

Экологическая физиология и биохимия гидробионтов

УДК 597.562:577.29.088:004

АНАЛИЗ ТЕРМИНОВ ГЕННОЙ ОНТОЛОГИИ ДЛЯ БЕЛКОВ С ВЫСОКОЙ ОСМОТИЧЕСКОЙ АКТИВНОСТЬЮ ИЗ СЫВОРОТКИ КРОВИ АТЛАНТИЧЕСКОЙ ТРЕСКИ *GADUS MORHUA*

З. М. Базарова¹, А. С. Васильев¹, П. А. Павлова³, И. Ю. Торопыгин², А. М. Андреева^{1,*}

¹Институт биологии внутренних вод им. И.Д. Папанина РАН,
152742 пос. Борок, Ярославская обл., Некоузский р-н; e-mail: *aam@ibiw.ru

²Институт биомедицинской химии им. В.Н. Ореховича,
119121 г. Москва, ул. Погодинская, 10, г. Москва

³Санкт-Петербургский государственный университет,
199034 г. Санкт-Петербург, Университетская наб., д. 7–9

Поступила в редакцию 04.04.2024

В работе представлены разные способы графической интерпретации терминов геной онтологии (ГО) для небольшого протеома из 20-ти белков с высокой осмотической активностью (ОАБ) из сыворотки крови атлантической трески *Gadus morhua*. Анализ терминов проводили двумя способами: 1) с помощью пакета topGO [v2.46.0, Alexa & Rahnenfuhrer, 2021], а визуализацию — с помощью пакета ggplot2; 2) с помощью базы данных UniProt и графическая визуализация результатов в Excel 2016. Полученные результаты показали, что основную часть ОАБ трески составляют внеклеточные “истинные” белки плазмы, среди которых идентифицированы фетуин, гемопексин и аполипопротеины. ОАБ задействованы в биологических процессах кроветворения, протеолиза и развития, а также регуляции обменных процессов. Их молекулярные функции включают транспорт, гидролазную активность и ингибирование протеиназной активности. При использовании пакетов topGO и ggplot2 программный код не смог корректно графически отобразить все полученные данные из-за небольшой выборки белков. Обработка данных через DB UniProt дает корректное представление о функциях и участии белков в биологических процессах.

Ключевые слова: атлантическая треска, сыворотка крови, осмотически активные белки, электрофорез, MALDI, геной онтология, биоинформатика.

DOI: 10.47021/0320-3557-2024-75-85

ВВЕДЕНИЕ

В предыдущей работе [Базарова и др., 2022 (Bazarova et al., 2022)] в сыворотке крови трески атлантической *Gadus morhua* L. нами были идентифицированы с помощью масс-спектрометрии MALDI (Matrix-assisted laser desorption/ionization) доминирующие по относительному содержанию белки с высокой осмотической активностью — гемопексины, ингибиторы протеиназ и аполипопротеины в составе липопротеинов высокой плотности. Их высокая осмотическая активность объясняется максимально высоким отрицательным зарядом, определяющим их локализацию в анодной области диск-электрофореграммы. Белки с таким высоким электроотрицательным потенциалом способны наиболее эффективно связывать неорганические катионы и диполи воды, а значит, и вносить максимальный вклад в онкотическое давление плазмы [Andreeva et al., 2023]. Помимо описанных четырех белков, в анодной области электрофореграммы сыворотки трески было обнаружено еще около десятка белков с высоким отрицательным зарядом, обозначенных нами как осмотически активные белки, или ОАБ. Для понимания того, как эти белки (вместе с описанными ранее доминирующими белками

сыворотки) участвуют в осморегуляции, необходимо, помимо идентификации, провести анализ их основных функций и локализации в организме с использованием терминов геной онтологии (Gene Ontology, GO).

Геной онтология является проектом биоинформатики, направленным на создание единого словаря терминов для классификации у разных организмов их функций и локализации продуктов генов. Традиционно исследования протеомов плазмы крови позвоночных охватывают большие массивы белков из ≥ 1000 ед. [Anderson et al., 2004]. Их анализ в терминах GO проводится, как правило, с использованием баз данных с графической функцией, например, баз данных KEGG (*Kyoto Encyclopedia of Genes and Genomes*), David (*Database for Annotation, Visualization and Integrated Discovery*; david.ncifcrf.gov) и некоторых других с использованием графических библиотек, например, ggplot2 [He et al., 2021; <http://www.springer.com/series/6991>]. В данной работе для анализа локализации и функций небольшой выборки из осмотически активных белков трески (20 ОАБ) мы попытались исполь-

зовать R библиотеку topGO, которая предоставляет инструменты для тестирования терминов GO, и пакет для графической визуализации ggplot2. Параллельно для анализа терминов использовали DB UniProt и графику в Excel 2016.

Цель работы — идентификация с помощью MALDI и анализ в терминах генной онтологии локализации и функций 20 осмотически

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Объект исследования. Использовали образец сыворотки крови от одного экземпляра трески атлантической *Gadus morhua* (самец; II стадия зрелости гонад) (Teleostei, Paracanthopterygii, Gadiformes) из выборки рыб, состоящей из 12 экз., отловленных в июне в районе мыса Картеш, Белое море. Показатели длины рыб ad и ab по Смиуту варьировали от 16.5 до 29.5 см и от 18.0 до 32.5 см соответственно.

Получение сыворотки крови. Для получения крови использовали каудотомию. Кровь из хвостового сосуда рыбы собирали в пластиковую пробирку. Далее отстаивали в холодильнике при 4°C в течение ночи, после чего жидкость над сгустком осторожно отбирали пипеткой и переносили в чистую пластиковую пробирку и использовали для электрофореза.

Электрофорез. Для разделения белков сыворотки крови использовали двумерный электрофорез в полиакриламидном геле (2D-PAGE). Для первого направления использовали диск-электрофорез в 7.5%-ном PAG (disk-PAGE), для второго — SDS-электрофорез в 12.5%-ном PAG (SDS-PAGE). Анодную область диск-электрофореграммы, где располагались белки с высокой осмотической активностью, локализовали по белку-“навигатору” трансферрину (Tf), который ограничивал осмотически активные белки (ОАБ) со стороны катода [Andreeva et al., 2023].

