



Проект по биоинформатике

Группа 2

Anopheles

Малярийные комары

Участник	Вид	Условия обитания	GC содержание
Либина Яна	<i>Anopheles gambiae</i>	Стоячие пресные водоёмы	45%
Рябов Олег	<i>Anopheles darlingi</i>	Реки: прибрежье, заводы	48%
Векшин Кирилл	<i>Anopheles stephensi</i>	Стоячие пресные водоёмы	45%
Нелипович Софья	<i>Anopheles albimanus</i>	Стоячие пресные водоёмы	49%
Синицына Владислава	<i>Anopheles arabiensis</i>	Мелкие, хорошо прогреваемые водоемы	44.5%
Лазарев Никита	<i>Anopheles sinensis</i>	Теплые пресные водоёмы	44%
Ромашов Федор	<i>Anopheles funestus</i>	Мелководные пресные заросшие водоёмы	41.5%
Скворцова Ирина	<i>Anopheles ziemanni</i>	Пресные воды, холмы	44%
Волянский Никита	<i>Anopheles coluzzii</i>	Медленно проточные или стоячие водоемы	44.5%
Старостин Ярослав	<i>Anopheles aquasalis</i>	Мангровые болота, прибрежные лагуны	48%

Вид	ID	Длина генома	Кол-во генов
Anopheles gambiae	GCF_943734735.2	264.5 Mb	12519
Anopheles darlingi	GCF_943734745.1	181.6 Mb	12393
Anopheles stephensi	GCF_013141755.1	243.5 Mb	15394
Anopheles albimanus	GCF_013758885.1	172.6 Mb	12945
Anopheles arabiensis	GCF_016920715.1	256.8 Mb	15763
Anopheles sinensis	GCA_000441895.2	220.8 Mb	19352
Anopheles funestus	GCF_943734845.2	250.7 Mb	14819
Anopheles ziemanni	GCF_943734765.1	270 Mb	14085
Anopheles coluzzii	GCF_943734685.1	262.6 Mb	14579
Anopheles aquasalis	GCF_943734665.1	170.5 Mb	12989

Параметры поиска вторичных структур

- Z-hunt:

!./zhunt3 12 8 12 genomic.fna, порог 400

- Z-DNABERT:

Стандартные параметры:

```
model = 'HG kouzine'  
model_confidence_threshold = 0.5  
minimum_sequence_length = 10
```

- Квадруплексы:

```
pattern="(?:G{3,5}[ATGC]{1,7}){3,}G{3,5}"  
pattern_minus="(?:C{3,5}[ATGC]{1,7}){3,}C{3,5}"
```

Число квадруплексов

Вид	Экзоны	Интроны	Промотеры	Downstream	Межгенные
<i>Anopheles gambiae</i>	1298	6051	566	119	10696
<i>Anopheles darlingi</i>					
<i>Anopheles stephensi</i>	3258	4856	381	78	5812
<i>Anopheles albimanus</i>	851	32	162	29	2167
<i>Anopheles arabiensis</i>	2246	10466	845	157	7151
<i>Anopheles sinensis</i>	945	1699	914	288	9243
<i>Anopheles funestus</i>	1073	5858	473	86	5964
<i>Anopheles ziemanni</i>	395	7831	273	285	12392
<i>Anopheles coluzzii</i>	1356	13441	620	520	12433
<i>Anopheles aquasalis</i>	1686	7497	574	186	9864

Доля квадруплексов

Вид	Экзоны	Интроны	Промотеры	Downstream	Межгенные
<i>Anopheles gambiae</i>	0.05	0.24	0.02	0.004	0.4
<i>Anopheles darlingi</i>					
<i>Anopheles stephensi</i>	0.082	0.247	0.026	0.009	0.425
<i>Anopheles albimanus</i>	0.148	0.005	0.028	0.005	0.376
<i>Anopheles arabiensis</i>	0.114	0.53	0.043	0.008	0.362
<i>Anopheles sinensis</i>	0.072	0.129	0.069	0.022	0.706
<i>Anopheles funestus</i>	0.084	0.476	0.037	0.007	0.466
<i>Anopheles ziemanni</i>	0.19	0.37	0.01	0.01	0.59
<i>Anopheles coluzzii</i>	0.0506	0.502	0.023	0.019	0.46
<i>Anopheles aquasalis</i>	0.065	0.29	0.02	0.007	0.38

Число Z-ДНК

Вид	Экзоны	Интроны	Промотеры	Downstream	Межгенные
<i>Anopheles gambiae</i>	57699	142425	17300	3866	107631
<i>Anopheles darlingi</i>					
<i>Anopheles stephensi</i>	175204	105828	14289	2425	69180
<i>Anopheles albimanus</i>					
<i>Anopheles arabiensis</i>	51298	93427	11627	1632	61428
<i>Anopheles sinensis</i>	21058	13155	9045	2091	71437
<i>Anopheles funestus</i>	2770	40686	6560	2646	28234
<i>Anopheles ziemanni</i>	13288	60152	5483	2963	68518
<i>Anopheles coluzzii</i>	6625	24083	1185	785	13869
<i>Anopheles aquasalis</i>	19282	15707	3262	353	13993

Доля Z-ДНК

Вид	Экзоны	Интроны	Промотеры	Downstream	Межгенные
<i>Anopheles gambiae</i>	0.19	0.47	0.06	0.01	0.36
<i>Anopheles darlingi</i>					
<i>Anopheles stephensi</i>	0.58	0.30	0.10	0.011	0.38
<i>Anopheles albimanus</i>					
<i>Anopheles arabiensis</i>	0.25	0.46	0.06	0.008	0.3
<i>Anopheles sinensis</i>	0.18	0.11	0.07	0.017	0.61
<i>Anopheles funestus</i>	0.3	0.44	0.07	0.03	0.3
<i>Anopheles ziemanni</i>	0.09	0.44	0.04	0.02	0.5
<i>Anopheles coluzzii</i>	0.16	0.59	0.03	0.02	0.34
<i>Anopheles aquasalis</i>	0.5	0.41	0.08	0.009	0.36

