

Федеральное государственное бюджетное учреждение науки
Лимнологический институт Сибирского отделения Российской академии наук



БОЛБАТ АЛЕКСАНДР ВАСИЛЬЕВИЧ

СТРУКТУРА И ЭВОЛЮЦИЯ МИТОХОНДРИАЛЬНОГО ГЕНОМА РЕЛИКТОВЫХ ПИЯВОК

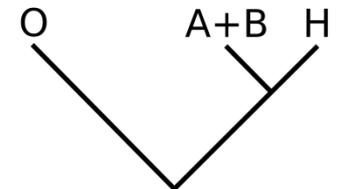
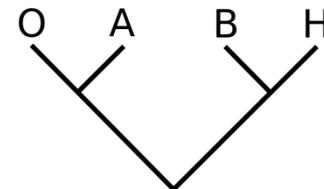
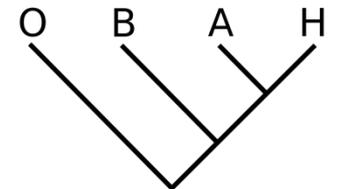
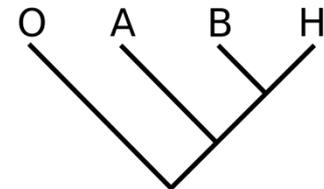
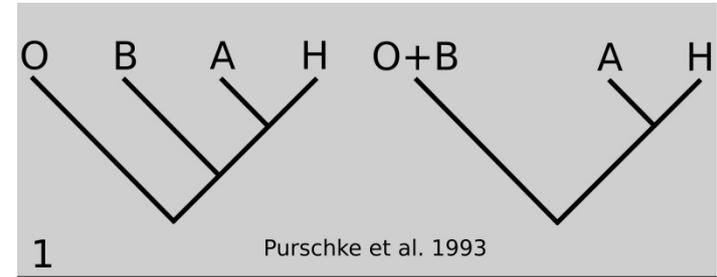
НАУЧНЫЙ ДОКЛАД ПО ДИССЕРТАЦИИ
НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ
КАНДИДАТА БИОЛОГИЧЕСКИХ НАУК
1.5.7 – ГЕНЕТИКА

Научный руководитель:
к.б.н. Кайгородова И. А.

Иркутск
2022

Актуальность работы

1. Реликтовые пиявкоподобные паразиты отряда Acanthobdellida остаются важным объектом исследований в эволюционной биологии и таксономических исследованиях. Научный интерес к ним поддерживается ввиду их предполагаемого промежуточного эволюционного положения между малощетинковыми червями (Oligochaeta) и пиявками (Hirudinea).
2. На протяжении многих лет вопрос происхождения и эволюционной истории акантобделлид не получалось решить с использованием морфологических данных и коротких маркерных последовательностей.
3. Митохондриальный геном животных является важным объектом для изучения ввиду его гаплоидного наследования и ограниченной рекомбинации, что позволяет с большей точностью реконструировать филогенетическую историю организма.



Цели и задачи

Цель – получение сведений о структуре полного митохондриального генома, степени генетической и таксономической дивергенции и эволюционной истории представителей отряда Acanthobdellida.

Для достижения цели были поставлены следующие задачи:

1. Получить образцы представителей акантобделлид из различных, географически разобщённых мест обитания.
2. С помощью методов секвенирования нового поколения получить геномные прочтения образцов и на их основе реконструировать структуру митохондриального генома акантобделлид.
3. Проанализировать и аннотировать полученные генетические последовательности.
4. Реконструировать на основе полных митогеномов эволюционную историю кольчатых червей, оценить степень генетической дивергенции и их применимость для молекулярной делимитации таксонов.

Биологический материал

Вид	Точка сбора	Номер в коллекции
<i>Acanthobdella peledina</i>	р. Питэльвен, Швеция	Б28
	р. Барбитай, Иркутская Область	Б33
	оз. Большой Харбей, Ямало-Ненецкий АО	Б35
	оз. Лабынкыр, республика Саха (Якутия)	Б52
	р. Рассоха, Иркутская Область	Б53
	р. Яна, республика Саха (Якутия)	Б55
<i>Paracanthobdella livanowi</i>	оз. Ажабачье, Камчатский Край	Б49
		Б56
<i>Codonobdella sp.</i>	оз. Байкал, залив Малое Море	Б45
		Б47
<i>Baicaloclepsis grubei</i>	оз. Байкал, залив Малое Море	Б57
<i>Baicaloclepsis echinulata</i>	оз. Байкал, пролив Ольхонские Ворота	Б58

Данные из международной базы SRA

Вид	Номер в базе SRA
<i>Erpobdella octoculata</i>	SRX9009198
<i>Glossiphonia complanata</i>	SRX8928147
<i>Haemopis sanguisuga</i>	SRX9009141
<i>Haemopis sanguisuga</i>	SRX9009400
<i>Lumbriculus variegatus</i>	SRX9009164
<i>Piscicola geometra</i>	SRX9009199
<i>Theromyzon tessulatum</i>	SRX8928146

Материалы и методы

- **ДНК** выделяли с использованием спин-колонок diaGene для культур клеток и фенол-хлороформным методом.
- **Секвенирование** проводили на Illumina NextSeq 500 с длиной прочтения 150 нуклеотидов в одну сторону.
- **Сборку** проводили с использованием MIRA v. 5. Митохондриальные контиги искали по гомологии с помощью программы BLAST среди контигов с наибольшим покрытием. Проблемные и повторяющиеся регионы были собраны из необработанных прочтений, извлечённых из данных секвенирования программой собственной разработки (fastq-grep).
- **Проверку качества** сборки проводили рекурсивным картированием прочтений на полученный геном с использованием Bowtie2.
- **Филогенетический анализ** проводили с использованием IQ-TREE и BEAST2 с моделью замен, рекомендованной jModelTest2, а также со схемой сегментации, рекомендованной PartitionFinder2. К полученным ультраметрическим деревьям был применён алгоритм видовой делимитации Generalized Mixed Yule Coalescent.

