

>> Mitochondrial retrograde signaling

Introduction

Человеческая митохондриальная ДНК кодирует 2 рРНК, 22 тРНК и 13 белков, всего 37 генов. Однако помимо них были найдены гены кодирующие 3 коротких пептида (Humanin, MOTS-C, Gau), предположительно участвующих в ретроградной сигнализации. Мы проанализировали эволюционную стабильность уже известных сигнальных пептидов и попытались выявить новых кандидатов на эту роль.

Human mtDNA encodes two rRNAs, 22 tRNAs, and 13 proteins subunits, 37 genes at all. Recently, genes encoding three short peptides have been discovered, the latter likely involved in retrograde signalling. We analyzed evolutionary stability of already known signal peptides and tried to find new ones.

Aims

Провести выравнивание генов несущих последовательности Humanin, MOTS-c и Gau, проследить их эволюционную историю. Идентифицировать таксоны имеющие эти пептиды
Найти все возможные открытые рамки считывания (OPC, ORF) в митохондриальных геномах разных видов, проанализировать их консервативность. Оценить стабильность лучших кандидатов.

Align genes harboring Humanin, MOTS-c and Gau reading frames. Reconstruct their evolutionary history. Identify taxa having these peptides. Find all possible sORFs in mitochondrial genomes of different origins. Analyze their conservation. Assess stability of candidate peptides.

