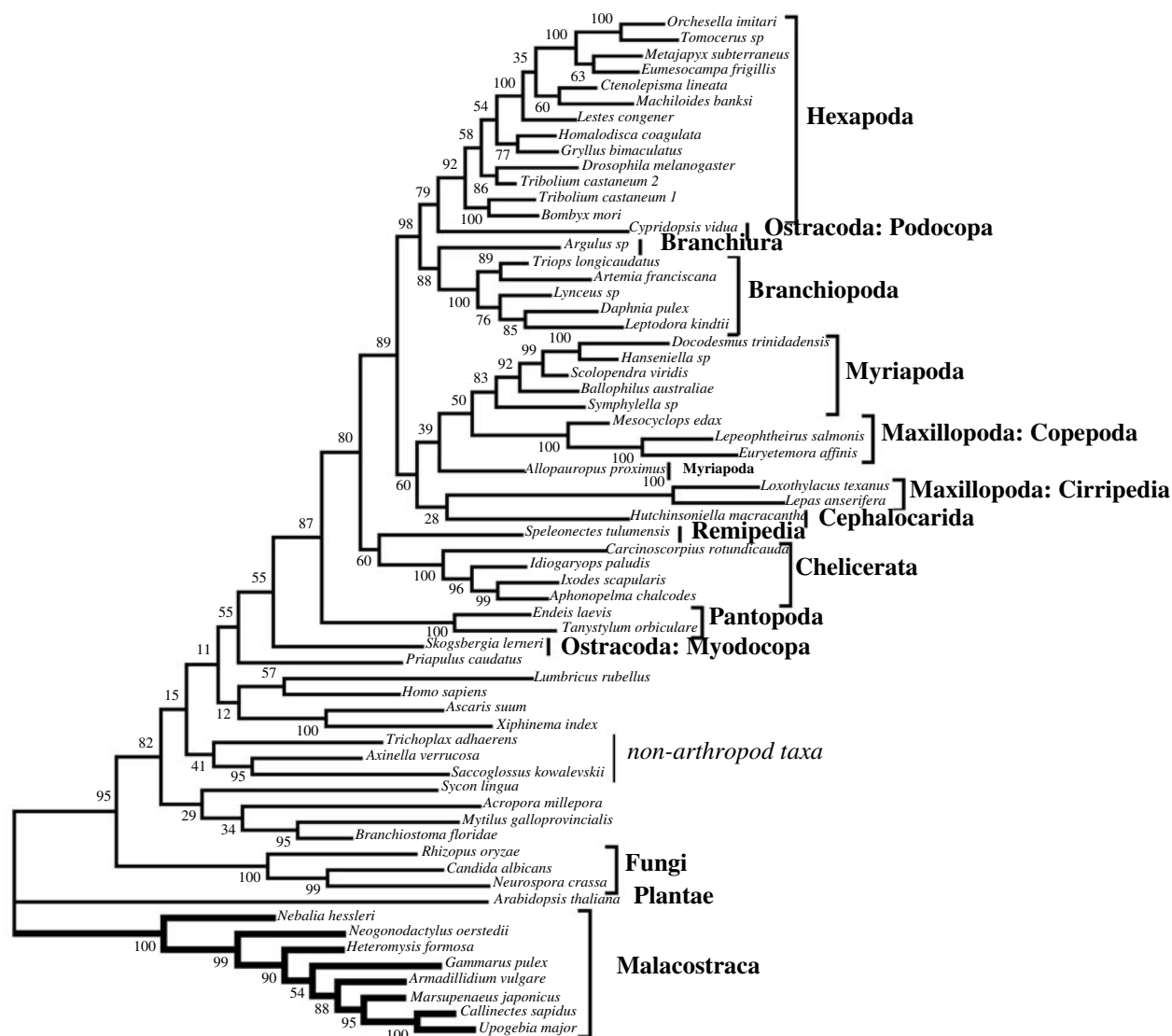
**Рис. 4.** Фрагмент (по видам) выравнивания С-концевой части белка фактора инициации 5А (eIF5A). Серым фоном выделен консенсус С-концевой последовательности из восьми аминокислот CX<sub>2</sub>CX<sub>3</sub>CX<sub>7</sub>TA<sup>V</sup>/L<sup>D</sup>/E<sup>K</sup> и стоп-кодона Нехарода и “Entomotraca”. Звездочками обозначены стоп-кодоны.

роль eIF5A. Имеются данные о его участии в запрограммированной клеточной смерти в ответ на действие патогенов у растений, а также в регуляции дифференцировки мышечных и нервных тканей животных. Короткая последовательность на С-конце eIF5A изменчива, однако же у Нехарода и “Ento-

mostraca” сохраняется присущий обоим этим группам вполне специфический мотив (рис. 4). В нашей выборке только eIF5A тихоходки *Richtersius coronifer* вполне подходит под свойственный Нехарода и “Entomotraca” консенсус, хотя последовательности и некоторых других видов также в определенной степени с ним сходны. В типе членистоногих Нехарода и “Entomotraca” ясно отличаются от Chelicerata и Malacostraca. Не исключено, что эти различия связаны со специфическими для каждой из двух последних групп эволюционными изменениями (аутопоморфиями). Например, С-концевой фрагмент eIF5A у хелицерных и пантопод на четыре аминокислотных остатка короче, чем у Нехарода и “Entomotraca”, или на 2–3 остатка короче, чем у большинства других животных. С-концевой фрагмент eIF5A у Malacostraca тоже короче на одну–две аминокислоты, чем у большинства животных, что свидетельствует, скорее всего, о его укорочении в процессе эволюции этого таксона. Однако эти обстоятельства не дают оснований считать мотив, свойственный Нехарода и “Entomotraca”, предковым для членистоногих в целом. Такое предположение противоречит требованию максимальной экономии, поэтому сходство С-концевой последовательности белка eIF5A представляется, скорее, синапоморфией Нехарода и “Entomotraca”.

**3. Белок eEF1A.** Фактор элонгации 1А (eEF1A) – многофункциональный, жизненно необходимый, умеренно консервативный белок. Его широко использовали в филогенетике, в том числе, для реконструкции родственных отношений в типе членистоногих [13, 56]. Мы обратили внимание на то, что “высшие” ракообразные (Malacostraca) обладают уникальным комплексом аминокислотных замен в белке EF1A, отличающем их не только от других членистоногих, но и всех других животных, а также грибов и растений. Тем не менее, фактор элонгации Malacostraca является ортологом EF1A и не относится к недавно описанному семейству EF-подобных (EFL) белков [57]. Malacostraca размещаются на дереве EF1A в его корне, отдельно от других животных, что является результатом артефакта притяжения длинных ветвей (рис. 5). Общеизвестно, что самые примитивные среди высших ракообразных – виды надотряда Phyllocarida [1, 3, 58]. Представитель филлокарид, *Nebalia hessleri*, имеет общие с остальными таксонами Malacostraca (Decapoda, Mysidacea, Amphipoda, Isopoda, Stomatopoda) аутопоморфии в белке EF1A и объединяется с ними на дереве.

