

УДК 577.122:616.65-006-07

База данных «Протеомика рака простаты»

С. С. Шишкин¹, Л. И. Ковалев¹, М. А. Ковалева¹, И. Н. Крахмалева¹, К. В. Лисицкая¹,
Л. С. Еремина¹, А. В. Иванов¹, Е. В. Герасимов¹, Э. Г. Садыхов¹, Н. Ю. Уласова²,
О. С. Соколова², И. Ю. Торопыгин³, В. О. Попов¹

¹Учреждение Российской академии наук Институт биохимии им. А.Н. Баха РАН, 119071, Москва, Ленинский проспект, 33

²Московский государственный университет им. М.В. Ломоносова, 119991, Москва, Ленинские горы, МГУ, 1

³Учреждение Российской академии медицинских наук НИИ биомедицинской химии им. В.Н. Ореховича РАМН, 119121, Москва, ул. Погодинская, 10

E-mail: shishkin@inbi.ras.ru

Поступила в редакцию 28.09.2010 г.

РЕФЕРАТ На основе результатов протеомных исследований тканей простаты от пациентов со злокачественными и доброкачественными опухолями, а также ряда культивируемых клеточных линий человека создан информационный ресурс – база данных «Протеомика рака простаты» («ПП», <http://ef.inbi.ras.ru>). «ПП» содержит семь взаимосвязанных модулей, в состав каждого из которых входят протеомные и биомедицинские данные о белках объекта исследований, представленные на четырех информационных уровнях. Основа модулей – синтетические двумерные карты (первый информационный уровень), представляющие собой обобщенные результаты фракционирования белков методом двумерного электрофореза. На каждой карте показаны белки, идентифицированные с помощью масс-спектрометрии. Второй информационный уровень образуют данные, полученные при изучении свойств каждого из идентифицированных белков. На третьем уровне собраны различные опубликованные материалы о каждом белке, включая биомедицинские сведения. На четвертом уровне представлены интернет-ссылки, связывающие описание белка в «ПП» с записями в публичных базах данных NCBI и UniProt. Всего в «ПП» представлена информация о 359 белках, среди которых оказались 17 потенциальных белковых маркеров рака простаты, в частности AGR2, представители семейств аннексинов, S100 и ряд других белков, включая новые белки PRO2675 и PRO2044. Созданная база данных может быть использована для решения широкого круга задач, в число которых входит изучение молекулярных основ этиологии и патогенеза заболеваний простаты, поиск новых диагностических маркеров и т.д.

КЛЮЧЕВЫЕ СЛОВА: протеомика, рак простаты, компьютерная база данных.

ВВЕДЕНИЕ

Первое десятилетие постгеномной эры ознаменовалось интенсивным развитием биоинформатики, расширением хорошо известных крупных баз данных (в частности, NCBI и UniProt), а также созданием во многих странах специальных информационных ресурсов, предназначенных для обеспечения разнообразных биомедицинских исследований [1–4]. В этом отношении на фоне информационных ресурсов, созданных в Ирландии (UCD-2DPAGE, <http://proteomics-portal.ucd.ie:8082/cgi-bin/2d/2d.cgi>) [2] и Индии (Human Proteinpedia, www.humanproteinpedia.org) [3], достижения отечественных ученых выглядят весьма скромными.

В настоящее время одним из наиболее актуальных направлений в биомедицинских исследованиях пред-

ставляется поиск эффективных биомаркеров рака простаты (РП), которые могли бы обеспечить разработку новых методов диагностики этого заболевания [5–8]. Значительное внимание к РП обусловлено тем, что за последние годы частота этого заболевания значительно возросла, в том числе и в России [9, 10], а в ряде стран РП вышел на первое место по частоте встречаемости в структуре мужской онкопатологии [11, 12]. В ранней диагностике РП в настоящее время центральное место принадлежит определению в крови одного из наиболее изученных биомаркеров – так называемого простат-специфического антигена (ПСА). Однако известно, что этот тест дает значительное количество ложноположительных и ложноотрицательных результатов, которые приводят не только к негативным клиническим послед-

ствиям, но и к большим финансовым потерям [5, 7]. Как следствие, в США и других западных странах развернулись поиски новых биомаркеров РП, которые особенно активизировались благодаря развитию протеомных и других постгеномных технологий [6, 8, 13].

На основе использования различных протеомных технологий с 2005 г. в Институте биохимии им. А.Н. Баха РАН в сотрудничестве с рядом других научных и медицинских учреждений также проводятся поиски новых биомаркеров РП [14, 15]. Для оптимизации этих исследований в 2009 г. был создан отечественный информационный ресурс «Протеомика рака простаты» («ПРП», <http://ef.inbi.ras.ru>), в котором суммированы экспериментальные материалы, собраны актуальные опубликованные данные с указанием источников, а также приведены ссылки на ряд других биомедицинских информационных ресурсов, представленных в сети Интернет. В данной статье описана структура новой расширенной версии «ПРП» и рассмотрены возможности использования этого информационного массива.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Биоматериалы – биоптаты и операционный материал тканей простаты больных РП ($n = 72$) и доброкачественной гиперплазией простаты (ДГП, $n = 69$) – предоставлены сотрудниками отделения урологии городской клинической больницы им. С.П. Боткина. Диагностика заболеваний осуществлялась с использованием клинических, гистологических и иммунохимических (определение уровня ПСА) методов. Проводимую с целью гистологической верификации диагноза трансректальную мультифокальную пункционную биопсию выполняли под контролем трансректальной ультрасонографии и получали от одного больного до 18 образцов тканей из различных зон простаты [16, 17]. Гистологически все случаи РП были охарактеризованы как аденокарцинома. Стадию рака по Глиссону определяли по общепринятой методике [16, 17].

Параллельно анализировали белки клеточных культур РС-3 (ACC 465), DU-145 (ACC 261), ВРН-1 (ACC 143), приобретенных в German Collection of Microorganisms and Cell Cultures (Германия), а также культивируемых клеток линии LNCaP, образцы которой получены от И.Г. Шемякина (ГНЦ прикладной микробиологии и биотехнологии, г. Оболensk). Клетки культивировали в среде RPMI-1640 с добавлением NEPEs, пирувата натрия, гентамицина и 20% эмбриональной телячьей сыворотки (ЭТС) как описано ранее [18]. Клетки выращивали с использованием культурального пластика («Costar», США и «Nunc», Дания) в CO₂-инкубаторе («Sanyo», Япония). Кроме того, ис-

следовали белки культивируемых клеток двух линий рабдомиосаркомы человека (A-204 и RD), приобретенных в НИИ вирусологии им. Д.И. Ивановского РАМН, а также белки культивируемых нормальных миобластов человека, любезно предоставленных Т.Б. Крохиной, выведенных как описано ранее [19].

