

Настоящее изобретение относится к новому белку INSP002, идентифицированному в настоящей заявке как секретируемый белок, т.е. как член семейства DAN, относящегося к суперсемейству цитокинов, имеющих в своей структуре цистиновые узлы, и к применению этого белка и последовательностей нуклеиновой кислоты кодирующего гена для диагностики, профилактики и лечения заболеваний.

Все цитируемые публикации, патенты и патентные заявки во всей своей полноте введены в настоящее описание посредством ссылки.

Предшествующий уровень техники

В настоящее время в области разработки лекарственных средств произошел крутой перелом, который положил начало новой эре функциональной геномики, пришедшей на смену старым методам. Термин "функциональная геномика" применяется к способам использования средств биоинформатики для определения функций исследуемых последовательностей белков. Такие средства становятся все более необходимыми, и это связано с тем, что оснащение научно-исследовательских лабораторий, занимающихся определением функций этих последовательностей белков, пока не дает возможности быстро обрабатывать все нарастающий поток данных для этих последовательностей.

Поскольку эффективность и точность методов биоинформатики возрастает, то эти методы быстро вытесняют стандартные методы биохимической характеристики. Действительно, современные средства биоинформатики, используемые для идентификации белков настоящего изобретения, могут давать окончательные результаты с достаточно высокой степенью достоверности.

Различные институты и коммерческие организации занимаются обработкой непрерывно поступающих данных о последовательностях, и на основе полученных результатов приходят к важным открытиям. Однако необходимость в идентификации и характеристике других генов и полипептидов, кодируемых этими генами, для их дальнейшего исследования и поиска новых лекарственных средств, все еще остается актуальной.

Общее описание секретируемых белков

Способность клеток продуцировать и секретируют внеклеточные белки, является главным фактором во многих биологических процессах. Ферменты, факторы роста, белки внеклеточного матрикса и молекулы, передающие сигналы, секретируются клетками посредством слияния секреторной везикулы с плазматической мембраной. В большинстве случаев, но не всегда, белки транспортируются в эндоплазматический ретикулум и в секреторные везикулы посредством сигнального пептида. Сигнальные пептиды представляют собой цис-активные последовательности, которые влияют на транспорт полипептидных цепей из цитоплазмы в мембрано-ассоциированный компартмент, такой как секреторная везикула. Полипептиды, которые нацелены на секреторные везикулы, либо секретируются во внеклеточный матрикс, либо удерживаются в плазматической мембране. Полипептиды, которые удерживаются в плазматической мембране, имеют один или более трансмембранных доменов. Примерами секретируемых белков, которые играют главную роль в функционировании клеток, являются цитокины, гормоны, белки внеклеточного матрикса (адгезивные молекулы), протеазы и факторы роста и дифференцировки.

Факторы роста представляют собой относительно большую группу полипептидов, которые обладают общей способностью индуцировать размножение клеток как *in vivo*, так и *in vitro*. Факторы роста отличаются от классических эндокринных гормонов, таких как инсулин или гормон роста, по двум своим важным механизмами действия. Во-первых, эндокринные гормоны обычно синтезируются в конкретных железах (таких как поджелудочная железа, в случае инсулина), тогда как факторы роста часто синтезируются в клетках и тканях различных типов. Во-вторых, классические эндокринные гормоны высвобождаются в физиологические жидкости в области их синтеза и переносятся в кровотоке в их ткань-мишень. Важные отличительные свойства факторов роста заключаются в том, что в большинстве случаев они действуют только в тех тканях, в которых они синтезируются (см. Health. J.K. (1993) Growth Factors, Oxford University Press, Oxford, UK, pp.15-33).

Хотя уровень сходства последовательностей факторов роста невысок, однако, по своему структурному и функциональному сходству, они могут быть подразделены на суперсемейства. Примерами таких суперсемейств являются:

- (a) гемопоэтические факторы роста, такие как гормон роста, IL-2, IL-4, G-CSF и CNTF, каждый из которых имеет структурный мотив в виде пучка из четырех спиралей;
- (b) члены семейства, имеющие бета-конформацию типа "клеверный лист", такие как IL-1-бета, IL-1-альфа, FGF и фактор роста кератиноцитов;
- (c) EGF-подобные факторы роста, такие как EGF и TGF-альфа, каждый из которых имеет иммуноглобулин-подобный домен; и
- (d) факторы роста, имеющие в своей структуре цистиновые узлы, такие как NGF, TGF-бета, PDGF и гликопротеиновые гормоны.