В качестве основы для поиска ОАБ в крови трески использовали параметры электрофоретической подвижности (Rf) как меры отрицательного заряда белка, небольшой молекулярной массы (Mr) и высокого содержания в плазме (сыворотке) крови [Andreeva et al., 2023]. Проявление осмотической активности этими белками мы рассматривали в формате разработанной для рыб “безальбуминовой” гипотезы капиллярного обмена [Andreeva, 2020].

Идентификация белков с помощью масс-спектрометрии MALDI. Для идентификации ОАБ использовали MALDI. Последовательность этапов подготовки белков к масс-спектрометрии описана ниже:

1) в пробирку с помещенным в нее фрагментом геля, содержащим индивидуальный бе-

активных белков из сыворотки крови трески атлантической *Gadus morhua* с использованием R-программирования — библиотеки topGO, с последующей графической визуализацией результатов анализа с помощью пакета ggplot2, а также обработки данных “вручную” с использованием базы данных UniProt с дальнейшей визуализацией в Excel 2016.

лок, добавляли отмывочный раствор из бикарбоната аммония (AmBic) и ацетонитрила (ACN) так, чтобы он покрывал кусочек геля;

2) пробирки (по числу исследуемых белков) помещали в термостат на 12–15 мин при температуре 35–45°C, меняя несколько раз отмывочный раствор для достижения полного обесцвечивания фрагмента геля;

4) после полного удаления отмывочного раствора в каждую пробирку добавляли раствор ACN,

5) после “затвердения” пробы вследствие ее обезвоживания под воздействием ACN, пробирки центрифугировали; ACN полностью удаляли с помощью пипетки;

6) отмывые фрагменты геля с белками подвергали трипсинолизу: трипсин разбавляли раствором AmBic и разливали в пробирки по 6–8 мкл в каждую;

7) помещали пробирки в термостат на ночь (минимум на 4 ч) с температурой 70°C;

8) в каждую пробирку добавляли по 8 мкл трифтороуксусной кислоты (ТФУ) и оставляли на 30 мин.;

9) на мишень сначала наносили по капле 2.5-дигидроксибензойную кислоту (Dhb), далее — раствор из каждой пробирки (с помощью пипетки) (рис. 1).

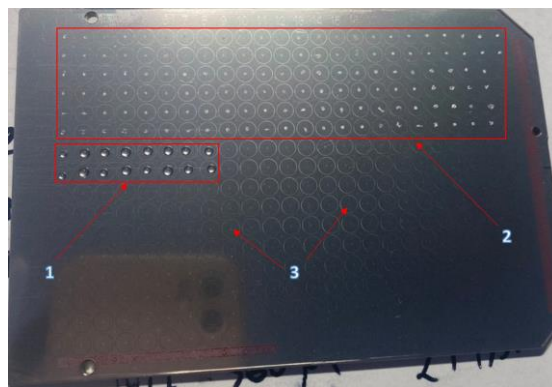


Рис. 1. Мишень для MALDI: 1 — пробы, не успевшие закристаллизоваться (капля Dhb + капля из пробирки); 2 — закристаллизовавшиеся пробы; 3 — свободные отсеки на мишени.

Fig. 1. Target for MALDI: 1 — samples that have not yet crystallized (a drop of Dhb + a drop from a test tube); 2 — crystallized samples; 3 — free compartments on the target.

Далее проводили MALDI-TOF MS и MS/MS масс-спектрометрию на MALDI — времяпролетном масс-спектрометре с УФ-лазером (336 нм) в режиме положительных ионов в диапазоне масс 500–8000 Да с калибровкой их по известным пикам аутолиза трипсина¹.

Следующим этапом был поиск кандидатов по полученным спектрам в базе Mascot. По результатам идентификации составляли список потенциальных белков-кандидатов, ранжированных в соответствии со значением показателя соответствия score, рассчитанного для каждого потенциального кандидата. Программа Mascot, после загрузки в нее полученного спектра, выдает список кандидатных белков по взятой пробе. Кандидатные белки подбираются программой по степени совпадения аминокислотных последовательностей — от белков с большими величинами score к белкам с меньшими величинами. Надежно идентифицированным считали белок с величиной score ≥ 80 –85.

Молекулярную массу кандидатов (Mr), рассчитанную в программе Mascot, обозначали как Mr_{calc}; молекулярную массу, рассчитанную в программе ONE-Dscan Ver 1.31 (Scanalytic Inc.) для белков из PAGE обозначали как Mr_{obs}.

Анализ локализации и функций кандидатов в терминах генной онтологии и графическая визуализация результатов с использованием

R-программирования — библиотеки topGO и пакета ggplot2.

Для анализа идентифицированных белков в терминах GO сохраняли последовательности в формате Fasta-файла и дальше использовали EggNOG² для поиска ортологов и получения GO терминов для каждого белка. Анализ GO терминов проводили с помощью пакета topGO (v2.46.0) с графической визуализацией в пакете ggplot2 [Alexa, Rahnenfuhrer, 2021]. Анализ терминов GO проводили в режиме “Биологические процессы”, “Молекулярные функции” и “Клеточная локализация”. Для оценки статистической значимости использовали критерий Колмогорова-Смирнова³ были выбраны процессы со значимостью $p < 0.05$, из которых отбирались те процессы, в которых было задействовано ≥ 10 генов, кодирующих белки, представленные в таблице 1 (раздел “Результаты”).

Для исследования кандидатов “вручную” использовали базу данных UniProt: каждый кандидатный белок по отдельности загружали в базу данных и анализировали по терминам генной онтологии (клеточной локализации, биологическим процессам и молекулярным функциям), полученные данные вносили в таблицу 2 (раздел Результаты) и далее работали в программе Excel 2016 для визуализации результатов.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Электрофоретическое разделение белков сыворотки крови трески. В 2D-SDS-PAGE анодная область электрофореграммы сыворотки была представлена 21 белком (ОАБ) (рис. 2), перечень которых дан в таблице 1.