Ранее монофилия филлокарид и остальных “высших” ракообразных была показана при анализе 18S рРНК [59] и подтверждена в исследованиях белоккодирующих генов, но наиболее наглядно общность происхождения всех Malacostraca заметна по аутопоморфному комплексу замен в EF1A [13, 56]. Казалось бы, наличие аутопоморфии у этого таксона никак не помогает решить вопрос о происхождении насекомых, сохраняющих плезиоморф-



**Рис. 5.** Байесово (MrBayes 3.1.2) дерево аминокислотных последовательностей фактора элонгации EF1A разных групп эукариот, включающее основные группы Arthropoda. Цифры у узлов дерева обозначают апостериорную вероятность этих узлов в процентах. Размещение ветви Malacostraca в корне дерева объясняется артефактом притяжения длинных ветвей (объяснения в тексте).

ное состояние вместе с другими членистоногими. Однако плезиоморфии в белке EF1A насекомых свидетельствует о том, что они не могли отделиться от ветви Malacostraca позже, чем произошло отделение филокаррид, как предполагали некоторые авторы [18].

### Митохондриальные гены

Первоначальный, подготовленный для анализа набор включал в себя аминокислотные последовательности 10 конкатенированных митохондриальных белков 178 видов Arthropoda, двух видов Onychophora и одного вида Priapulida, взятых в качестве множественной внешней группы. Построенное по

этому набору PhymI-дерево (не приведено) показало, что имеется значительная межвидовая гетерогенность скорости эволюции митохондриальных белков, которая, скорее всего, привела к объединению аномально быстро эволюционировавших последовательностей некоторых видов насекомых и хелицерных в одну группу в основании дерева. Влияние этого фактора изменяет также положение на дереве и других последовательностей, дающих аномально длинные ветви, например, некоторых видов ракообразных.

Для уменьшения гетерогенности набора по скорости эволюции мы удалили аномально быстро эволюционирующие последовательности и вариабельные участки, трудно поддающиеся выравнива-



нию. Полученное выравнивание содержало 82 последовательности и имело 2361 позиций. На дереве, построенном по этому набору с помощью программы MrBayes 3.1.2 (не приведено), как и на дереве последовательностей рибосомных белков (рис. 1), с апостериорной вероятностью 0.95 выделяется группировка Pancrustacea (все Hexapoda + все Crustacea). Многоножки (Muriaroda) примыкают к основанию этой группы, поддерживая концепцию Mandibulata, тогда как на дереве рибосомных белков они объединяются с хелицеровыми. Немногочисленные в наборе группы видов (Cirripedia, Collembola и Diplura) последовательно ответвляются в основании Pancrustacea между многоножками и основной группой ракообразных. Принципиальное отличие дерева митохондриальных белков от дерева ядерных рибосомных белков заключается в том, что на нем Hexapoda не монофилетичны (и это предполагает неоднократное возникновение “гексаподного” плана строения), а также в том, что жаброногие ракообразные (Branchiopoda) не являются ближайшей к Hexapoda группой, как на дереве рибосомных белков.

Невысокий уровень апостериорной вероятности некоторых групп на дереве митохондриальных белков указывает, что были найдены топологии, альтернативные к итоговой байесовой, но с незначимым отличием правдоподобия. Отслеживая ход байесова анализа в двух параллельных запусках, мы обнаружили, что 50%-ные значения поддержки некоторых групп (основной клады Crustacea, группы Collembola + Diplura + Crustacea + Insecta *sensu stricto*, а также группы Muriaroda + все Crustacea + все Hexapoda) определяются перемещением отдельных групп последовательностей. В случае Crustacea это связано с объединением нескольких быстро эволюционирующих последовательностей ракообразных с ногохвостками и с перемещением последовательности ракушковых (тоже быстро эволюционирующей) к хелицеровым, что сказывается также и на поддержке второй группировки. В случае группировки Mandibulata 50%-ная поддержка определяется объединением многоножек с хелицеровыми во многих вариантах деревьев.

После удаления шести быстро эволюционирующих последовательностей было получено дерево, по топологии отличающееся от предыдущего тем, что в нем объединяются многоножки с хелицеровыми (64%-ная поддержка); оно отличается также тем, что уровень статистической поддержки основных групп гораздо выше (рис. 6а). В остальной части это дерево характеризуется таким же положением основных групп, как и обсуждавшееся выше дерево, полученное по набору из 82 видов.