Факторы роста являются внеклеточными молекулами, и для осуществления своего биологического действия, они взаимодействуют со специфическими высокоаффинными рецепторами, локализованными на плазматических мембранах клеток-мишеней. Характеризация молекул ряда различных рецепторов факторов роста показала, что они могут быть разделены на определенные семейства: тирозинкиназные рецепторы, ассоциированные с G-белком рецепторы, состоящие из семи трансмембранных доменов, и

серин/треонинкиназные рецепторы. Тирозинкиназные рецепторы отличаются тем, что они имеют внеклеточный домен, трансмембранный домен и внутриклеточный домен, обладающий тирозинкиназной активностью. Серин/треонинкиназные рецепторы факторов роста имеют сходство с тирозинкиназными рецепторами, содержащими внеклеточный домен, трансмембранный домен и внутриклеточный домен. Внутриклеточный домен обладает присущей ему серин/треонинкиназной активностью.

Нарушение регуляции факторов роста приводит к ряду патологических состояний, включая, но, не ограничиваясь ими, онкологические заболевания (Bartucci M. et al. (2001) *Cancer Res.* Sep. 15; 61(18):6747-54, Dias S. et al. (2001) *Proc. Natl. Acad. Sci. USA.* Sep.11;98(19):10857-62, Djavan B. et al. (2001) *World J. Urol.* 19(4):225-33), воспалительные заболевания (Fiocchi C. (2001) *J.Clin.Invest.* Aug;108(4):523-6, Hodge S. et al. (2001) *Respirology.* Sep.6(3):205-211, Fenwick S.A. et al. (2001) *J. Anat.* Sep. 199(Pt 3):231-40), неврологические заболевания (Cooper J.D. et al. (2001) *Proc. Natl. Acad. Sci. USA* 98(18):10439-44, Fahnestock M. et al. (2001) *Mol. Cell. Neurosci* 18(2):210-20) и метаболические заболевания (Vickers M.H. et al. (2001) *Endocrinology* 142(9):3964-73).

Суперсемейство, имеющее в структуре цистиновые узлы

Типичная структура, наблюдаемая у суперсемейства с цистиновыми узлами, основана на присутствии 6 цистиновых остатков, образующих 3 дисульфидных связи. Две из этих дисульфидных связей образуют "кольцеобразную" структуру, которая пронизана третьей дисульфидной связью (Sun et al. 1995). Домены "цистиновый узел" часто обнаруживают более чем 6 цистиновых остатков. Дополнительные цистиновые остатки обычно используются для образования других дисульфидных связей в домене "цистиновый узел" или межцепевых дисульфидных связей во время димеризации.

Это суперсемейство с "цистиновыми узлами" подразделяется на подсемейство, которое включает гликопротеиновые гормоны (например, фолликулостимулирующий гормон), белки трансформирующего фактора роста-бета (TGF-бета) (например, белок морфогенеза кости 4), белки, подобные тромбоцитарному фактору роста (PDGF-подобные белки) (например, тромбоцитарный фактор роста A), факторы роста нервных клеток (NGF) (например, нейротропный фактор головного мозга) и семейство отобранных методом дифференциального скрининга абберрантных генов в нейробластоме (DAN) (например, cerberus). Суперсемейство DAN включает Cerl, Cerberus, Caronte, Dnm/Gremlin, PRDC, DAN Dante и CeCanl (Massague et al., *Genes Dev.* 2000 Mar. 15;14(6):627-44; Massague & Wotton, *EMBO J.* 2000 Apr. 17;19(8):1745-54).

Предполагается, что члены подсемейства DAN способны модулировать действие членов подсемейства белков TGF-бета (Pearce et al., *Dev.Biol.* 1999 May 1;209(1):98-110). Более конкретно, возможно, что члены подсемейства DAN способны модулировать действие белков морфогенеза кости (BMP) в процессе ее развития.