Анализ локализации и функций ОАБ в терминах генной онтологии и графическая визуализация результатов. В программе R-studio представленные в таблице 2 белки были оценены по показателям клеточной локализации (рис. 3а), молекулярным функциям (рис. 3б) и участию в биологических процессах (рис. 3с).

Анализ локализации и функций кандидатов в терминах генной онтологии с помощью базы данных UniProt и графическая визуализация результатов в Excel 2016. Анализ проведен с помощью UniProt: каждый кандидат загружали в DB и выбирали информацию по его участию в клеточной локализации, биологических процессах и молекулярной функции. Полученную информацию вносили в таблицу 2, группировали данные из каждой колонки по отдельности (объединяли в функциональные группы) и далее графически представляли с помощью Excel 2016 (рис. 4а, б, с).

ОБСУЖДЕНИЕ

Результаты показали, что основную часть из двух десятков ОАБ трески составляют внеклеточные “истинные” белки плазмы, то есть белки, выполняющие в системе кровообращения специфические для нее функции: аполипопротеины (A-I, A-IV), гемопексины и ингибиторы протеиназ. Последние включают ингиби-

торы цистеиновых протеиназ — фетуины, и ингибиторы сериновых протеиназ — серпины, например, alpha-1-antitrypsin. Перечисленные “истинные” белки описаны также другими авторами не только у трески [Bohne-Kjersem et al., 2009; Kaneko et al., 2015], но и у других видов костистых рыб [Concha et al., 2004] и у млекопитающих [Anderson et al., 2004].

¹ <http://www.matrixscience.com>

² <http://eggno-mapper.embl.de/>

³ <https://www.codecamp.ru/blog/kolmogorov-smirnov-test-r/>

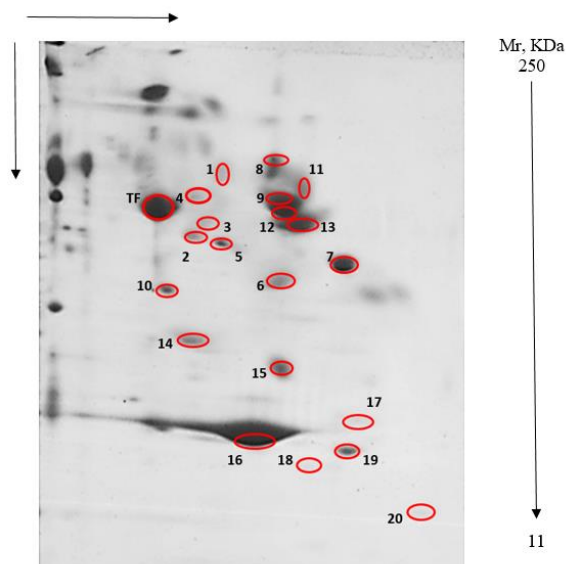


Рис. 2. 2D-SDS-PAGE сыворотки крови трески атлантической *Gadus morhua*: I направление (горизонтальная стрелка) — диск-электрофорез, II направление (вертикальная стрелка) — SDS-электрофорез. Tf — трансферрин — “навигатор” для поиска белков с высокой осмотической активностью; ОАБ (1–20) — осмотически активные белки; М, kDa — молекулярная масса белков в килодальтонах.

