

БИОИНФОРМАТИКА

УДК 575

©Коллектив авторов

ПЕПТИДНЫЕ РЕЦЕПТОРЫ, СОПРЯЖЕННЫЕ С G-БЕЛКАМИ, И ИХ ЭНДОГЕННЫЕ ЛИГАНДЫ В ГЕНОМЕ ЧЕЛОВЕКА

М.Е. Левченко¹, В.В. Поройков¹, М. Канехиса²

¹ГУ НИИ Биомедицинской химии РАМН им. В.Н. Ореховича, 119121 Москва,
ул. Погодинская, 10, тел. (095) 247-3029, факс (095) 245-0857,
эл. почта: vvp@ibmh.msk.su

²Центр биоинформатики, Институт химических исследований,
Университет Киото, Япония

Семейство пептидных рецепторов, сопряженных с G-белками, а также их лигандов - эндогенных пептидов - вовлечено в регуляцию многих важных физиологических процессов в организме, вследствие чего они являются привлекательными мишенями для фармакологических исследований и конструирования лекарств (drug design). В связи с завершением черновой расшифровки генома человека стало возможным получение исчерпывающей картины всех генов, кодирующих как пептидные рецепторы, так и соответствующие им пептиды. С применением стандартных методов биоинформатики (поиск по гомологии, кластерный анализ, построение профилей и т.п.) предпринята попытка идентификации всех пептидных рецепторов, сопряженных с G-белками, и их лигандов в геноме человека. В результате проведенного исследования мы показали, что вероятное количество пептидных рецепторов в геноме человека составляет 218, а пептидов (пептидных предшественников) - 126 последовательностей. Идентифицировано 12 новых потенциальных пептидных рецепторов, сопряженных с G-белками, и 10 новых пептидных предшественников. По гомологии с известными последовательностями нами определена их вероятная функция. Проведена классификация всех вероятных пептидных рецепторов на основе их лигандной специфичности. Обсуждается лиганд-рецепторное соответствие пептидов и их рецепторов в геноме человека.

Ключевые слова: черновая последовательность генома человека, рецепторы, сопряженные с G-белками, пептид, пептидный рецептор

ВВЕДЕНИЕ. Суперсемейство пептидных рецепторов, сопряженных с G-белками (G-protein coupled receptors, в дальнейшем называемых GPCR), является одним из наиболее обширных белковых семейств, участвующих в регуляции большого числа физиологических процессов в клетке [1]. Эти рецепторы имеют сходное строение, для которого характерно наличие трансмембранного домена, семь раз пронизывающего мембрану. Однако в функциональном отношении эти рецепторы существенно различаются, так как

ПЕПТИДНЫЕ РЕЦЕПТОРЫ В ГЕНОМЕ ЧЕЛОВЕКА

GPCR разных групп участвуют в распознавании и проведении внеклеточных сигналов самой различной физической и химической природы, включая свет, запах, заряженные ионы, нуклеотиды, липиды, протеины и т. д. Пептидные рецепторы, лигандами которых служат эндогенные пептиды, составляют одно из подсемейств этой обширной группы рецепторов. Семейство пептидных рецепторов, в свою очередь, представляет собой достаточно большую группу белков, включая рецепторы для нейропептидов, хемокинов, опиоидов и др. Пептидные рецепторы экспрессируются почти во всех тканях организма, но их функционирование особенно важно для регуляции нервной, иммунной и эндокринных систем [2], что указывает на их ключевую роль в нормальной и патологической физиологии организма [3-5]. Данная группа пептидных рецепторов представляет интерес для исследователей не только в силу их большого биологического значения. В противоположность рецепторам других групп GPCR-семейства, в случае пептидных рецепторов оба компонента лиганд-рецепторного аппарата - как рецептор, так и его лиганд (биологический пептид) - кодируются геномом, что предоставляет возможность анализа лиганд-рецепторного аппарата в комплексе, с точки зрения как рецепторной, так и лигандной его компоненты одновременно. Это обеспечивает возможность оценки баланса лигандов и рецепторов в геноме путем сопоставления количества пептидных рецепторов, кодируемых геномом, с количеством кодируемых геномом пептидов (пептидных предшественников), соответствующих этим рецепторам. Последнее стало возможным в настоящее время благодаря завершению черновой расшифровки нуклеотидных последовательностей генома человека [6]. Поэтому нами была поставлена задача выявления всех вероятных пептидных рецепторов, сопряженных с G-белками, и соответствующих им лигандов (эндогенных пептидов) в геноме человека. Вследствие того, что как биологические пептиды, так и соответствующие им рецепторы представляют собой привлекательные мишени для фармакологических исследований и конструирования лекарств, настоящая работа может служить основой для таких разработок, а также дальнейших физиологических и клинических исследований пептидной сигнальной системы человека.