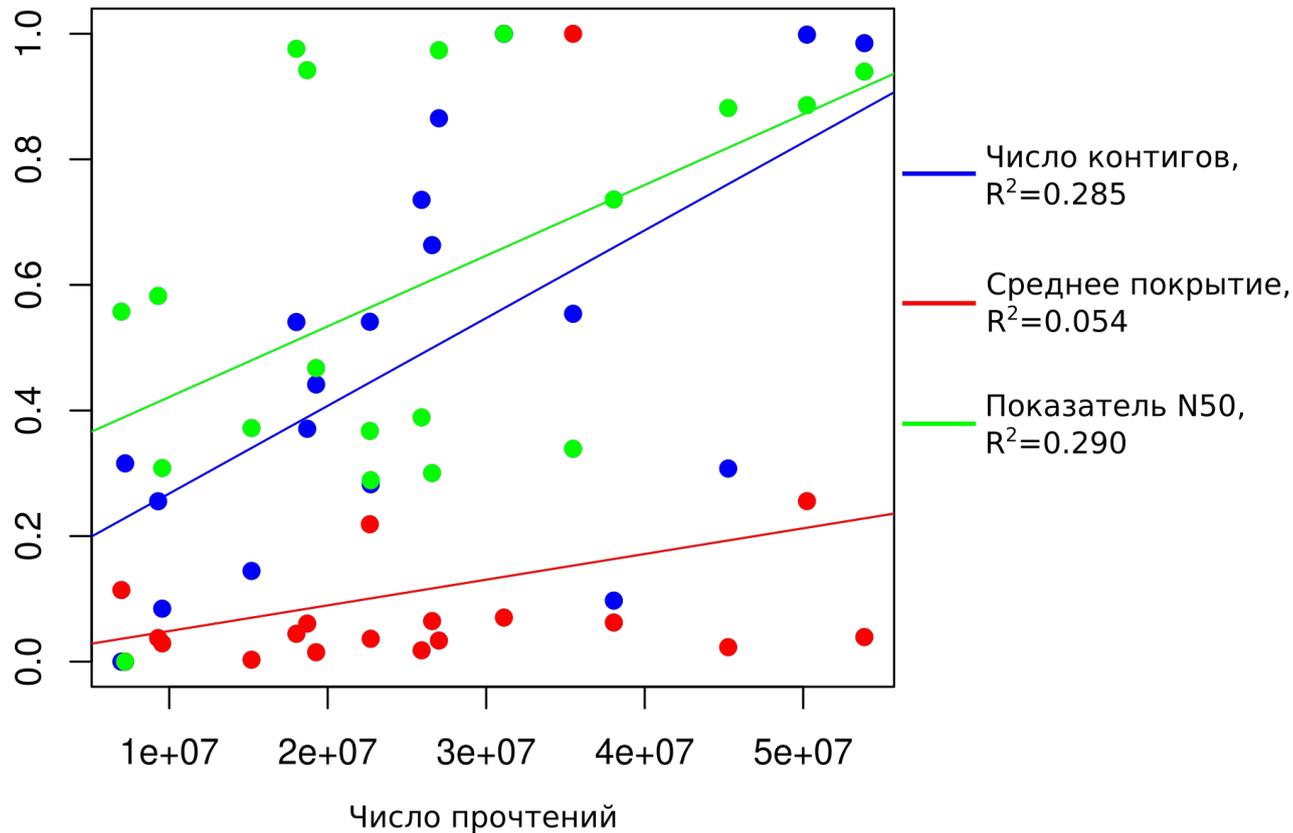
Статистические характеристики геномной сборки

Вид	Номер в коллекции	Число контигов	Среднее покрытие	N50	Размер митогенома
<i>Acanthobdella peledina</i>	Б28	491 622	26.48	459	18 528
	Б33	380 282	195.28	440	16 709*
	Б35	643 295	70.36	996	17 499
	Б52	642 489	226.28	896	16 388*
	Б53	634 709	44.26	943	17 271*
	Б55	380 058	48.59	975	16 547*
<i>Paracanthobdella livanowi</i>	Б49	566 108	39.60	973	15 500
	Б56	125 634	63.70	764	15 411
<i>Codonobdella sp.</i>	Б45	282 539	62.30	945	14 486
	Б47	323 026	23.93	528	14 573**
<i>Baicaloclepsis grubei</i>	Б57	450 235	65.65	381	14 751
<i>Baicaloclepsis echinulata</i>	Б58	152 709	13.92	444	15 120

Статистические характеристики геномной сборки

Вид	Номер SRA	Число контигов	Среднее покрытие	N50	Размер митогенома
<i>Erpobdella octoculata</i>	SRX9009198	231 899	41.93	371	15 580
<i>Glossiphonia complanata</i>	SRX8928147	69 873	107.25	607	15 468
<i>Haemopis sanguisuga</i>	SRX9009141	216 384	42.89	629	14 530
<i>Haemopis sanguisuga</i>	SRX9009400	250 961	11.32	117	14 462**
<i>Lumbriculus variegatus</i>	SRX9009164	246 237	30.71	892	15 883
<i>Piscicola geometra</i>	SRX9009199	387 593	852.10	415	14 788
<i>Theromyzon tessulatum</i>	SRX8928146	118 265	35.86	388	15 913

Зависимость показателей сборки от числа прочтений



Порядок генов

Polychaeta

Urechis	cox1	cox2	Pro	Asp	atp8	Thr	nd4L	nd4	Met	Asn	Gly	nd2	Tyr	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	12S	16S	Val	Ser	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Cys	Glu
Marenzelleria	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	Gln	Trp	Ser	Arg	cox3	nd6	cytb	atp6	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Ala	Met	12S	Val	16S	Leu	Leu	nd1	Ile	Lys	nd3	Ser	nd2
Alitta	cox1	Asn	cox2	Gly	Tyr	atp8	Met	Asp	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	x	12S	Val	16S	Leu	Ser	Ala	Leu	nd1	Ile	Lys	nd3	Ser	nd2
Glycera	cox1	Asn	cox2	Asp	atp8	Tyr	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Gly	Met	12S	Val	16S	Leu	Ser	Ala	Leu	nd1	Ile	Lys	nd3	Ser	nd2
Manayunkia	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Ser	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	16S	Val	Leu	Ala	nd1	Ile	Lys	nd3	Ser	nd2	
Pista	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2
Clymenella	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	Lys	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	nd3	Ser	nd2
Galathealinum	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2

Oligochaeta (почвенные)

Lumbricus	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2
-----------	------	-----	------	-----	------	-----	-----	------	-----	-----	------	-----	------	-----	-----	-----	-----	-----	-----	-----	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Oligochaeta (пресноводные)

Nais	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Glu	Pro	nd4L	nd4	Phe	Thr	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Tubifex	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Olavius	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	Trp	atp6	nd3	nd6	Ala	Leu	nd1	Ile	Ser	cytb	His	nd5	Phe	Lys	Arg	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	nd2		
Lumbriculus	cox1	Asn	cox2	Asp	atp8	Gly	cox3	Trp	atp6	cytb	Trp	His	Gln	Tyr	nd6	Tyr	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2

Acanthobdellida

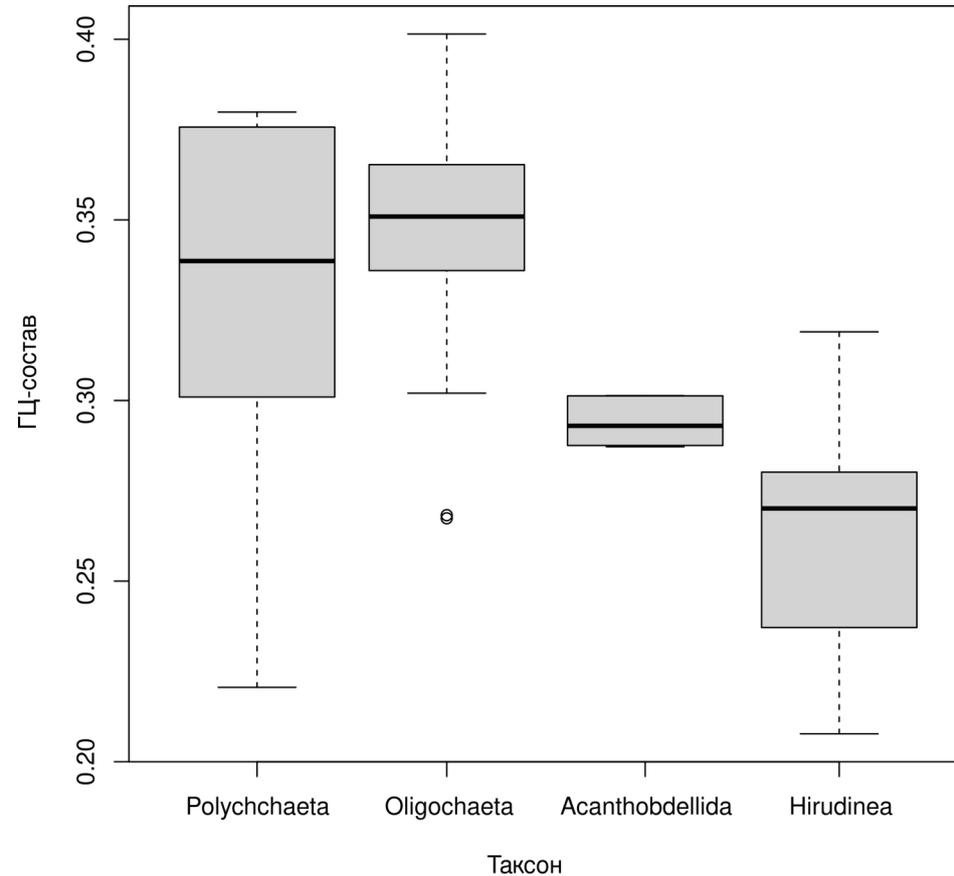
Acanthodella	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2
Paracanthobdella	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2