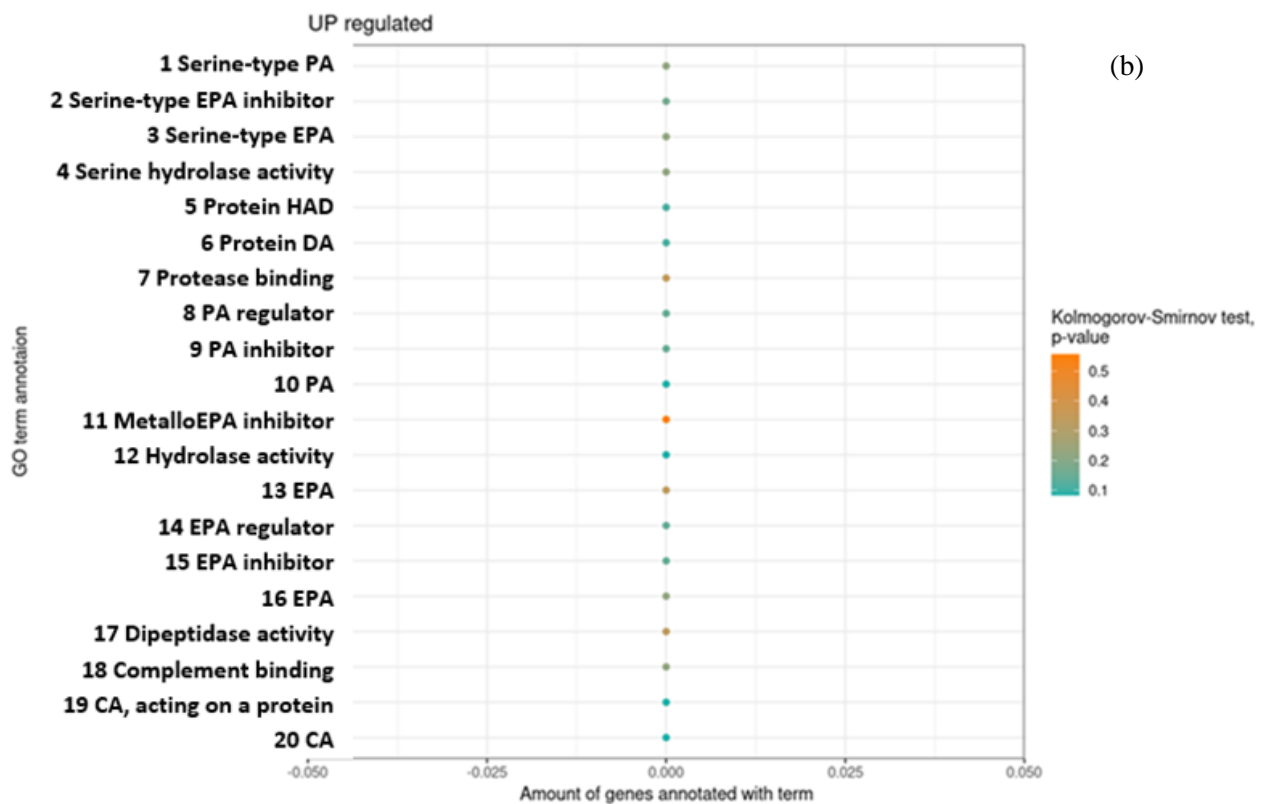
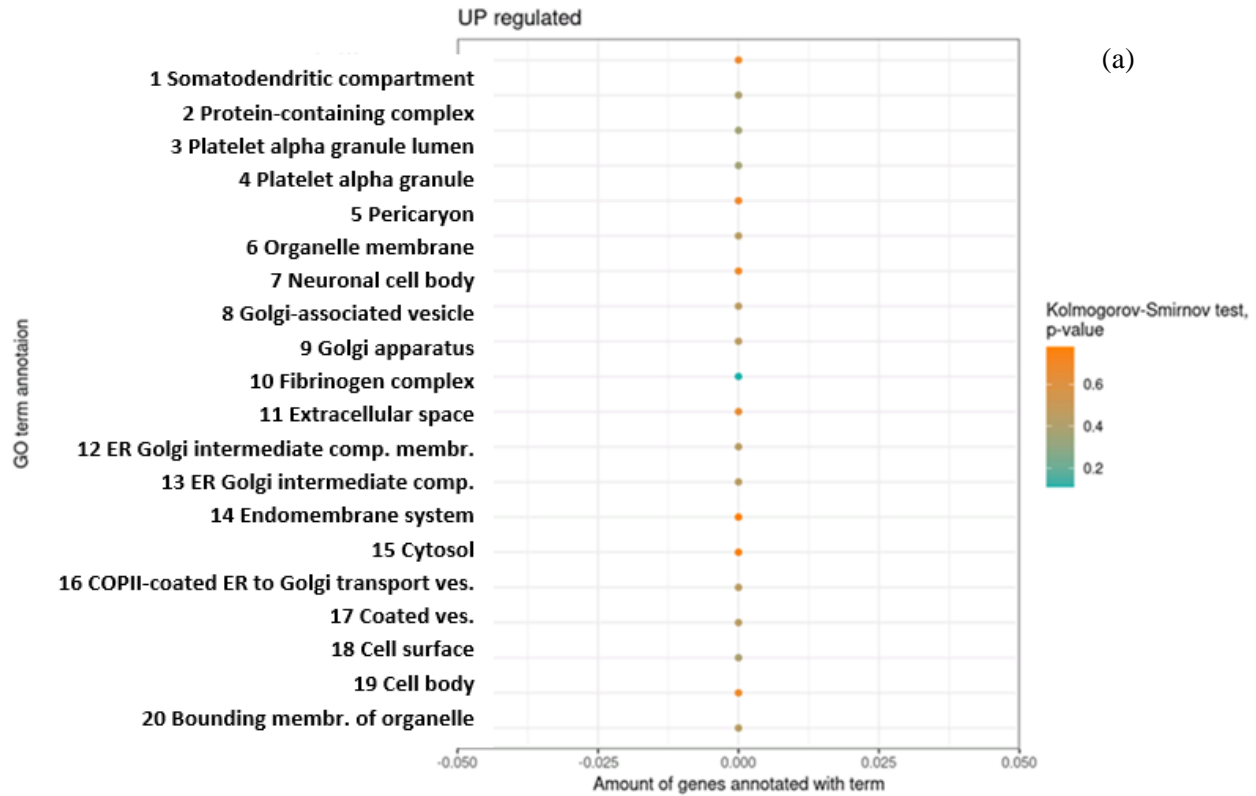
Fig. 2. 2D-SDS-PAGE of blood serum of Atlantic cod *Gadus morhua*: I direction (horizontal arrow) — disk electrophoresis, II direction (vertical arrow) — SDS electrophoresis. Tf — transferrin — “navigator” for searching for proteins with high osmotic activity; OAB (1–20) — osmotically active proteins; M, kDa — molecular weight of proteins in kilodaltons.

Таблица 1. Идентификация трансферрина и осмотически активных белков из сыворотки крови трески атлантической
Table 1. Identification of transferrin and osmotically active proteins from Atlantic cod blood serum

№	Candidate Protein	Mascot Search Results	Mr _{calc} /Mr _{obs} , Da	Score
Tf	serotransferrin [<i>Gadus morhua</i>]	XP_030218894.1	73516/70000	280
1	alpha-2-antiplasmin [<i>Gadus morhua</i>]	XP_030236562.1	54874/72000	117
2	alpha-2-HS-glycoprotein-like [<i>Gadus morhua</i>]	XP_030228808.1	32309/54500	103
3	complement factor B-like isoform X2 [<i>Gadus morhua</i>]	XP_030216206.1	72002/55000	81
4	alpha-2-HS-glycoprotein-like [<i>Gadus morhua</i>]	XP_030228808.1	32309/54500	95
5	beta-Ala-His dipeptidase-like isoform X1 [<i>Lepisosteus oculatus</i>]	XP_015212653.1	54227/54000	128
6	stonustoxin subunit beta-like [<i>Gadus morhua</i>]	XP_030206226.1	28266/40000	102
7	alpha-1-antitrypsin homolog [<i>Gadus morhua</i>]	XP_030201052.1	47082/44000	173
8	fetuin-B-like [<i>Gadus morhua</i>]	XP_030228824.1	53059/72000	96
9	hemopexin-like [<i>Gadus morhua</i>]	XP_030199795.1	48611/60000	194
10	myosin, light polypeptide 3, skeletal muscle [<i>Hippoglossus stenolepis</i>]	XP_035005709.1	16677/36000	62
11	kininogen-1-like isoform X2 [<i>Gadus morhua</i>]	XP_030227991.1	38746/68000	49
12	hemopexin-like [<i>Gadus morhua</i>]	XP_030210065.1	48611/55000	149
13	hemopexin-like [<i>Gadus morhua</i>]	XP_030210065.1	47494/45000	140
14	myosin heavy chain, fast skeletal muscle-like [<i>Anabas testudineus</i>]	XP_026220654.1	222030/31000	88
15	myosin heavy chain, fast skeletal muscle-like [<i>Notolabrus celidotus</i>]	XP_034540808.1	221203/30000	81
16	apolipoprotein A-I-2-like [<i>Gadus morhua</i>]	XP_030236469.1	29769/26000	260
17	apolipoprotein A-I-2-like [<i>Gadus morhua</i>]	XP_030236469.1	29769/26000	232
18	apolipoprotein A-I-like [<i>Gadus morhua</i>]	XP_030216412.1	27807/26000	129
19	apolipoprotein A-IV-like [<i>Gadus morhua</i>]	XP_030203246.1	28610/25000	182
20	coactosin-like protein [<i>Channa argus</i>]	KAF3689264.1	16060/17000	75

Примечание. Обозначение (Tf) и нумерация (1–17) белков даны в соответствии с рис. 1.