Приготовление белковых экстрактов, их фракционирование методом двумерного электрофореза по О'Фарреллу, визуализацию белков окрашиванием Кумасси голубым R-250 и азотнокислым серебром и анализ полученных двумерных электрофореграмм (ДЭ) выполняли как описано ранее [20, 21]. Кроме того, применяли технологию двумерного электрофореза, включающую изоэлектрофокусирование с использованием иммобилизового градиента pH (IPG-PAGE) и комплекта оборудования Ettan IPGphor 3 фирмы «GE Healthcare» (Швеция) по протоколу фирмы-производителя. Идентификацию белков методами MALDI-TOF MS (matrix-assisted laser desorption/ionization time-of-flight mass spectrometry) и MS/MS (tandem mass spectrometry) проводили на MALDI-времяпролетном масс-спектрометре Ultraflex («Bruker», Германия) с УФ-лазером (336 нм) в режиме положительных ионов в диапазоне масс 500–8000 Да с калибровкой их по известным пикам аутолиза трипсина и с последующим применением программы Mascot, опция Peptide Fingerprint («Matrix Science», США) [21, 22]. Белки идентифицировали путем поиска совпадений значений экспериментальных масс с массами белков, аннотированных в базах данных (NCBI Protein и SwissProt/TrEMBL). Точность измеренных моноизотопных масс в рефлекто-моде после докалибровки по пикам аутолиза трипсина составляла 0.005%, точность измеренных масс фрагментов – ± 1 Да. Для выявления гипотетических белков, соответствующих фрагментам идентифицированных с помощью MALDI-TOF MS полноразмерных белков – продуктов соответствующих генов, проводили тандемную масс-спектрометрию. Молекулярные массы белковых фракций определяли с использованием наборов высокоочищенных рекомбинантных белков SM0661 (10–200 кДа) и SM0671 (10–170 кДа) («Fermentas», США). Денситометрию ДЭ и/или их отдельных фрагментов проводили после сканирования (сканер Epson expression 1680) или съемки на цифровую фотокамеру (Nikon 2500 или Canon PowerShot A1000 IS). Компьютерную обработку изображений с денситометрией белковых фракций выполняли с помощью пакета программ Melanie ImageMaster, версий 6 и 7 («Genebio», Швейцария).

Сбор и обработку данных для заполнения многоуровневой компьютерной базы данных «Протеомика рака простаты» осуществляли с использованием

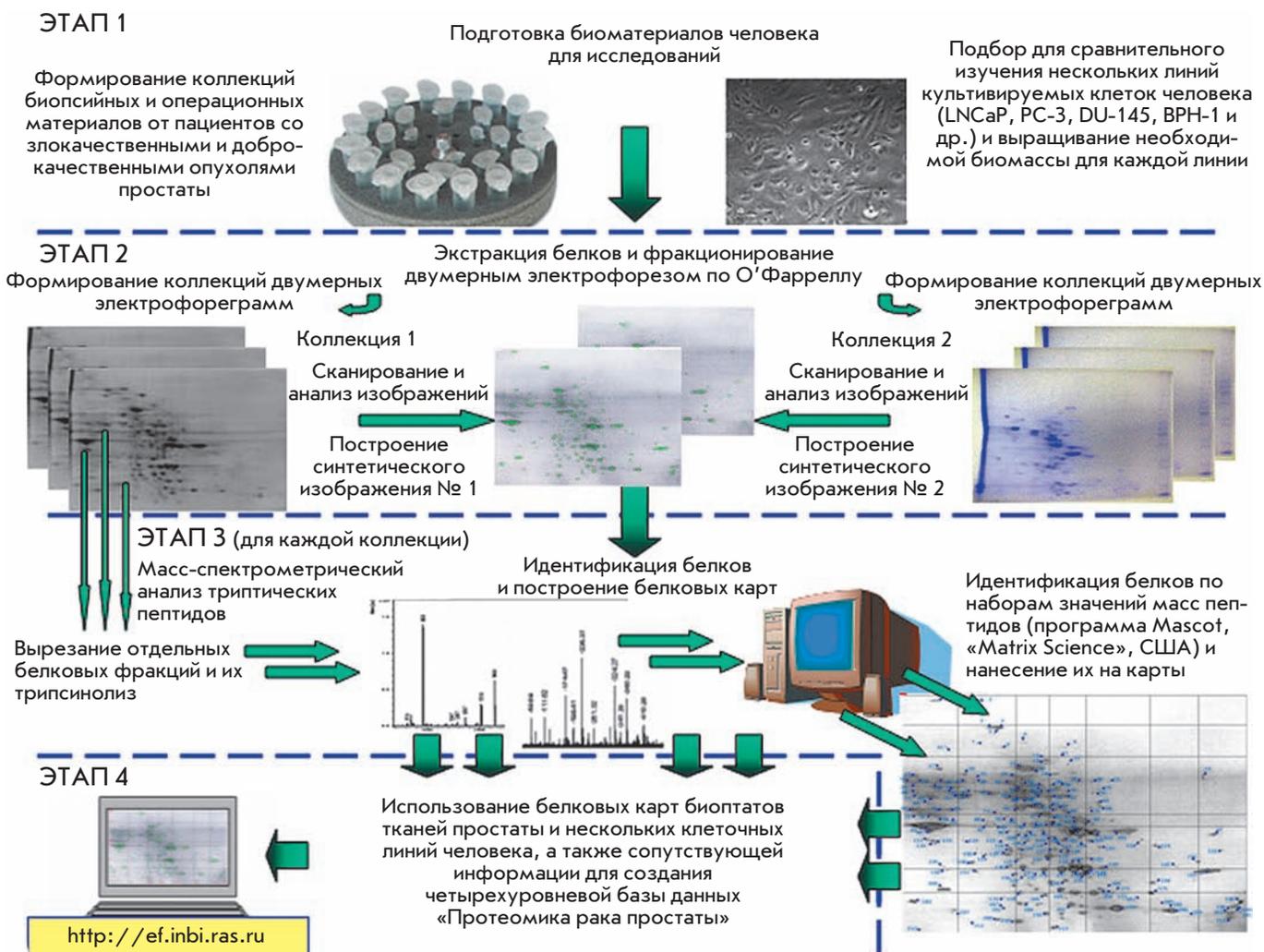


Рис. 1. Основные этапы протеомного изучения белков образцов тканей простаты, полученных от больных злокачественными и доброкачественными опухолями, а также белков из нескольких линий культивируемых клеток человека.

программ MapThis!, Molly Pinguin Software и пакета программного обеспечения Mozilla Firefox, а также других программных средств, в частности, входящих в набор Microsoft Office. Поставленные задачи выполняли при помощи интерактивной базы данных на основе СУБД MySQL с соответствующей Web-структурой, позволяющей в режиме on-line вносить и редактировать данные с любого компьютера, подключенного к сети Интернет. Статистическую обработку результатов проводили с использованием пакетов программ BIOSTAT и Microsoft Office Excel 2003.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В соответствии с традиционной стратегией протеомных исследований, которая сложилась в конце XX

века, формирование отечественного информационного ресурса «Протеомика рака простаты» осуществлялось в виде нескольких последовательных этапов системного изучения белков в образцах тканей простаты, полученных от больных злокачественными и доброкачественными опухолями (рис. 1) (по [23, 24]). Кроме того, как видно из рис. 1, параллельно изучены белки из нескольких линий культивируемых клеток человека.