Hirudinea

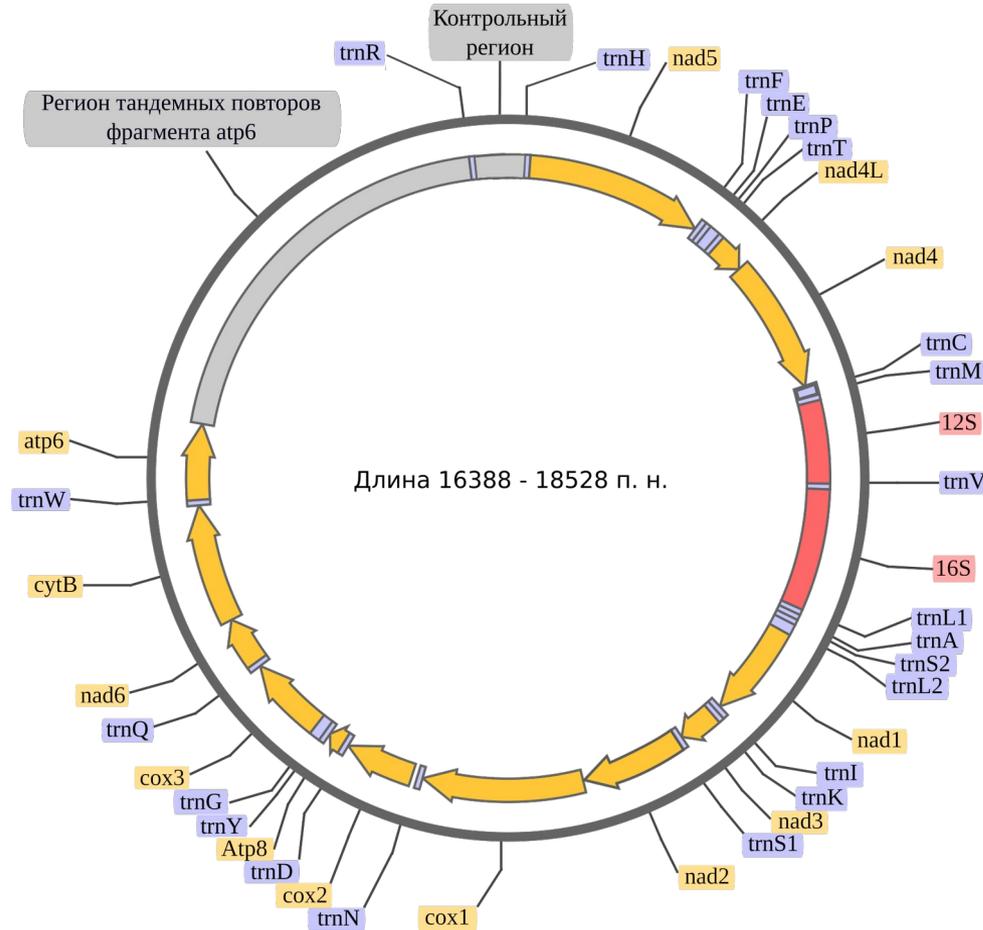
Glossiphonia	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Baicalocleipsis	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Theromyzon	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	His	Arg	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	Ser	nd2
Piscicola	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Codonobdella	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Erpobdella	cox1	Asn	cox2	Asp	atp8	Tyr	Gly	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Erpobdella*	cox1	Asn	cox2	Asp	atp8	Gly	Tyr	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ser	Ala	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Haemopsis	cox1	Asn	cox2	Asp	atp8	Gly	Tyr	cox3	Gln	nd6	cytb	Trp	atp6	X	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Whitmania	cox1	Asn	cox2	Asp	atp8	Gly	Tyr	cox3	Gln	nd6	cytb	Trp	atp6	Arg	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Cys	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Ile	Lys	nd3	Ser	nd2	
Haemadipsa	cox1	Asn	cox2	Asp	atp8	Gly	Tyr	cox3	Gln	nd6	cytb	Arg	Trp	atp6	His	nd5	Phe	Glu	Pro	Thr	nd4L	nd4	Met	12S	Val	16S	Leu	Ala	Ser	Leu	nd1	Lys	Ile	nd3	Cys	Ser	nd2	

* Erpobdella_octoculata_KC688270;

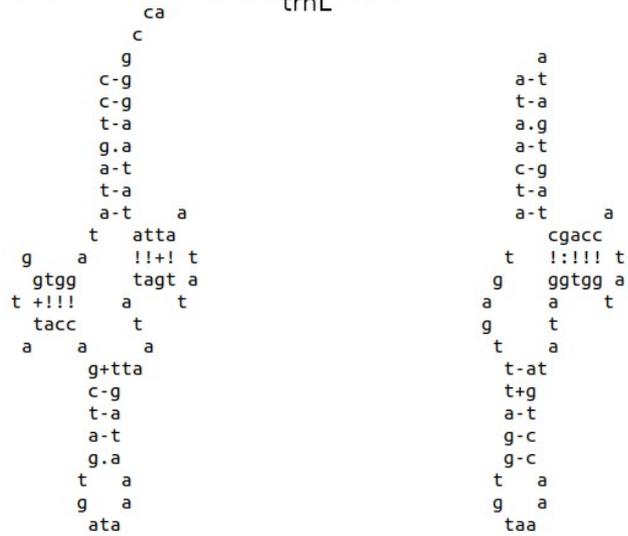
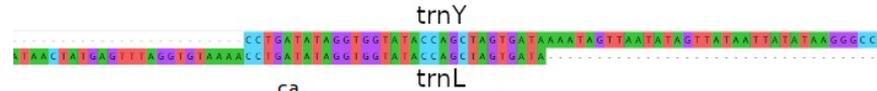
Доля ГЦ-оснований в разных группах КОЛЬЧЕЦОВ