МЕТОДИКА. Принимая во внимание существенную разницу в структуре пептидных рецепторов и их лигандов, мы проводили их поиск с использованием различных подходов. Общая схема проделанных процедур представлена на рисунке 1.

Поиск рецепторов. Процедура поиска новых пептидных рецепторов в геноме человека состояла в следующем. Из баз данных SWISS-PROT (release 39.0) [7] и TrEMBL (release 21.12) [7], содержащих хорошо изученные аминокислотные последовательности, были отобраны все полностью расшифрованные аминокислотные последовательности пептидных рецепторов всех биологических видов. Полученный набор аминокислотных последовательностей (498 записей, включая 108 белков человека) был использован как базис для построения оригинального профиля подсемейства пептидных рецепторов, сопряженных с G-белками, с использованием скрытых цепей Маркова (см. ниже). Под термином "подсемейство пептидных рецепторов, сопряженных с G-белками", здесь подразумеваются только те рецепторы, которые принадлежат к семейству А (родопсиноподобных) GPCR-рецепторов (в соответствии с классификацией Kolakowski [8] и Horn et al. [9]). Пептидные рецепторы, принадлежащие к семейству В (так называемые секретиноподобные рецепторы), которые активируются большими пептидами, такими как кальцитонин или антидиуретический гормон, нами не рассматривались вследствие существенной разницы в структуре и низкой гомологии аминокислотных последовательностей как у рецепторов, так и у самих пептидов. Все отобранные для построения профиля пептидные GPCR из SWISS-PROT и TrEMBL были подвергнуты процедуре множественного выравнивания при помощи программы Clustal W [10].

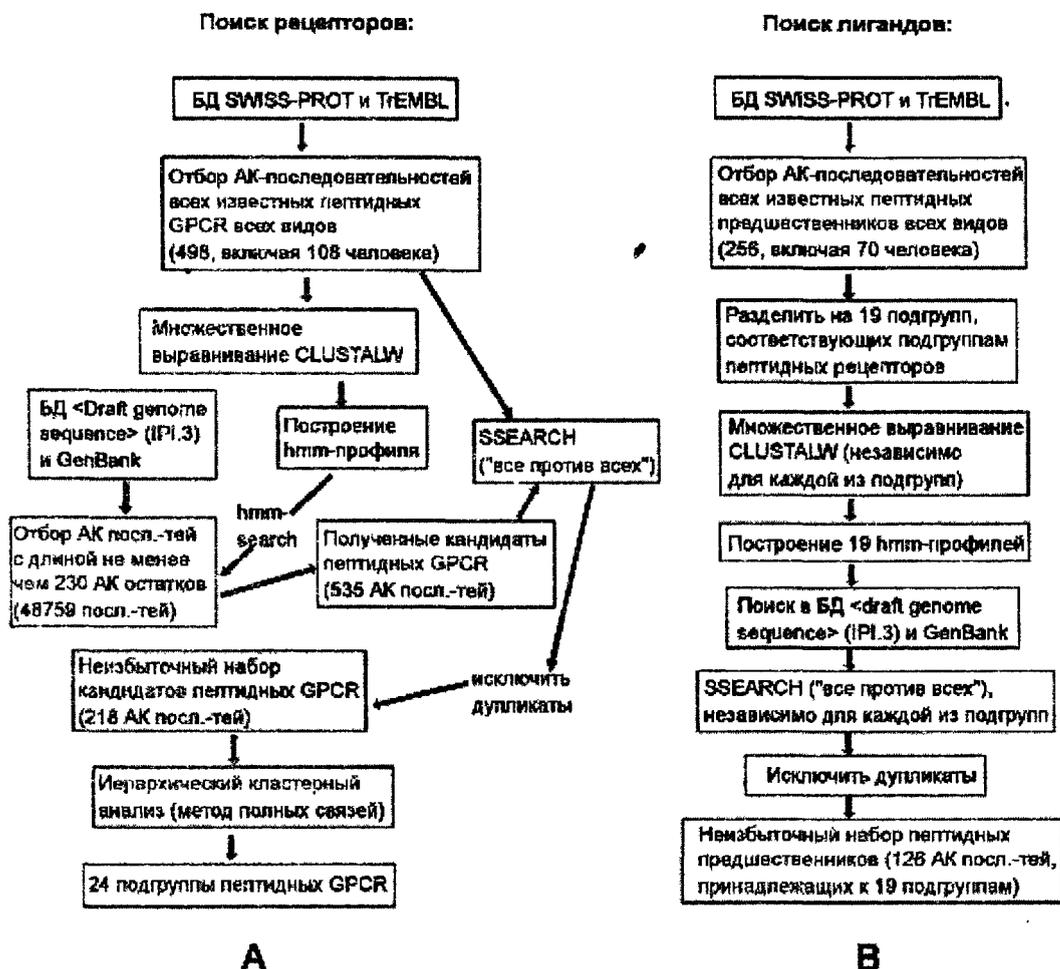


Рисунок 1.