Note. The designation (Tf) and numbering (1–17) of proteins are given as in fig. 1.



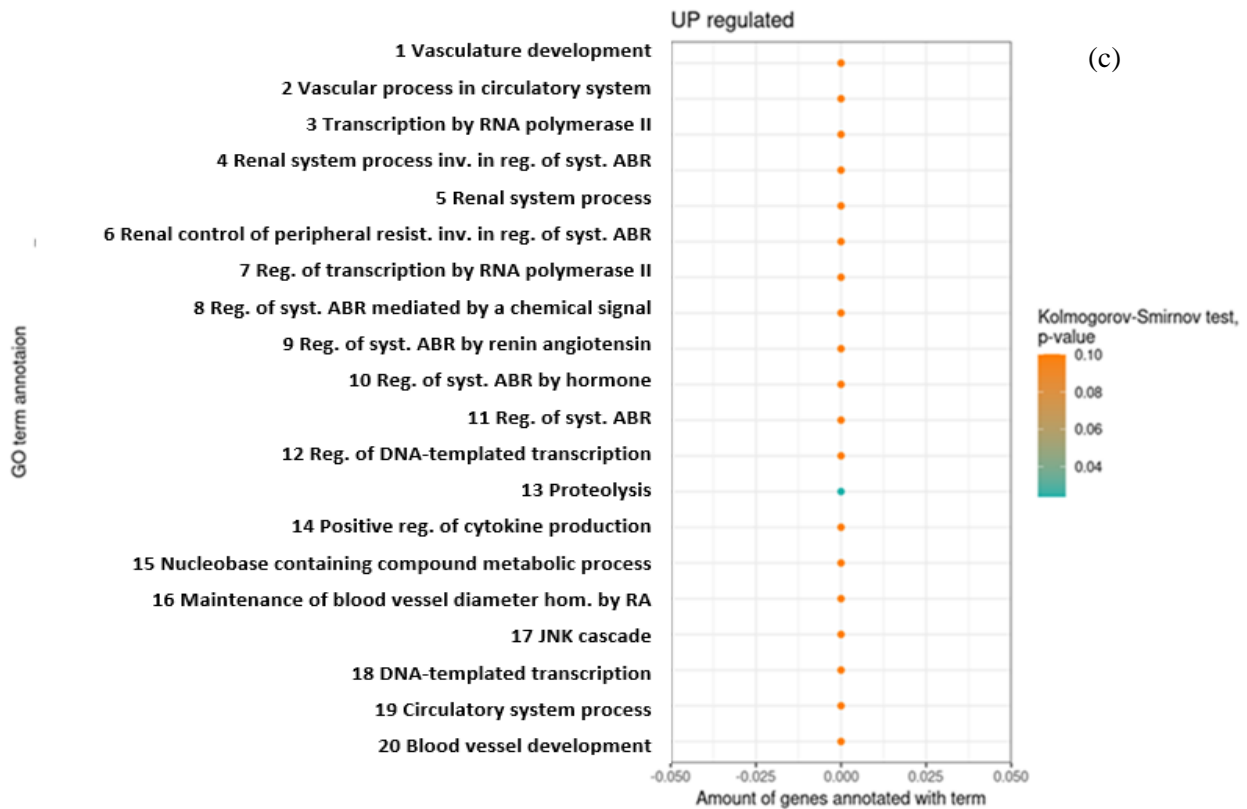


Рис. 3. Графическая визуализация анализа результатов геной онтологии осмотически активных белков сыворотки трески *Gadus morhua* с помощью пакета ggplot2. По оси x — число генов, кодирующих белки в терминах GO; по оси y — показатели клеточной локализации (а), молекулярных функций (б) и биологических процессов (с). Цветовая индикация обозначает уровень значимости показателя в соответствии с критерием Колмогорова-Смирнова: бирюзовым цветом выделены наиболее значимые, оранжевым — менее значимые функции. Сокращения: ER — endoplasmic reticulum, comp. — compartment, membr. — membrane, ves. — vesicle, reg. — regulation, ABR — arterial blood pressure, resist. — resistance, syst. — systemic, hom. — homeostasis, RA — renin-angiotensin, inv. — involved, PA — peptidase activity, EPA — endopeptidase activity, HDA — homodimerization activity, DA — dimerization activity, CA — catalytic activity.

Fig. 3. Graphical visualization of gene ontology analysis of osmotically active serum proteins from cod *Gadus morhua* using the ggplot2 package. The x -axis is the number of genes encoding proteins in GO terms; along the y -axis — indicators of cellular localization (a), molecular functions (b) and biological processes (c). The color indication indicates the level of significance of the indicator in accordance with the Kolmogorov-Smirnov criterion: the most significant functions are highlighted in turquoise, less significant functions are highlighted in orange. Abbreviations: ER — endoplasmic reticulum, comp. — compartment, membr. — membrane, ves. — vesicle, reg. — regulation, ABR — arterial blood pressure, resist. — resistance, syst. — systemic, hom. — homeostasis, RA — renin-angiotensin, inv. — involved, PA — peptidase activity, EPA — endopeptidase activity, HDA — homodimerization activity, DA — dimerization activity, CA — catalytic activity.

