

## ПОИСК ВОЗМОЖНЫХ ГОМОЛОГОВ ХЕМОКИНОВ ПОЗВОНОЧНЫХ У ПРОКАРИОТ

Хемокины — это обширная группа цитокинов, объединённая сходным строением и способностью распознавать родопсинподобные рецепторы. Основными их функциями в организме позвоночных являются:

- регуляция распределения клеток (прежде всего лимфоцитов) по лимфоидным органам;
- активация иммунокомпетентных клеток и привлечение их в очаг воспаления [3].

Хемокины позвоночных характеризуются быстрой эволюцией [6]. Их возникновение и эволюцию обычно связывают с ранними позвоночными [6], и именно в пределах типа Позвоночные сосредоточены основные исследования эволюции данных белков [4, 7]. У беспозвоночных и примитивных хордовых они на сегодняшний день не обнаружены [4].

Однако имеются данные, что в эволюционном плане функция хемокинов гораздо шире, чем только обеспечение хемотаксиса иммунокомпетентных клеток, и их первичная роль — обеспечение хемотаксиса клеток различных типов [5, 7]. В свете этого была выдвинута гипотеза, что эволюционные предшественники хемокинов могли возникнуть у одноклеточных организмов. Кроме того, сообщается о недавнем обнаружении у беспозвоночных (в том числе простейших) молекул, сходных с интерлейкинами позвоночных [2], что указывает на актуальность поиска за пределами типа Позвоночные также гомологов хемокинов.

При этом вопрос о наличии или отсутствии гомологов хемокинов у одноклеточных организмов является в настоящее время неизученным.

**Цель исследования:** изучение вопроса о возможных гомологах хемокинов у бактерий и архей.

### Материалы и методы

Поиск белков, гомологичных хемокинам млекопитающих, производился с помощью сервиса UniProt (<http://www.uniprot.org/>), в частности, размещённой на этом сайте программы BLAST. Назначением программ серии BLAST является поиск локальных гомологий между аминокислотными или нуклеотидными последовательностями.

В качестве последовательностей-запросов при первичном поиске использовались хемокины человека и позвоночных животных.

Из числа обнаруженных таким способом последовательностей отбирались те, у которых участок гомологии занимал большую часть длины белка-субъекта (найденного при поиске белка) и белка-запроса, в идеальном случае — почти всю длину данных белков. Так как BLAST ищет локальную гомологию, и гомология на коротких участках встречается чаще, чем на протяжённых, то факт обнаружения протяжённого участка гомологии расценивался как достаточно достоверное сходство (меньше вероятность случайного совпадения). Допускалось превышение длины последовательности-запроса над длиной последовательности-субъекта максимум в 2–3 раза, что служило дополнительным фактором надёжности. Помимо длины участка гомологии и самой аминокислотной последовательности белка при анализе сходства учитывались и вычисленные программой BLAST показатели идентичности.

Для каждой найденной последовательности:

- определялся филогенетический профиль и предпринималась попытка определить функцию методом гомологий (с помощью той же программы BLAST);
- производился поиск гомологов в типе Позвоночные или классе Млекопитающие с целью первичной оценки достоверности сходства и вероятности филогенетического родства.

На втором этапе найденные последовательности комбинировались с последовательностями хемокинов, по этим комбинациям генерировалось множественное выравнивание и строились филогенетические деревья — с помощью программ Clustal W2 ([http://www.ebi.ac.uk/Tools/phylogeny/clustalw2\\_phylogeny/](http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/)), WebPrank (<http://www.ebi.ac.uk/goldman-srv/webprank/>) и Cobalt ([http://www.stva.ncbi.nlm.nih.gov/tools/cobalt/re\\_cobalt.cgi?](http://www.stva.ncbi.nlm.nih.gov/tools/cobalt/re_cobalt.cgi?)).

С помощью программы BLAST NCBI ([https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE\\_TYPE=BlastSearch&BLAST\\_SPEC=blast2seq&LINK\\_LOC=blasttab](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=blasttab)) строились филогенетические деревья по парным выравниваниям. Назначение филогенетических деревьев в данном исследовании — анализ эволюционных отношений между исследуемыми белками прокариот и хемокинами.