Схема последовательного применения ряда методов биоинформатики для идентификации G-белок связанных пептидных рецепторов (А) и соответствующих им пептидных предшественников (В) в геноме человека. Детальные пояснения - см. в тексте.

Полученный в результате процедуры выравнивания файл с выровненными последовательностями пептидных рецепторов был использован как основа для построения профиля данного семейства рецепторов при помощи метода скрытых цепей Маркова, реализованного в программе Pfam [11]. После того, как был рассчитан профиль, нами был подготовлен набор аминокислотных последовательностей человека для поиска в нем вероятных пептидных GPCR. Для этого из баз данных GenBank (release 124.0) [12] и черновой последовательности генома человека (Ensembl [13], IPI.1 data set) были отобраны все аминокислотные последовательности человека длиной не менее 230 аминокислотных остатков (что соответствует минимальной длине GPCR). Таким образом, был получен набор из 48 759 аминокислотных последовательностей человека для поиска в нем кандидатов всех пептидных рецепторов в геноме человека при помощи построенного ранее профиля. Порог E-value (expectation value - ожидаемое число совпадений с исходной оценкой, не превышающей заданной) для этой процедуры был определен экспериментальным путем, то есть на основе результатов предварительного поиска рецепторных кандидатов с использованием рассчитанного нами профиля в нескольких наборах данных с известными аминокислотными последовательностями. При этом то минимальное значение

ПЕПТИДНЫЕ РЕЦЕПТОРЫ В ГЕНОМЕ ЧЕЛОВЕКА

E -value $\approx 10^{-5}$, при котором отобранные кандидаты еще принадлежат к группе пептидных рецепторов, было принято нами за пороговое значение (при большем значении этой величины отбираемые при помощи профиля кандидаты уже не относятся к пептидным рецепторам, хотя и принадлежат к группе GPCR). В результате проведенного поиска было получено 535 аминокислотных последовательностей - вероятных пептидных рецепторов. Но так как этот набор последовательностей избыточен, то для устранения дубликатов полученные данные были объединены с извлеченными ранее массивами из SWISS-PROT и TrEMBL, после чего для каждой пары из этого объединенного набора последовательностей (метод "все против всех") были рассчитаны коэффициенты сходства (similarity scores) при помощи программы SSEARCH 3.1, которая является реализацией алгоритма Smith-Waterman [14]. В результате этой процедуры из полученного набора кандидатов пептидных рецепторов, сопряженных с G-белками, были устранены двойные записи. Полученный таким образом не избыточный набор данных содержит 218 аминокислотных последовательностей, пригодных для дальнейшего анализа. Причем этот набор содержит 12 новых (не идентифицированных ранее) вероятных пептидных GPCR. С целью классификации всех полученных в результате поиска кандидатов пептидных рецепторов, сопряженных с G-белками, а также для определения функции 12 не идентифицированных ранее вероятных рецепторов был проведен иерархический кластерный анализ (метод полных связей) на основе полученных ранее коэффициентов гомологии. В результате этой процедуры было установлено 24 группы пептидных рецепторов с различной лигандной специфичностью.

Поиск лигандов пептидных рецепторов Биологически активные пептиды являются продуктами ферментативного расщепления так называемых пептидных предшественников. Одна последовательность пептидного предшественника может содержать единственный биологически активный пептид, несколько копий одного и того же пептида, либо несколько различных биологически активных пептидов. Длина молекулы предшественника колеблется в пределах от 70 до 600 аминокислотных остатков, а длина молекулы биологически активного пептида, полученного в результате ферментативного расщепления молекулы предшественника, составляет от 3 до 200 аминокислотных остатков. Таким образом, значительная вариабельность длины, внутренней организации и низкая степень гомологии аминокислотных последовательностей делают очень сложным прямой поиск кандидатов пептидных предшественников в базах данных с использованием традиционных методов, таких как BLAST [15]. В отличие от рецепторов, сопряженных с G-белками, которые имеют такую отличительную черту, как наличие в структуре характерного трансмембранного домена, что делает их поиск стандартными методами относительно несложным, поиск пептидных предшественников требует специального подхода. В настоящей работе был использован комбинированный метод, основанный на отборе последовательностей пептидных предшественников, соответствующих полученным в результате кластерного анализа подгруппам пептидных рецепторов, и построении оригинальных профилей для каждой из подгрупп предшественников по отдельности. Такой подход, как построение профиля для каждой из подгрупп по отдельности, неприменим в случае рецепторов, так как чрезвычайно высокая степень гомологии между пептидными GPCR существенно усложняет определение порога чувствительности. Таким образом, сначала были собраны все известные аминокислотные последовательности пептидных предшественников всех биологических видов из баз данных SWISS-PROT и TrEMBL отдельно для каждой из подгрупп пептидных рецепторов, полученных ранее в результате кластерного анализа. Несмотря на то, что число подгрупп пептидных рецепторов равно 24, число групп пептидных предшественников включает в себя лишь 19 подгрупп. Этот факт может быть объяснен тем, что некоторые лиганды пептидных рецепторов (эндогенные пептиды) либо просто не имеют предшественников, кодируемых в