Первый этап этой стратегии – составление коллекции ДЭ белковых препаратов (не менее 50), полученных при фракционировании нескольких десятков биоптатов или образцов тканей простаты различных больных (не менее 30 человек). На рис. 2а представлена типичная ДЭ белков биопсийного образца тканей простаты с РП. При изучении белков клеточных

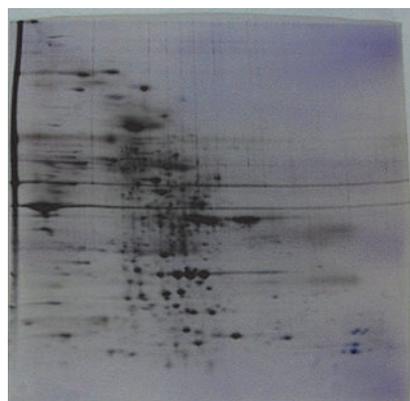
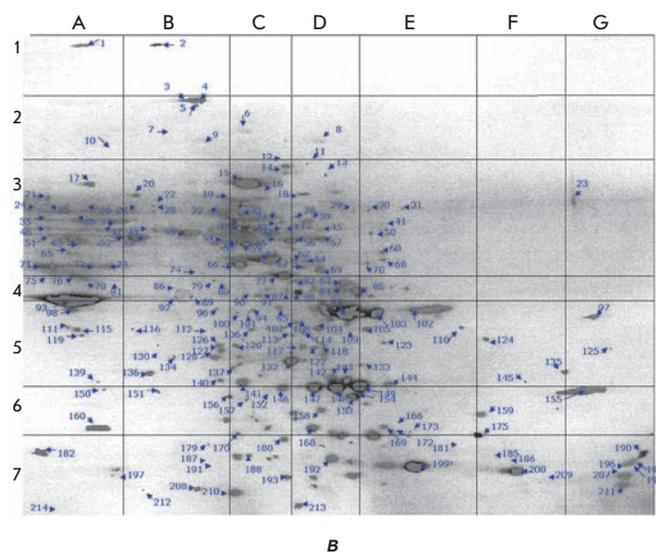
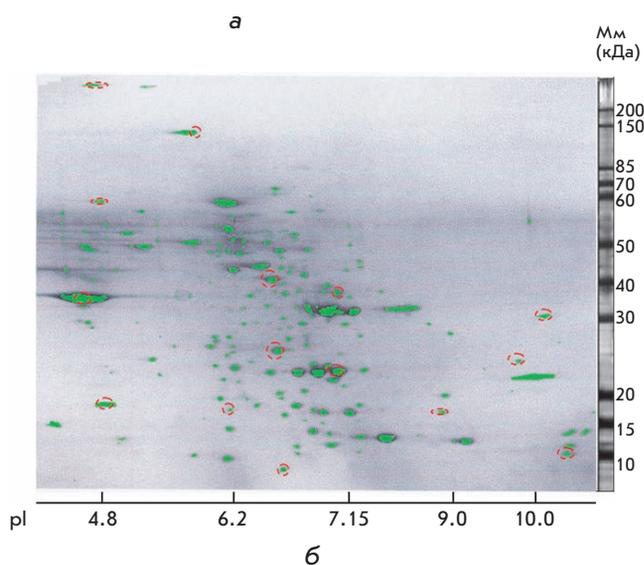


Рис. 2. Результаты протеомного изучения белков образцов раковых тканей простаты. а – Типичная ДЭ белков биопсийного образца тканей простаты с РП. б – Синтетическое изображение ДЭ белков биопсийного образца тканей простаты с РП; фракции, служившие реперными точками, выделены красными овалами, справа показано расположение маркеров молекулярных масс, внизу – результаты определения изоэлектрических точек. в – Двумерная белковая карта, построенная на основе синтетического изображения ДЭ белков биопсийного образца тканей простаты с РП; синими стрелками показаны фракции, охарактеризованные по электрофоретическим свойствам.



линий, учитывая однородность анализируемого материала, соответствующие коллекции формировали из 20 ДЭ.

Далее распределение белковых фракций на каждой электрофореграмме документировали в виде изображения, которое регистрировали и сохраняли в виде графического файла формата *.tif. Полные изображения ДЭ и (в некоторых случаях) их отдельных участков получали по результатам сканирования и/или по данным цифровой фотографии. Адекватность отобранных для последующего анализа ДЭ оценивали в ходе предварительного сравнения результатов фракционирования белков. С этой целью использовали метод компьютерного «наложения изображений» [23, 24].

Вторым этапом стало построение синтетических двумерных карт белков исследуемых объектов. Изображения ДЭ, имевшиеся в каждой коллекции, стандартизировали с помощью пакета программ Melanie ImageMaster по 15 выбранным реперным точкам, соответствовавшим четко идентифицируемым «мажорным» белковым фракциям. На рис. 2б в качестве

примера показаны соответствующие реперные точки на ДЭ белков из образцов тканей простаты.

Затем каждое изображение анализировали по методу Камингса [25] с некоторыми модификациями [20, 24]. Основой этого анализа стал общий принцип, предусматривавший разделение изображений на 49 условных прямоугольных участков (фрагментов), границы которых образовывали стандартным образом проведенные шесть горизонтальных и шесть вертикальных линий, а также края самой электрофореграммы. Точки для проведения горизонтальных линий находили с помощью специальных белков – маркеров молекулярных масс, которые наносили на каждую гелевую пластину перед проведением фракционирования во втором направлении (SDS-электрофорез в пластине градиентного полиакриламидного геля). Таким образом, белковые фракции, располагающиеся на соответствующих горизонтальных линиях, будут иметь одинаковые значения молекулярных масс. Для проведения условных вертикальных линий применяли разные белковые маркеры, для которых предварительно определяли

значения pI [20, 24]. Как следствие, каждое анализируемое изображение оказалось фрагментированным на 49 прямоугольных участков, на которых располагалось обычно не более 10 белковых фракций (лишь на четырех участках количество фракций превысило 20). Проведенная фрагментация изображений существенно облегчила последующее сопоставление изображений и построение синтетических двумерных карт.

Такую работу провели с 60 изображениями наиболее качественных ДЭ белков, полученных из гиперплазийных образцов, и с 70 электрофореграммами белков из раковых образцов. Результаты сравнения стандартизированных изображений ДЭ и при ДГП, и при РП показали, что в исследуемых образцах стабильны координаты пятен не менее чем 95% белковых фракций. Количественные и/или качественные вариации затрагивали не более 5% белковых фракций. По-видимому, вариабельность этих фракций могла быть обусловлена как непосредственными генетическими причинами (например, однонуклеотидным полиморфизмом), так и различиями в уровнях экспрессии соответствующих генов, а также особенностями тканевого состава исследуемых образцов и выраженностью патологического процесса.

Вывод о стабильности координат пятен позволил сформировать двумерные карты белков тканей простаты при ДГП и РП. Сформированные двумерные карты были подвергнуты сравнительному пофрагментному анализу. Паттерны распределения белковых фракций, представленных на двумерных картах ДГП и РП, оказались весьма схожими. Различие состояло в том, что на карте РП присутствовало около двух десятков белковых фракций, которые на соответствующих участках карты ДГП присутствовали или в значительно меньшем количестве, или не выявлялись вовсе. Эти белковые фракции рассматривались как особенно перспективные для дальнейшего изучения, результаты которого будут приведены ниже. В целом, проведенный анализ позволил построить суммарную синтетическую двумерную карту белков предстательной железы человека, на которую удалось нанести более 200 белковых фракций с параметрами Мм 8.5–450 кДа и pI 4.5–11.5 (рис. 2в). Каждая из этих фракций получила индивидуальный семизначный номер, в котором первые четыре цифры представляли собой значения десятичного логарифма экспериментально установленной молекулярной массы этой фракции, три последние – значения изоэлектрической точки в соответствии с номенклатурой, разработанной ранее [20, 24].