Семейства эпигенетических генов

1. ALKBH1 (PF13532) - DNA demethylation, RNA demethylation
2. BAP1 (PF01088) - Histone deubiquitination
3. HDAC10 (PF00850) - Histone acetylation
4. EEF1AKMT4 (PF08241) - Protein methylation
5. JADE1 (PF10513, PF13831, PF13832) - Histone acetylation
6. SIRT7 (PF02146) - Histone acetylation
7. SMYD1 (PF00856, PF01753) - Histone methylation
8. TDG (PF03167) - DNA hydroxymethylation
9. TSSK6 (PF00069) - Histone phosphorylation
10. ZNHIT1 (PF04438) - Histone acetylation

Эпигенетические гены по семействам

- Геном: *Anopheles gambiae* (Либина Яна)

Семейство	Функция	Найденные гены	Кол-во генов
ALKBH1	DNA demethylation	LOC11175806, LOC1272107, ...	6
BAP1	Histone deubiquitination	LOC133393025, LOC1280562	2
HDAC10	Histone acetylation	LOC11175620, LOC1271709, ...	6
EEF1AKMT4	Protein methylation	LOC1275267, LOC1271115, ...	31
JADE1	Histone acetylation	LOC1270953, LOC1271324, ...	16
SIRT7	Histone acetylation	LOC1273018, LOC1271515, LOC1281716	3
SMYD1	Histone methylation	LOC1277297, LOC1277307, ...	56
TDG	DNA hydroxymethylation	LOC11175560	1
TSSK6	Histone phosphorylation	LOC4576161, LOC133393867, ...	190
ZNHIT1	Histone acetylation	LOC1278102 , LOC1269056, LOC1280794	3

Таблицы с распределением найденных структур

- Геном: *Anopheles gambiae* (Либина Яна)
- https://github.com/yalibina/bioinf_project_ind2025

Участок	Число квадруплексов в участке	Доля квадруплексов в участке	Число предсказаний Zhunt в участке	Доля предсказаний Zhunt в участке	Число предсказаний Z-DNABERT в участке	Доля предсказаний Z-DNABERT в участке
Exons	1298	0.05	57699	0.19	16153	0.18
Introns	6051	0.24	142425	0.47	28792	0.41
Promoters (1000 bp from TSS)	566	0.02	17300	0.06	5761	0.08
Downstream (200 bp)	119	0.004	3866	0.01	872	0.009
Intergenic	10696	0.40	107631	0.36	29716	0.41

Таблицы с распределением найденных структур

- Геном: *Anopheles gambiae* (Либина Яна)
- https://github.com/yalibina/bioinf_project_ind2025

Фоновое распределение:	
Участок	Доля в геноме
Exons	0.18
Introns	0.43
Promoters (1000 up from TSS)	0.06
Downstream (200 bp)	0.01
Intergenic	0.40

Таблицы с распределением найденных структур

- Геном *Anopheles gambiae*

Участок	Число участков с квадруплексом	Доля участков с квадруплексом	Число участков с предсказанием Zhunt	Доля участков с предсказанием Zhunt	Число участков с предсказанием Z-DNABERT	Доля участков с предсказанием Z-DNABERT
Exons	1120	0.02	23153	0.35	7030	0.33
Introns	3621	0.07	17518	0.32	8139	0.31
Promoters (1000 bp from TSS)	566	0.04	7998	0.13	3687	0.14
Downstream (200 bp)	119	0.008	3068	0.2	813	0.15
Intergenic	2574	0.22	6807	0.57	2546	0.58

Anopheles darlingi (Рябов Олег)

- Идентификатор в базе: GCF_943734745.1
- Длина генома: 181.6 Mb
- Число генов: 12 393
- Число найденных квадруплексов: -
- Число найденных Z-DNA: -



Эпигенетические гены по семействам

- Геном: Anopheles darlingi (Рябов Олег)

Семейство	Функция	Найденные гены	Кол-во генов
ALKBH1	DNA demethylation	LOC125956712, ...	3
BAP1	Histone deubiquitination	LOC125958170, ...	3
UBE2D3	Histone modification write	LOC125950237, ...	23
EEF1AKMT4	Protein methylation	LOC125953901, ...	27
PRMT8	Histone modification write	LOC125953569, ...	17
SAP18	Histone modification erase cofactor	LOC125950479	1
SIRT2	Histone modification erase, Histone modification write cofactor	LOC125948918, ...	2
SUDS3	Histone modification erase cofactor	LOC125949030	4
TSSK6	Histone phosphorylation	LOC125949014, ...	168
ZNHIT1	Histone acetylation	LOC125959484, ...	4

Таблицы с распределением найденных структур

- Геном; *Anopheles darlingi*

Anopheles stephensi (*Векшин Кирилл*)

- Идентификатор в базе:
- Длина генома: 243.5 Mb
- Число генов: 15 394
- Число найденных квадруплексов: 17318
- Число найденных Z-DNA: 218776



Таблицы с распределением найденных структур

- Геном: *Anopheles stephensi*

Таблица 1, количества попаданий

Участок	Число квадруплексов	Доля квадруплексов	Число предсказаний Zhunt	Доля предсказаний Zhunt	Число предсказаний ZDNABERT	Доля предсказаний ZDNABERT
Exons	3258	0.22	175204	0.49	42358	0.56
Introns	4856	0.33	105828	0.29	28741	0.38
Promoters (1000 up from TSS)	381	0.026	14289	0.04	5124	0.068
Downstream (200 bp)	78	0.005	2425	0.007	647	0.009
Intergenic	5812	0.40	69180	0.19	31285	0.41

Таблицы с распределением найденных структур

- Геном: *Anopheles stephensi*

Таблица 2, количества участков с попаданиями

Участок	Число участков с квадруплексом	Доля участков с квадруплексом	Число участков с предсказаниями Zhunt	Доля участков с предсказаниями Zhunt	Число участков с предсказаниями ZDNABERT	Доля участков с предсказаниями ZDNABERT
Exons	2147	0.21	48231	0.15	11247	0.18
Introns	3278	0.32	28714	0.09	8912	0.14
Promoters (1000 bp from TSS)	429	0.042	7845	0.024	3218	0.051
Downstream (200 bp)	92	0.009	1532	0.0047	418	0.0066
Intergenic	4236	0.42	18724	0.058	9231	0.15