Митогеном *A. peledina*



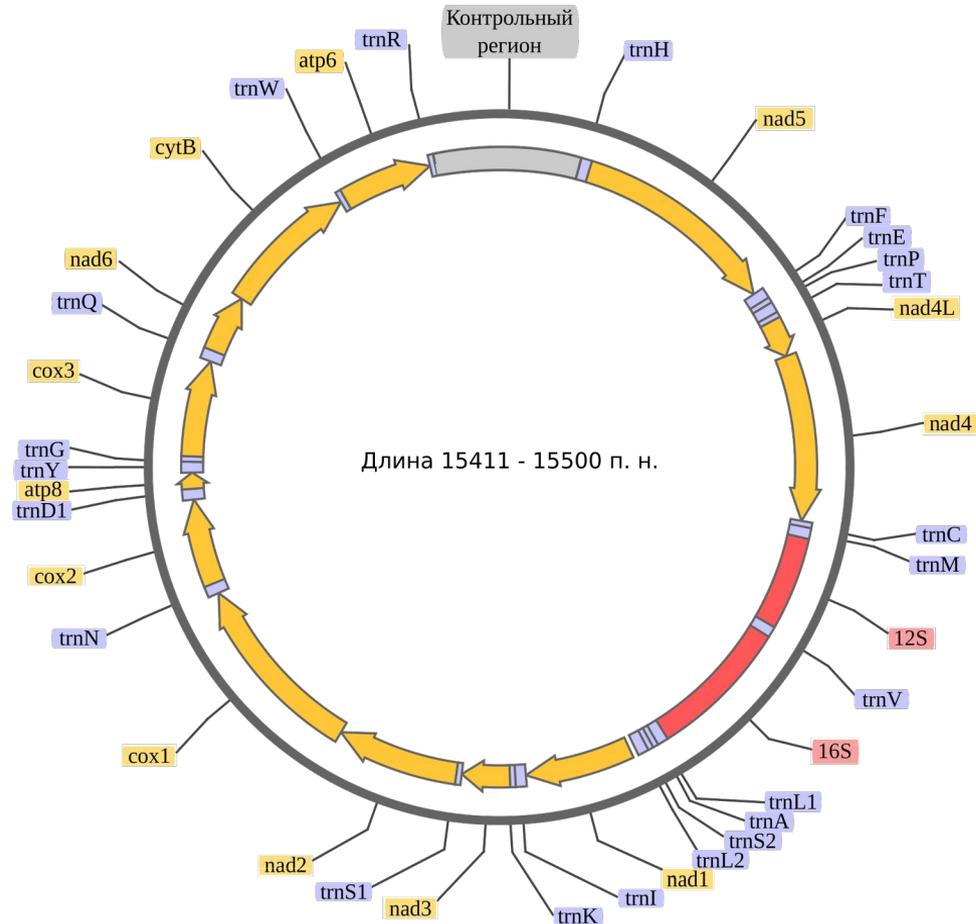
тРНК-подобные структуры *A. peledina*



mtRNA-Tyr(ata)
66 bases, %GC = 30.3
Sequence c[8450,8515]

D-loop mtRNA-Leu(taa)
55 bases, %GC = 34.5
Sequence c[8485,8539]

Митогеном *P. livanowi*



тРНК-подобные структуры *P. livanowi*

```

      c
      t-a
      t+g
      t-a
      a-t
      a-t
      t-a
      a.gtta
      t      t
      tt a      a
      c aaaa      a
      a !!!!      t
      g tttt      a
      g a      a
      a-tgaa
      t.t
      g-c
      t-a
      t-a
      t t
      t t
      gtg
  
```

TV-loop mRNA-His(gtg)
61 bases, %GC = 18.0
Sequence c[77,137]

```

      a
      t-a
      a-t
      g-c
      t.t
      t-a
      t-a
      t-at
      t      t
      tgg a      t
      a aca      g
      t !!!      a
      t tgt      t
      tta c      a
      tt-ac
      a-t
      a-t
      t-a