Аналогичным образом были построены другие синтетические карты белков культивируемых клеток человека, только их строили с использованием суще-

ственно меньшего количества изображений ДЭ в соответствующих коллекциях, как отмечалось выше.

Соответственно, каждая из построенных синтетических карт представляла собой информационный массив, отражающий полученные результаты об электрофоретических свойствах белковых фракций (в виде их распределения в системе прямоугольных координат) в каждом из изучавшихся объектов. Эти карты в виде графических файлов формата *.jpg с разрешением не менее 300 пикселей на дюйм составили первый уровень представления собранных материалов в формируемой компьютерной базе данных. Содержащаяся в них информация послужила основой для следующих этапов исследований и дальнейшего обобщения сведений об отдельных белках. Таким образом, синтетические карты стали своеобразными модулями, позволяющими формализовать и характеризовать биохимические свойства изучаемых белков. В целом, к настоящему времени в «ПРП» имеется семь соответствующих модулей (табл. 1). Для работы с модулями в «ПРП» предусмотрена специальная панель, позволяющая переходить от одного модуля к другому. Кроме того, модульные двумерные карты могут масштабироваться, и пользователь получает возможность под визуальным контролем пометать различные белки на картах, создавая специальные ссылки («кнопки») для перехода на следующие информационные уровни – второй, третий и четвертый, содержащие сведения об отдельных изучавшихся белках. В «ПРП» встроена также программа автоматического пересчета координат (по осям фракционирования в первом и вто-

Таблица 1. Модули в «ПРП» и идентифицированные в них белки

Модули – синтетические карты белков в исследуемых объектах (метод фракционирования)	Идентифицированные белки
Белки биоптатов предстательной железы (рак и гиперплазия)	165
Белки клеток LNCaP (IEF-PAGE*)	60
Белки клеток LNCaP (IPG-PAGE)	18
Белки клеток PC-3 (IEF-PAGE*)	25
Белки клеток ВРН-1 (IEF-PAGE*)	24
Белки клеток рабдомиосаркомы (IEF-PAGE*)	29
Белки нормальных миобластов человека (IEF-PAGE*)	38

*В модификации [20].

Общая схема организации базы данных «ПРП»
(на примере трех модулей с перекрестными ссылками)

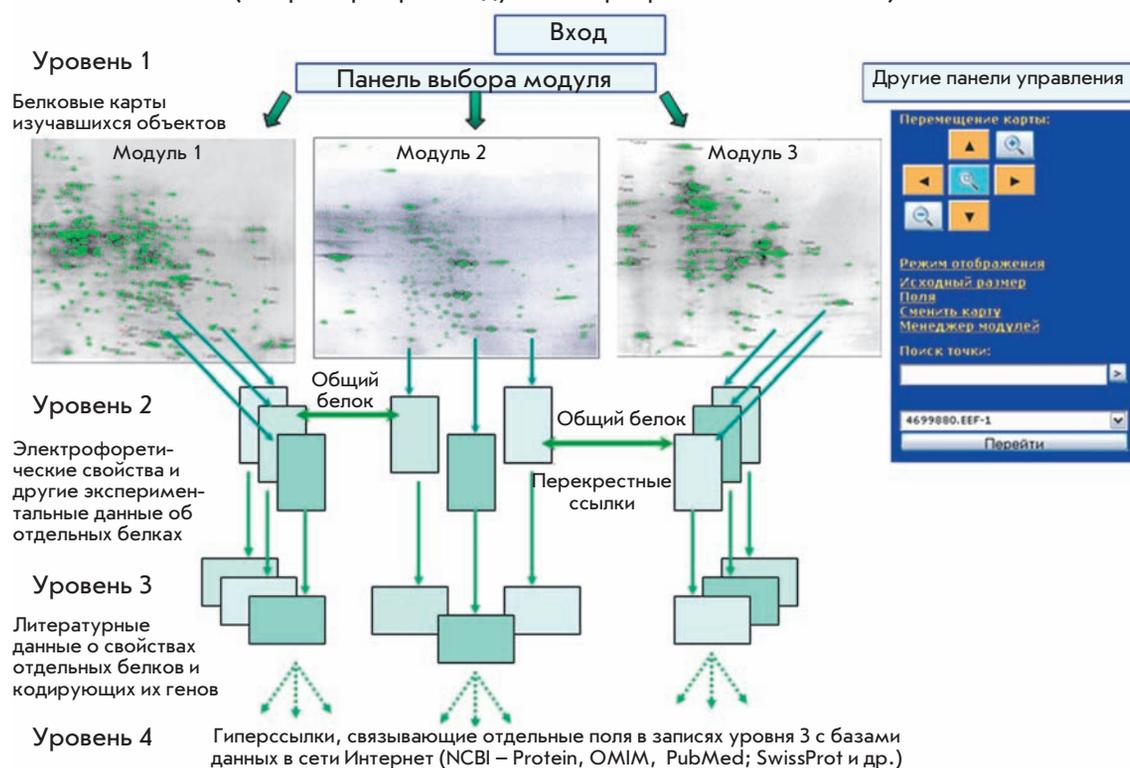


Рис. 3. Общая схема организации базы данных «Протеомика рака простаты» («ПРП»).

ром направлениях), осуществляемого при движении курсора по карте. Общая схема организации «ПРП» представлена на рис. 3.

Третьим этапом протеомных исследований, результаты которых были использованы при создании «ПРП», стали работы по идентификации отдельных белковых фракций. Белки идентифицировали в основном методами масс-спектрометрии. Результаты идентификации белков по отдельным модулям суммированы в табл. 1.

Как видно из табл. 1, общее количество идентифицированных белков в «ПРП» достигло 359. Среди них удалось выявить многие известные белки: ферменты гликолиза (глицеральдегид-3-фосфат-дегидрогеназа, триозофосфатизомераза и т.д.) и других метаболических процессов; цитоскелетные (актин, трансгелины и т.д.) и митохондриальные (порины, супероксиддисмутаза и т.д.) белки. Оказалось, что некоторые из идентифицированных белков в изучаемых объектах представлены несколькими изоформами, например трансгелины, что описано ранее [21, 22].

Особое внимание при идентификации, естественно, было обращено на те белковые фракции, которые качественно или количественно различались в образцах тканей простаты с ДГП и РП. Ранее уже сообщалось о предварительных результатах по иден-

тификации с помощью сравнительного протеомного анализа одного из потенциальных биомаркеров РП – белка AGR2 [14], недавно аналогичные данные получены для белка Dj-1 [26]. В целом, удалось выявить 17 потенциальных биомаркеров РП, часть из которых новые. Краткие сведения о потенциальных биомаркерах РП представлены в табл. 2. В качестве примера на рис. 4 показаны результаты масс-спектрометрической идентификации одного из новых потенциальных биомаркеров РП – белка PRO2675, в первичной структуре которого присутствует альбуминоподобный домен.