Таблица 2. Анализ локализации и функций кандидатов в терминах геной онтологии с помощью базы данных UniProt

Table 2. Analysis of location and function of candidates in terms of gene ontology using the UniProt database

№	Candidate Protein	Cellular Component	Molecular Function	Biological Process
1	alpha-2-antiplasmin [<i>Gadus morhua</i>]	Cc	Serine-type endopeptidase inhibitor activity	Negative regulation of fibrinolysis
2	alpha-2-HS-glycoprotein-like [<i>Gadus morhua</i>]	Cc	Cysteine-type endopeptidase inhibitor activity	Negative regulation of endopeptidase activity
3	complement factor B-like isoform X2 [<i>Gadus morhua</i>]	Ex	Calcium ion binding	Animal organ development; cell development; system development
4	alpha-2-HS-glycoprotein-like [<i>Gadus morhua</i>]	Cc	Cysteine-type endopeptidase inhibitor activity	Negative regulation of endopeptidase activity

№	Candidate Protein	Cellular Component	Molecular Function	Biological Process
5	beta-Ala-His dipeptidase-like isoform X1 [<i>Lepisosteus oculatus</i>]	Ex	Metal ion binding; metallopeptidase activity	Proteolysis
6	stonustoxin subunit beta-like [<i>Gadus morhua</i>]	Ex	Potassium channel regulator activity; toxin activity	Hemolysis in another organism
7	alpha-1-antitrypsin homolog [<i>Gadus morhua</i>]	Ex	Serine-type endopeptidase inhibitor activity	Negative regulation of peptidase activity
8	fetuin-B-like [<i>Gadus morhua</i>]	Ex	Cysteine-type endopeptidase inhibitor activity; endopeptidase inhibitor activity; metallopeptidase inhibitor activity	Binding of sperm zona pellucida; negative regulation of endopeptidase activity; single fertilization
9	hemopexin-like [<i>Gadus morhua</i>]	Ex	Metallopeptidase activity; zinc ion binding	Proteolysis
10	myosin, light polypeptide 3, skeletal muscle [<i>Hippoglossus stenolepis</i>]	Cc	Calcium ion binding	Muscle contraction
11	kininogen-1-like isoform X2 [<i>Gadus morhua</i>]	Ex	Cysteine-type endopeptidase inhibitor activity	Blood coagulation; inflammatory response; vasolidation
12	hemopexin-like [<i>Gadus morhua</i>]	Ex	Metallopeptidase activity; zinc ion binding	Proteolysis
13	hemopexin-like [<i>Gadus morhua</i>]	Ex	Metallopeptidase activity; zinc ion binding	Proteolysis
14	myosin heavy chain, fast skeletal muscle-like [<i>Anabas testudineus</i>]	Cc	Actin filament binding; ATP binding; cytoskeletal motor activity	System development
15	myosin heavy chain, fast skeletal muscle-like [<i>Notolabrus celidotus</i>]	Cc	Actin filament binding; ATP binding; cytoskeletal motor activity	System development
16	apolipoprotein A-I-2-like [<i>Gadus morhua</i>]	Ex	Lipid binding	Lipid transport; lipoprotein metabolic process
17	apolipoprotein A-I-2-like [<i>Gadus morhua</i>]	Ex	Lipid binding	Lipid transport; lipoprotein metabolic process
18	apolipoprotein A-I-like [<i>Gadus morhua</i>]	Ex	Lipid binding	Lipid transport; lipoprotein metabolic process
19	apolipoprotein A-IV-like [<i>Gadus morhua</i>]	Ex	Lipid binding	Lipid transport; lipoprotein metabolic process
20	coactosin-like protein [<i>Channa argus</i>]	Ex	Actin binding	Regulation of actin filament polymerization

Сокращения: Ex — extracellular component (внеклеточное расположение), Cc — cellular component (клеточное расположение).

Среди внеклеточных белков, не относящихся к категории “истинных” белков плазмы, мы обнаружили субъединицу стонустоксина (stonustoxin subunit beta-like) — мощного цитолитического яда, описанного у каменной рыбы *Synanceia verrucosa* (Perciformes) [Tang et al., 2023].

Другой группой белков, обнаруженных среди ОАБ трески являются внутриклеточные белки в виде фрагментов миозина, гликопротеина alpha-2-HS-glycoprotein-like и антиплазмина alpha-2-antiplasmin.

Согласно проведенному анализу GO, ОАБ сыворотки трески задействованы в ряде биологических процессов: регуляции развития и об-

менных процессов, кровообращения, репродукции и липидного обмена. По молекулярным функциям: ОАБ участвуют в транспорте, гидролазной активности, ингибировании протеиназ, процессах иммунной защиты, регуляции и мышечной активности цитоскелета.

Обращает внимание на себя тот факт, что среди перечисленных функций ОАБ отсутствует осмотическая, между тем, как нахождение белков в анодной области диск-электрофорограммы предполагает их высокий отрицательный заряд и, как следствие, высокую осмотическую активность [Andreeva, 2020; Andreeva et al., 2023].



Рис. 4. Графическая визуализация анализа результатов генной онтологии осмотически активных белков сыворотки трески *Gadus morhua* с помощью Excel 2016: показатели клеточной локализации (a), молекулярных функций (b) и биологических процессов (c).

Fig. 4. Graphical visualization of the analysis of the results of gene ontology of osmotically active serum proteins of the cod *Gadus morhua* using Excel 2016: indicators of cellular localization (a), molecular functions (b) and biological processes (c).

Это позволяет предположить, что белки из анодной зоны подвижности выполняют в плазме (сыворотке) крови, прежде всего, свои основные функции (транспорта, иммунной защиты, регуляции обменных процессов), а осмотическую активность они проявляют по причине особой структурной организации, удовлетворяющей критериям осмотической активности. В крови млекопитающих максимально высокую осмотическую активность проявляет специализированный белок — сывороточный

альбумин. Поскольку у трески, как и у значительного числа костистых рыб, этот белок отсутствует [Bohne-Kjersem et al., 2009; Shugang Li et al., 2017] вследствие его потери в ходе эволюции (Noel et al., 2010), то вероятно, что его функцию взяли на себя другие белки. “Безальбуминовая” гипотеза капиллярного обмена предлагает рассматривать в качестве заменивших альбумин белков множественные ОАБ, характеризующиеся высокой осмотической активностью [Andreeva, 2020].