Methods

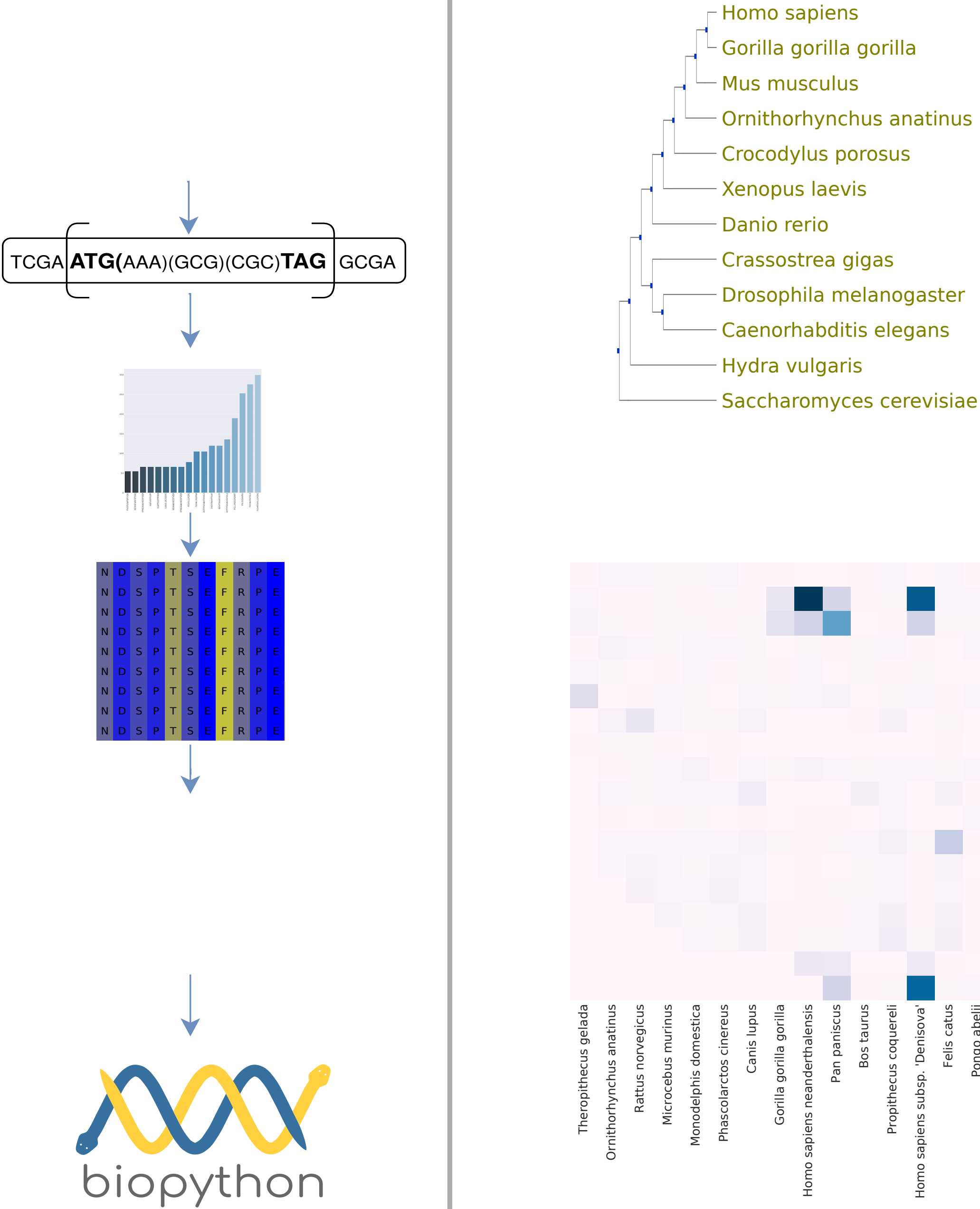


рис. 2 Тепловая карта, показывающая количество общих ORF в парах видов. (Более яркая окраска показывает большее количество общих ORFs)

fig. 2 Heatmap showing the number of common ORFs in pairs of species (More bright coloring indicates more joint ORFs)

Выбор полных митохондриальных геномов из базы данных NCBI. Использовались 68 видов из различных таксономических групп.

Поиск коротких ORF в трех рамках на прямой и обратной цепи ДНК

Отбор наиболее распространённых ORF.

Выравнивание выбранных участков, содержащих ORF (ClustalW)

Поиск похожих пептидов в базах данных (BLAST).

Анализ параметров ORF (наличие последовательности Шайна-Дальгарно (ШД), наличие шпилек РНК, GC-состав, dN/dS).

Предсказание стабильности транслируемых пептидов.