Для каждого идентифицированного белка (отмеченного «кнопкой» на соответствующей двумерной карте, как указано выше) формировался второй информационный уровень, представлявший собой стандартизованную систему из 15 полей для записи текстовой и графической информации, полученной в собственных исследованиях данной белковой фракции. В этой системе четыре поля предназначены для общих характеристик белка, шесть полей – для сведений о результатах идентификации и пять полей – для дополнительной информации. В качестве примера основные заполненные поля второго информационного уровня для одного из потенциальных биомаркеров РП – белка NANS (синтазы

Таблица 2. Потенциальные биомаркеры РП, включенные в модуль «Белки биопатов предстательной железы (гиперплазия, рак)» и другие модули «ПРП»

Универсальные номера*	Белок (некоторые синонимы и символ в «ПРП»)	Номера записей в базах данных (NCBI** и <i>Swiss-Prot</i>)	Дополнительная информация в «ПРП» и некоторые подтверждающие ссылки***
5653580	Комплекс легких цепей ферритина (К-(L)F)	182516, <i>P02792</i>	[15]
4785508 (4799550)	Шаперонин (HSPD1)	31542947, NP_002147, 118190, <i>P10809</i>	Обнаружен в клетках рабдомиосаркомы; {Bindukumar B. et al. 2008, 18646040}
4716560 (4756612)	Белок-дисульфидизомераза (ER60)	7437388, <i>P30101</i>	[15]
4531685	N-ацетилнейраминат-фосфат-синтаза (NANS)	12652539, AAN00008, NP_061819, 605202, <i>Q9NR45</i>	Обнаружен в клетках рабдомиосаркомы; [15]
4502675	Аннексин 2, изоформа 2 (ANXA2-i2)	4757756, NP_004030 151740, <i>P07355</i>	{Shiozawa Y. et al. 2008, 18636554; Hastie C. et al. 2008, 18211896}
4454692	Неизвестный белок PRO2675, содержащий альбуминовый домен (PRO2675)	7770217	[15]
4447605	Белок 29 эндоплазматического ретикулума, изоформа 1 (ERp29)	5803013, NP_006808, 602287, <i>P30040</i>	{Myung J.K. et al. 2004, 15598346}
4352630 (4342630)	Белок Dj-1 (Dj-1)	50513593, ISOA_A, 606324, <i>Q99497</i>	{Bindukumar B. et al. 2008, 18646040}
4356607 (4344615)	Белок Dj-1, электрофоретическая изоформа (Dj-1-ei)	31543380	
4336712 (4301795)	Простатический связывающий белок (нейрополипептид h3, PEBP1)	21410340, AAN31102, 604591, <i>P30086</i>	[15]; {Li et al. 2008, 18161940; Woods Ignatoski K.M. et al. 2008, 18722266}
4286750 (4290620)	NM23В-белок, нуклеозиддифосфаткиназа В	4505409, NP_002503, 156491, <i>P22392</i>	{Johansson B. et al. 2006, 16705742}
4255880	Безымянный белок (NEDO human cDNA sequencing project, tissue type=«testis») (NEDO)	21758704, BAC05360	
4204630	Белок, связывающий жирные кислоты, изоформа 5 (E-FABP)	30583737, AAP36117, 605168, <i>Q01469</i>	[15]; {Morgan E.A. et al. 2008, 18360704}
4279900	AGR2 (AGR2)	37183136, AAQ89368, 606358, <i>Q4JM47</i>	[14, 15]; {Zhang J.S. et al. 2005, 15834940; Zweitzig D.R. et al. 2007, 17694278}
41811130	Гистон H3 семейства 3A (H3f3a)	55665435	[15]
4161675	Неизвестный белок PRO2044, содержащий альбуминовый домен (PRO2044)	6650826	[15]
4021610	S100 кальцийсвязывающий белок A11 (S100A11)	12655117, AAN01410, 603114, <i>P31949</i>	{Rehman I. et al. 2004, 15668896; Schaefer K.L. et al. 2004, 15150091}

*В скобках приведены номера модулей «Белки клеток LNCaP, модификация IEF-2DE».

**Номера из баз данных NCBI приведены в следующем порядке: Protein, Genbank и/или Nucleotide, OMIM.

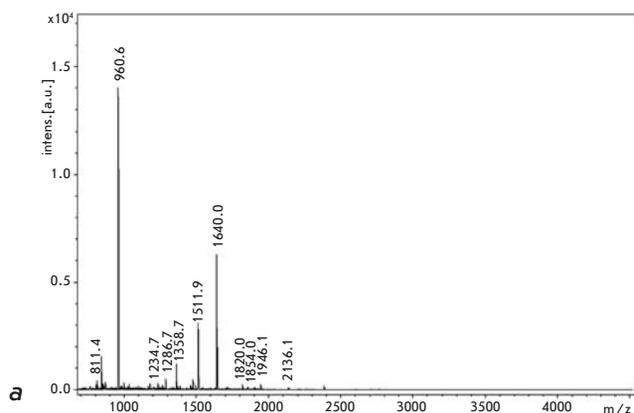
***В квадратных скобках – номера ссылок из Списка литературы, в фигурных скобках – ссылки и номера публикаций, имеющиеся в Интернет-базе данных PubMed.

Примечание. Жирным выделены обнаруженные новые для простаты потенциальные биомаркеры.

N-ацетилфосфатнейраминовой кислоты) – представлены на *рис. 5*.

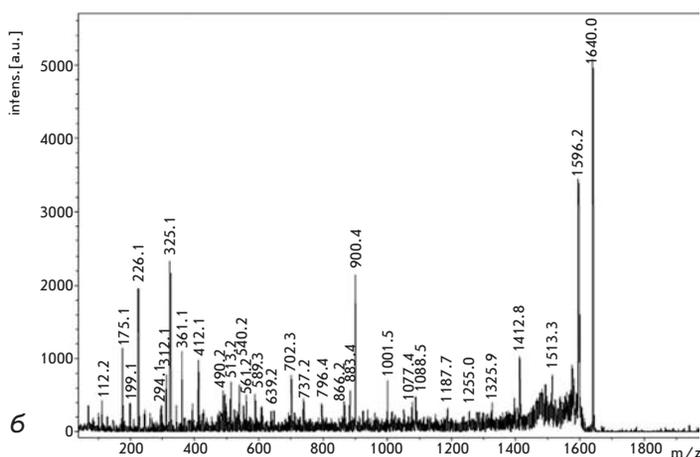
Поскольку один и тот же белок мог присутствовать в нескольких изучавшихся объектах, то на втором информационном уровне имеется возможность с помощью управляющей панели создавать перекрестные ссылки между одинаковыми белками в разных модулях. Такая панель для белка Dj-1 показана на *рис. 6*.

