HCVMAP: БАЗА ДАННЫХ ПО ФУНКЦИОНАЛЬНОМУ КАРТИРОВАНИЮ БЕЛКОВ ВИРУСА ГЕПАТИТА С

[1]Соболев Б.Н., [2]Оленина Л.В., [1]Кураева Т.Е., [1]Колесанова Е.Ф., [1]Рудик А.В., [1]Поройков В.В., [1]Арчаков А.И.

[1]НИИ биомедицинской химии им. В.Н. Ореховича РАМН, Москва

[2]Институт биологии гена РАН, Москва

Представлена база данных по функциональному картированию белков вируса гепатита С – одного из наиболее распространенных этиологических агентов, вызывающих вирусные гепатиты. Ресурс предоставляет сведения по структурным особенностям антигенных участков в аминокислотных последовательностях, относящимся к разным вариантам вируса, что позволяет анализировать особенности структурно-функциональных взаимосвязей.

Введение

В связи с быстрым увеличением числа расшифрованных бактериальных и вирусных геномов возникли предпосылки для принципиально нового подхода к созданию вакцин – т.н. “обратной вакцинологии” [1]. Сегодня исследователи могут отбирать потенциальные иммуногенные компоненты уже на стадии компьютерного анализа генных последовательностей инфекционных агентов. Особенно перспективным представляется этот подход в тех случаях, когда инфекционный агент отличается выраженной генетической изменчивостью, которая ведет к быстрому появлению вариантов с новыми антигенными свойствами. В связи с этим возникла необходимость в создании специализированных информационных ресурсов по картированию вирусных белков, которые бы содержали данные о структурном разнообразии или консервативности выявленных функциональных и антигенных участков. Подобные ресурсы позволят отработать новые подходы к определению антигенной специфичности, наметить перспективные молекулярные мишени для создания новых вакцинных препаратов, а также средств иммунодиагностики и химиотерапии.

Характерным примером инфекционного агента, проявляющего высокую изменчивость, является вирус гепатита С (ВГС), для которого характерно большое число вариантов, постоянно увеличивающееся за счет идентификации все новых и новых геномных последовательностей. Вирус гепатита С (ВГС) является основным этиологическим фактором парентерально распространяемых вирусных гепатитов человека. Общее количество инфицированных ВГС приближается к 300 млн. человек и составляет около 3% населения земного шара [2]. В России и других странах число выявленных случаев ВГС инфекции увеличивается с каждым годом в несколько раз [3]. ВГС представляет серьезную угрозу здоровью человечества, тем самым обуславливая насущную необходимость изучения вируса, диагностики и лечения ВГС инфекции, а также вакцинопрофилактики.

Нами разработана и составлена база данных по функциональному картированию белков, кодируемых геномами разных вариантов ВГС – HCVMAP. Представленные в ней наборы данных были использованы нами при определении потенциальных иммуногенных участков, наиболее перспективных для разработки вакцинных конструкций [4]. С 2002 года база данных HCVMAP представлена в Интернет (www.ibmh.msk.su/hcvmap). Сегодня HCVMAP содержит структурные и функциональные данные по антигенным участкам белков ВГС. В ближайшее время наша база данных будет представлять данные и по другим функциональным характеристикам (рецепторной специфичности, устойчивости к лекарственным препаратам и т.д.).

Исходные данные

Первичные наборы данных включают аминокислотные последовательности белков ВГС и сведения по их функциональным характеристикам.

Исходные аминокислотные последовательности взяты из общедоступных баз данных Swiss-Prot и TrEmble и получены, главным образом, за счет «компьютерной трансляции» генных нуклеотидных последовательностей. Напомним, что геном ВГС содержит единственную рамку считывания, которая кодирует т.н. полипротеин-предшественник длиной более 3000 аминокислотных остатков. Этот белок расщепляется на 3 структурных белка, формирующих вирусную частицу, и 6 неструктурных белков, обеспечивающих жизненный цикл вируса (схема на странице www.ibmh.msk.su/hcvmap/profile_frame.php). Аминокислотные последовательности представлены символьными строками, соответствующими полному белку-предшественнику или его фрагменту. К концу марта 2004 года число неидентичных последовательностей собранных в HCVMAP, составило около 9000. Каждая из соответствующих записей содержит также список идентификаторов Swiss-Prot/TrEmbl, которые соответствуют точным копиям или подстрокам данной последовательности.