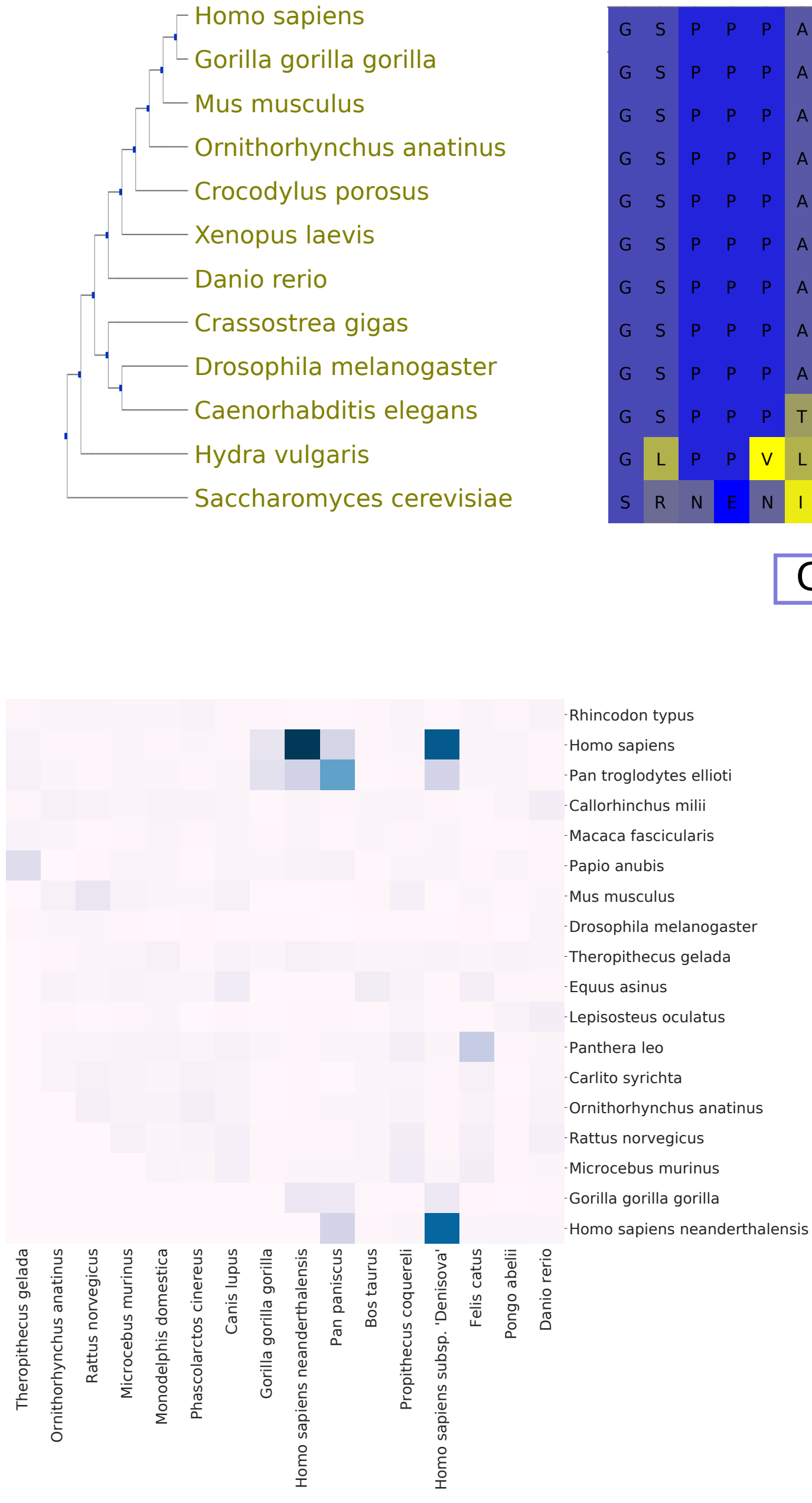
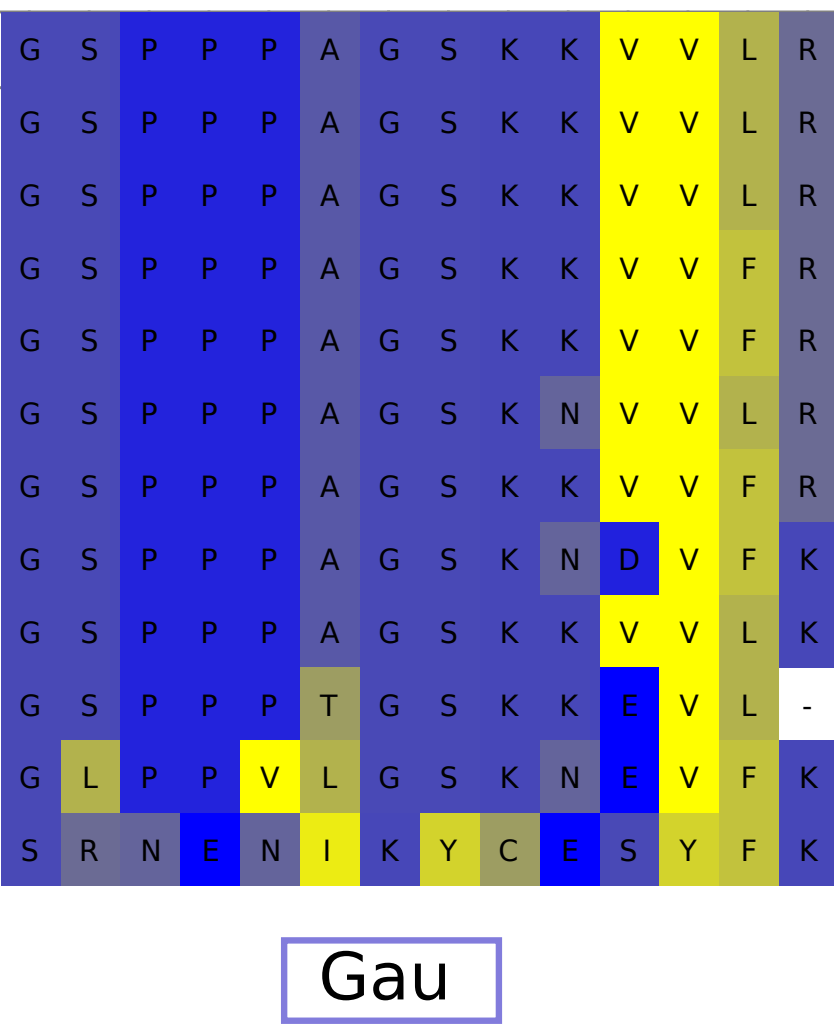


рис. 2 Тепловая карта, показывающая количество общих ORF в парах видов. (Более яркая окраска показывает большее количество общих ORFs)

fig. 2 Heatmap showing the number of common ORFs in pairs of species (More bright coloring indicates more joint ORFs)



