

DNA modification

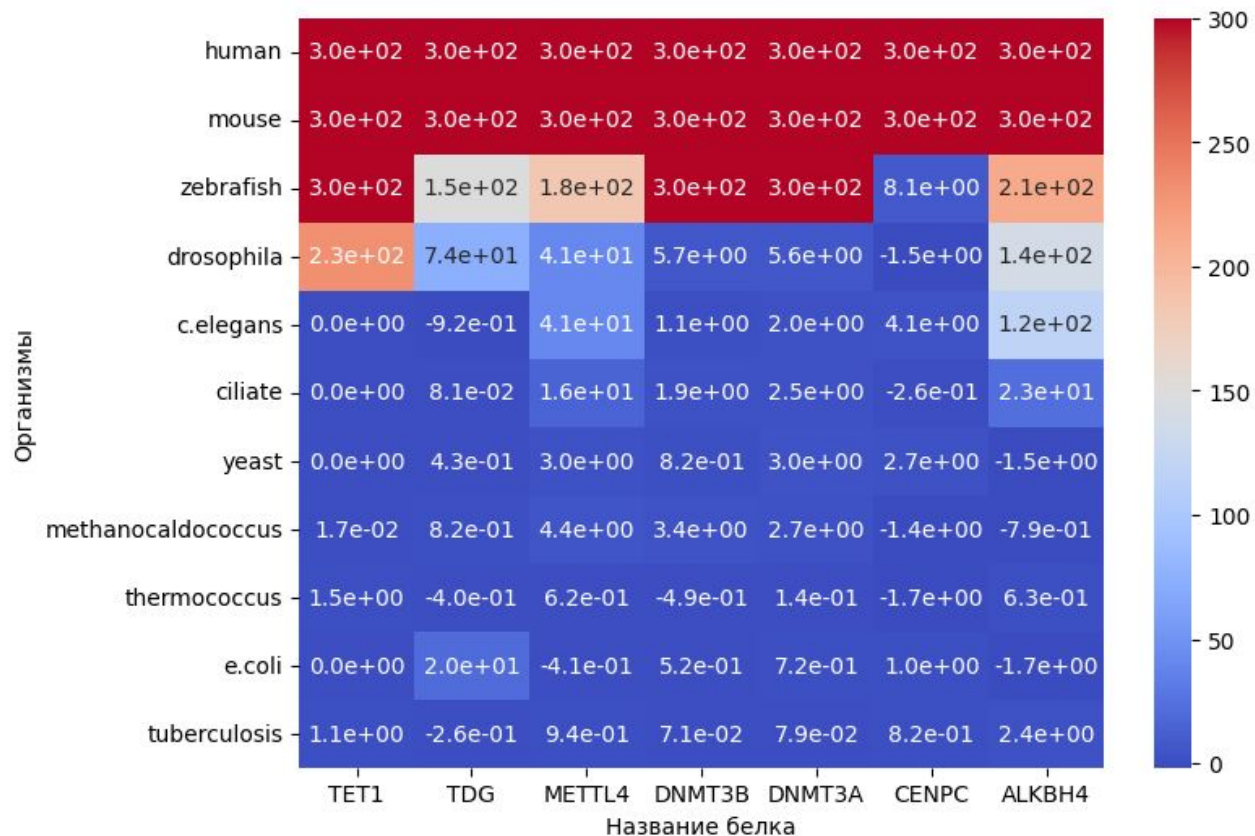
Выбранные белки и краткая информация о них

ФИО	Белок	Количество изоформ	Идентификатор белка	Длина	Домены
Иванова А.С.	TET1	12	NP_001393294.1	2165	TET_JBP
Щербакова Е.А.	TDC	5	ENST00000266775.13	406	UDG-like super family
Жардецкий Е. Н.	DNMT3A	8	NP_072046.2	912	PWWP_DNMT3A, ADDz_Dnmt3a
Божедомов А. Я.	DNMT3B	8	ENST00000328111.6	853	PWWP_DNMT3B, ADDz_Dnmt3b
Жумлякова С.Д.	AICDA				
Поздняков А. С.	CENPC	3	NP_001803.2	943	CENPC_N, CENPC_mid, cupin_RmlC-like
Шестаков Д.В.	ALKBH4	2	ENST00000292566.4	302	Fe2OG dioxygenase
Мельник В. К.	METTL4	2	NP_073751.3	472	MT-A70

Таблица выбора дальних гомологов белков

ФИО	Белок	Организм-гомолог	Идентификатор белка-гомолога	Длина	Домены
Иванова А.С.	TET1	drosophila	NP_647750	2056	TET_JBP
Щербакова Е.А.	TDG	zebrafish	XP_005163071.1	404	UDG-like super family
Жардецкий Е. Н.	DNMT3A	zebrafish	NP_001315096.1	978	PWWP_DNMT3A, ADDz_Dnmt3a, Dcm super family
Божедомов А. Я.	DNMT3B	zebrafish	NP_001020621.1	816	PWWP_DNMT3B, ADDz_Dnmt3b
Жумлякова С.Д.	AICDA				
Поздняков А. С.	CENPC	zebrafish	XP_005162783.1	1129	cupin_RmlC-like, PTZ00449, PLN03237
Шестаков Д.В.	ALKBH4	drosophilia	NP_609314.1	304	Fe2OG dioxygenase
Мельник В.К.	METTL4	zebrafish	XP_689178.3	450	MT-A70

Тепловая карта



Дальний гомолог: TET1

У нескольких организмов обнаруживаются белки, для которых e-value не равняется 0, однако значение не достаточно мало: thermococcus (e-value = 0.21) и tuberculosis (e-value = 0.29). Поэтому наиболее дальний гомолог наблюдается у дрозофилы.