Функциональные характеристики отбираются из литературных источников с использованием доступных библиографических ресурсов (PubMed и др.). В базу данных включаются главным образом сведения, которые оцениваются при экспертном анализе полнотекстовых документов.

Обработанные данные

В эту группу входят выровненные аминокислотные последовательности и функциональные (антигенные) участки, картированные по выровненным последовательностям.

Для первоначального выравнивания последовательностей используется программа ClustalW. Ручная редакция выровненных последовательностей производится с помощью пакета ProMSED, любезно предоставленного А.М. Ерошкиным [5]. Эта процедура выполнена на представительном наборе последовательностей с уровнем идентичности менее 85%. Остальные последовательности присоединятся к этому шаблонному выравниванию с использованием собственной подпрограммы.

Данные по антигенному картированию собраны в таблицах B- и T-эпитопов, которые находятся на отдельных страницах:

B-epitopes (www.ibmh.msk.su/hcvmap/bepitope1.php);

T-epitopes (www.ibmh.msk.su/hcvmap/tepitope.php).

Таблицы содержат последовательности пептидов, с разметкой их положения в белках ВГС (нумерация по полной последовательности полипротеина-предшественника - за исключением мимотопов). Каждая запись содержит короткую функциональную характеристику участка. Большинство представленных эпитопов специфично для человека, однако некоторые из них определены в опытах с животными моделями.

Совмещение выровненного набора последовательностей с функциональными картами позволяет “вырезать” выровненные фрагменты, покрывающие данный функциональный участок.

Работа с таблицами и навигация

Просматривая таблицу эпитопов пользователь может вызывать страницу с литературной ссылкой, использовав значок “книжки”. Вызванная страница содержит реферат соответствующей статьи и гиперссылку на библиографическую систему PubMed. Щелчок мыши на строке, отображающей пептидный фрагмент, вызовет выровненный набор гомологичных участков с частотами встречаемости отдельных пептидов в наборе

последовательностей. Здесь можно вызвать список последовательностей, содержащих данный пептид, и выбрать из этого списка последовательность, в которой указаный фрагмент будет выделен цветом.

На отдельной странице Profile (www.ibmh.msk.su/hcvmap/profile_frame.php) можно просмотреть профиль встречаемости аминокислотных остатков в выровненных последовательностях. Навигация может осуществляться с помощью схемы полипротеина-предшественника.

Заключение

За последние два десятилетия был идентифицирован ряд вирусов, проявляющих высокую изменчивость, так что каждый из них представлен множеством геномов. Сопоставление этих последовательностей особенно результативно тогда, когда оно производится с привлечением функциональных характеристик, отражающих различия или сходство между отдельными вариантами вируса и, таким образом, дополняет другие ресурсы, посвященные этой проблеме (http://hepatitis.ibcp.fr/, http://s2as02.genes.nig.ac.jp/, http://hcv.lanl.gov/content/hcv-db/). Предоставляемые материалы могут использоваться для выявления структурных признаков, характерных для различных серологических групп. Соответствующие инструментальные средства планируется разместить на сайте ресурса.

Работа поддержана междисциплинарной программой Министерства промышленности, науки и технологий РФ "Вакцины нового поколения и медицинские диагностические системы будущего"

Литература

1. *Mora, M., Veggi, D., Santini, L., Pizza, M, Rappuoli, R.* Reverse vaccinology // Drug Discov. Today., 2003, v. 15, p. 459-464.
2. *Rosen H.R., Gretch D.R.* Hepatitis C virus: current understanding and prospects for future therapies // Molecular Medicine Today, 1999, v. 5, p. 393-399
3. *Балаян М.С., Михайлов М.И.* Энциклопедический словарь – вирусные гепатиты.- М.:Амипресс, 1999. - 304с.
4. *Olenina, L.V., Nikolaeva, L.I., Sobolev, B.N., Blokhina, N.P., Archakov, A.I., Kolesanova, E.F.* (2002) Mapping and characterization of B cell linear epitopes in the conservative regions of hepatitis C virus envelope glycoproteins // J. Viral. Hepat., 2002, v. 9, p. 174-182.
5. *Frolov, A.S., Pika, I.S., Eroshkin, A.M.* ProMSED: protein multiple sequence editor for Windows 3.11/95 // 1997, Comput. Appl. Biosci., 13, 243-248.