Gau

MOTS-c

Humanin

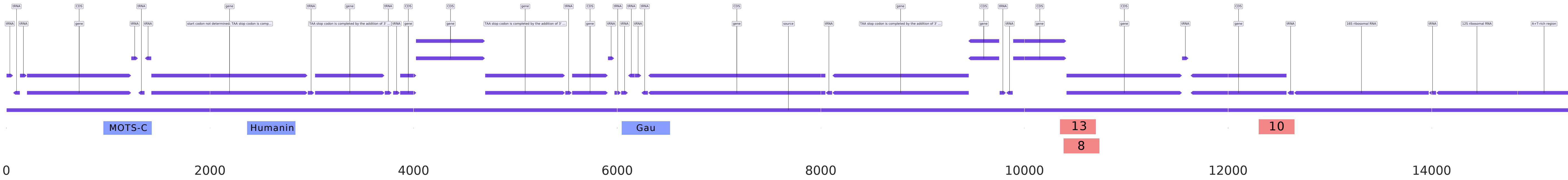
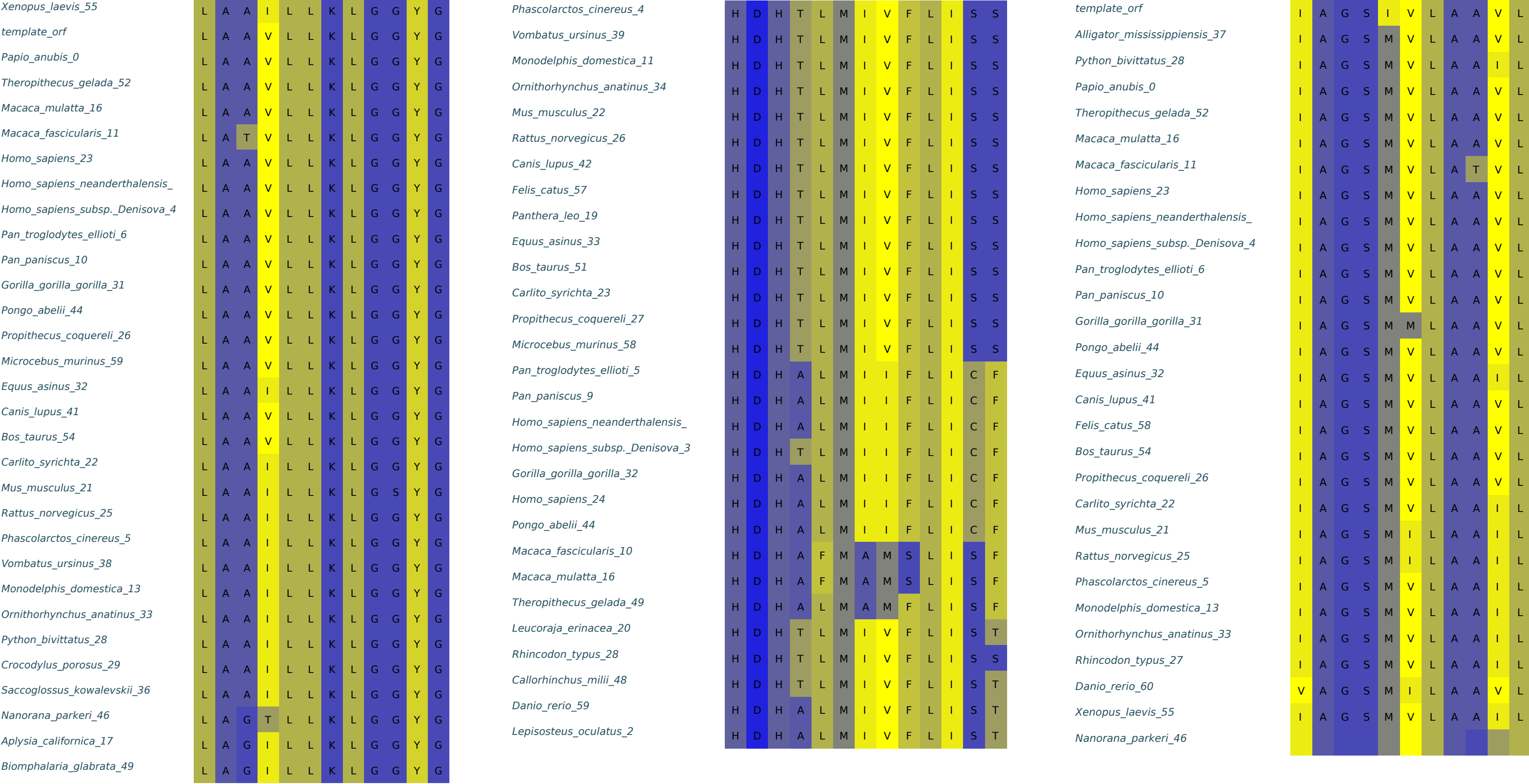
рис 1. выравнивание 16S rRNA с Humanin, cox1 с Gau и 12S rRNA с MOTS-C, дендрограмма с таксономическими группами