геноме (что известно, например, в случае опиоидных или N-формил пептидных рецепторов), либо же лиганды этих рецепторов в настоящее время неизвестны (семейство хемокино-подобных рецепторов). По алгоритму, описанному выше для пептидных рецепторов, были построены профили для каждой из подгрупп пептидных предшественников, после чего был проведен поиск кандидатов пептидных предшественников в черновой последовательности генома человека (PI.1 data set) и базе данных GenBank (E-value < 0,05). На заключительной стадии анализа для каждой подгруппы найденных вероятных кандидатов пептидных предшественников был проведен поиск гомологии последовательностей при помощи программы SSEARCH для исключения дубликатов.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ. Полученное в результате проведенного нами анализа полное число функциональных пептидных рецепторов, сопряженных с G-белками, в геноме человека составляет 218. Этот набор белков является не избыточным и включает в себя 12 впервые идентифицированных кандидатов из черновой последовательности генома. Кластерный анализ всего набора данных по рецепторам указывает на наличие четкой тенденции к разделению пептидных рецепторов в соответствии с их лигандной специфичностью. Нами были получены 24 подгруппы пептидных рецепторов, сопряженных с G-белками, соответствующих различным классам биологических пептидов (см. рис. 2). Более детальная информация представлена на сайте <http://web.kuicr.kyoto-u.ac.jp/~maria/peptide-cluster.htm>). В результате визуального анализа полученной дендрограммы была выявлена принадлежность 12 впервые идентифицированных в настоящей работе вероятных рецепторов к соответствующим подгруппам семейства пептидных рецепторов, идентифицированным в результате кластерного анализа (например, подгруппа опиоидных или галаниновых рецепторов). Впервые идентифицированные нами кандидаты GPCR имеют высокую степень гомологии с известными ранее пептидными рецепторами какой-либо из полученных 24 подгрупп (коэффициент Smith-Waterman, рассчитанный программой SSEARCH, колеблется в пределах от ~600 до ~2000). Таким образом, отнесение новых кандидатов пептидных рецепторов к какой-либо из подгрупп рецепторов, обладающих разной лигандной специфичностью, позволяет определить вероятную лигандную специфичность новых рецепторов на основе гомологии их последовательностей с хорошо изученными белками. Идентификационные номера новых кандидатов рецепторов

Таблица 1. Новые пептидные рецепторы, сопряженные с G-белками, идентифицированные в геноме человека

Идентификационный номер из набора данных PI 1	Вероятная биологическая функция
IGI M1 ctg14846 47	Adrenocorticotropic hormone receptor
IGI M1 ctg13993 13	Adrenomedullin receptor
IGI M1 ctg13260 4	Angiotensin type 1 receptor
IGI M1 ctg15894 34	C-C Chemokine type 6 receptor
IGI M1 ctg14493 1	C-C Chemokine type 9 receptor
IGI M1 ctg13238 8	C-X-C Chemokine type 4 receptor
IGI M1 ctg12387 1	Chemokine/ like receptor 1 receptor
IGI M1 ctg12638 24	Galanin receptor
IGI M1 ctg25478 12	Neuropeptide Y type 4 receptor
IGI M1 ctg12905 3	Neuropeptide Y type 6 receptor
IGI M1 ctg13541 2	Opioid type K receptor
IGI M1 ctg16334 16	Opioid receptor

и их возможная лигандная специфичность приведены в таблице 1.