Эпигенетические гены по семействам

- Геном: *Anopheles stephensi* (Векшин Кирилл)

Семейство	Функция	Найденные гены	Кол-во генов
ALKBH1	DNA demethylation	LOC118502523...	3
BAP1	Histone deubiquitination	LOC118512888...	4
HDAC10	Histone acetylation	LOC118510456 ...	6
EEF1AKMT4	Protein methylation	LOC1275267 ...	46
JADE1	Histone acetylation	LOC1270953...	48
SIRT7	Histone acetylation	LOC1273018...	6
SMYD1	Histone methylation	LOC1277297 ...	77
TDG	DNA hydroxymethylation	LOC118504424	1
TSSK6	Histone phosphorylation	LOC4576161...	536
ZNHIT1	Histone acetylation	LOC1278102...	8

Anopheles sinensis (Лазарев Никита)

- Идентификатор в базе:
- Длина генома: 220.8 Mb
- Число генов: 19 352
- Число найденных квадруплексов: 13089
- Число найденных Z-DNA: 116786



Таблицы с распределением найденных структур

Первая таблица, количество попаданий:

Region	G4	zhunt	zdnabert	all	G4_fraction	zhunt_fraction	zdnabert_fraction
exon	945	5365	21058	57780	0.072198	0.051961	0.180313
intron	1699	3648	13155	38436	0.129804	0.035331	0.112642
upstream	914	3144	9045	19352	0.069830	0.030450	0.077449
downstream	288	744	2091	19352	0.022003	0.007206	0.017905
intergenic	9243	90350	71437	28937	0.706165	0.875052	0.611691
all	13089	103251	116786	NaN	NaN	NaN	NaN

Таблицы с распределением найденных структур

Вторая таблица, количество участков с попаданиями:

Region	num_regions	with_G4	G4_fraction	with_zhunt	zhunt_fraction	with_zdnabert	zdnabert_fraction
exon	57780	882	0.015265	356	0.006161	13782	0.238525
intron	38436	1141	0.029686	205	0.005334	4855	0.126314
upstream	19352	939	0.048522	260	0.013435	7298	0.377119
downstream	19352	347	0.017931	89	0.004599	2270	0.117301
intergenic	28937	5080	0.175554	3688	0.127449	16458	0.568753

Anopheles albimanus (Нелипович Софья)

- Идентификатор в базе: GCF_013758885.1
- Длина генома: 172.6 Mb
- Число генов: 12 945
- Число найденных квадруплексов: 5760
- Число найденных Z-DNA:



Таблицы с распределением найденных структур

- Геном: *Anopheles albimannus*

Участок	Число квадруплексов	Доля квадруплексов	Число предсказаний Zhun	Доля предсказаний Zhun	Число предсказаний ZDNABERT	Доля предсказаний ZDNABERT
Exons	851	0.148				
Introns	32	0.005				
Promoters (1000 up from TSS)	162	0.028				
Downstream (200 bp)	29	0.005				
Intergenic	2167	0.376				

Anopheles arabiensis (Синицына Владислава)

- Идентификатор в базе: GCF_016920715.1
- Длина генома: 256.8 Mb
- Число генов: 15 763
- Число найденных квадруплексов: 19765
- Число найденных Z-DNA: 202915



Эпигенетические гены по семействам

- Геном: *Anopheles arabiensis*

Проверяемое семейство	Найденный ген
TSSK6	LOC120904707
BAP1	LOC120900743
ZNHIT	LOC120893243
EEF1AKMT4	LOC120893343
ALKBH1	LOC120893648
ATAD2	LOC120893485
CRUM2	LOC120893478
CLOCK	LOC120893672
HS71A	LOC120893250
TDRKH	LOC120893722

Таблицы с распределением найденных структур

Геном: *Anopheles arabiensis*

Таблица 1.

Участок	Число квадруплексов	Доля квадруплексов	Число предсказаний Zhunt	Доля предсказаний Zhunt	Число предсказаний ZDNABERT	Доля предсказаний ZDNABERT
exon	2246	0.1136	1304264	0.1957	51298	0.2528
intron	10466	0.5295	3361662	0.5044	93427	0.4604
promoters	845	0.0428	334950	0.0503	11627	0.0573
downstream	157	0.0079	63108	0.0095	1632	0.0080
intergenic	7151	0.3618	2103124	0.3156	61428	0.3027

Таблица 2.

Участок	Число участков с квадруплексом	Доля участков с предсказанным квадруплексом	Число участков предсказаний Zhunt	Доля участков предсказаний Zhunt	Число участков предсказаний ZDNABERT	Доля участков предсказаний ZDNABERT
exon	17112	0.0298	207129	0.3607	184014	0.3205
intron	15708	0.0888	59166	0.3345	46545	0.2632
promoters	1536	0.0502	17080	0.5586	14516	0.4747
downstream	328	0.0107	5236	0.1712	3066	0.1003
intergenic	2306	0.1825	7330	0.5802	6220	0.4923

Anopheles funestus (Ромашов Федор)

- Идентификатор в базе:
- Длина генома: 250.7 Mb
- Число генов: 14 819
- Число найденных квадруплексов: 12804
- Число найденных Z-DNA: 95212 (ZDNABERT),
147986 (ZHUNT)



Эпигенетические гены по семействам

- Геном: *Anopheles funestus*

Проверяемое семейство	Гены
ALKBH1	LOC125763436
BAP1	LOC125772298
ZNHIT1	LOC125762023
EEF1AKMT4	LOC125769691
TSSK6	LOC125764077
HDAC10	LOC125771668
JADE1	LOC125771123, LOC125764873, LOC125771123
SIRT7	LOC125764665
SMYD1	LOC125766226, LOC125770930
TDG	LOC125764691

Таблицы с распределением найденных структур

• Геном *Anopheles funestus*

Участок	Число квадруплексов	Доля квадруплексов	Число предсказаний Zhunt	Доля предсказаний Zhunt	Число предсказаний ZDNABERT	Доля предсказаний ZDNABERT
Exons	1073	0.083802	33957	0.229461	27700	0.299421
Introns	5858	0.457513	73798	0.498682	40686	0.439792
Promoters (1000 bp from TSS)	473	0.036942	8828	0.059654	6560	0.070910
Downstream (200 bp)	86	0.006717	2031	0.013724	2646	0.028602
Intergenic	5964	0.465792	44124	0.298163	28234	0.305193