Было обнаружено, что члены подсемейства DAN действуют как антагонисты белков морфогенеза кости (BMP), которые являются членами подсемейства TGF-бета, входящего в суперсемейство, имеющее домены "цистиновый узел" (Stanley et al., *Mech.Dev.* 1998 Oct, 77(2):173-84; Massague et al. 2000 (supra); Massague J. & Wotton D, 2002 (supra)). Мономеры BMP подвергаются гомо- или гетеродимеризации посредством связывания своих доменов "цистиновый узел", после чего они взаимодействуют с рецепторами клеточной поверхности. Считается, что члены подсемейства DAN способны связываться с BMP посредством их собственных доменов "цистиновый узел". Это приводит к предотвращению связывания BMP с его природным партнером по димеризации, в результате чего, BMP теряет свою способность взаимодействовать со своим сигнальным рецептором, присутствующим на поверхности клетки. Эксперименты, конкретно направленные на DAN, Cerl и DRM, показали, что они ингибируют действие BMP4 (Pearce et al., 1999 (см. выше)).

Более ясное понимание функции cerberus было достигнуто в результате исследования на связывание (Piccolo S. et al. *Nature* 1999 Feb. 25;397(6721):707-10). В первых функциональных исследованиях, проводимых на cerberus, использовали белок cerberus *Xenopus laevis* (cer). Микроинъекция мРНК cerberus *Xenopus* в эмбрионах *Xenopus* показала, что белок cer индуцирует образование эктопических головок в передней энтодерме организатора Спэмманна (Bouwmeester et al. *Nature* 1996 Aug. 15;382(6592):595-601; Bouwmeester T. *Int., J. Dev. Biol.* 200, 145(1 Spec No):251-8). Исследования на связывание, проведенные Piccolo и сотрудниками; показали, что белок cerberus *Xenopus* связывается с белками Nodal, BMP и Wnt через независимые сайты и ингибирует их действие. Более конкретно, эти исследователи обнаружили, что белок cerberus обладает высокоспецифичной аффинностью и ингибирующим действием по отношению к Xnr-1 (члену семейства Nodal), BMP4 (члену семейства BMP) и Xwmt-8 (члену семейства Wnt). В этой работе была установлена взаимосвязь cerberus, а, следовательно, и других членов семейства DAN с путями эволюции и дифференцировки тканей.

Склерозин, кодируемый геном SOST, также является членом подсемейства DAN (Brumkow et al., 2001, *Am. J. Hum. Genet.* 68:577-589). Ген SOST ассоциируется со склеростеозом, т.е. склерозирующей дисплазией кости, наследуемой по аутосомно-рецессивному типу. Фенотип, ассоциируемый со склеростеозом, представляет собой прогрессирующее разрастание скелетных костей, которое приводит к гигантизму, деформации лица и к ущемлению седьмого и восьмого черепных нервов (Brumkow et al. 2001 (см. выше)). Взаимосвязь между склеростеозом и SOST была установлена посредством картирования по го-

мозиготности семей, подверженных указанному заболеванию. Brumkow и сотрудники идентифицировали сходство между фенотипом, ассоциированным со склеростеозом, и эффектами, ассоциированными с другими членами подсемейства DAN. Аргумент в пользу такой взаимосвязи является еще более убедительным, если предположить, что склеростеоз может быть обусловлен потерей негативного регулятора для члена подсемейства TGF-бета, а более конкретно, BMP.

Поэтому, идентификация секретируемых белков и, в частности, факторов роста, таких как члены суперсемейства, имеющие в своей структуре цистиновые узлы, в частности, члены подсемейства DAN, является крайне необходимой для лучшего понимания основных путей, которые приводят к развитию приведенных выше патологических состояний и ассоциированных с ними заболеваний, приведенных выше, и для разработки более эффективных методов генотерапии или лекарственной терапии указанных заболеваний.

Описание изобретения

Настоящее изобретение основано на обнаружении функций белка INSP002, действующего как секретируемый белок, в частности, как секретируемый белок подсемейства DAN, входящего в суперсемейство, имеющего "цистиновые узлы" в своей структуре.