Поиск кандидатов пептидных предшественников с использованием построенных нами оригинальных профилей выявил, что количество пептидных

ПЕПТИДНЫЕ РЕЦЕПТОРЫ В ГЕНОМЕ ЧЕЛОВЕКА

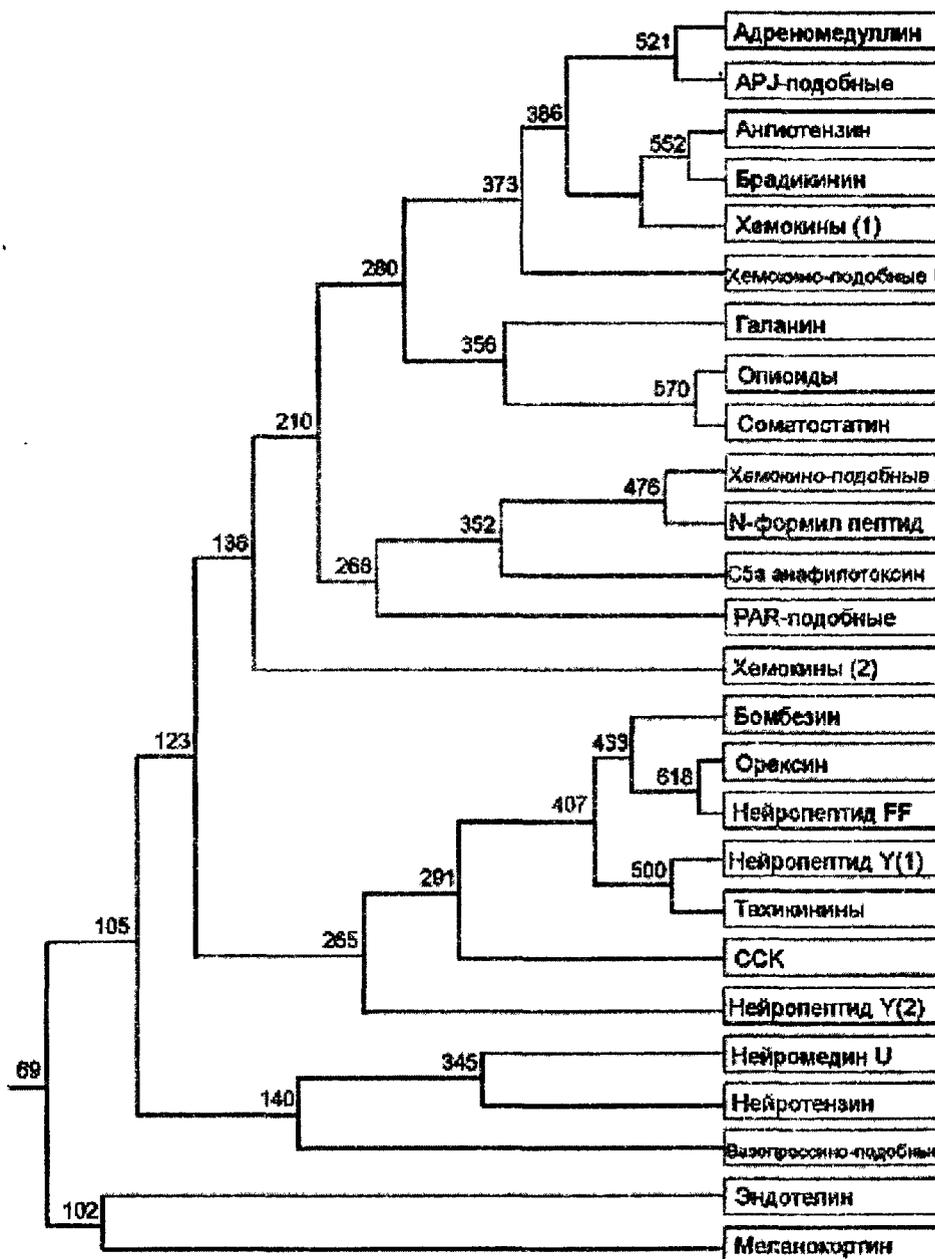


Рисунок 2.

Семейство пептидных рецепторов, сопряженных с G-белками, в геноме человека. Представлены результаты иерархического кластерного анализа, основанного на гомологии последовательностей (метод полных связей); указанные цифры обозначают статистические оценки локального выравнивания. Цифрами в круглых скобках после названия рецепторных групп обозначены номера подгрупп тех групп рецепторов, которые при кластеризации разбились на подгруппы.