ЗАКЛЮЧЕНИЕ

Использование пакета topGO позволило, в целом, определить перечень молекулярных функций ОАБ трески, их участие в биологических процессах и локализацию в организме (внеклеточные, внутриклеточные белки). Однако, графическая визуализация с помощью пакета ggplot2 не позволила “привязать” тот или иной белок к конкретной локализации, функции и процессам, в связи с чем, потребовалась ручная “привязка” каждого из 20 ОАБ к той или иной функции и локализации. Кроме того, мы заметили, что результат графической визуализации не отражает число генов, кодирующих кандидатные белки. Об этом судили по некорректной индикации данных по оси x: заданный параметр количества генов, кодирующих кандидаты, был представлен диапазоном величин от -0.05 до 0.05 , и все точки (цветовые индикаторы) оказались расположенными в ко-

ординате 0, что не соответствует действительности. Мы полагаем, что это произошло по причине использования нами небольшого протеома из двух десятков белков, поскольку в данной программе чем больше выборка белков, тем меньше вероятность ошибки. Программный код отбирает процессы, за которые отвечают минимум 10 генов, кодирующих белки, а так как у нас выборка составляет всего 20 белков, то программа, вероятно, не смогла выполнить корректный анализ по данному параметру. По этой причине используемый нами на небольшой выборке белков способ графической интерпретации результатов не является исчерпывающим. Анализ ГО через базу данных UNIPROT, хотя и занимает достаточно много времени, но дает корректные результаты для небольших выборок белков.

Работа выполнена в рамках госзадания №121050500046-8.

СПИСОК ЛИТЕРАТУРЫ

- Базарова З.М., Торопыгин И.Ю., Васильев А.С. и др. Поиск и идентификация осмотически активных белков в сыворотке крови атлантической трески *Gadus morhua* // Труды Института биологии внутренних вод им. И.Д. Папанина РАН, 2022. Вып. 99(102). С. 88–92. DOI: 10.47021/0320-3557-2022-88-92.
- Alexa A., Rahnenfuhrer J. (2021) Gene set enrichment analysis with topGO // <http://www.mpi-sb.mpg.de/~alexa>
- Anderson N.L., Polanski M., Pieper R. et al. The human plasma proteome: a nonredundant list developed by combination of four separate sources // *Molecular & Cellular Proteomics*. 2004. Vol. 3. P. 311–326. DOI: 10.1074/mcp.M300127-MCP200.
- Andreeva A.M. Structural Organization of Plasma Proteins as a Factor of Capillary Filtration in Pisces // *Inland Water Biology*. 2020. Vol. 13. № 4. P. 664–673. DOI: 10.1134/S1995082920060036.
- Andreeva A.M., Bazarova Z.M., Toropygin I.Yu. et al. Serum Osmotically Active Proteins in the Atlantic Cod *Gadus morhua* // *Journal of Evolutionary Biochemistry and Physiology*. 2023. Vol. 59. P. 325–336. DOI: 10.31857/S004445292302002X.
- Bohne-Kjersem A., Skadsheim A., Goksoyr A., Grosvik B.E. Candidate biomarker discovery in plasma of juvenile cod (*Gadus morhua*) exposed to crude North Sea oil, alkyl phenols and polycyclic aromatic hydrocarbons (PAHs) // *Marine Environmental Research*. 2009. Vol. 68(5), Iss. 5. P. 268–277. DOI: 10.1016/j.marenvres.2009.06.016.
- Concha M.I., Smith V.J., Castro K. et al. Apolipoproteins AI and A-II are potentially important effectors of innate immunity in the teleost fish *Cyprinus carpio* // *European journal of biochemistry*. 2004. Vol. 271, Iss. 14. P. 2984–2990. DOI: 10.1111/j.1432-1033.2004.04228.x.
- He Y., Yu H., Zhao H. et al. Transcriptomic analysis to elucidate the effects of high stocking density on grass carp (*Ctenopharyngodon idella*) // *BMC Genomics*. 2021. Vol. 22. 620. DOI: 10.1186/s12864-021-07924-4.
- Kaneko K., Nishii S., Izumi Y. et al. Proteomic Analysis after Sequential Extraction of Matrix Proteins in Urinary Stones Composed of Calcium Oxalate Monohydrate and Calcium Oxalate Dihydrate // *Analytical Sciences*. 2015. Vol. 31(9). P. 935–942. DOI: 10.2116/analsci.31.935.

- Liu M., Yan X., Zhang M., Li X., Li S., Jing M. Influence of Human Papillomavirus Infection on the Natural History of Cervical Intraepithelial Neoplasia 1: A Meta-Analysis // *BioMed Research International*. 2017. Vol. 2017. 8971059. DOI: 10.1155/2017/8971059.
- Noel E.S., Reis M., Arai Z., Ober E.A. Analysis of the Albumin/ α -Fetoprotein/Afamin/Group specific component gene family in the context of zebrafish liver differentiation // *Gene Expression Patterns*. 2010. Vol. 10(6). P. 237–243. DOI: 10.1016/j.gep.2010.05.002.
- Tang T., Huang Y., Peng C., Liao Y., Lv Y., Shi Q., Gao B. A Chromosome-Level Genome Assembly of the Reef Stonefish (*Synanceia verrucosa*) Provides Novel Insights into Stonustoxin (sntx) Genes // *Molecular Biology and Evolution*. 2023. Vol. 40(10). msad215. DOI: 10.1093/molbev/msad215.
- Электронный ресурс. URL: <http://www.springer.com/series/6991>
- Электронный ресурс. URL: <http://eggnog-mapper.embl.de/>
- Электронный ресурс. URL: <http://www.matrixscience.com>
- Электронный ресурс. URL: <https://vniiz.org/science/publication/article-383/conf90-article-17>
- Электронный ресурс. URL: <https://www.codecamp.ru/blog/kolmogorov-smirnov-test-r/>
- Электронный ресурс. URL: <https://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html>