      t-a
      t t
      t a
      ta
  
```

TV-loop mRNA-?(Stop|Leu)(ta)
58 bases, %GC = 15.5
Sequence c[177,234]

```

      c
      t-a
      t+g
      g-c
      a a
      a-t
      a-t
      a-t c
      t ttac
      at !!!! c
      a aatg a
      t t a
      gc g
      a t
      g+ta
      g+t
      t-a
      a-t
      g-c
      t a
      t a
      ttt
  
```

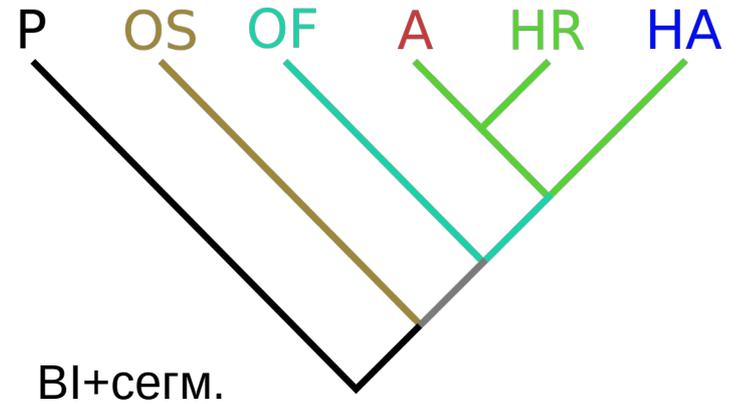
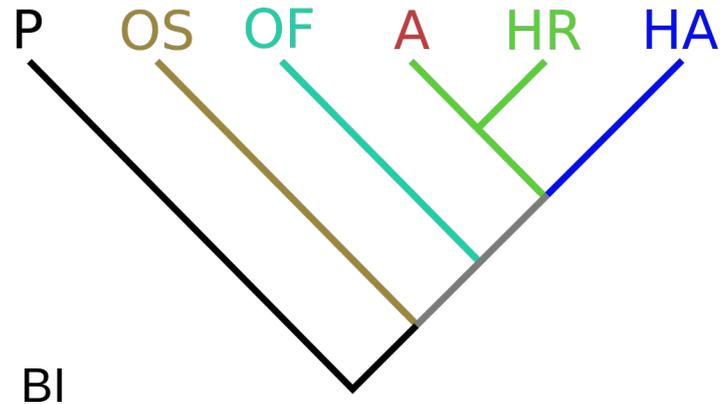
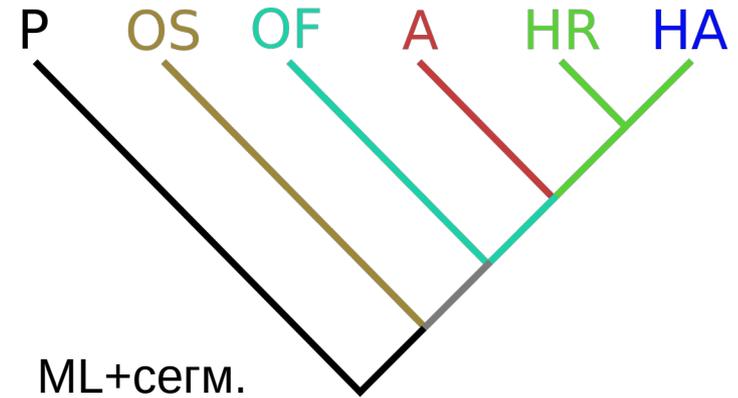
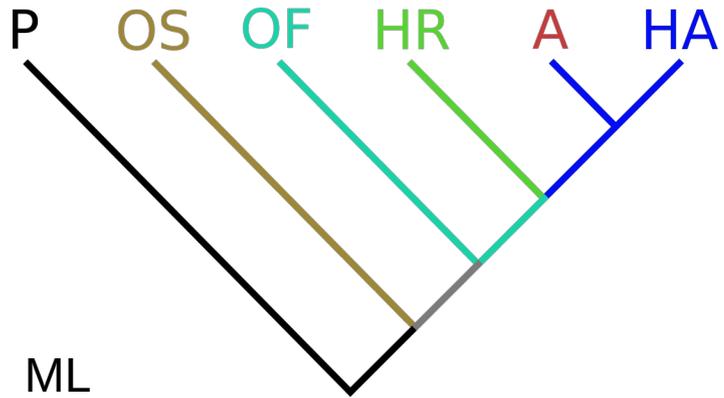
D-loop mRNA-Lys(ttt)
56 bases, %GC = 26.8
Sequence c[525,580]

```

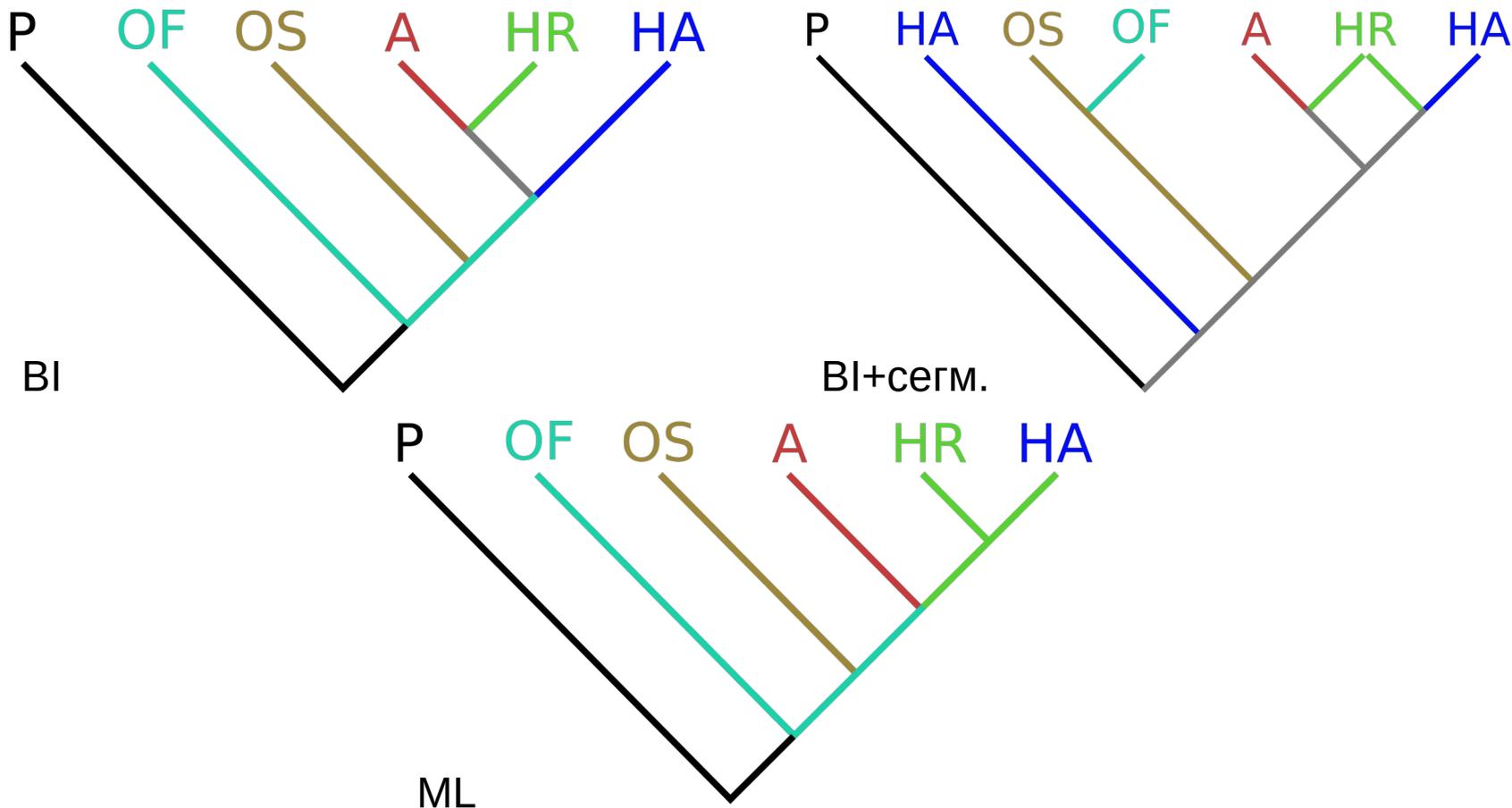
      a
      t-a
      a-t
      t-a
      t-a
      a-t
      c-g      ca
      tat atggc a
      at !:!!! a
      g ttccg g
      g g      tg
      at c
      ta g
      t-at
      g-c
      g.a
      a.g
      c t
      c a
      c a
      gat
  