предшественников в геноме человека составляет 126, причем это число включает в себя 10 впервые открытых в данной работе кандидатов. Эти результаты представлены в таблице 2. Общее количество пептидов и их рецепторов в геноме человека, разделенных в результате кластерного анализа на функциональные подгруппы, приведено в таблице 3 (полный список всех идентификационных номеров как рецепторов так и пептидов представлен в Интернете на сайте

Таблица 2. Новые пептидные предшественники, идентифицированные в геноме человека

Идентификационный номер из набора данных IPI 1	Вероятная биологическая функция
IGI M1 ctg12246 51	Bombesin precursor
IGI M1 ctg14493 14	ССК precursor
IGI M1 ctg3461 111	Chemokine precursor
IGI M1 ctg3461 117	Chemokine precursor
IGI M1 ctg14101 9	Chemokine precursor
IGI M1 ctg14439 43	Chemokine precursor
IGI M1 ctg3461 31	Chemokine precursor
IGI M1 ctg13740 44	Chemokine precursor
IGI M1 ctg17129 22	Orexin precursor
IGI M1 ctg14601 23	Tachykinin precursor

<http://web.kuicr.kyoto-u.ac.jp/~maria/supplement.htm>).

В результате проведенного иерархического кластерного анализа всех 218 идентифицированных пептидных рецепторов, сопряженных с G-белками, была получена дендрограмма, представленная на рисунке 2. Было выявлено 24 группы рецепторов, строго соответствующих различным пептидным лигандам. Но при анализе представленной дендрограммы обращает на себя внимание тот факт, что две из этих 24 групп, а именно группы рецепторов хемокинов и нейропептидов Y, разбились каждая дополнительно еще на две подгруппы, члены которых имеют большее сходство с пептидными рецепторами из других семейств, чем с рецепторами из своей группы. При более детальном рассмотрении каждой из образовавшихся подгрупп (см. <http://web.kuicr.kyoto-u.ac.jp/~maria/peptide-cluster.htm>) можно установить, что в случае хемокиновых рецепторов в первую подгруппу входят рецепторы типа C-C 1-5, C-C 8, C-C X и C-X3-C. Рецепторы типа C-C 6-11, C-X-C 3-5, а также рецепторы интерлейкина-8 и BONZO-рецепторы составляют вторую подгруппу. В случае же рецепторов нейропептида Y (NPY) в отдельную подгруппу обособлены только рецепторы типа NPY2, вторую же подгруппу составляют оставшиеся рецепторы данной группы NPY1 и NPY3-6. Данный факт достаточно сложно интерпретировать с функциональной точки зрения, но можно предположить, что такое разбиение на подгруппы, свидетельствующее о высокой степени гомологии аминокислотных последовательностей с другими членами семейства пептидных рецепторов, может указывать на наличие не выявленных пока эволюционных или функциональных взаимосвязей между данными подтипами рецепторов.

Также при анализе полученной дендрограммы выявлено, что наибольшим родством обладают близкие в функциональном отношении группы рецепторов: например, ангиотензины и брадикинины, опиоиды и соматостатин, орексин и нейропептид FF. Противоположная картина наблюдается лишь для производных меланокортина и эндотелина, имеющих низкую степень гомологии аминокислотных последовательностей как друг с другом, так и с остальными пептидными рецепторами, сопряженными с G-белками, что может указывать на структурно-функциональную специфичность рецепторов данных групп.

Посредством визуального анализа полученной дендрограммы была выявлена принадлежность 12 впервые идентифицированных в настоящей работе вероятных рецепторов к соответствующим подгруппам семейства пептидных рецепторов, идентифицированным в результате кластерного анализа. Полный список названий идентифицированных в настоящей работе рецепторов и их вероятные биологические функции приведены в таблице 1.

Поиск кандидатов пептидных предшественников лигандов GPCR с использованием оригинальных профилей выявил, что их количество в геноме человека составляет 126, причем это число включает в себя 10 впервые открытых в данной работе кандидатов. Полный список названий впервые

ПЕПТИДНЫЕ РЕЦЕПТОРЫ В ГЕНОМЕ ЧЕЛОВЕКА

идентифицированных в настоящей работе пептидных предшественников и их вероятные биологические функции представлены в таблице 2.

После идентификации всех возможных кандидатов пептидных рецепторов и пептидных предшественников в геноме человека мы сопоставили количество пептидных рецепторов, сопряженных с G-белками, с количеством соответствующих им лигандов для того, чтобы ответить на вопрос, существует ли взаимно-однозначное соответствие между ними на геномном уровне. Полученные нами результаты показывают, что вероятное полное число функциональных пептидных рецепторов в геноме человека составляет 218, а число пептидных предшественников - 126. Но несмотря на то, что число полученных в результате кластерного анализа подгрупп пептидных рецепторов равно 24, число групп пептидных предшественников содержит только 19 подгрупп (см. выше раздел "Поиск лигандов пептидных рецепторов"). Таким образом, принимая во внимание тот факт, что для пяти подгрупп пептидных рецепторов не найдено соответствующих им пептидных предшественников, среднее отношение числа всех предполагаемых пептидных рецепторов к числу их лигандов составляет приблизительно 3:2 (табл. 3). Но как легко заметить (см. табл. 3), более, чем в