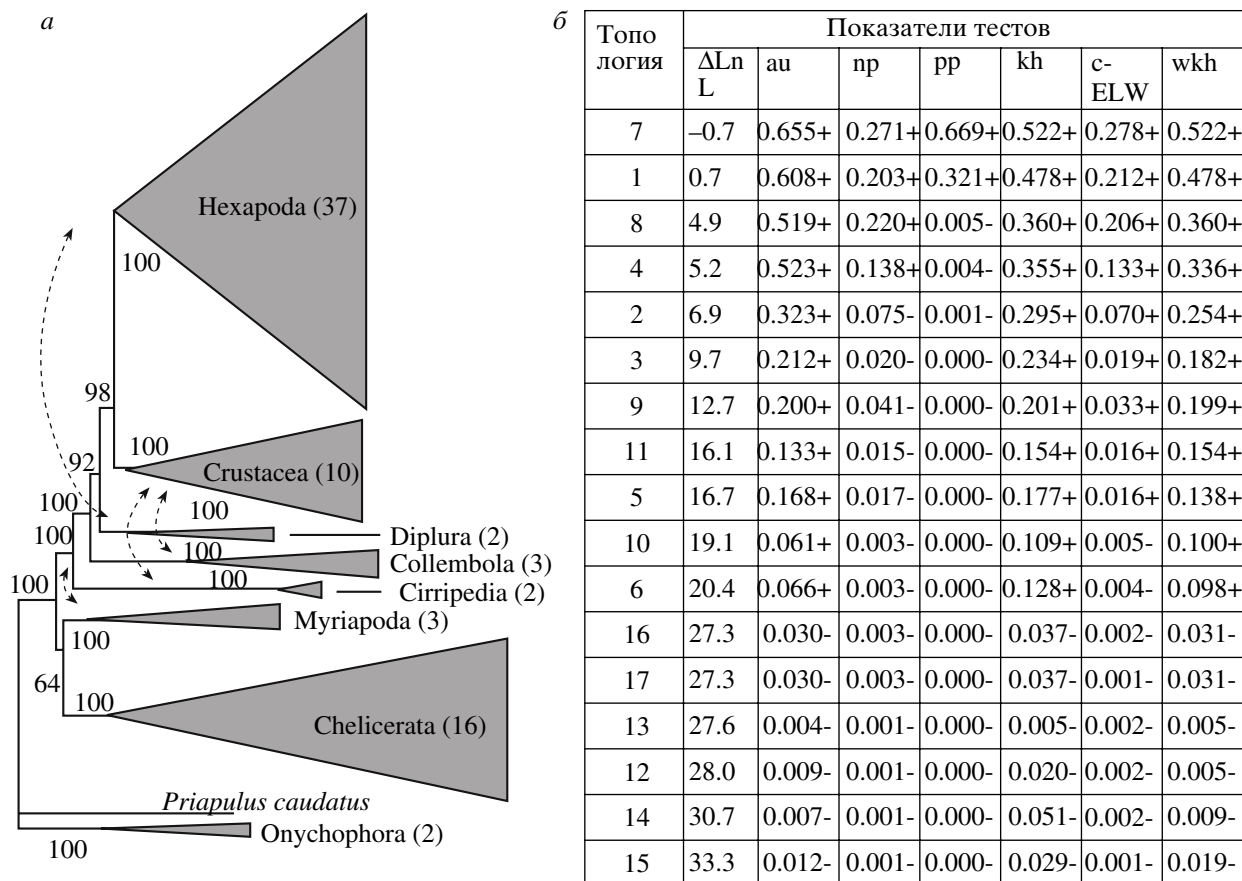
Совместимость наших данных с некоторыми возможными филогенетическими гипотезами оценивали с помощью статистических тестов с использованием параметрического бутстрэп-анализа. То-

пологию дерева на рис. 6а искусственно изменяли, как описано в подписях к рисунку. Получившиеся 16 вариантов топологий сравнивали с исходной байесовой топологией с помощью программы TREE-PUZZLE и последующей статистической оценкой в программе CONSEL. Результаты анализа приведены на рис. 6б. В первом столбце таблицы цифрами обозначены варианты сравниваемых топологий, расположенные в порядке увеличения их отличий от байесовой топологии. Во втором столбце приведены значения отличий сравниваемых топологий от исходной по значению логарифма максимального правдоподобия (Ln Lh). В следующих шести столбцах приведены статистические показатели отличий сравниваемых топологий от исходной, полученные разными тестами. Показатель менее 0.05 указывает на достоверность отличия. Согласно этим тестам, в первой строке таблицы оказывается топология 7, отличающаяся от исходной байесовой топологии (1) тем, что многоножки объединяются с хелицеровыми, однако разница между этими двумя вариантами статистически недостоверна. Следующие девять вариантов деревьев (от 2-го до 11-го, исключая 7-ой) не отличаются достоверно от исходного дерева по большинству тестов, за исключением самых строгих (столбцы pr и pp). Все эти варианты характеризуются монофилией основных групп Crustacea (Malacostraca + Branchiopoda) при различном положении Diplura и Collembola. Остальные шесть вариантов, характеризующиеся сближением Hexapoda с жаброногими ракообразными (Branchiopoda) и, следовательно, парафилией ракообразных, независимо от положения многоножек, достоверно отличаются от лучшей топологии по всем тестам.

Таким образом, анализ последовательностей митохондриальных белков показывает несовместимость его результатов с гипотезами, сближающими монофилетическую группу Hexapoda с Branchiopoda, т.е. подразумевающими происхождение гексапод от общего с жаброногими ракообразными предка. Вместе с тем достоверно не исключаются варианты, в которых все виды Hexapoda и Crustacea объединены в монофилетические группы, однако варианты, в которых Collembola помещены в основание Pancrustacea, по статистическим показателям выглядят предпочтительнее.

## ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Беглый обзор полученных в настоящей работе результатов показывает их очевидное противоречие. Ранее в публикациях других авторов предлагались многочисленные взаимоисключающие филогенетические гипотезы об эволюции членистоногих в целом и Pancrustacea, в частности, часть из которых обосновывались анализом молекулярных данных. В этой ситуации необходимо сделать выбор, каким данным или результатам какого анализа следует доверять в большей степени. Рассмотрим несколько



**Рис. 6.** *a* – Байесово (MrBayes 3.1.2) дерево аминокислотных последовательностей митохондриальных белков 73 видов Arthropoda и 3 видов в качестве внешней группы. Цифры у узлов дерева указывают величины статистической поддержки этих узлов в процентах. Монофилетические группировки представлены в виде затененных треугольников, размеры которых пропорциональны числу входящих в них видов (основание) и длине ветвей (высота). Цифры в скобках у названия крупных группировок обозначают число входящих в них видов. Пунктирные стрелки показывают наблюдаемые перемещения групп при изменении таксономической выборки. *б* – Результаты статистической оценки значимости отличий некоторых возможных вариантов деревьев методом параметрического бутстрэпа с помощью программ TREE-PUZZLE и CONSEL. Варианты деревьев приведены в скобочном формате: 1 – ((Cirr,(Coll,(Dipl,(Hexa,Crus))))), (Myri,Cheli); 2 – ((Cirr,(Dipl,(Coll,(Hexa,Crus))))), (Myri,Cheli); 3 – ((Cirr,(Coll,(Dipl,(Hexa,Crus))))), (Myri,Cheli); 4 – ((Cirr,(Hexa,Crus),(Dipl,Coll))), (Myri,Cheli); 5 – (((Hexa,(Crus,Cirr),(Dipl,Coll))), (Myri,Cheli); 6 – ((Coll,(Dipl,(Hexa,(Crus,Cirr))))), (Myri,Cheli); 7 – (Cheli,(Myri,(Cirr,(Coll,(Dipl,(Hexa,Crus)))))); 8 – (Cheli,(Myri,(Coll,(Dipl,(Hexa,(Crus,Cirr)))))); 9 – (Cheli,(Myri,(Dipl,(Hexa,Coll),(Crus,Cirr))))); 10 – (Cheli,(Myri,(Coll,(Hexa,Dipl),(Crus,Cirr))))); 11 – (Cheli,(Myri,(Hexa,(Dipl,Coll),(Crus,Cirr))))); 12 – ((Cirr,(Coll,(Dipl,(Hexa,Bran),Mala))), (Myri,Cheli); 13 – (((Cirr,(Coll,(Dipl,(Hexa,Bran),Mala))), (Myri,Cheli); 14 – ((Cirr,(Mala,(Bran,(Coll,(Dipl,Hexa))))), (Myri,Cheli); 15 – (Cheli,(Myri,(Cirr,(Mala,(Bran,(Coll,(Dipl,Hexa)))))); 16 – (Cheli,(Myri,(Cirr,(Mala,(Bran,(Hexa,(Coll,Dipl)))))); 17 – ((Cirr,(Mala,(Bran,(Hexa,(Coll,Dipl))))), (Myri,Cheli)). Знаки плюс справа от показателей означают, что эти варианты достоверно не отличаются от исходного; знаки минус указывают на статистически значимые отличия.