```

D-loop mRNA-Ile(gat)
62 bases, %GC = 38.7
Sequence c[12626,12687]

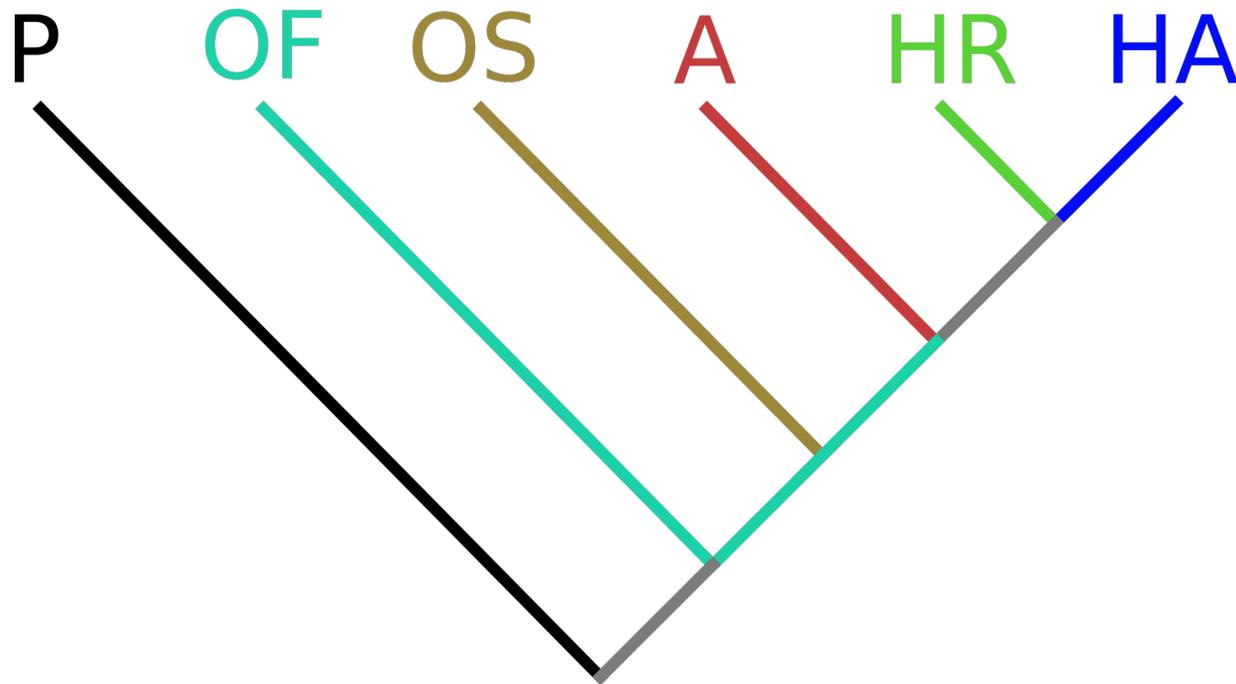
Кладограммы на основе гена *cox1*



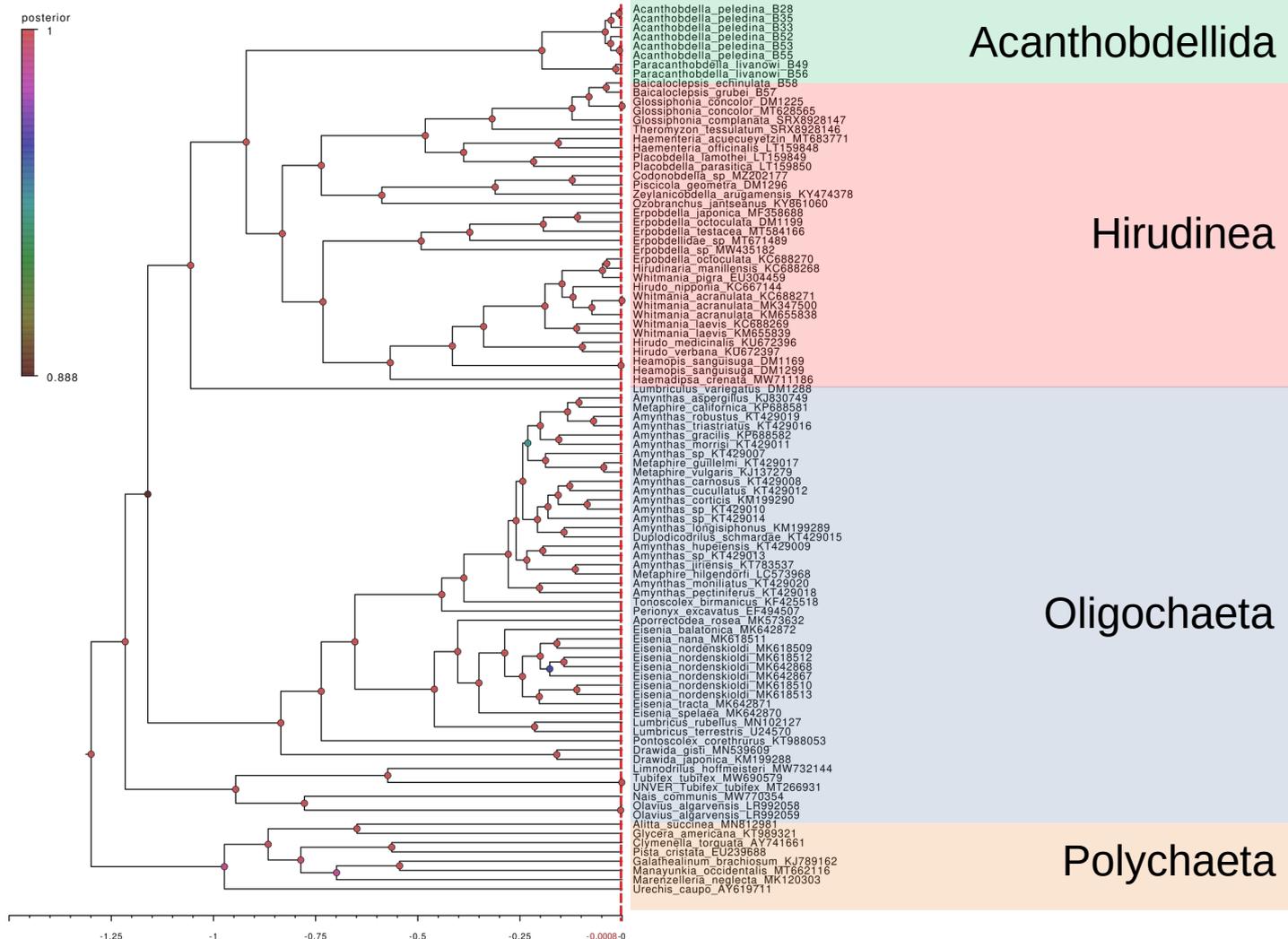
Кладограммы на основе гена *12S*



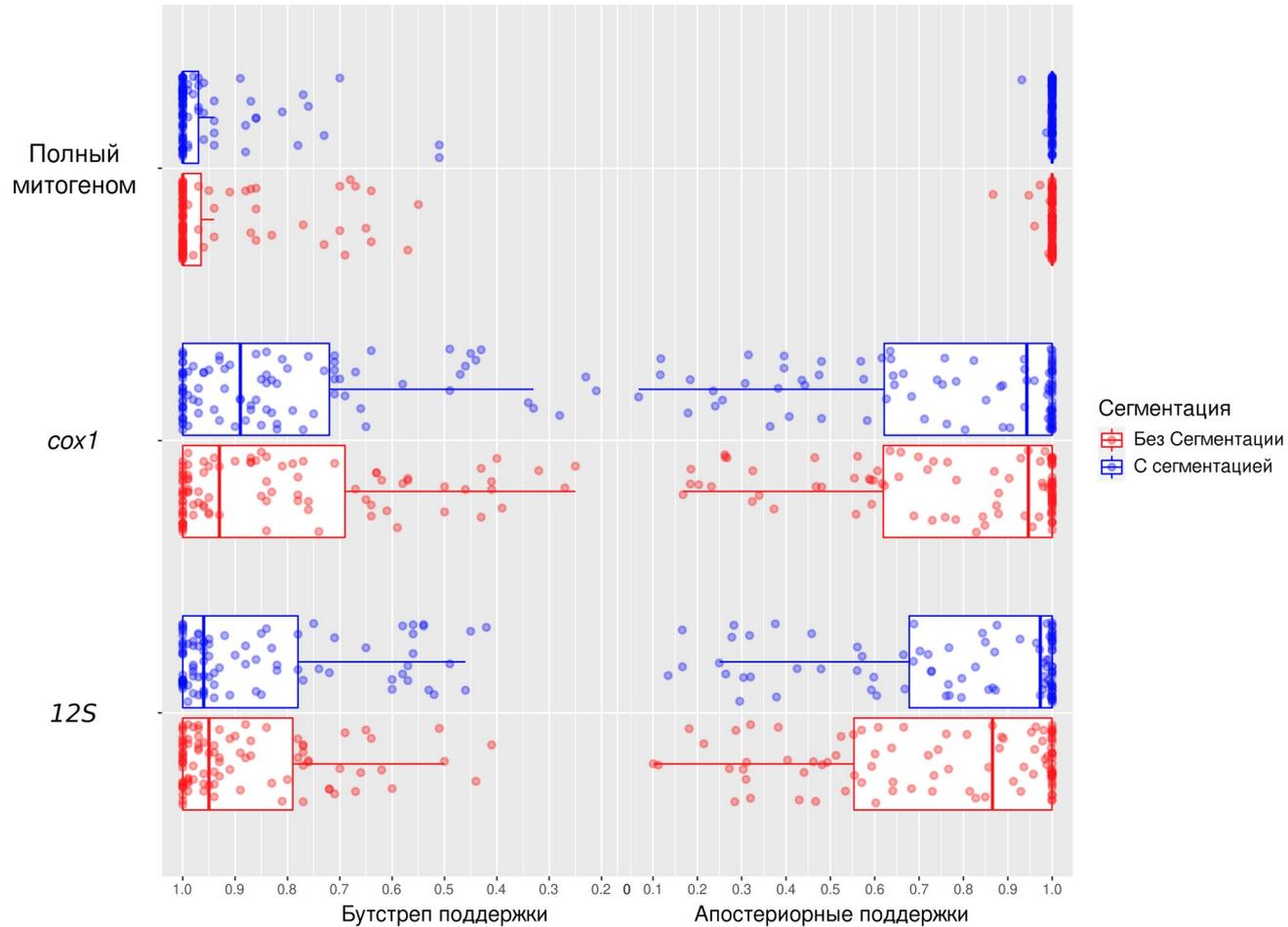
Кладограмма на основе полных митогеномов