Отбирались последовательности, родство которых с хемокинами, по результатам данного этапа, оказывалось наиболее вероятным. Такой же отбор производился по результатам каждого последующего этапа с целью обнаружить последовательности, наиболее достоверно родственные хемокинам.

На третьем этапе с помощью программы BLAST NCBI проводилась заключительная проверка родства найденных бактериальных белков между собой и с хемокинами с помощью метода гомологий и построения филогенетического дерева по парным выравниваниям. Также на этом этапе анализировалось распределение совпадений по каждой последовательности-запросу с целью выяснения, имеется ли гомология в пределах одного участка или сходство данного бактериального белка и хемокинов более глобальное.

С помощью этой же программы строились филогенетические деревья по парным выравниваниям.

На четвёртом этапе для белков, родство которых с хемокинами оказывалось наиболее вероятным, предсказывались вторичная структура в программе JPred ([http://www.compbio.dundee.ac.uk/jpred4/index\\_up.html](http://www.compbio.dundee.ac.uk/jpred4/index_up.html)) и расположение дисульфидных остатков с помощью сервера DiANNA (<http://clavius.bc.edu/~clotelab/DiANNA/>). Производился анализ результатов данных вычислений на предмет сходства со вторичной структурой и распределением дисульфидных связей в молекулах хемокинов.

Производился также поиск возможных гомологов данных белков в царствах Грибы и Растения и проверка найденных гомологов на наличие сходства с хемокинами с использованием сервера UniProt. Данный поиск производился с целью выяснения, имеются ли эволюционные провалы между группами белков, найденными на предыдущих этапах.

В качестве заключительной серии экспериментов выполнялся поиск переменных и консервативных (то есть эволюционно устойчивых) участков белков с помощью сервера ConSurf (<http://consurf.tau.ac.il/>).

### Результаты и обсуждение

По результатам первого этапа были отобраны следующие белки эубактерий, показывающие гомологию с хемокинами позвоночных (далее в качестве названия белка используется идентификатор UniProt, так как эти белки не охарактеризованы и не имеют названий):

Идентификатор белка в UniProt	Вид бактерии
X5DFU9	<i>Draconibacterium orientale</i>
A0A0N1LTS4	<i>Phodopseudomonas</i> sp. AAP120
L1PHK4	<i>Actinomyces</i> sp. oral taxon 181 str. F0379
R7LMN2	<i>Clostridium</i> sp. CAG:729
A0A0A6DN53	<i>Janthinobacterium lividum</i>

A0A0J6H647	<i>Pseudomonas lini</i>
A0A059DJG3	<i>Hyphomonas</i> sp. CY54-11-8
A0A0Q7WNI9	<i>Mesorhizobium</i> sp. Root554

Были обнаружены также гомологи последнего белка (A0A0Q7WNI9), принадлежащие представителям домена Archaea (род *Methanosarcina*) — A0A0P0J056, A0A0E3NE26, A0A0E3H885.

При этом показатели идентичности для протяжённых выравниваний составляли порядка 20 — 30 %, что ожидаемо при таких больших эволюционных расстояниях (для сравнения, гомология между СС- и СХС-хемокинами меньше 30 % [6]).

Белки A0A0Q7WNI9 и A0A0P0J056, а также их ближайшие гомологи отличаются от остальных найденных бактериальных белков тем, что имеют в своём составе два соседствующих остатка цистеина на участке гомологии с хемокинами, при этом данные остатки включаются в выравнивание. Данные белки и их гомологи были объединены под названием соответственно: «группа A0A0Q7WNI9» и «группа A0A0P0J056».

Результаты построения филогенетического дерева в программе ClustalW2 на основе множественного выравнивания в WebPRANK оказались неустойчивы, что исключало возможность их использования в исследовании. Построение филогенетического дерева в программе Cobalt дало более устойчивые результаты (вероятно, благодаря наличию ограничения показателя максимального различия последовательностей, за счёт чего программа отбрасывает недостоверные сходства). Эти результаты были использованы в исследовании.