технических обстоятельств, которые могли стать возможными причинами артефактов в реконструкции деревьев и привести к неадекватности применяемых к ним статистических тестов.

В первую очередь, весьма очевидна разница в общем числе признаков, использованных нами в анализе ядерных генов (более 11 тысяч позиций) и митохондриальных генов (немногим более 2 тыс. позиций после удаления переменных блоков). Результаты, полученные на основе большей выборки признаков, вызывают больше доверия. Из этих соображений, можно было бы выбрать дерево, представленное на

рис. 1, как обоснованное большим числом признаков. Однако оценка выборок по размеру оправдана, если нет причин подозревать вклада неслучайных факторов. К сожалению, у нас нет достаточных оснований исключить влияние систематических ошибок, скорее наоборот. Дерево на рис. 1 не является ультраметрическим. Например, ветви, ведущие на нем к представителям двукрылых, являются длиннейшими во всей выборке Hexapoda. Другими словами, эти виды двукрылых накопили больше замен в аминокислотных последовательностях рибосомных белков относительно общего предка, чем все другие

виды *Neharoda*, взятые в сравнение. Это не случайность, а результат более быстрой молекулярной эволюции в линии *Diptera*, как было показано ранее [60]. Байесово дерево на рис. 1 не является дистанционным: программа MrBayes использует символично-ориентированный алгоритм построения дерева. Он, тем не менее, избавляет далеко не от всех ошибок, связанных с разной скоростью эволюции. В частности, размещение двукрылых в основании дерева насекомых, а прямокрылых и термитов – в кроне (рис. 1), кажется невероятным из опыта зоологии, но понятно как артефакт, связанный с разницей длин ветвей. По аналогичной причине (ввиду разницы в скоростях эволюции, а не вследствие действительных родственных отношений) более длинные ветви *Malacostraca* могли бы сместиться на дереве ближе к корню в большей степени, нежели чуть более короткие ветви *Branchiopoda*.

Вместе с тем, было бы неоправданным отдать предпочтение дереву митохондриальных белков на основании наблюдаемой неоднородности скорости эволюции ядерных генов рибосомных белков, так как митохондриальные белки проявляют еще большую неоднородность в скоростях эволюции между разными филогенетическими линиями. Это следует из разницы длин ветвей дерева (данные не представлены) и присутствия артефактов, связанных с объединением совершенно разных таксонов членистоногих, имеющих общим только то, что ветви, ведущие к ним на дереве, оказываются более длинными (т.е. скорость молекулярной эволюции их митохондриальных белков была выше). Любопытно, что митохондриальные белки двукрылых (*Diptera*) по скорости эволюции не “выбиваются” из нормы и правильно размещаются в кроне дерева насекомых. Напротив, митохондриальные белки хоботных и перепончатокрылых испытывали аномально быструю эволюцию. На деревьях ведущие к ним экстремально длинные ветви группируются вместе, причем этот кластер длинных ветвей способен перемещаться к хелицерным, объединяясь с длинными ветвями некоторых клещей.

Другая особенность набора митохондриальных белков, которую можно трактовать как пример его недостаточной разрешающей способности, – невозможность выявления с его помощью монофилии *Neharoda*. Причем, если *Diplura* на сконструированном дереве занимают обособленное положение или же группируются с насекомыми (в зависимости от таксономической выборки и параметров, использованных при построении дерева), то группировка ногохвосток с насекомыми отвергается во всем исследованном диапазоне параметров. Последний результат согласуется с опубликованными результатами анализа митохондриальных белков [61, 62], но противоречит другим молекулярным реконструкциям, неизменно подтверждающим родство ногохвосток и насекомых [49, 63, 64].

Важным обстоятельством при анализе данных по многим генам является выбор адекватных моделей замен и правильная оценка гетерогенности сайтов по скорости эволюции. Мы обнаружили существенные различия между рибосомными белками в отношении этих параметров. Можно, конечно, вычислить значения параметров, усредненные для конкатенированного выравнивания, но кажется более предпочтительным применять к отдельным частям адаптированные значения. Только немногие из программ реконструкции филогении предусматривают такую возможность. Мы использовали одну из них, MrBayes 3.1.2, применив команду *partition*. Однако при статистических оценках достоверности отличий значения правдоподобия для различных топологий современные программы не позволяют использовать специфические параметры для отдельных генов в составе общего выравнивания. Это заставляет с известной осторожностью относиться к результатам таких оценок.

Не менее, если не более, проблематичен вопрос о специфичности параметров не для отдельных генов или белков в составе конкатенированного выравнивания, а для отдельных таксонов применительно к одному белку. Общие соображения, что ортологичные белки обычно выполняют одну и ту же функцию в клетках, скажем, жуков и пауков, и поэтому замена, например, гистицинов на глицины без потери функциональности тех же самых белков тех и других видов равновероятна, только отчасти правдоподобны. Вполне возможно, что длинные ветви дерева отражают, по крайней мере, в отдельных случаях, именно изменившийся до определенной степени спектр допустимых замен. В отношении митохондриальных белков установлено, что общая для эукариот модель (матрица весов замен и стационарные частоты аминокислот) эволюции митохондриальных генов имеет некоторую специфичность для млекопитающих, членистоногих и даже *Pancrustacea* [62, 65, 66]. Две последние модели описаны относительно недавно и не предусмотрены в программе MrBayes 3.1.2. Таким образом, при построении дерева митохондриальных белков мы вынуждены жертвовать или оптимизированной моделью замещений, или возможностью независимой оптимизации параметров  $\Gamma$ -распределения. При этом остается открытым вопрос о возможности реализации специфических моделей замещений в эволюции групп меньшего ранга, чем *Pancrustacea*.