Подавляющее большинство из 359 идентифицированных фракций составляли известные белки (и/или их электрофоретические изоформы), о которых в литературе и различных базах данных накоплено множество разных сведений. Некоторые из этих сведений, подобранные с учетом задач «ПРП», составили третий информационный уровень. Этот уровень представляет собой стандартизированную систему из 23



Начало – Конец	Наблюдаемые Mr	Расчетные Mr	Последовательность пептида
1 - 16	1706.77	1705.82	MPADLPSLAADFVESK + Oxidation (M)
27 - 39	1639.82	1638.79	DVFLGMFLYEYAR + Oxidation (M)
76 - 92	2045.05	2044.09	VFDEFKPLVEEPQNLK
106 - 113	960.52	959.56	FQNALLVR
118 - 131	1511.74	1510.84	VPQVSTPTLVEVSR
249 - 263	1763.73	1762.77	AVMDDFAAFVEKCK + Oxidation (M)

a'



b

#	Immon.	a	a*	a ^o	b	b*	b ^o	Seq	v	w	w*	y	y*	y ^o	#
1	101.11	101.11	84.08		129.10	112.08		K							15
2	72.08	200.18	183.15		228.17	211.14		V	1467.78	1480.80		1511.84	1494.82	1493.83	14
3	70.07	297.23	280.20		325.22	308.20		P	1370.73	1369.73		1412.77	1395.75	1394.76	13
4	101.07	425.29	408.26		453.28	436.26		Q	1242.67	1241.67		1315.72	1298.70	1297.71	12
5	72.08	524.36	507.33		552.35	535.32		V	1143.60	1156.62		1187.66	1170.64	1169.65	11
6	60.04	611.39	594.36	593.38	639.38	622.36	621.37	S	1056.57	1055.57		1088.59	1071.57	1070.58	10
7	74.06	712.44	695.41	694.42	740.43	723.40	722.42	T	955.52	968.54	970.52	1001.56	984.54	983.55	9
8	70.07	809.49	792.46	791.48	837.48	820.46	819.47	P	858.47	857.47		900.51	883.49	882.50	8
9	74.06	910.54	893.51	892.53	938.53	921.50	920.52	T	757.42	770.44	772.42	803.46	786.44	785.45	7
10	86.10	1023.62	1006.59	1005.61	1051.61	1034.59	1033.60	L	644.34	643.34		702.41	685.39	684.40	6
11	72.08	1122.69	1105.66	1104.68	1150.68	1133.66	1132.67	V	545.27	558.29		589.33	572.30	571.32	5
12	102.05	1251.73	1234.70	1233.72	1279.73	1262.70	1261.72	E	416.23	415.23		490.26	473.24	472.25	4
13	72.08	1350.80	1333.77	1332.79	1378.79	1361.77	1360.78	V	317.16	330.18		361.22	344.19	343.21	3
14	60.04	1437.83	1420.80	1419.82	1465.83	1448.80	1447.82	S	230.12	229.13		262.15	245.12	244.14	2
15	129.11							R	74.02	73.03		175.12	158.09		1

b'

1 MPADLPSLAA DFVESKDVCK NYAEAKDVFL GMFLYEYARR HPDYSVLLLL RLAKTYETTL
 61 EKCCAAADPH EYAKVFDEF KPLVEEPQNL IKQNCSELFQ LGEYKFQNAL LVRYTKKVPQ
 121 VSTPTLVEVS RNLGKVGSKC CKHPEAKRMP CAEDYLSVVL NQLCVLHEKT PVSDRVTKCC
 181 TESLVNRRPC FSALEVDETY VPKEFNAETF TFHADICTLS EKERQIKKQT ALVELVKHKP
 B 241 KATKEQLKAV MDDFAAFVEK CCKADDKETC FAEFGKLLVA ASQAALGL

Рис. 4. Результаты масс-спектрометрической идентификации белка PRO2675. а, а' – Масс-спектр триптических пептидов, полученный методом MALDI-TOF MS (а), и идентификация пептидов с помощью программы Mascot (а'). б, б' – Масс-спектр одного из триптических пептидов, полученный методом MALDI-TOF MS/MS (б), и его идентификация с помощью программы Mascot (б'). в – Аминокислотная последовательность белка PRO2675 (по записям AAF69644.1 GI:7770217 в базе данных Protein, NCBI); красными буквами показаны аминокислотные остатки выявленных пептидов; серым выделен пептид, последовательность которого установлена методом MALDI-TOF MS/MS; подчеркнутые участки последовательности соответствуют альбуминовому домену.

ЭКСПЕРИМЕНТАЛЬНЫЕ СТАТЬИ

Экспериментальные данные (уровень 2)

Общие характеристики					
Фракция (белок)	N-ацетилнейраминат-фосфат-синтаза, NANS (N-acetylneuraminic acid phosphate synthase) Редактировать				
Локализация на ДЭ, номер	D4, 4531685 Редактировать				
Мм, кДа	38.0 Редактировать				
pI (наблюдаемая величина)	6.85 Редактировать				
Идентификация					
Метод	MALDI TOF MS Редактировать				
Совпадение масс выявленных триптических пептидов с последовательностью белка, %	58% Редактировать				
Анализ выявленных триптических пептидов программы Mascot («Matrix Science», США)	Начало-Конец	Наблюдаемые Mr	Расчетные Mr		Последовательность пептида
	66 – 74	1064.56	1063.55	1063.57	-0.02 0 ALERP YTSK
	87 – 96	1331.59	1330.58	1330.61	-0.03 0 HLEFSDQYR
	132 – 145	1540.68	1539.68	1539.72	-0.04 0 VGSGDTNNFPYLEK
	150 – 166	1919.79	1918.78	1918.83	-0.04 0 GRPMVISSGMQSM DTMK 4 Oxidation (M)
	199 – 205	866.44	865.43	865.45	-0.03 0 VISEYQK
	206 – 233	2767.42	2766.41	2766.50	-0.08 0 LFPPDIPIGYSGHETGIAISVAVALGAK
	247 – 264	1866.88	1865.87	1865.91	-0.04 0 GSDHSASLEPGE LAELVR
	301 – 315	1693.83	1692.82	1692.87	-0.05 0 IPEGTILTMDMLTVK 2 Oxidation (M)
	321 – 333	1448.71	1447.70	1447.73	-0.04 0 GYPPE DIFNLV GK
	335 – 355	2401.08	2400.07	2400.14	-0.06 0 VLVTV EEDDTIMEELVDNHGK Oxidation (M)
	Редактировать				
Иллюстрация-1	Редактировать				
Аминокислотная последовательность (после удаления сигнальных последовательностей)	0001 MPLELELC PG RWVGGQHPCF IIAEIGQNHQ GDLDVAKRMI RMAKECGADC 0051 AKFQKSELEF KFNRRALERP YTSKHSWGKT YGENKRHLEF SHDQYRELQR 0101 YAEVGIFFT ASGMDEMAVE FLHELNV PFF KVGSGDTNNF PYLEKTAKKG 0151 RPMVISSGMQ SMDTMKQVYQ IVKPLNPNFC FLQCTSAYPL QPEDVNL RVI 0201 SEYQKLF PDI PIGYSGHETG I AISVAVAL GAKVLERHIT LDKTWKGS DH 0251 SASLEPGE LA ELVRSVRLVE RALGSPTRQL LPCEMACNEK LGKSVVAKVK 0301 IPEGTILTMD MLTVKVGEPK GYPPE DIFNLV GK VLVTV EEDDTIMEELV 0351 DNHGKKIKS				
	Редактировать				
Комментарий к аминокислотной последовательности	Аминокислотная последовательность белка NANS. Красным цветом выделены участки последовательности, соответствующие выявленным триптическим пептидам Редактировать				
Дополнительная информация по идентификации					
Дополнительная информация	MS/MS Fragmentation of VGSGDTNNFPYLEK Found in gi 8453156, N-acetylneuraminic acid phosphate synthase [Homo sapiens] Match to Query 16: 1539.695159 from(1540.702435,1+) Редактировать				
Тандемная масс-спектрометрия	Monoisotopic mass of neutral peptide Mr(calc): 1539.72 Ions Score: 22 Expect: 8.7				

Рис. 5. Основные поля второго информационного уровня для белка NANS.