Результаты построения филогенетического дерева в программе BLAST NCBI также оказались устойчивыми и пригодными для анализа. По результатам второго этапа в рассмотрении остались белки: X5DFU9, A0A0J6H647 (и их гомологи), группа A0A0Q7WNI9, группа A0A0P0J056, так как результаты данного этапа свидетельствовали в пользу родства данных белков хемокинам позвоночных. По результатам третьего этапа из рассмотрения был исключен A0A0J6H647 и его гомологи. Построенные деревья свидетельствовали в пользу родства группы A0A0Q7WNI9 и группы A0A0P0J056.

На четвёртом этапе удалось установить, что белки группы A0A0Q7WNI9 и группы A0A0P0J056, вероятно, схожи по своей вторичной структуре (данные JPred), характеризующейся N-концевым  $\alpha$ -спирализованным участком и  $\beta$ -структурой остальной части белка. При расчёте вторичной структуры хемокинов в JPred было обнаружено сходство рассчитанной структуры хемокинов с описанными выше белками. Отличие структур состоит в наличии на С-конце хемокинов  $\alpha$ -спирализованного участка, в пределах которого UniProt BLAST также не обнаруживает гомологии первичной структуры с белками группы A0A0Q7WNI9 и группы A0A0P0J056. Этот участок, вероятно, необходим для связывания с гликозаминогликанами [6], что важно для создания пространственного градиента хемокина [3]. В свете выдвинутой гипотезы было сделано предположение, что он мог быть относительно поздним образованием — приспособлением хемокинов к функционированию в развитой соединительной ткани. Белок X5DFU9, по результатам аналогичного моделирования, оказался резко отличным по вторичной структуре от хемокинов (в нём преобладают  $\alpha$ -спирали), на основании чего он был также исключён из рассмотрения.

Распределение дисульфидных связей в белках группы A0A0Q7WNI9 (данные DiANNA) оказалось достаточно переменным. Сделать какой-либо определённый вывод на основе этих данных не представляется возможным.

Обнаружение гомологичных друг другу белков у археев (метаносарцин), зубактерий и позвоночных — отдалённо родственных друг другу организмов — естественным образом ставит вопрос о наличии подобных гомологов и у других таксонов живых организмов.

В пределах царств Растения и Грибы обнаруживались единичные белки, проявляющие отдалённую гомологию первичной и иногда вторичной структуры как с белками группы

A0A0Q7WNI9 и группы A0A0P0J056, так и с хемокинами (также единичные сходства). Среди данных белков оказались среди всего прочего, гидрофобины грибов, а также два белка с невыясненным филогенетическим профилем (что связано с отсутствием данных об их ближайших гомологах): A0A0E0E7Z9 (*Oryza meridionalis*), A0A0L0S5U7 (*Allomyces nasostogunus*). Гидрофобины грибов, помимо сходны с хемокинами тем, что являются небольшими (порядка 100 аминокислот) и секреторными белками [1]. Но их функция совершенно иная — они функционируют как поверхностно-активные вещества, входя в состав клеточной стенки гриба [1]. Такое большое различие функций снижает вероятность того, что эти белки родственны хемокинам позвоночных, однако полностью исключить такую возможность на сегодняшний день нельзя. Этот вопрос требует дальнейшего изучения.

Что касается белков A0A0E0E7Z9 и A0A0L0S5U7, то отдалённость гомологии и отсутствие данных о филогенетических профилях этих белков пока не позволяют сделать каких-либо выводов. Впрочем, неполными являются и данные о филогенетическом профиле белков группы A0A0P0J056 — имеются только данные об их наличии у археев рода *Methanosarcina*.

Таким образом, имеется эволюционный провал между белками группы A0A0Q7WNI9, белками группы A0A0P0J056 и хемокинами. Данные о белках, которые могут его заполнить, пока немногочисленны. Вопрос эволюционных провалов на сегодняшний день является самым малоисследованным моментом выдвинутой гипотезы, поэтому планируется продолжить исследования в этой области.