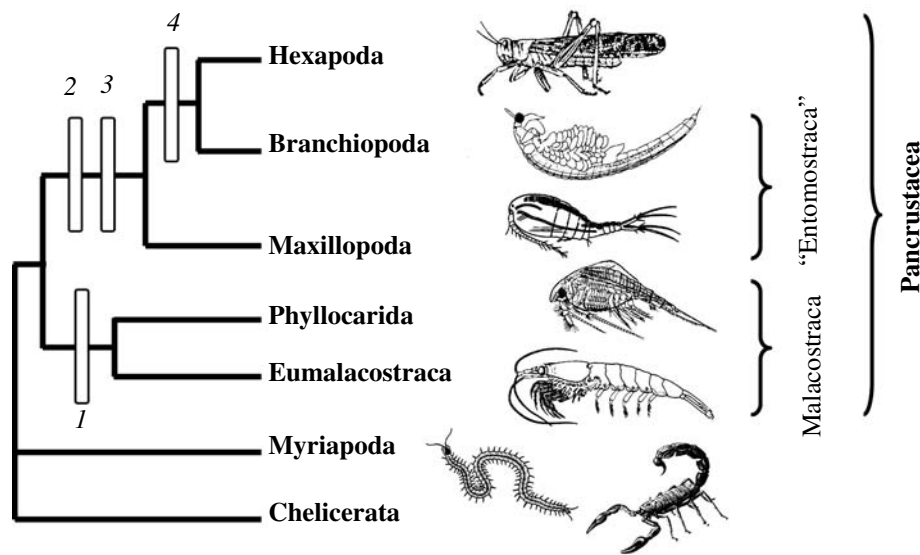
Другое важное обстоятельство – различия в таксономических выборках. В настоящее время выборка митохондриальных генов членистоногих таксономически значительно богаче: уже известны митохондриальные геномы у многих ключевых таксонов членистоногих, сведения по ядерным генам которых ограничены генами рибосомных РНК и немногих белков. Более полная таксономическая выборка важна не только тем, что из нее можно узнать положение большего числа таксонов, но и тем, что она

сильно влияет на точность вычисления параметров молекулярной эволюции, таких как скорости эволюции отдельных сайтов и др. Эти параметры скажутся на положении многих ветвей сконструированного дерева, поэтому более полная таксономическая выборка нередко оказывает даже большее влияние на точность реконструкции, чем число взятых в сравнение признаков [67, 68]. Конечно, недостаточность таксономической выборки геномных данных современных таксонов носит временный характер, и в скором будущем ситуация изменится. По-видимому, параллельно будут возрастать вычислительные возможности: в настоящее время матрицу (число видов, умноженное на число признаков) в несколько раз большего размера, чем использованная для получения дерева на рис. 1, невозможно проанализировать методом правдоподобия или Байеса с должным числом параметров ни на одном суперкомпьютере.

Сложность выбора модели аминокислотных замен поднимает вопрос, не лучше ли для реконструкции филогении использовать нуклеотидные последовательности [69]. Имеются экспериментальные данные в пользу целесообразности учета нуклеотидных последовательностей митохондриальных белоккодирующих генов при сравнении таксонов насекомых ранга выше семейств [70]. Наша позиция по этому вопросу совершенно определенная. Мы считаем, на основании общих теоретических соображений, что учет синонимичных замен необходим для реконструкции филогении близких форм, но при сравнении далеких форм он лишь является дополнительным источником ошибок. Имеется крайне мало причин, по которым насыщение синонимическими мутациями не может быть достигнуто за несколько десятков миллионов поколений. Таким образом, сходства по позициям вырожденности кодонов (кроме регуляторных сайтов в ДНК/мРНК) у видов, разошедшихся в эволюции более 100 млн. лет назад, возникают вследствие реверсий и повторных замен, учет которых скорее ухудшит точность реконструкции. Это сходство может также усиливаться из-за систематических ошибок, а именно, конвергенций под действием факторов, контролирующих нуклеотидный состав сам по себе, что особенно актуально для митохондриальной ДНК с ее очень смещенным нуклеотидным составом [71]. Поэтому для реконструкции филогении далеких видов мы используем только аминокислотные последовательности белков (или нуклеотидные последовательности генов, кодирующих РНК со сложными функциями, таких как рРНК). Первые насекомые известны из карбона и вряд ли могли возникнуть ранее девона [11, 25, 72], тогда как дивергенция ракообразных и вовсе датируется кембрием [73]. По причине большого возраста этих групп мы игнорируем нуклеотидные последовательности белоккодирующих генов в нашем исследовании.

Все высказанные соображения негативного характера не отвергают возможности уже сейчас формулировать более обоснованные филогенетические гипотезы. Тысячи доступных признаков – позиций выравнивания – дают надежду найти среди них признаки, мало подверженные гомоплазиям и пригодные для кладистического анализа. Подобный пример дает рибосомный белок S28: эволюция обсуждаемых в нем признаков (рис. 3) поляризуется однозначно, и сходство Hexapoda и “Entomostraca” в аминокислотной последовательности, без сомнения, синапоморфное. Остается вопрос, не может ли это сходство быть конвергентным. Здесь ситуация не такая простая. Замены  $K \rightarrow V$  и  $D \rightarrow G$  происходят со “средней” частотой, если отвлечься от эволюции конкретных сайтов и ориентироваться на модель rtREV (которая обеспечивает наилучшее приближение характера аминокислотных замещений в RpS28). Замена аминокислот  $R \rightarrow K$  вообще относится к самым распространенным: из 190 возможных типов реципрокных замен, согласно модели rtREV [46], только три наблюдаются чаще. Если исходить из этих соображений, может показаться весьма вероятным, что сходство по этим остаткам приобретено случайно. Однако применить статистику модели rtREV к отдельно взятым позициям вряд ли допустимо. По-видимому, в протеоме, в целом, замены лизина на аргинин и обратно очень часто оказываются нейтральными, что и обеспечивает их высокую частоту согласно модели rtREV. Но в положении 27 белка RpS28 большой частоты таких замен не наблюдается: наоборот, состояние этих трех признаков очень стабильно у Hexapoda и “Entomostraca”, с одной стороны, и в остальных группах, сохраняющих плезиоморфное состояние, с другой. Исключения легко описываются в рамках экономных сценариев. Например, аутоапоморфии двукрылых в инфраотряде Culicomorpha легко объяснить последовательным замещением  $Gly32 \rightarrow Asn32 \rightarrow Ser32$  (нумерация остатков по белку RpS28 *Drosophila melanogaster*). Этот сценарий полностью согласуется с известными, хорошо обоснованными филогенетическими отношениями среди Culicomorpha [74, 75]. Из характера эволюции следует, что предпочтение аргинина или лизина в положении 27 белка RpS28 связано с сильными функциональными (по-видимому, ковариационными) ограничениями в разных группах членистоногих.