Таблицы с распределением найденных структур

- Геном: *Anopheles funestus*

Участок	Число участков с квадруплексом	Доля участков с предсказанным квадруплексом	Число участков с предсказаниями Zhunt	Доля участков с предсказанным Zhunt	Число участков с предсказаниями ZDNABERT	Доля участков с предсказанным ZDNABERT
Exons	2759	0.014267	47937	0.247882	105759	0.546880
Introns	3327	0.053163	15937	0.254662	33231	0.531008
Promoters (1000 bp from TSS)	412	0.028667	6311	0.439118	8509	0.592054
Downstream (200 bp)	80	0.005566	1908	0.132758	6980	0.485667
Intergenic	1392	0.123656	5552	0.493204	7218	0.641201

Anopheles ziemanni (Скворцова Ира)

- Идентификатор в базе: GCF_943734765.1
- Длина генома: 270.0 Mb
- Число генов: 14 085
- Число найденных квадруплексов: 21 176
- Число найденных Z-DNA: 150404



Таблицы с распределением найденных структур

Геном: Anopheles ziemanni

	Число квадруплексов	Доля квадруплексов	Число предсказаний Z-DNABERT	Доля предсказаний Z-DNABERT	Число предсказаний Z_HUNT	Доля предсказаний Z_HUNT
Exons	395	0.019025	13288	0.097404	1352	0.072634
Introns	7831	0.377179	60152	0.440926	7808	0.419469
Promoters (1000 bp from TSS)	273	0.013149	5483	0.040191	530	0.028473
Downstream (200 bp)	285	0.013727	2963	0.021719	569	0.030568
Intergenic	12392	0.596860	68518	0.502250	9878	0.530676

	Число участков с квадруплексом	Доля участков с квадруплексом	Число участков с Z-DNABERT	Доля участков с Z-DNABERT	Число участков с Z_HUNT	Доля участков с Z_HUNT
Exons	339	0.120469	6485	0.375311	762	0.324117
Introns	1747	0.620824	5083	0.294172	827	0.351765
Promoters (1000 bp from TSS)	244	0.086709	2586	0.149661	336	0.142918
Downstream (200 bp)	255	0.090618	2023	0.117079	364	0.154828
Intergenic	229	0.081379	1102	0.063777	62	0.026372

Эпигенетические гены по семействам

- Геном: *Anopheles ziemanni* (Скворцова Ира)

Семейство	Функция	Найденные гены
ACINU	RNA modification	LOC131287203
ACTB	Chromatin remodeling cofactor	LOC131288930
ALKBH	Histone modification	LOC131291176
ANP32B	Histone chaperone	LOC131283953
APEX1	DNA modification cofactor	LOC131286663
ASXL3	Scaffold protein, Polycomb group (PcG) protein	LOC131285432
BMI1	Polycomb group (PcG) protein	LOC131286991
TSSK6	Histone modification write	LOC131286170
BAP1	Histone modification erase, Polycomb group (PcG) protein	LOC131284642
EEF1AKMT4	Protein modification	LOC131285456

Anopheles coluzzii (Волянский Никита)

- Идентификатор в базе: GCF_943734685.1
- Длина генома: 262.6 Mb
- Число генов: 14 616
- Число найденных квадруплексов: 28370
- Число найденных Z-DNA: 46547



Таблицы с распределением найденных структур

- Геном: *Anopheles coluzzii*

Участок	Число квадруплексов	Доля квадруплексов	Число предсказаний Zhunt	Доля предсказаний Zhunt	Число предсказаний ZDNABERT	Доля предсказаний ZDNABERT
Exons	1356	0.050684	53389	0.089970	6625	0.163888
Introns	13441	0.502392	321418	0.541649	24083	0.595760
Promoters (1000 bp from TSS)	620	0.023174	11487	0.019358	1185	0.029314
Downstream (200 bp)	520	0.019436	8174	0.013775	785	0.019419
Intergenic	12433	0.464716	255825	0.431113	13869	0.343088

Таблицы с распределением найденных структур

- Геном: *Anopheles coluzzii*

Участок	Число участков с квадруплексом	Доля участков с предсказанным квадруплексом	Число участков предсказаний Zhun	Доля участков предсказаний Zhunt	Число участков предсказаний ZDNABERT	Доля участков предсказаний ZDNABERT
Exons	3330	0.018805	8353	0.014076	10576	0.059725
Introns	2863	0.199359	1139	0.079312	1866	0.129935
Promoters (1000 bp from TSS)	346	0.024093	517	0.036	942	0.065594
Downstream (200 bp)	313	0.021795	467	0.032519	733	0.051041
Intergenic	2171	0.174406	807	0.06483	1374	0.110379

Эпигенетические гены по семействам

- Геном: *Anopheles coluzzii* (Волянский Никита)

Семейство	Найденные гены
ALKBH1	LOC120952103, LOC120952630, ...
CELF1	LOC120956733, LOC120961680...
METTL16	LOC120955374, LOC120955370
MBNL1	LOC120949013, LOC120959764...
HSPA1A	LOC120955774, LOC120951877, ...
BAP1	LOC120947726, LOC120958686...
ZNHIT1	LOC120956161, LOC120953083...
EEF1AKMT4	LOC120958582, LOC120957077...
TSSK6	LOC120953657, LOC120953657, ...
MAGOH	LOC120958031

Anopheles aquasalis (Старостин Ярослав)

- Идентификатор в базе:
- Длина генома: 170.5 Mb
- Число генов: 12 989
- Число найденных квадруплексов: 611355
- Число найденных Z-DNA: 4517716, 646927 (ZDNABERT)



Эпигенетические гены по семействам

- Геном: *Anopheles aquasalis* (Старостин Ярослав)