Организм: drosophila

Белок: NP_647750

Длина TET: 2165 **Длина белка-гомолога:** 2056

E-Value: 2.53e-101

Набор доменов в белке-гомологе: TET_JBP

Описание: <https://pubmed.ncbi.nlm.nih.gov/25936838/> деметилирует ДНК, метилированную в 6-м положении аденина

Дальний гомолог TDG

Организм: zebrafish

Белок: XP_005163071.1 (G/T mismatch-specific thymine DNA glycosylase isoform)

Длина TDG: 406 **Длина белка-гомолога:** 404

E-Value: 7.85e-151

Набор доменов в белке-голомоле: UDG-like super family

Описание: <https://pubmed.ncbi.nlm.nih.gov/23880758/>,
<https://pubmed.ncbi.nlm.nih.gov/24722243/> - снижение уровней 5mC и 5hmC в дедифференцированных клетках связано с активным деметилированием в процессе регенерации плавников у зебрафиш.

Дальний гомолог DNMT3A

Организм: zebrafish

Белок: NP_001315096.1

Длина DNMT3A: 912 **Длина белка-гомолога:** 978

E-Value: 0*

Набор доменов в белке-гомологе: PWWP_DNMT3A, ADDz_Dnmt3a,
Dcm super family

*следующий слайд

Поиск дальнего гомолога: DNMT3A

При запуске blast'a оказалось, что только у дрозофилы обнаруживаются белки, для которых e-value достаточно мал, но не равняется нулю (NP_788726.2, NP_001097893.1 и NP_001262888.1; evalue = 2.76e-06).

Эти белки представляют собой различные изоформы six-banded [*Drosophila melanogaster*], связанные с активацией метил-связывания хроматина. Превышение размера исследуемого DNMT3A в полтора раза (1583, 1583, 1606 соответственно) и [указания](#) на гомологичность данных белков скорее с MBD5 позволяют сделать вывод, что они нам не подходят.

Следующим претендент и очевидный гомолог - NP_001315096.1 у данио-рерио (evalue = 0). Представитель [семейства](#) DNMT3 у рыб, чуть длиннее изучаемого белка (978 против 912), метилирует ДНК, модулирует [функцию](#), связанную с эмоциональным контролем, социальным взаимодействием и познанием. (домены: PWWP_DNMT3A, ADDz_Dnmt3a, Dcm super family).

Дальний гомолог DNMT3B

Организм: zebrafish

Белок: NP_001020621.1

Длина DNMT3B: 853, **Длина белка-гомолога:** 816

E-Value: 0

Набор доменов в белке-гомологе: PWWP_DNMT3B, ADDz_Dnmt3b

Описание: метелирует ДНК, экспрессируются в головном мозге, глоточных дугах, зачатках грудных плавников, кишечнике и плавательном пузыре, может играть роль в регуляции экспрессии генов во время процессов регенерации. ([источник](#))

Поиск дальнего гомолога CENPC

Организм: zebrafish

Белок: XP_005162783.1

Длина CENPC: 943

Длина белка-гомолога: 1129

E-Value: 2.93e-04

Набор доменов в белке-гомологе: cupin_RmlC-like, PTZ00449, PLN03237

Описание: Предполагается, что он обеспечивает центромерную ДНК-связывающую активность. Ортологичен человеческому CENPC (центромерный белок C) ([источник](#))

Поиск дальнего гомолога ALKBH4

Организм: drosophila

Белок: NP_609314.1

Длина ALKBH4: 302

Длина белка-гомолога: 304

E-Value: 1.79e-60

Набор доменов в белке-гомологе: -

Описание: Активирует окислительную активность ДНК-деметилазы. Участвует в кроветворении личиночных лимфатических узлов. Экспрессируется в нескольких структурах, включая зачаток передней энтодермы; эмбриональная/личиночная средняя кишка и т.д. Ортологичен ALKBH4 человека (гомолог 4 alkB, лизиндеметилаза). ([источник](#))

Дальний гомолог METTL4

Организм: zebrafish

Белок: XP_689178.3

Длина CENPC: 472

Длина белка-гомолога: 450

E-Value: 3.58e-80

Набор доменов в белке-гомологе: MT-A70

Описание: Предполагается активация метилтрансферазной активности связывания нуклеиновых кислот. Ортологичен METTL4 человека (метилтрансфераза 4, N6-аденозин)

https://www.ncbi.nlm.nih.gov/gene/?term=XP_689178.3

Вывод

В общем, эти данные подтверждают гипотезу о том, что многие из этих белков могли появиться и эволюционировать в более высоких организмах (drosophila, zebrafish)

Однако белки, такие как TDG, которые имеют высокие значения $-\log(E\text{-value})$ даже у прокариот, могут быть связаны с более древними и фундаментальными процессами, такими как ремонт и модификации связанных с DNA, уже на очень ранних этапах эволюции жизни.

С другой стороны, белки, такие как TET1, DNMT3B и DNMT3A, представлены только в позвоночных, что может указывать на их специфическую роль в эпигенетическом регулировании у более сложных организмов.