Рис. 6. Управляющая панель с перекрестными ссылками для белка Dj-1 из модуля «Белки биоптатов предстательной железы (рак и гиперплазия)».

Белки-аналоги в других модулях

Модуль	Точка	Перейти	Отсоединить
Белки клеток LNCaP, IPG-2DE	4301630.Dj-1	Перейти	Отсоединить
Белки клеток LNCaP, модификация IEF-2DE	4342635.Dj-1	Перейти	Отсоединить
Белки клеток рабдомиосаркомы	4342690.Dj-1	Перейти	Отсоединить
Белки клеток PC-3	4342630.Dj-1	Перейти	Отсоединить
Добавить			

полей для записи текстовой и графической информации. Из них 12 полей предназначены для сведений о самом белке, шесть – о кодирующем его гене, три – о различных проявлениях полиморфизма и два поля – для специально отобранных ссылок на публи-

кации об этом белке как общего характера, так и онкологической направленности (рис. 7).

Поля для текстовой информации третьего уровня обеспечены возможностью вставки гиперссылок, связывающих эти поля с различными базами данных

Литература	
Общие публикации	<p>PMID: 9790916 1. Thompson, D.A. Weigel, R.J. hAG-2, the human homologue of the <i>Xenopus laevis</i> cement gland gene XAG-2, is coexpressed with estrogen receptor in breast cancer cell. <i>Biochem. Biophys. Res. Commun.</i> 1998. V. 251(1). P. 111–116.</p> <p>PMID: 12975309 2. Clark, H.F., Gurney, A.L., Abaya, E., Baker, K., Baldwin, D., Brush, J., Chen, J., Chow, B., Chui, C., Crowley, C., Currell, B., Deuel, B., Dowd, P., Eaton, D., Foster, J., Grimaldi, C., Gu, Q., Hass, P.E., Heldens, S., Huang, A., Kim, H.S., Klimowski, L., Jin, Y., Johnson, S., Lee, J., Lewis, L., Liao, D., Mark, M., Robbie, E., Sanchez, C., Schoenfeld, J., Seshagiri, S., Simmons, L., Singh, J., Smith, V., Stinson, J., Vagts, A., Vandlen, R., Watanabe, C., Wieand, D., Woods, K., Xie, M.H., Yansura, D., Yi, S., Yu, G., Yuan, J., Zhang, M., Zhang, Z., Goddard, A., Wood, W.L., Godowski, P., Gray, A. The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. <i>Genome Res.</i> 2003. V. 13(10). P. 2265–2270.</p> <p>PMID: 15340161 3. Zhang Z, Henzel WJ. Signal peptide prediction based on analysis of experimentally verified cleavage sites. <i>Protein Sci.</i> 2004. V. 13(10). P. 2819–2824.</p> <p>Редактировать</p>
Публикации онкологической направленности	<p>PMID: 12592373 1. Fletcher G.C., Patel S., Tyson K., Adam P.J., Schenker M., Loader J.A., Daviet L., Legrain P., Parekh R., Harris A.L., Terrett J.A. hAG-2 and hAG-3, human homologues of genes involved in differentiation, are associated with oestrogen receptor-positive breast tumours and interact with metastasis gene C4.4a and dystroglycan. <i>Br. J. Cancer.</i> 2003. V. 88(4). P. 579–585.</p> <p>PMID: 15532095 2. Kristiansen G., Pilarsky C., Wissmann C., Kaiser S., Bruemendorf T., Roepcke S., Dahl E., Hinzmann B., Specht T., Pervan J., Stephan C., Loening S., Dietel M., Rosenthal A. Expression profiling of microdissected matched prostate cancer samples reveals CD166/MEMD and CD24 as new prognostic markers for patient survival. <i>J. Pathol.</i> 2005. V. 205(3). P. 359–376.</p> <p>PMID: 15867376 3. Liu D., Rudland P.S., Sibson D.R., Platt-Higgins A., Barraclough R. Human homologue of cement gland protein, a novel metastasis inducer associated with breast carcinomas. <i>Cancer Res.</i> 2005. V. 65(9). P. 3796–3805.</p> <p>PMID: 15958538 4. Smirnov D.A., Zweitzig D.R., Foulk B.W., Miller M.C., Doyle G.V., Pienta K.J., Meropol N.J., Weiner L.M., Cohen S.J., Moreno J.G., Connelly M.C., Terstappen L.W., O'Hara S.M. Global Gene Expression Profiling of Circulating Tumor Cells. <i>Cancer Res.</i> 2005. V. 65. P. 4993–4997.</p> <p>PMID: 15834940 5. Zhang J.S., Gong A., Cheville J.C., Smith D.J., Young C.Y. AGR2, an androgen-inducible secretory protein overexpressed in prostate cancer. <i>Genes Chromosomes Cancer.</i> 2005. V. 43(3). P. 249–259.</p> <p>PMID: 17022460 6. Kovalev L.I., Shishkin S.S., Khasigov P.Z., Dzeranov N.K., Kazachenko A.V., Kovaleva M.A., Toropygin I.Yu., Mamykina S.V. Identification of AGR2 protein, a novel potential cancer marker, using proteomics technologies. <i>Prikl Biokhim Mikrobiol.</i> 2006. V. 42(4). P. 480–484. (Ковалев Л.И., Шишкин С.С., Хасигов П.З., Дзеранов Н.К., Казаченко А.В., Ковалева М.А., Торопыгин И.Ю., Мамыкина С.В. Идентификация белка AGR2 – нового потенциального маркера рака с использованием протеомных технологий. <i>Приклад. биохимия и микробиол.</i> 2006. Т. 42. № 4. С. 480–484).</p> <p>PMID: 18829536 7. Ramachandran V., Arumugam T., Wang H., Logsdon C.D. Anterior gradient 2 is expressed and secreted during the development of pancreatic cancer and promotes cancer cell survival. <i>Cancer Res.</i> 2008. V. 68(19). P. 7811–7818.</p> <p>Редактировать</p>

Рис. 7. Поля третьего уровня для специально отобранных ссылок на публикации о белке AGR2.

в Интернете, в частности Protein, OMIM и PubMed, входящими в базу знаний NCBI, а также SwissProt. Благодаря этому был сформирован четвертый информационный уровень, позволяющий пользователю оперативно и эффективно привлекать материалы ряда современных международных баз данных, содержащих, в частности, результаты секвенирования генома человека.