Проверяемое семейство	Название гена
PAK2 (PF05028)	LOC126570243
PRMT7 (PF06325)	LOC126576335
PRDM2 (PF00856)	LOC126571906
PRMT9 (PF13414)	LOC126581796
PRKAA1 (PF00069)	LOC126574753
PRDM4 (PF00096)	LOC126576043
PPM1G (PF00076)	LOC126577504
SMARCD1 (PF02201)	LOC126574654
SPEN (PF13893)	LOC126571025
TAF5 (PF00400)	LOC126573068

Таблицы с распределением найденных структур

Геном *Aopheles aquasalis*

Таблица 1, количества попаданий

Участок	Число квадруплексов	Доля квадруплексов	Число предсказаний Zhun	Доля предсказаний Zhun	Число предсказаний ZDNABERT	Доля предсказаний ZDNABERT
Exons	591	0.009	2265	0.037	9447	0.15
Introns	6858	0.138	3687	0.074	14937	0.301
Promoters (1000 bp from TSS)	574	0.044	746	0.005	3262	0.253
Downstream (200 bp)	186	0.014	113	0.008	353	0.027
Intergenic	9864	0.897	5265	0.47	13993	0.36

Таблица 2, количества участков с попаданиями

Участок	Число участков с квадруплексом	Доля участков с предсказанным квадруплексом	Число участков с предсказаниями Zhun	Число участков с предсказаниями Zhun	Число участков с предсказаниями ZDNABERT	Доля участков с предсказаниями ZDNABERT
Exons	591	0.009	1150	0.01	4390	0.07
Introns	6858	0.138	742	0.014	2774	0.05
Promoters (1000 bp from TSS)	574	0.044	385	0.003	2166	0.168
Downstream (200 bp)	186	0.014	101	0.0007	325	0.025
Intergenic	9864	0.897	299	0.027	1742	0.158

Групповая часть

Параметры выравнивания и построения деревьев

MUSCLE Alignment Options

Option	Setting
GAP PENALTIES	
Gap Open	<input checked="" type="checkbox"/> -2.90
Gap Extend	<input checked="" type="checkbox"/> 0.00
Hydrophobicity Multiplier	<input checked="" type="checkbox"/> 1.20
MEMORY/ITERATIONS	
Max Memory in MB	<input checked="" type="checkbox"/> 2048
Max Iterations	<input checked="" type="checkbox"/> 16
ADVANCED OPTIONS	
Cluster Method (Iterations 1,2)	<input checked="" type="checkbox"/> UPGMA
Cluster Method (Other Iterations)	<input checked="" type="checkbox"/> UPGMA
Min Diag Length (Lambda)	<input checked="" type="checkbox"/> 24

Buttons at the bottom: Help, Reset, Cancel, OK

M11: Analysis Preferences

Phylogeny Reconstruction

Option	Setting
ANALYSIS	
Scope	→ All Selected Taxa
Statistical Method	→ Neighbor-joining
PHYLOGENY TEST	
Test of Phylogeny	→ Bootstrap method
No. of Bootstrap Replications	→ 100
SUBSTITUTION MODEL	
Substitutions Type	→ Amino acid
Model/Method	→ p-distance
RATES AND PATTERNS	
Rates among Sites	→ Uniform Rates
Gamma Parameter	→ Not Applicable
Pattern among Lineages	→ Same (Homogeneous)
DATA SUBSET TO USE	
Gaps/Missing Data Treatment	→ Pairwise deletion
Site Coverage Cutoff (%)	→ Not Applicable
SYSTEM RESOURCE USAGE	
Number of Threads	→ 7