Таблица 3. Общее количество биологических пептидов и их рецепторов в геноме человека *

	Название рецептор-лигандной подгруппы	Общее количество рецепторов (включая новые)	Количество только новых кандидатов рецепторов	Общее количество лигандов (включая новые)	Количество только новых кандидатов лигандов
01	Adrenomedullin	2	1	2	-
02	Angiotensin	6	1	5	-
03	APJ-like	1	-	2	-
04	Bombesin	3	-	4	1
05	Bradykinin	5	-	2	-
06	C5a anaphylatoxin	4	-	3	-
07	CCK	7	-	3	1
08	Chemokines & IL-8	69	3	69	6
09	Chemokine-like I	6	1	-	-
10	Chemokine-like II	3	-	-	-
11	Endothelin	9	-	5	-
12	N-formyl peptide	7	-	-	-
13	Galanin	4	1	3	-
14	Melanocortin	16	1	2	-
15	Neuromedin U	6	-	1	-
16	Neuropeptide Y	14	2	4	-
17	Neurotensin	2	-	2	-
18	Opioid	14	2	-	-
19	Orexine	3	-	2	1
20	Neuropeptide FF	2	-	1	-
21	PAR-like	8	-	-	-
22	Somatostatin	10	-	2	-
23	Tachykinin	11	-	7	1
24	Vasopressin-like	6	-	7	-
	Общее число	218	12	126	10

*Подгруппы рецепторов, не имеющих соответствующих им пептидных предшественников (см. текст), заштрихованы. Сокращения: CCK, cholecystokinin; IL-8, interleukin-8; PAR, proteinase-activated receptor.

половине случаев число рецепторов приблизительно соответствует числу лигандов внутри каждой отдельной подгруппы, то есть их отношение очень близко к пропорции 1:1. Особенно интересен случай хемокиново-интерлейкиновой группы, где количество рецепторов этой обширной белковой группы строго

соответствует количеству пептидных предшественников (69 лиганд-рецепторных пар). Однако, примерно в трети случаев число рецепторов значительно превалирует над числом лигандов (например, в случае подгрупп меланокортина, соматостатина, нейропептида Y и так далее). Возможно несколько объяснений данного феномена. Прежде всего, не может игнорироваться тот факт, что одна последовательность предшественника может содержать несколько копий одного и того же пептида, либо несколько разных биологических пептидов. Например, все меланокортиновые пептиды имеют общий предшественник (проопиомеланокортин), процессинг которого ведет к образованию четырех зрелых пептидов разной природы: адренокортикотропин (ACTH), α -MSH, β -MSH и Lys- γ 2-MSH [16]. Другим аналогичным примером может служить предшественник соматостатина, ферментативное расщепление которого приводит к образованию двух зрелых биологически активных продуктов (соматостатин-14 и соматостатин-28) [17]. Подобных примеров существует немало. Но только лишь один этот факт не может служить достаточным объяснением феномена "недостаточности" пептидных предшественников в геноме, так как существует немало пептидных предшественников, содержащих в себе всего лишь один пептид. Другим вероятным объяснением такого явления может служить тот факт, что не исключены случаи действия одного пептидного лиганда на различные подтипы пептидных рецепторов. В таком случае, относительно малое количество пептидных предшественников может быть оправдано с точки зрения множественности действия пептидов. И, наконец, еще одним возможным (и, вероятно, наиболее убедительным) объяснением количественного превалирования числа рецепторов над числом лигандов могут быть технические трудности в поиске пептидов, упомянутые выше. Наш метод поиска выявил несколько новых кандидатов пептидных предшественников, которые обнаруживают высокую степень гомологии с известными ранее последовательностями. Но в силу объективных препятствий, затрудняющих поиск пептидов в геноме, можно предположить, что некоторые кандидаты не были найдены и что представленный в настоящей работе список пептидных предшественников неполон, и имеются некоторые неидентифицированные гены пептидных предшественников.

В результате проведенной нами работы мы показали, что вероятное количество функциональных пептидных рецепторов, сопряженных с G-белками, в геноме человека составляет 218, а число соответствующих им лигандов - пептидных предшественников - 126 белков. Это число включает в себя 12 новых пептидных рецепторов и 10 новых пептидных предшественников. Выполнена кластеризация рецепторов и их лигандов, позволившая установить вероятную биологическую функцию впервые идентифицированных в настоящем исследовании белков. Сопоставлено количество пептидных рецепторов, сопряженных с G-белками, с количеством соответствующих им лигандов, что позволило нам сделать выводы о лиганд-рецепторном соответствии на геномном уровне. Полученные результаты могут использоваться для дальнейших исследований пептидной сигнальной системы человека.