База данных «ППП» выполнена в виде интерактивного Web-ресурса на основе СУБД MySQL, расположенного по адресу <http://ef.inbi.ras.ru>, и соответственно доступна для посещения с любого компьютера, подключенного к сети Интернет с использованием браузеров типа Mozilla Firefox и Microsoft Internet Explorer. Вместе с тем, обеспечено и разграничение

прав доступа на три категории: «Гость», «Менеджер», «Администратор». Для каждой категории четко определена сфера возможных работ с «ППП». В частности, пользователи с правом доступа «Менеджер» обладают возможностями оперативно вносить записи в имеющиеся поля второго и третьего уровней, а также корректировать соответствующие записи, а пользователи с правом доступа «Администратор», кроме того, могут расширять базу данных за счет создания новых дополнительных модулей и новых функциональных элементов. Пользователь с правами «Гостя» имеет доступ к просмотру всех полей базы данных без возможности их редактирования.

Таким образом, в результате проведенных исследований создан оригинальный многомодульный

отечественный информационный ресурс «Протеомика рака простаты», в котором суммированы данные о белках, присутствующих в тканях простаты при ДГП и РП, а также о белках из нескольких клеточных линий человека, что открывает широкие возможности для дальнейшей работы с протеомной и другой биохимической информацией. Можно надеяться, что использование «ПРП» биохимики и другими специалистами, занятыми решением био-

медицинских проблем РП, будет способствовать повышению эффективности проводимых исследований и, в частности, активизирует поиски новых биомаркеров этого заболевания. ●

Исследования проводились при поддержке Департамента науки и промышленной политики города Москвы (государственные контракты № 8/3-373н-08 и 8/3-375н-08).

СПИСОК ЛИТЕРАТУРЫ

- Gottlieb B., Beitel L.K., Wu J.H., Trifiro M. // Hum. Mutat. 2004. V. 23. P. 527–533.
- Westbrook J.A., Wheeler J.X., Wait R., Welson S.Y., Dunn M.J. // Electrophoresis. 2006. V. 27. P. 1547–1555.
- Kandasamy K., Keerthikumar S., Goel R., Mathivanan S., Patankar N., Shafreen B., Renuse S., Pawar H., Ramachandra Y.L., Acharya P.K., Ranganathan P., Chaerkady R., Keshava Prasad T.S., Pandey A. // Nucleic Acids Res. 2009. V. 37 (Database issue). P. D773–D781.
- Vizcaino J.A., Cote R., Reisinger F., Barsnes H., Foster J.M., Rameseder J., Hermjakob H., Martens L. // Nucleic Acids Res. 2010. V. 38 (Database issue). P. D736–D742.
- Stamey T.A., Caldwell M., McNeal J.E., Nolley R., Hemenez M., Downs J. // J. Urol. 2004. V. 172. P. 1297–1301.
- Zhang J.S., Gong A., Cheville J.C., Smith D.I., Young C.Y. // Genes Chromosomes Cancer. 2005. V. 43. № 3. P. 249–259.
- Lim L.S., Sherin K. // Am. J. Prev. Med. 2008. V. 34. № 2. P. 164–170.
- Leman E.S., Getzenberg R.H. // J. Cell Biochem. 2009. V. 108. № 1. P. 3–9.
- Злокачественные новообразования в России в 2008 году. Заболеваемость и смертность / Ред. Чиссов В.И., Старинский В.В., Петрова Г.В. М.: МЕДпресс-информ, 2008. С. 18.
- Злокачественные новообразования в России и странах СНГ в 2002 г. ГУ РОНЦ им. Н.Н. Блохина РАМН / Ред. Давыдов М.И., Аксель Е.М. М.: МИА, 2004. 256 с.
- Jemal A., Siegel R., Ward E., Murray T., Xu J., Smigal C., Thun M.J. // CA Cancer J. Clin. 2007. V. 57. № 1. P. 43–66.
- Maddams J., Brewster D., Gavin A., Steward J., Elliott J., Utley M., Muller H. // Br. J. Cancer. 2009. V. 101. № 3. P. 541–547.
- Примроуз С., Тваймен Р. Геномика. Роль в медицине. М.: БИНОМ. Лаборатория знаний, 2008. 277 с.
- Ковалев Л.И., Шишкин С.С., Хасигов П.З., Дзеранов Н.К., Казаченко А.В., Ковалева М.А., Торопыгин И.Ю., Мамыкина С.В. // Прикл. биохим. и микробиол. 2006. Т. 42. № 4. С. 480–484.
- Шишкин С.С., Дзеранов Н.К., Тотров К.И., Казаченко А.В., Ковалев Л.И., Еремина Л.С., Ковалева М.А., Торопыгин И.Ю. // Урология. 2009. № 1. С. 56–58.
- Коган М.И., Лоран О.Б., Петров С.Б. Радикальная хирургия рака предстательной железы. М.: Гэотар-Медиа, 2006. 392 с.
- Шишкин С.С., Ковалев Л.И., Ковалева М.А., Крахмалева И.Н., Еремина Л.С., Макаров А.А., Лисицкая К.В., Лоран О.Б., Велиев Е.И., Охриц В.Е. Проблемы ранней диагностики рака простаты и возможности применения новых потенциальных биомаркеров. (Информационно-методическое письмо). М.: ООО Оригинальная компания, 2009. 45 с.
- Черников В.Г., Терехов С.М., Крохина Т.Б., Шишкин С.С., Смирнова Т.Д., Лунга И.Н., Аднорал Н.В., Ребров Л.Б., Денисов-Никольский Ю.И., Быков В.А. // Бюлл. эксп. биол. мед. 2001. Т. 131. № 6. С. 680–682.
- Крохина Т.Б., Шишкин С.С., Раевская Г.Б., Ковалев Л.И., Ершова Е.С., Черников В.Г., Мирочник В.В., Бубнова Е.Н., Кухаренко В.И. // Бюлл. эксп. биол. мед. 1996. Т. 122. № 9. С. 314–317.
- Kovalyov L.I., Shishkin S.S., Efimochkin A.S., Kovalyova M.A., Ershova E.S., Egorov T.A., Musalyamov A.K. // Electrophoresis. 1995. V. 16. P. 1160–1169.
- Еремина Л.С., Ковалев Л.И., Шишкин С.С., Торопыгин И.Ю., Буракова М.И., Ковалева М.А., Макаров А.А., Дзеранов Н.К., Казаченко А.В., Тотров К.И., Кононков И.В., Лоран О.Б. // Вопр. биол. мед. фарм. химии. 2007. № 3. С. 49–52.
- Ковалева М.А., Ковалев Л.И., Еремина Л.С., Макаров А.А., Буракова М.В., Торопыгин И.Ю., Серебрякова М.В., Шишкин С.С., Арчаков А.И. // Биомедицинская химия. 2008. Т. 54. № 4. С. 420–434.
- Anderson N.G., Anderson L. // Electrophoresis. 1996. V. 17. P. 443–453.
- Шишкин С.С., Ковалев Л.И., Громов П.С. Функциональная геномика человека и протеомика, как раздел функциональной геномики // Многоликость современной генетики человека. М.-Уфа: Гилем, 2000. С. 17–50.
- Cummings D. // Clin. Chem. 1982. V. 28. P. 782–789.
- Loran O.B., Veliev E.I., Okhrizts V.E., Lisitskaya K.V., Eremina L.S., Kovalyov L.I., Kovalyova M.A., Shishkin S.S. // Eur. Urol. Suppl. 2010. V. 9. № 2. P. 309.