Buttons at the bottom: Help, Cancel, OK

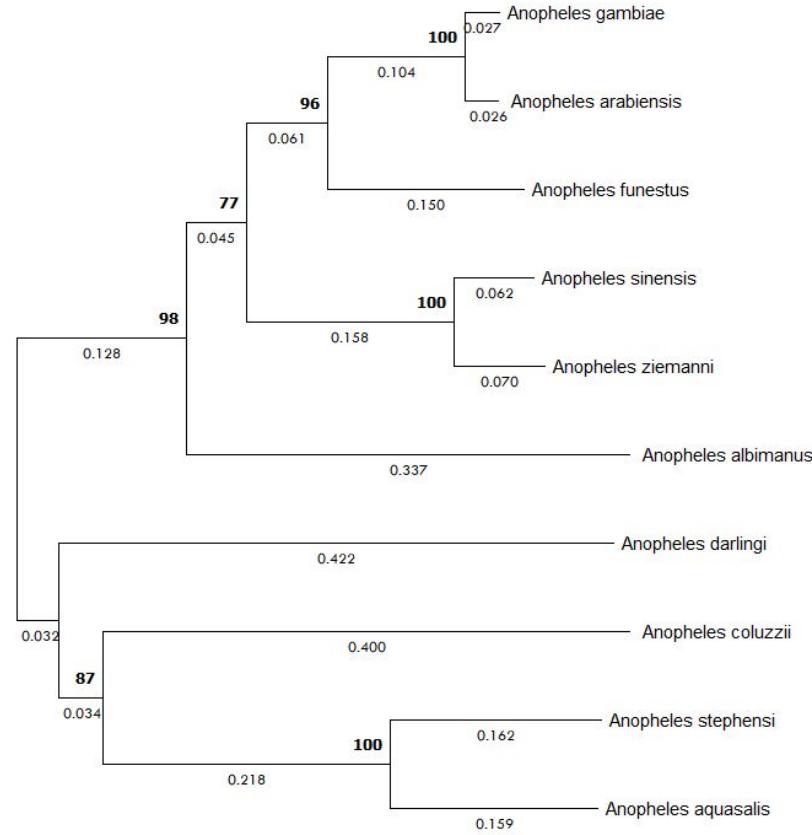
ALKBE1

(PF13532) - DNA demethylation, RNA demethylation

Protein Sequences	
Species/A	Sequence
1. gene=1	- M F Q T A F K Y Y - K S R N P P P S F E D - V L L I G S D H P H - L K P V R L G N I D - S O - C F A G L L S P S E W K V H E L T S R P G - L L V V A N P F T A E A
2. KFB49	- M F Q T A F K Y Y - K A R K P P P S F E N - V L I V G D D S Y P - F Q A I E A N L T G - G T W D A A S F R P V G E W R I F E V A T R P G - L F V I A N P F T P A E
3. gene=1	- M - - - - - V T P L R T A I A I - - - - - H R K C M T T G L A I R G W Q M L R A P L Q - Q L C H T K A A I E C V
4. XP_059	- M F Q C A F K Y Y - K A K K P P P S L N H - V L L V G D D S Y P - F Q A I E A N F I S - G T W D A S V F R P V G E W Q I F R V T T R P G - L F V I A N P F T P A E
5. XP_040	- M I W Q N F - - - A V O N C P P S I Y L P D F I K E E E C A I - M Q A V D K T - P P R R W T -
6. XP_040	- M F Q A A F K Y Y - K S R N P P P S F E D - V L L I G S D H P H - L K P V Q L G N I D - S O - C F A G L L S P S E W K V H E L T S R P G - L V V V A N P F T A E A
7. gene=1	- M L I S K V G A A K K I G K K V K R C C O L L I A K D A G I O T F E E P T E F - L V V C N A G L S T G L E D E R L L K E V L L H V G Q V D A I L M P P G K S Y C Y L R L A S N G A A A T V Y R A M I G I C P L C O D G A V I L L A F C C G F P C C L
8. XP_049	- M F Q T S F K Y Y - K S R N P V P T F E D - V F V I G M K H P H - L Q P V Q L N H I D - G G - N I L G L H P T T E W K V Y E L S T R P G - L L V L A N P F T C E V
9. gene=1	- M T I S K N L E K K I G K K I K R C C O L L I E K D A G T O F C D D P T D F - I V I C N A G L S T G L Q Q E A L L A E A L P H - G P I Q Q V R F P P G K S Y C F L H C T S K A S A H A V Y M A L N G V C T I G Q D G A V I L L A F C N A L P P C I
10. XP_035	- S Q L L D R F C K S F K Y Y - K A N Q P P P G L D D - V I Q I E D Q S D K R L E P V T L P T N G - L R F P G L L S P E C W R T Y R V S T R P G - L L V V A -