Для анализа С-концевой последовательности белка eIF5A, вариации которой обусловлены не только заменами, но и делециями/инсерциями, невозможно применить какие-либо разработанные статистические модели. Поскольку характерному для Hexapoda и “Entomostraca” мотиву, в разной степени, удовлетворяют еще несколько последовательностей за пределами этой группы, нельзя исключить, что сходство Hexapoda и “Entomostraca” также возникло независимо. Однако тот факт, что Hexapoda + “Entomostraca” не только похожи, но и



**Рис. 7.** Филогенетическое дерево членистоногих. 1 – аутоапоморфии Malacostraca в факторе элонгации EF1A; 2, 3 – синапоморфии Hexapoda и “Entomostraca” в белках RpS28 и eIF5A; 4 – синапоморфии в умеренно консервативных областях белков Hexapoda и Branchiopoda.

консервативны по этому признаку, более согласуется с гипотезой специфической функциональной роли этой последовательности в их белке eIF5A и ее наследования от общего предка.

Признаками белков RpS28 и eIF5A не ограничивается список синапоморфий Hexapoda + “Entomostraca” в консервативных сайтах. В кодируемых ядром белках мы обнаружили и другие подобные признаки, поддерживающие иерархически организованные группировки Hexapoda + Branchiopoda и Hexapoda + “Entomostraca”, что согласуется с деревом на рис. 7. Несмотря на большое число просмотренных признаков, только очень немногие оказались связанными с консервативными сайтами. Анализ выравнивания митохондриальных белков не выявляет каких-либо наглядных признаков в сайтах с редкими заменами, которые поддерживали бы топологии 6 или 7. По-видимому, истинный или ложный филогенетический сигнал, обеспечивающий конструирование разрешенного дерева (рис. 6), основан, главным образом, на эволюционных событиях в переменных сайтах, служащих мишенями неоднократных повторных замен, и гомоплазии в которых кажутся более вероятными.

Совокупность морфологических данных не позволяет сделать однозначного выбора в пользу монофилии ракообразных относительно насекомых или родства последних с Branchiopoda или Malacostraca, хотя последний вариант выглядит более предпочтительным [17]. В связи с этим особое значение приобретают доказательства монофилии Phyllocarida и других Malacostraca. Молекулярные данные, по сути, отвергают гипотезу происхождения Hexapoda от Malacostraca. Единственная возможность монофи-

лии Malacostraca и Hexapoda, которую можно теперь допускать, – это отделение последних до расхождения линий Phyllocarida и Eumalacostraca. Однако таксону Malacostraca с включенными в его состав филлокаридами можно дать крайне бедный диагноз, оставляя в нем, по сути, только признаки расчленения тела. Филлокариды проявляют значительное сходство с жаброногими ракообразными (Branchiopoda), что позволяло Шраму (F.R. Schram) объединить их в один класс “Phyllopoda” [76] (который, как сейчас выясняется благодаря молекулярным признакам, является полифилетическим). Допустив, что предки Hexapoda отделились до расхождения Phyllocarida и Eumalacostraca, мы вынуждены приписать их общему предку плезиоморфные характеристики “Phyllopoda” (за исключением, разве что, листовидности конечностей), а это делает равно возможным приписать ему и признаки плезиоморфных Branchiopoda. Приведенное рассуждение доказывает отсутствие противоречия между филогенетической гипотезой, объединяющей Hexapoda и Branchiopoda (рис. 1 и рис. 7), и совокупными данными о морфологической эволюции Pancrustacea. Имеются достаточно веские основания считать фильтрационный способ питания “Entomostraca” и Phyllocarida первичным [16]. В таком случае грасперный (путем хватания) механизм пищедобывания, свойственный многим Malacostraca [47] и сближающий их в этом отношении с насекомыми, развивается в их эволюции независимо, что и обуславливает определенное сходство насекомых и высших ракообразных.

Новый, современный виток филогенетических исследований основывается на геномных данных о десятках и сотнях, в перспективе – тысячах генов, и

берет в анализ десятки тысяч признаков. По-видимому, усовершенствование техники анализа со временем выведет нас и на новый уровень достоверности реконструируемой филогении. Но пока мы видим, что само по себе увеличение числа анализируемых признаков не снимает старых проблем филогенетики, например, создаваемых тем обстоятельством, что скорость эволюции в разных линиях неодинакова. Конечно, в эволюции молекул невозможна ситуация полного стаза, известного из морфологической эволюции, но разница в скоростях все же остается столь отличающейся от ожиданий “молекулярных часов” и столь значительной, что приводит к искажениям реконструируемой топологии. Более того, систематические ошибки при росте выборки признаков проявляют тенденцию к усилению искажений. Например, разница в длине ветвей Diptera и остальных насекомых на рис. 1 не так велика. По нашему личному опыту, современные программы, использующие алгоритмы максимального правдоподобия, в состоянии правильно разместить ветви с подобной относительной разницей в длине при филогенетических реконструкциях по отдельно взятому гену. Однако, с учетом очень большого числа анализируемых признаков, различия длин ветвей дерева на рис. 1 нельзя признать случайными. В результате алгоритм построения дерева, отдавая предпочтение гипотезе о большей скорости эволюции в линии Diptera, одновременно запрашивает и большее время для ее независимой эволюции (т. е. предполагает более раннее ответвление Diptera на дереве) для того, чтобы объяснить накопленные видами Diptera отличия от реконструируемого предка Hexapoda. По-видимому, новая геномная филогенетика потребует не только специальных программных средств работы с базами данных (сбора контигов и химер, тестирования ортологичности, выбраковки контаминаций), но и новых моделей и рабочих программ непосредственно для реконструкции деревьев.

В работе использованы вычислительные ресурсы Суперкомпьютерного комплекса НИВЦ МГУ “Чебышев” (<http://parallel.ru/cluster>) и комплекса The Biportal Университета Осло ([www.biportal.uio.no](http://www.biportal.uio.no)) для построения деревьев, а также платформа распределенных вычислений в National University of Ireland Maynooth in Ireland (<http://distributed.cs.nuim.ie>) для поиска оптимальных моделей аминокислотных замещений.

Работа получила финансовую поддержку Российского фонда фундаментальных исследований (08-04-01746, 09-04-01150, 09-04-92741) и Федерального агентства по науке и инновациям для ведущих научных школ (НШ-1275.2008.4).