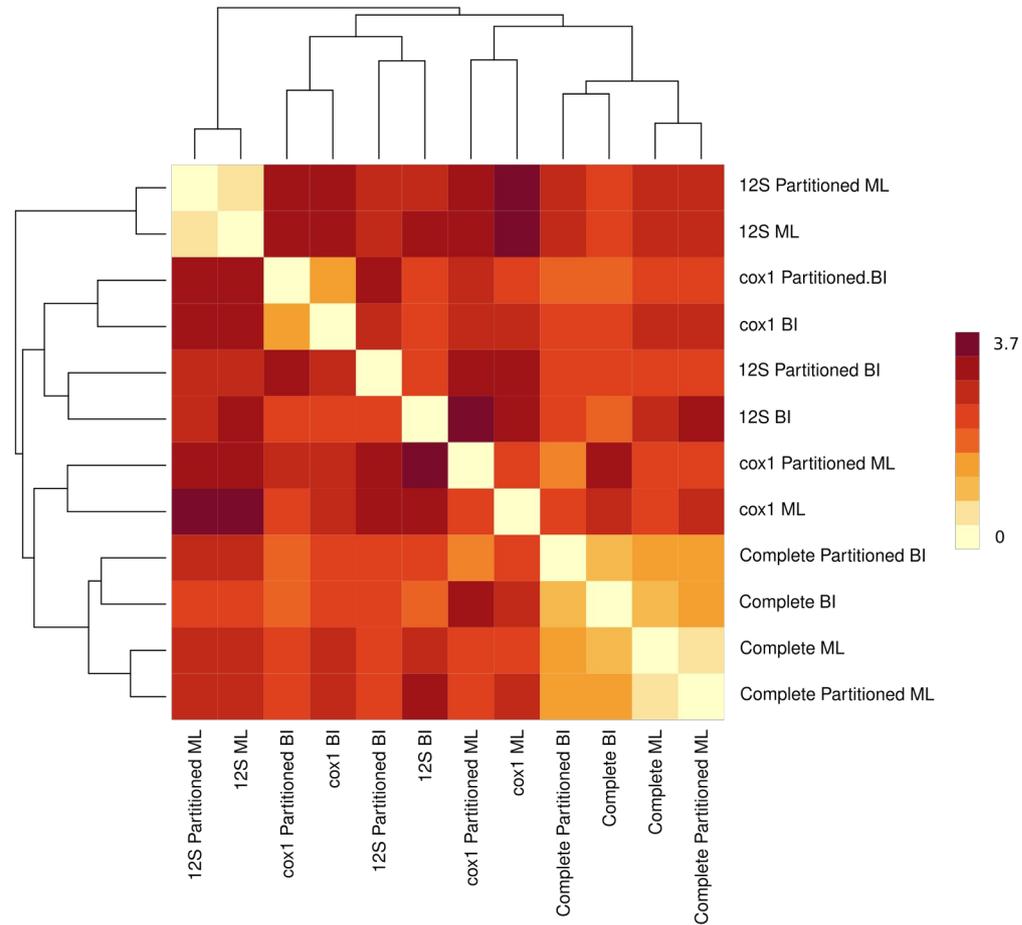
Древо на основе полных митогеномов



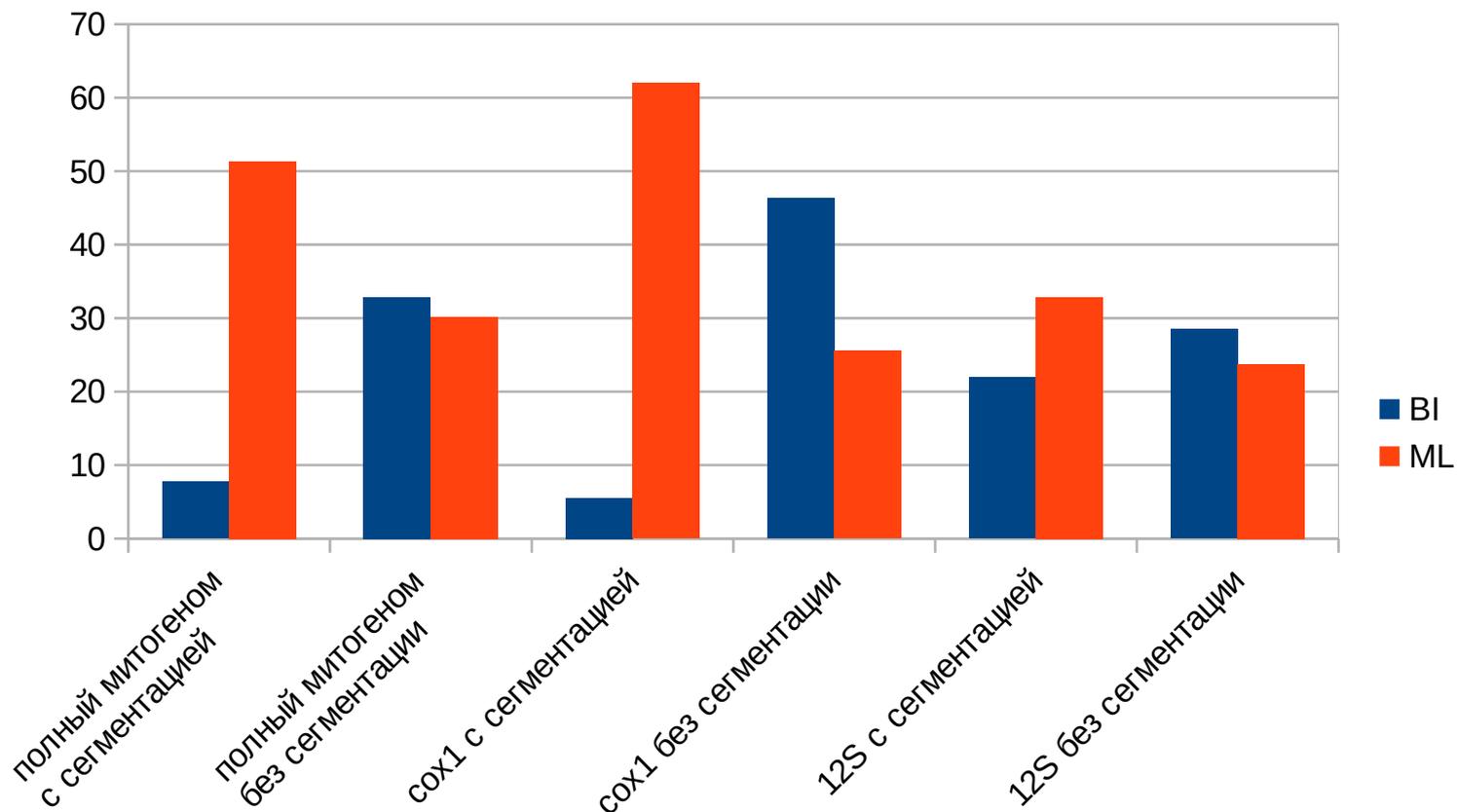
Сравнение поддержек филогений



Различия топологий деревьев



Сумма длин ветвей деревьев



Соответствие реконструированных филогенетических деревьев различным гипотезам об эволюции аннелид

Гипотеза \ Деревья	МГ		МГ		cox1		cox1		12S		12S	
	сегм.				сегм.				сегм.			
	ml	bi	ml	bi	ml	bi	ml	bi	ml	bi	ml	bi
Предок Clitellata – водный организм	+	+	+	+	-	-	-	-	+	-	+	+
Предок Acanthobdellida и Hirudinea – водный организм	+	+	+	+	+	+	+	+	+	-	+	+
Монофилетичность Rhynchobdellida	+	+	+	+	-	-	+	*	-	-	-	+
Монофилетичность Arhynchobdellida	+	+	+	+	+	+	+	+	**	-	**	+
Acanthobdellida – сестринская группа Hirudinea	+	+	+	+	-	-	-	-	**	-	**	-

Примечания: МГ – митогеном; «+» – древо соответствует гипотезе; «-» – древо не соответствует гипотезе; «*» – древо в целом соответствует гипотезе за исключением положения *O. jantseanus*; «**» – древо в целом соответствует гипотезе за исключением положения *O. algarvensis*.