ALKBEH1

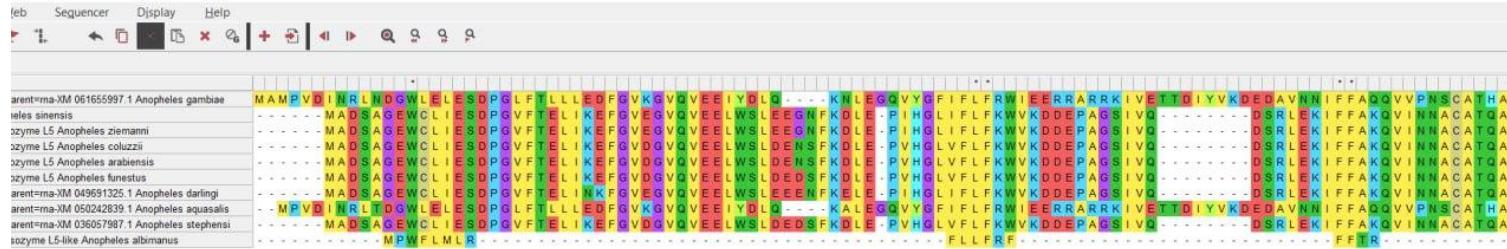
(PF13532) - DNA demethylation



0.10

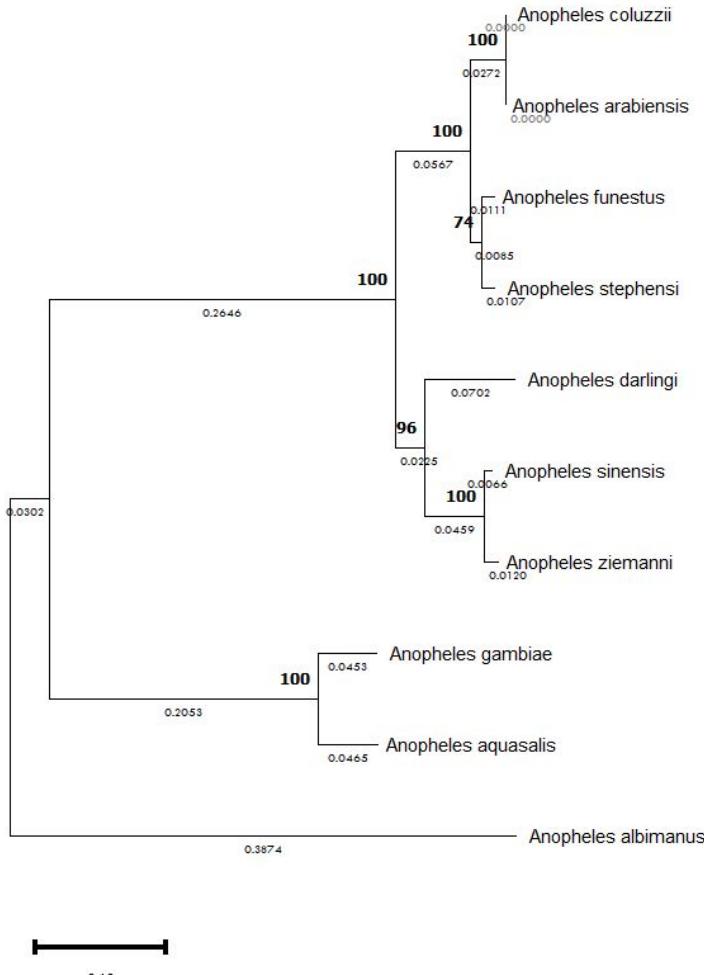
BAP1

(PF01088) - Histone deubiquitination



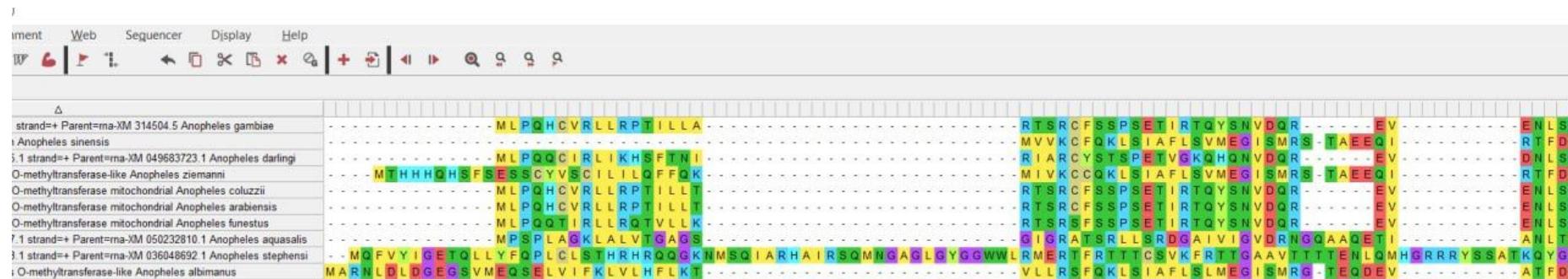
BAP1

(PF01088) - Histone deubiquitination



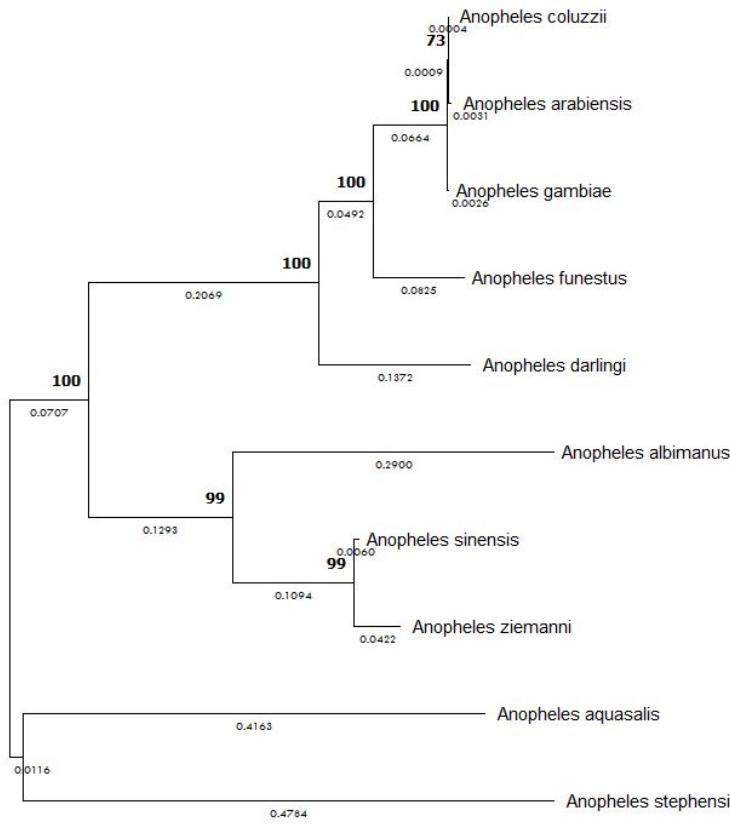
EEF1AKMT4

(PF08241) - Protein methylation



EEF1AKMT4

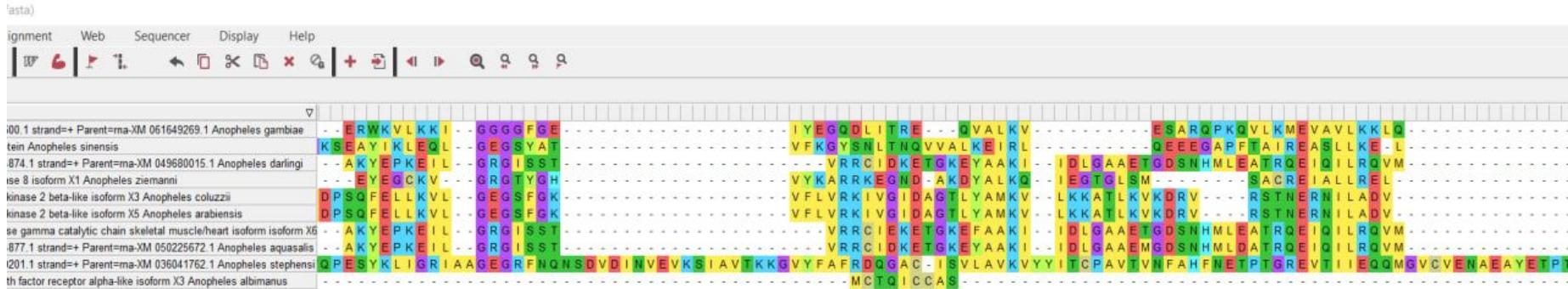
(PF08241) - Protein methylation



0.10

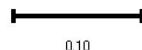
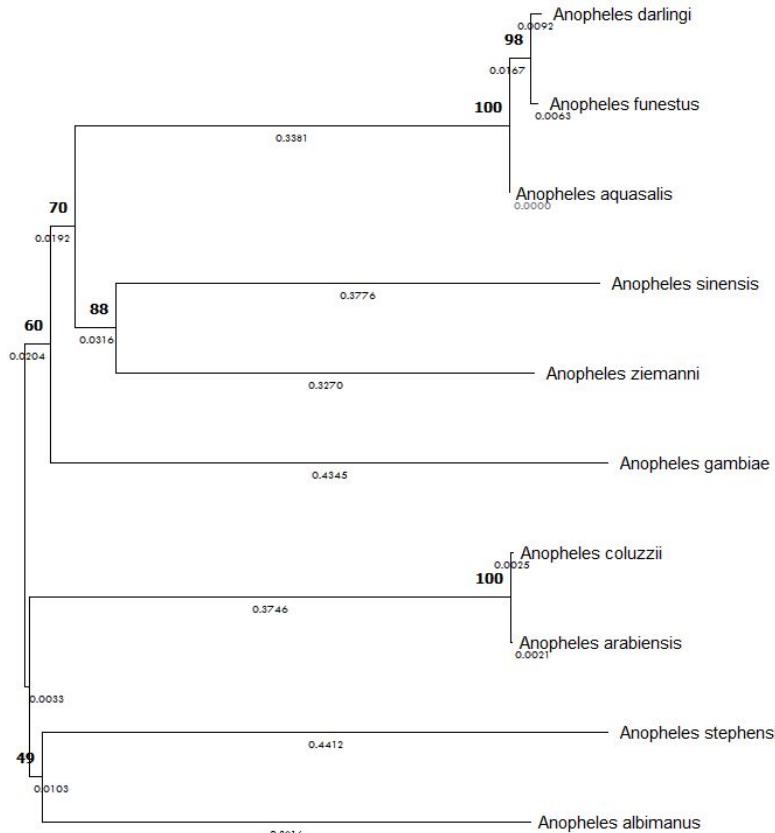
TSSK6