fig 1. alignment of the 16S rRNA region encoding Humanin, cox1 region Gau, and 12S rRNA encoding MOTS-C. dendrogram with taxonomy groups

Number of found ORFs	Predicted peptide	BLAST gene name	SD score	Srtand	Frame	Left border	Right border	Peptide lenght	Instability index	Stem loop
253	YVLISGAVFAI	cytochrome c oxidase subunit 2	0.0	1	1	6504	6555	17	19.1706	CTCC
136	IVLHDITYVVAHFHYVLS	cytochrome c oxidase subunit I	0.0	1	2	6616	6670	18	17.1667	CAACT
91	NVIVTAHAFVIFII	cytochrome c oxidase subunit 2	0.0	1	2	5686	5728	14	8.9071	ACGAT
91	DVOTRAYFTSAT	cytochrome oxidase subunit 1	0.0	1	2	2299	2344	15	22.0267	AGCTC
78	YVVAHFHYVLS	cytochrome c oxidase subunit I	0.0	1	0	7124	7178	18	17.1667	TAGCAAAT
78	HDITYVVAHFHYVLS	cytochrome c oxidase subunit I	0.0	1	0	7124	7178	18	17.1667	TAGCAAAT
66	NGAAIKGSFVQR	hypothetical protein DKP78_14280	0.000138	1	2	2536	2596	20	16.435	ACGAA
66	LA AVLKLG GYG	hypothetical protein	2.6819e-05	1	1	10678	10914	12	-8.85	CCATT
55	NDIKIVAFSTSSQ	NADH dehydrogenase subunit 5	0.0	1	0	12650	12692	14	18.6071	ATCT
45	HDHTLIIVFLISS	hypothetical protein	0.0	1	2	7081	7120	13	11.7154	GAGCTAAT
36	VELERFLNWWLL	hypothetical protein	0.00028	-1	3	1597	1693	32	7.8344	CCAATT
36	IVLANSLSLIVLHDITY	cytochrome oxidase subunit I	7.26e-07	1	0	6386	6440	18	9.2167	CGTAG
36	IAGSVLA AVL	hypothetical protein E155_019795	0.0	1	0	10664	10700	12	38.7083	ACTAGT
36	FAACEAAVGLAL	NADH dehydrogenase subunit 4L	0.0	1	0	10664	10700	12	38.7083	ACTAGT

Таб. 1 Таблица с предполагаемыми пептидами и их характеристиками. GC-состав и dN/dS не были включены из-за низкой информативности

Table 1. with possible peptides and their characteristics. GC-percent and dN/dS weren't included because of their low informativity



Conclusions and perspectives

<< Выравнивания последовательностей Humanin, MOTS-C и Gau оказались консервативными в группах Mammalia, Primates и Animalia соответственно.
<< По выравниваниям областей, содержащих консервативные ORF, были отобраны вероятные кандидаты на роль пептидов ретроградной сигнализации.
<< Для ORF были проанализированы основные параметры (наличие последовательностей Шайна-Дальгарно, шпильки в 5'-UTR (5'-нетранслируемая область РНК), GC-состав, dN/dS, стабильность).

Планируется проанализировать большее количество ORF среди уже найденных на аннотированных последовательностях и провести поиск в межгенных областях, выделить наиболее значимые параметры для предсказания пептидов ретроградной сигнализации, а также проанализировать данные RNAseq для подтверждения наличия соответствующих мРНК в митохондриях.

Acknowledgements

Мы благодарим Дмитрия Алексеевича Кнорре за идею проекта, Михаила Сергеевича Гельфанда за помощь в его развитии.
Также благодарим Ивана Кулаковского и Дмитрия Пензара за идеи о возможных путях проверки результатов и помощь в технических вопросах.