Выводы

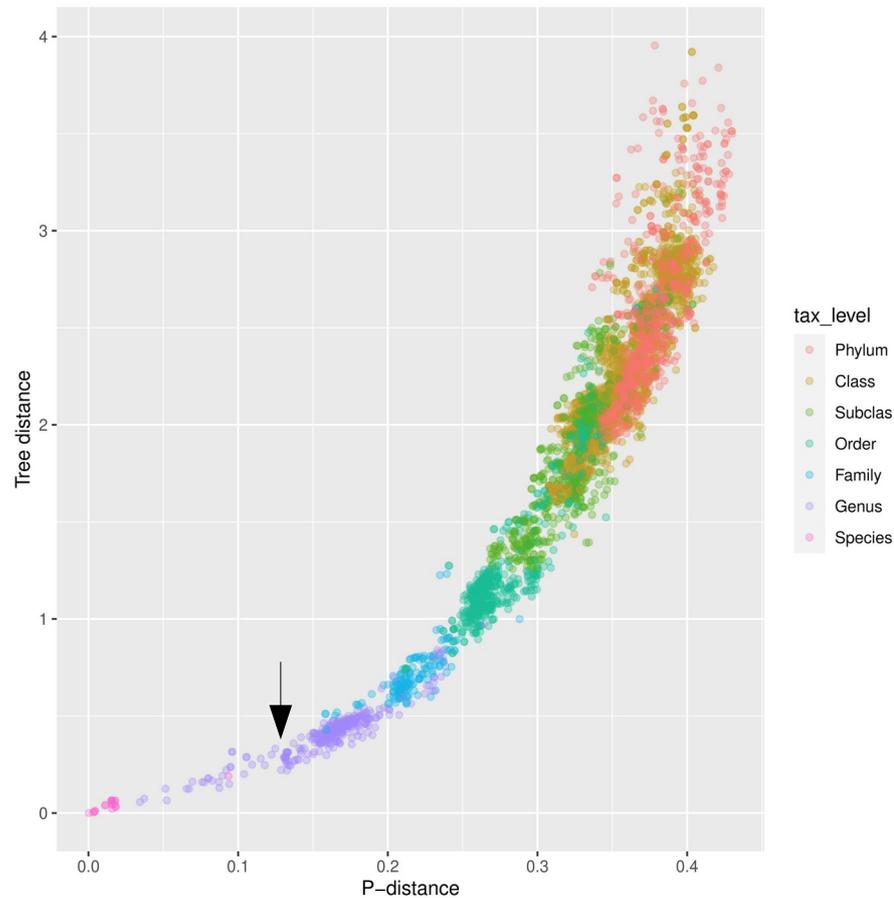
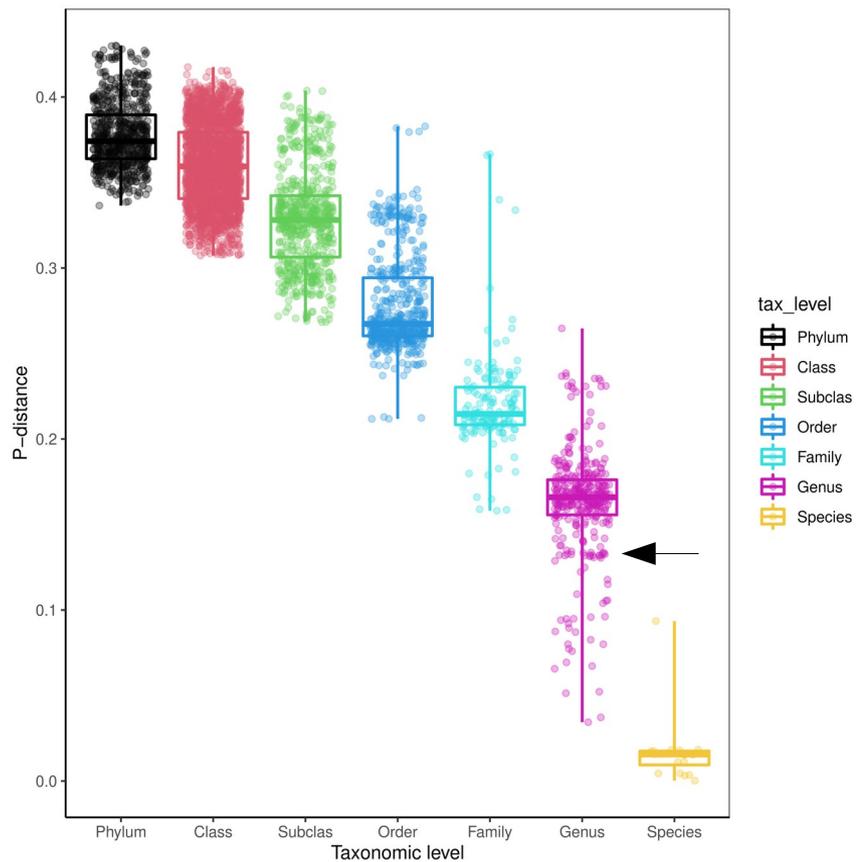
- 1) Впервые на основе данных секвенирования нового поколения реконструированы 12 митохондриальных геномов реликтовых (6 образцов *A. peledina* и 2 образца *P. livanowi*) и эндемичных байкальских пиявок (*Baicalocleipsis echinulata*, *Baicalocleipsis grubei* и два образца *Codonobdella* sp.).
- 2) Впервые реконструированы семь полных митогеномов олигохет и пиявок (*Lumbriculus variegates*, *Erpobdella octoculata*, *Glossiphonia complanata*, *G. concolor*, *Theromyzon tessulatum*, *Piscicola geometra* и 2 образца *Haemopsis sanguisuga*), необработанные геномные прочтения которых были получены из международной базы SRA.
- 3) Показано, что длина митохондриальных геномов камчатских образцов реликтовых пиявок *Paracanthobdella livanowi* составляет 15 411 и 15 500. Различия в длине нуклеотидных последовательностей не затрагивают функциональных генов и связаны с индивидуальной изменчивостью в контрольном регионе митогенома.
- 4) Выявлено, что размер митохондриальных геномов реликтового вида *Acanthobdella peledina*, имеющего обширный ареал, варьирует от 16 388 до 18 528 п. н. в зависимости от длины и количества tandemных повторов псевдогенных последовательностей *atp6* и *trnR* у образцов географически разрозненных популяций. Впервые у кольцецов обнаружен регион с tandemными повторами псевдогенных последовательностей *atp6* и *trnR*, который является уникальной структурной особенностью вида *A. peledina*.
- 5) Набор основных функциональных митохондриальных генов остаётся крайне консервативным у Clitellata. Однако порядок расположения этих генов в митогеномах, по-видимому, имеет таксон-специфический паттерн на уровне семейств.
- 6) Установлено, что длина и природа используемого для реконструкции филогении геномного фрагмента, а также метод обработки данных оказывают влияние на топологию и длины ветвей полученных филограмм. Использование полных митохондриальных геномов демонстрирует увеличение стабильности топологий филограмм, что отражается в повышении статистических поддержек узлов, по сравнению с короткими фрагментами маркерных генов вне зависимости от метода реконструкции и сегментации набора данных.
- 7) Реконструкция эволюционной истории на основе полных митохондриальных геномов подтвердила гипотезу о древних щетинконосных пиявках (*Acanthobdellida*) как промежуточной форме между *Oligochaeta* и *Hirudinea*.

СПАСИБО ЗА ВНИМАНИЕ

Защищаемые положения

- 1) Митохондриальный геном реликтовых пиявок представлен 37 функциональными генами, характерными для большинства многоклеточных животных, при этом геном *Acanthobdella peledina*, вида широко распространенного в Северной Евразии, содержит уникальные структурные элементы в виде варибельных тандемных повторов.
- 2) Использование полных митохондриальных геномов для реконструкции эволюционной истории акантобделлид стабильно подтверждают гипотезу об их происхождении от олигохетного предка и сестринском положении относительно пиявок.

Генетические дистанции на разных таксономических уровнях



Вывод

- С помощью статистического анализа генетических дистанций показано, что виды реликтовых пиявок принадлежат одному роду *Acanthobdella*.