(PF00069) - Histone phosphorylation



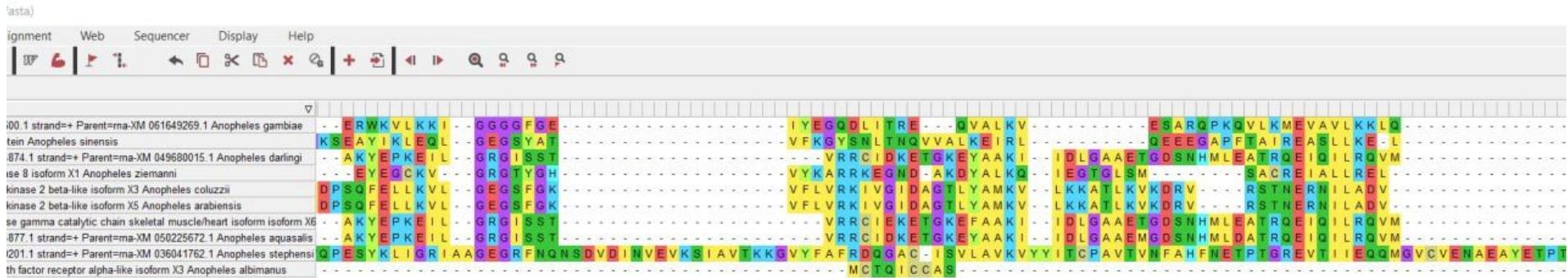
TSSK6

(PF00069) - Histone phosphorylation



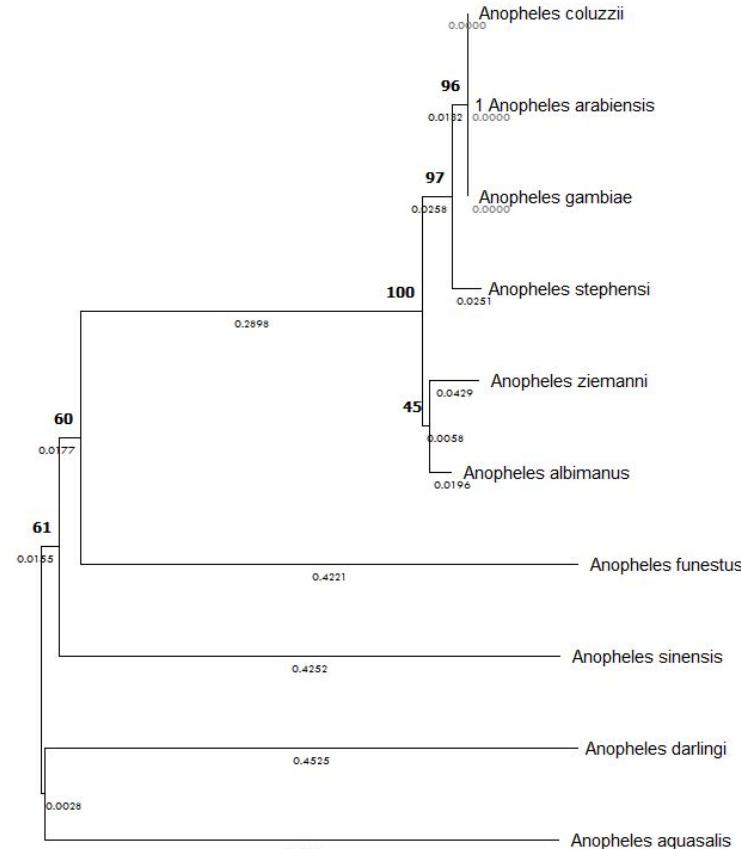
ZNHIT1

(PF04438) - Histone acetylation



ZNHIT1

(PF04438) - Histone acetylation



0.10

Спасибо за внимание!



bombardini malarini

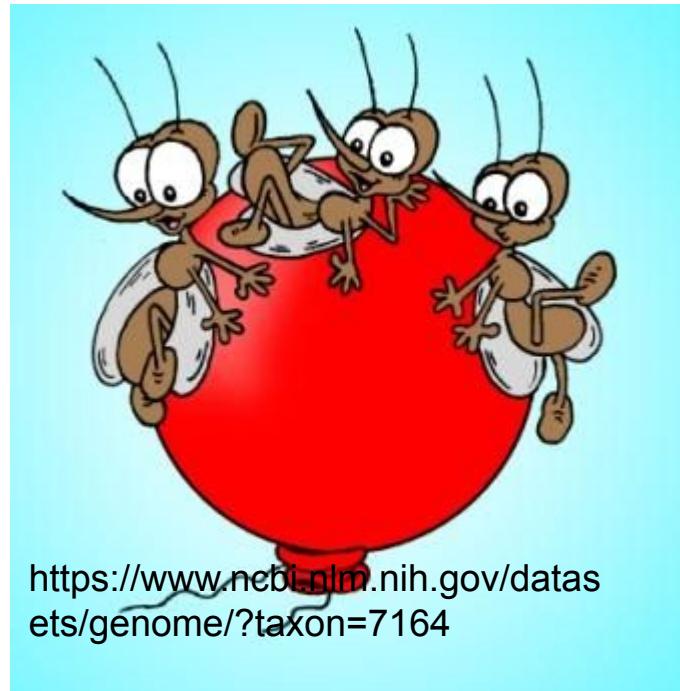
Докладывала команда:

Комарики на воздушном шарике

Наш девиз:

За семью и за дворы

Закусают комары!



<https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=7164>