Настоящая работа была осуществлена при финансовой поддержке министерства образования, науки, спорта и технологии Японии, Японского общества поощрения науки и Корпорации науки и технологии Японии. Компьютерные ресурсы были предоставлены Центром биоинформатики института химических исследований Университета Киото, Япония.

ЛИТЕРАТУРА

- 1 Bockaert J., Pin J.P. (1999) *EMBO J.*, **18**, 1723-1729
- 2 Karelin A.A., Blishchenko E.Yu., Ivanov V.T. (1998) *FEBS Lett.*, **428**, 7-12
- 3 Tseng L.F. (2001) *Trends Pharmacol Sci.*, **22**, 623-630

ПЕПТИДНЫЕ РЕЦЕПТОРЫ В ГЕНОМЕ ЧЕЛОВЕКА

- 4 Yamada K., Santo-Yamada Y., Wada E., Wada K. (2002) *Mol. Psychiatry*, **7**, 113-117
- 5 Seifert R., Wenzel-Seifert K. (2001) *J. Biol. Chem.*, **276**, 42043-42049
- 6 Lander E.S., Linton L.M., Birren ., Nusbaum C., Zody M.C., Baldwin J., et al. (2001) *Nature*, **409**, 860-921
- 7 <http://www.expasy.ch/sprot/sprot-top.html>
- 8 Kolakowski L.F. Jr. (1994) *Receptors Channels*, **2**, 1-7
- 9 Horn F. (1998) *Nucleic Acids Res.*, **26**, 275-279
- 10 Thompson J.D., Higgins D.G., Gibson T.J. (1994) *Nucleic Acids Res.*, **22**, 4673-4680
- 11 Bateman A., Birney E., Cerruti L., Durbin R., Etwiller L., Eddy S.R., et al. (2002) *Nucleic Acids Res.*, **30**, 276-280
- 12 <http://www.ncbi.nlm.nih.gov/Genbank/index.html>
- 13 <http://www.ensembl.org>
- 14 Smith T.F., Waterman M.S. (1981) *J Mol Biol*, **147**, 195-197
- 15 Muggleton S.H., Bryant C.H., Srinivasan A., Wittaker A., Topp S., Rawlings C. (2001) *J. Comput. Biol.*, **8**, 493-521
- 16 Versteeg D.H., Van Bergen P., Adan R.A., De Wildt D.J. (1998) *Eur. J. Pharmacol.*, **360**, 1-14
- 17 Patel Y.C. (1999) *Front. Neuroendocrinol.*, **20**, 157-198

Поступила 23.01.2004

G-PROTEIN COUPLED PEPTIDE RECEPTORS AND THEIR ENDOGENOUS LIGANDS IN THE HUMAN GENOME

M.E. Levchenko¹, V.V. Poroikov¹, M. Kanehina²

¹V.N. Orekhovich Institute of Biomedical Chemistry Rus. Acad. Med. Sci., Pogodinskaya Str., 10, Moscow, 119121, Russia; tel: +7 095 247-3029; Fax: +7 095 245-0857

E-mail: vvp@ibmh.msk.su

²Bioinformatics Center, Institute for Chemical Research, Kyoto University, Japan

The G-protein coupled peptide receptors as well as their ligands, endogenous peptides, are involved in regulation of many important physiological processes in the organism and therefore represent attractive targets for pharmaceutical investigation and drug design. With the completion of the human draft genome sequencing, it has become possible to take a comprehensive picture of all genes encoding both peptide receptors and peptides themselves. In the present study a first attempt has been made to carry out a comprehensive analysis of G-protein coupled peptide receptors and their respective endogenous peptide ligands in the human genome. We searched the genome sequence by means of sequential application of standard bioinformatical methods (such as homology search, hierarchical cluster analysis, building of hmm-profiles etc.) with the goal of identifying all the components of peptide ligand/receptor system in the human genome. As a result of this search it was concluded that the probable number of functional peptide receptors in the human genome is 218, and the probable peptide precursors' number is 126 amino acid sequences. These two groups include, respectively, 12 novel G-protein coupled peptide receptors and 10 novel peptide precursors, discovered in the present study. The probable biological functions of newly discovered candidates were determined based on the sequence similarity to the earlier known proteins. Classification of all peptide GPCRs and their ligands based on the ligand specificity was performed for all probable G-protein coupled peptide receptors. The issue of ligand-receptor specificity in the human genome is also discussed.

Key words: human draft genome sequence, bioinformatics, GPCRs, peptide, peptide receptor