### СПИСОК ЛИТЕРАТУРЫ

1. Рупперт Э.Э., Фокс Р.С., Барнс Р.Д. 2008. *Зоология беспозвоночных. Том 3. Членистоногие*. М.: Академия.
2. Клюге Н.Ю. 2000. *Современная систематика насекомых*. СПб.: Лань.
3. Догель В.А. 1981. *Зоология беспозвоночных*. Ред. Полянский Ю.И. М.: Высш. шк.
4. Turbeville J.M., Pfeifer D.M., Field K.G., Raff R.A. 1991. The phylogenetic status of arthropods, as inferred from 18S rRNA sequences. *Mol. Biol. Evol.* **8**, 669–686.
5. Ballard J.W.O., Ballard O., Olsen G.J., Faith D.P., Odgers W.A., Rowell D.M., Atkinson P. 1992. Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods. *Science*. **258**, 1345–1348.
6. Friedrich M., Tautz D. 1995. Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. *Nature*. **376**, 165–167.
7. Averof M., Akam M. 1995. Insect-crustacean relationships: insights from comparative developmental and molecular studies. *Phil. Trans. R. Soc. B*. **256**, 183–235.
8. Giribet G., Carranza S., Bagnù J., Riutort M., Ribera C. 1996. First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Mol. Biol. Evol.* **13**, 76–84.
9. Boore J.L., Lavrov D.V., Brown W.M. 1998. Gene translocation links insects and crustaceans. *Nature*. **392**, 667–668.
10. Aleshin V.V., Petrov N.B. 1999. Implicaciones del gen 18S ARNr en la evolución y filogenia de los Arthropoda. In: *Evolución y Filogenia de Arthropoda. Boletín de la Sociedad Entomológica Aragonesa*, núm. 26. Eds Melic A., de Haro J.J., Méndez M., Ribera I. Zaragoza: SEA, p. 177–196.
11. Glenner H., Thomsen P.F., Hebsgaard M.B., Sorensen M.V., Willerslev E. 2006. The origin of insects. *Science*. **314**, 1883–1884.
12. Budd G.E., Telford M.J. 2009. The origin and evolution of arthropods. *Nature*. **457**, 812–817.
13. Regier J.C., Shultz J.W., Kambic R.E. 2005. Pancrustacean phylogeny: hexapods are terrestrial crustaceans and maxillopods are not monophyletic. *Proc. R. Soc. B*. **272**, 395–401.
14. Giribet G., Richter S., Edgecombe G.D., Wheeler W.C. 2005. The position of crustaceans within Arthropoda – evidence from nine molecular loci and morphology. *Crustacean Issues*. **16**, 307–330.
15. Zrzavý J., Štys P. 1997. The basic body plan of arthropods: insights from evolutionary morphology and developmental biology. *J. Evol. Biol.* **10**, 353–367.
16. Dohle W. 2001. Are the insects terrestrial crustaceans? A discussion of some new facts and arguments and the proposal of the proper name Tetraconata for the monophyletic unit Crustacea + Hexapoda. *Ann. Soc. Entomol. Fr. (N.S.)*. **37**, 85–103.
17. Schram F.R., Jenner R.A. 2001. The origin of Hexapoda: a crustacean perspective. *Ann. Soc. Entomol. Fr. (N.S.)*. **37**, 243–264.
18. Павлов В.Я. 2000. *Периодическая система членистых*. М.: Изд-во ВНИРО.
19. Duman-Scheel M., Patel N.H. 1999. Analysis of molecular marker expression reveals neuronal homology in distantly related arthropods. *Development*. **126**, 2327–2334.

20. Richter S. 2002. The Tetraconata concept: hexapod-crustacean relationships and the phylogeny of Crustacea. *Org. Divers. Evol.* **2**, 217–237.
21. Harzsch S., Hafner G. 2006. Evolution of eye development in arthropods: Phylogenetic aspects. *Arthropod Struct. Dev.* **35**, 319–340.
22. Regier J.C., Shultz J.W., Ganley A.R., Hussey A., Shi D., Ball B., Zwick A., Stajich J.E., Cummings M.P., Martin J.W., Cunningham C.W. 2008. Resolving arthropod phylogeny: exploring phylogenetic signal within 41 kb of protein-coding nuclear gene sequence. *Syst. Biol.* **57**, 920–938.
23. Stollewerk A., Chipman A.D. 2006. Neurogenesis in myriapods and chelicerates and its importance for understanding arthropod relationships. *Integrat. Comp. Biol.* **46**, 195–206.
24. Bäcker H., Fanenbruck M., Wägele J.W. 2008. A forgotten homology supporting the monophyly of Tracheata: The subcoxa of insects and myriapods re-visited. *Zool. Anz.* **247**, 185–207.
25. Жерихин В.В., Пономаренко А.Г., Расницын А.П. 2008. *Введение в палеоэнтомологию*. М.: Т-во научн. изд. КМК.
26. Baurain D., Brinkmann H., Philippe H. 2007. Lack of resolution in the animal phylogeny: closely spaced cladogeneses or undetected systematic errors? *Mol. Biol. Evol.* **24**, 6–9.
27. Parkinson J., Whitton C., Schmid R., Thomson M., Blaxter M. 2004. NEMBASE: a resource for parasitic nematode ESTs. *Nucleic Acids Res.* **32**, D427–D430.
28. Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–33402.
29. Edgar R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797.
30. Hall T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98.
31. Jobb G., von Haeseler A., Strimmer K. 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol. Biol.* **4**, 18.
32. Friedman N., Ninio M., Pe'er I., Pupko T. 2002. A structural EM algorithm for phylogenetic inference. *J. Comput. Biol.* **9**, 331–353.
33. Roure B., Rodriguez-Ezpeleta N., Philippe H. 2007. SCAFoS: a tool for selection, concatenation and fusion of sequences for phylogenomics. *BMC Evol. Biol.* **7**, Suppl. 1, S2.
34. Jameson D., Gibson A.P., Hudelot C., Higgs P.G. 2003. OGRE: a relational database for comparative analysis of mitochondrial genomes. *Nucleic Acids Res.* **31**, 202–206.
35. Talavera G., Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577.
36. Guindon S., Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704.
37. Huelsenbeck J.P., Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* **17**, 754–755.
38. Keane T.M., Creevey C.J., Pentony M.M., Naughton T.J., McInerney J.O. 2006. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* **6**, 29.
39. Keane T.M., Naughton T.J., McInerney J.O. 2007. MultiPhyl: a high-throughput phylogenomics webserver using distributed computing. *Nucleic Acids Res.* **35**, W33–W37.
40. Felsenstein J. 2005. *PHYLIP (Phylogeny Inference Package) version 3.6*. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
41. Page R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357–358.
42. Schmidt H.A., Strimmer K., Vingron M., von Haeseler A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics.* **18**, 502–504.
43. Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508.
44. Shimodaira H., Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics.* **17**, 1246–1247.
45. Van de Peer Y., De Wachter R. 1993. TREECON: a software package for the construction and drawing of evolutionary trees. *Comput. Appl. Biosci.* **9**, 177–182.
46. Dimmic M.W., Rest J.S., Mindell D.P., Goldstein R.A. 2002. rtREV: an amino acid substitution matrix for inference of retrovirus and reverse transcriptase phylogeny. *J. Mol. Evol.* **55**, 65–73.
47. Whelan S., Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691–699.
48. Jones D.T., Taylor W.R., Thornton J.M. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275–282.
49. Mallatt J., Giribet G. 2006. Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Mol. Phylogenet. Evol.* **40**, 772–794.
50. Kashiyama K., Seki T., Numata H., Goto S.G. 2009. Molecular characterization of visual pigments in Branchiopoda and the evolution of opsins in Arthropoda. *Mol. Biol. Evol.* **26**, 299–311.
51. Kishino H., Hasegawa M. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29**, 170–179.
52. Rota-Stabelli O., Telford M.J. 2008. A multi criterion approach for the selection of optimal outgroups in phylogeny: Recovering some support for Mandibulata over Myriochelata using mitogenomics. *Mol. Phylogenet. Evol.* **48**, 103–111.
53. Павлинов И.Я. 2005. *Введение в современную филогенетику*. М.: Тов. науч. изд. КМК.