REFERENCES

- Alexa A., Rahnenfuhrer J. (2021) Gene set enrichment analysis with topGO // <http://www.mpi-sb.mpg.de/~alexa>
- Anderson N.L., Polanski M., Pieper R. et al. The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Molecular & Cellular Proteomics*, 2004, vol. 3, pp. 311–326. doi: 10.1074/mcp.M300127-MCP200.
- Andreeva A.M. Structural Organization of Plasma Proteins as a Factor of Capillary Filtration in Pisces. *Inland Water Biology*, 2020, vol. 13, no. 4, pp. 664–673. doi: 10.1134/S1995082920060036.
- Andreeva A.M., Bazarova Z.M., Toropygin I.Yu. et al. Serum Osmotically Active Proteins in the Atlantic Cod *Gadus morhua*. *Journal of Evolutionary Biochemistry and Physiology*, 2023, vol. 59, pp. 325–336. doi: 10.31857/S004445292302002X.
- Bazarova Z.M., Toropygin I.Yu., Vasiliev A.S. et al. Search and identification of osmotically active proteins in the blood serum of the atlantic cod *Gadus Morhua*. *Inland Water Biology*, 2022, vol. 99, pp. 88–92. doi: 10.47021/0320-3557-2022-88-92 (In Russian)
- Bohne-Kjersem A., Skadsheim A., Goksoyr A., Grosvik B.E. Candidate biomarker discovery in plasma of juvenile cod (*Gadus morhua*) exposed to crude North Sea oil, alkyl phenols and polycyclic aromatic hydrocarbons (PAHs). *Marine Environmental Research*, 2009, vol. 68(5), iss. 5, pp. 268–277. doi: 10.1016/j.marenvres.2009.06.016.
- Concha M.I., Smith V.J., Castro K. et al. Apolipoproteins AI and A-II are potentially important effectors of innate immunity in the teleost fish *Cyprinus carpio*. *European journal of biochemistry*, 2004, vol. 271, iss. 14, pp. 2984–2990. doi: 10.1111/j.1432-1033.2004.04228.x.
- He Y., Yu H., Zhao H. et al. Transcriptomic analysis to elucidate the effects of high stocking density on grass carp (*Ctenopharyngodon idella*). *BMC Genomics*, 2021, vol. 22, 620. doi: 10.1186/s12864-021-07924-4.
- Kaneko K., Nishii S., Izumi Y. et al. Proteomic Analysis after Sequential Extraction of Matrix Proteins in Urinary Stones Composed of Calcium Oxalate Monohydrate and Calcium Oxalate Dihydrate. *Analytical Sciences*, 2015, vol. 31(9), pp. 935–942. doi: 10.2116/analsci.31.935.
- Liu M., Yan X., Zhang M., Li X., Li S., Jing M. Influence of Human Papillomavirus Infection on the Natural History of Cervical Intraepithelial Neoplasia 1: A Meta-Analysis. *BioMed Research International*, 2017, vol. 2017, 8971059. doi: 10.1155/2017/8971059.
- Noel E.S., Reis M., Arai Z., Ober E.A. Analysis of the Albumin/ α -Fetoprotein/Afamin/Group specific component gene family in the context of zebrafish liver differentiation. *Gene Expression Patterns*, 2010, vol. 10(6). pp. 237–243. doi: 10.1016/j.gep.2010.05.002.
- Tang T., Huang Y., Peng C., Liao Y., Lv Y., Shi Q., Gao B. A Chromosome-Level Genome Assembly of the Reef Stonefish (*Synanceia verrucosa*) Provides Novel Insights into Stonustoxin (sntx) Genes. *Molecular Biology and Evolution*, 2023, vol. 40(10), msad215. doi: 10.1093/molbev/msad215.
- Electronic resource: URL: <http://www.springer.com/series/6991>
- Electronic resource: URL: <http://eggnog-mapper.embl.de/>
- Electronic resource: URL: <http://www.matrixscience.com>
- Electronic resource: URL: <https://vniiz.org/science/publication/article-383/conf90-article-17>
- Electronic resource: URL: <https://www.codecamp.ru/blog/kolmogorov-smirnov-test-r/>
- Electronic resource: URL: <https://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html>

ANALYSIS OF GENE ONTOLOGY TERMS FOR PROTEINS WITH HIGH OSMOTIC ACTIVITY FROM BLOOD SERUM OF THE ATLANTIC COD *GADUS MORHUA*

Z. M. Bazarova¹, A. S. Vasilyev¹, P. A. Pavlova³, I. Yu. Toropygin^{1,2}, A. M. Andreeva^{1,*}

¹*Papanin Institute for Biology of Inland Waters, Russian Academy of Sciences,
152742 Borok, Russia; e-mail: *aam@ibiw.ru*

²*V.N. Orekhovich Research Institute of Biomedical Chemistry, Russian Academy of Medical Sciences,
119121 Moscow, Pogodinskaya st., 10, Russia*

³*St. Petersburg State University, 199034 St. Petersburg, Universitetskaya embankment 7–9, Russia*

Revised 04.04.2024

The paper presents different ways to graphically interpret gene ontology (GO) terms for a small proteome of 20 proteins with high osmotic activity (OAP) from the blood serum of the Atlantic cod *Gadus morhua*. Term analysis was carried out in two ways — 1) using the topGO package (v2.46.0, Alexa & Rahnenfuhrer, 2021), and visualization using the ggplot2 package; 2) using the UniProt database and graphical visualization of the results in Excel 2016. The results obtained showed that the main part of cod OAP consists of extracellular “true” plasma proteins, among which fetuin, hemopexin and apolipoproteins were identified. OAP are involved in the biological processes of hematopoiesis, proteolysis and development, as well as the regulation of metabolic processes. Their molecular functions include transport, hydrolase activity and inhibition of proteinase activity. When using the topGO and ggplot2 packages, the program code was unable to correctly graphically display all the data obtained due to the small sample of proteins. Data processing through DB UniProt gives a correct idea of the functions and participation of proteins in biological processes.

Keywords: Atlantic cod, blood serum, osmotically active proteins, electrophoresis, MALDI, gene ontology, bioinformatics