Анализ белка A0A0Q7WNI9 и его гомологов с помощью ConSurf показал, что два соседствующих остатка цистеина в нём максимально консервативны, то есть эволюционно максимально устойчивы. Это имеет место и для хемокинов [6]. Наиболее вариабельным оказался N-концевой участок. У хемокинов аналогичный участок является индивидуальным для каждого конкретного хемокина [3]. Это означает сходство распределения консервативных и вариабельных участков аминокислотной последовательности между белками группы A0A0Q7WNI9 и хемокинами.

По результатам большей части этапов исследования достаточную степень сходства с хемокинами обнаруживает белок A0A059DJG3. Но методом гомологий удалось установить, что это, скорее всего, тиюэстераза, притом имеющая гомологи-тиуюэстеразы среди белков растений, грибов и животных. Если учесть данные о функции, то родство данного белка с хемокинами выглядит маловероятным. На основе анализа участка гомологии было установлено, что он ограничен и захватывает аминокислотные остатки в пределах отрезка 80 — 130 последовательности A0A059DJG3. Аналогичные результаты, подтверждающие наличие небольшого участка гомологии, были получены для прокариотических тиюэстераз-гомологов данного белка. Было сделано предположение, что это сходство на уровне фрагмента. Отчасти оно подтверждается тем, что данный участок наиболее гомологичен также ряду тиюэстераз архей, и, таким образом, может обладать повышенной консервативностью, но этот вопрос нуждается в дополнительной исследовании.

### Выводы

1. Обнаружены две возможно родственные друг другу группы белков бактерий и архей (группа A0A0Q7WNI9 и группа A0A0P0J056).
2. Возможно, эти группы белков родственны хемокинам позвоночных. Но на сегодняшний день вопрос об их родстве окончательно не решён.
3. Обнаружен белок A0A059DJG3, который обнаруживает сходство с хемокинами на уровне фрагмента. Возможно, он относится к группе тиюэстераз, имеющих гомологию с хемокинами на уровне фрагмента.
4. Между белками группы A0A0Q7WNI9, белками группы A0A0P0J056 и хемокинами имеются эволюционные провалы, на сегодняшний день до конца не

заполненные, хотя имеются немногочисленные данные о белках, которые могут заполнить эти провалы. Исследования в этой области продолжаются.

#### Литература

1. Белозерская Т. А. Гидрофобины грибов: структура и функции/ Т. А. Белозерская// Микология и фитопатология. — 2001. — Том 35, вып. 1. — С. 3–11.
2. Ройт А. Иммунология. Пер. с англ. /А. Ройт, Дж. Бростофф, Д. Мейл. — М.: Мир, 2000. — 592 с.
3. Ярилин, А. А. Иммунология: учебник/А. А. Ярилин. — М.: ГЭОТАР-Медиа, 2010. — 752 с.
4. DeVries, Mark E. Defining the Origins and Evolution of the Chemokine/Chemokine Receptor System/Mark E. DeVries, Alyson A. Kelvin, Luoling Xu, Longsi Ran, John Robinson and David J. Kelvin// The Journal of Immunology. — 2006 — Vol. 176, No. 1. — P. 401-415.
5. Huising, Mark O. Molecular evolution of CXC chemokines: extant CXC chemokines originate from the CNS/Mark O. Huising, René J. M. Stet, Corine P. Kruiswijk, Huub F. J. Savelkoul, B. M. Lidy Verburg-van Kemenade// Trends in Immunology. — 2003 — Volume 24, Issue 6. — P. 306-312.
6. Paul, William E. Fundamental Immunology/William E. Paul. — Philadelphia: Lippincott Williams & Wilkins Publishers, 2003. — 1701 p.
7. Zlotnik, Albert. The chemokine and chemokine receptor superfamilies and their molecular evolution/Albert Zlotnik, Osamu Yoshie, Hisayuki Nomiyama//Genome Biology. — 2006 — Volume 7, Issue 12 — Article 243.