54. Алешин В.В., Константинова А.В., Михайлов К.В., Никитин М.А., Петров Н.Б. 2007. Нужно ли много генов для филогенетического дерева? *Биохимия*. **72**, 1610–1623.
55. Zanelli C.F., Valentini S.R. 2007. Is there a role for eIF5A in translation? *Amino Acids*. **33**, 351–358.
56. Regier J.C., Shultz J.W. 1998. Molecular phylogeny of arthropods and the significance of the Cambrian 'explosion' for molecular systematics. *Am. Zool.* **38**, 918–928.
57. Keeling P.J., Inagaki Y. 2004. A class of eukaryotic GT-Pase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1alpha. *Proc. Natl. Acad. Sci. USA*. **101**, 15380–15385.
58. Заренков Н.А. 1983. *Членистоногие. Ракообразные*. Часть II. М.: Из-во Моск. Ун-та.
59. Spears T., Abele L.G. 1998. Crustacean phylogeny inferred from 18S rDNA. In: *Arthropod Relationship. The Syst. Assoc. Special Vol. Ser. 55*. Eds Fortey R.A., Thomas R.H. London: Chapman & Hall, pp. 169–187.
60. Savard J., Tautz D., Lercher M.J. 2006. Genome-wide acceleration of protein evolution in flies (Diptera). *BMC Evol. Biol.* **6**, 7.
61. Nardi F., Spinsanti G., Boore J.L., Carapelli A., Dallai R., Frati F. 2003. Hexapod origins: monophyletic or paraphyletic? *Science*. **299**, 1887–1889.
62. Carapelli A., Liò P., Nardi F., van der Wath E., Frati F. 2007. Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evol. Biol.* **7**, Suppl. 2, S8.
63. Luan Y.X., Mallatt J.M., Xie R.D., Yang Y.M., Yin W.Y. 2005. The phylogenetic positions of three Basal-hexapod groups (protura, diplura, and collembola) based on ribosomal RNA gene sequences. *Mol. Biol. Evol.* **22**, 1579–1592.
64. Timmermans M.J., Roelofs D., Mariën J., van Straalen N.M. 2008. Revealing pancrustacean relationships: phylogenetic analysis of ribosomal protein genes places Collembola (springtails) in a monophyletic Hexapoda and reinforces the discrepancy between mitochondrial and nuclear DNA markers. *BMC Evol. Biol.* **8**, 83.
65. Adachi J., Hasegawa M. 1996. Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* **42**, 459–468.
66. Abascal F., Posada D., Zardoya R. 2007. MtArt: a new model of amino acid replacement for Arthropoda. *Mol. Biol. Evol.* **24**, 1–5.
67. Rosenberg M.S., Kumar S. 2003. Taxon sampling, bioinformatics, and phylogenomics. *Syst. Biol.* **52**, 119–124.
68. Heath T.A., Zwickl D.J., Kim J., Hillis D.M. 2008. Taxon sampling affects inferences of macroevolutionary processes from phylogenetic trees. *Syst. Biol.* **57**, 160–166.
69. Goremykin V., Moser C. 2009. Классификация генов семейства ERF арабидопсиса с помощью Байесовского анализа. *Молекуляр. биология*. **43**, (статья в настоящем номере).
70. Fenn J.D., Song H., Cameron S.L., Whiting M.F. 2008. A preliminary mitochondrial genome phylogeny of Orthoptera (Insecta) and approaches to maximizing phylogenetic signal found within mitochondrial genome data. *Mol. Phylogenet. Evol.* **49**, 59–68.
71. DeSalle R., Freedman T., Prager E.M., Wilson A.C. 1987. Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *J. Mol. Evol.* **26**, 157–164.
72. Историческое развитие класса насекомых. 1980. Под ред. Родендорфа Б.Б., Расницына А.П. *Тр. Палеонтол. Ин-та АН СССР*. Т. 175. М.: Наука.
73. Walossek D., Müller K.J. 1998. Cambrian "Orsten"-type arthropods and the phylogeny of Crustacea. In: *Arthropod Relationship. The Syst. Assoc. Special Vol. Ser. 55*. Eds Fortey R.A., Thomas R.H. London: Chapman & Hall, pp. 139–153.
74. Miller B.R., Crabtree M.B., Savage H.M. 1997. Phylogenetic relationships of the Culicomorpha inferred from 18S and 5.8S ribosomal DNA sequences (Diptera: Nematocera). *Insect Mol. Biol.* **6**, 105–114.
75. Pawlowski J., Szadziewski R., Kmieciak D., Fahrni J., Bittar G. 2003. Phylogeny of the infraorder Culicomorpha (Diptera: Nematocera) based on 28S RNA gene sequences. *Syst. Entomol.* **21**, 167–178.
76. Schram F.R. 1986. *Crustacea*. Oxford: Oxford University Press.