В первом аспекте настоящее изобретение относится к полипептиду, который

(i) содержит аминокислотную последовательность, представленную в SEQ ID NO:2, SEQ ID NO:4 или SEQ ID NO:7;

(ii) представляет собой ее фрагмент, который обладает функцией секретируемого белка, предпочтительно, функцией члена суперсемейства цитокинов с цистиновыми узлами, более предпочтительно, члена подсемейств DAN, или который имеет общую с полипептидом (i) антигенную детерминанту; или

(iii) представляет собой функциональный эквивалент (i) или (ii).

Предпочтительно, полипептид первого аспекта настоящего изобретения:

(i) содержит аминокислотную последовательность, представленную в SEQ ID NO:6 или SEQ ID NO:8;

(ii) представляет собой его фрагмент, который обладает функцией секретируемого белка, предпочтительно, функцией члена суперсемейства цитокинов с цистиновыми узлами, более предпочтительно, члена подсемейств DAN, либо который имеет общую с полипептидами (i) антигенную детерминанту; или

(iii) представляет собой функциональный эквивалент (i) или (ii).

В соответствии с другим вариантом своего первого аспекта, настоящее изобретение относится к полипептиду, который

i) состоит из аминокислотной последовательности, представленной в SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7 или SEQ ID NO:8;

(ii) представляет собой его фрагмент, который обладает функцией секретируемого белка, предпочтительно, функцией члена суперсемейства цитокинов с цистиновыми узлами, более предпочтительно, члена подсемейств DAN, или который имеет общую с полипептидом (i) антигенную детерминанту; или

(iii) представляет собой функциональный эквивалент (i) или (ii).

Полипептид, имеющий последовательность, представленную в SEQ ID NO:2, будет далее называться "полипептид экзона 1 INSP002". Полипептид, имеющий последовательность, представленную SEQ ID NO:4, будет далее называться "полипептид экзона 2 INSP002". Полипептид, имеющий последовательность, представленную SEQ ID NO:6, будет далее называться "полипептидом INSP002". Первые 22 аминокислоты полипептида экзона 1 INSP002 представляют собой сигнальный пептид, а последовательности полипептида INSP002 без сигнальной последовательности представлены в SEQ ID NO:7 и SEQ ID NO:8. Полипептид, имеющий последовательность, представленную в SEQ ID NO:7, будет далее называться "полипептид экзона 1 INSP002 без сигнального пептида". Полипептид, имеющий последовательность, представленную в SEQ ID NO:8, будет далее называться "полипептидом INSP002 без сигнального пептида".

В соответствии с другим вариантом своего первого аспекта, настоящее изобретение относится к полипептиду, который

i) содержит или состоит из аминокислотной последовательности, представленной в SEQ ID NO:14;

(ii) представляет собой его фрагмент, который обладает функцией секретируемого белка, предпочтительно, функцией члена суперсемейства цитокинов с цистиновыми узлами, более предпочтительно, члена подсемейств DAN, или который имеет общую с полипептидом (i) антигенную детерминанту; или

(iii) представляет собой функциональный эквивалент (i) или (ii).

Полипептид, имеющий последовательность, представленную в SEQ ID NO:14, представляет собой вариант полипептида INSP002. Он идентичен полипептиду INSP002, за исключением того, что по сравнению с полипептидом INSP002, содержит две аминокислотные делеции в положениях 107 и 108 и одну аминокислотную замену в положении 110. Полипептид, имеющий последовательность, представленную в SEQ ID NO:14, далее будет называться "вариантом полипептида INSP002".

Предпочтительно полипептид первого аспекта настоящего изобретения функционирует как член суперсемейства цитокинов, имеющих в своей структуре цистиновые узлы, предпочтительно как член

подсемейства DAN. Термин "цитокин с цистиновыми узлами" хорошо известен, и специалист с помощью одного из известных анализов, может легко установить, функционирует ли данный полипептид как член суперсемейства цитокинов, имеющих цистиновые узлы в своей структуре.

В частности, специалист с помощью одного из известных анализов может легко установить, функционирует ли данный полипептид как член подсемейства DAN, и является ли он антагонистом членов суперсемейства TGF-бета, в частности, антагонистом BMP. Для того чтобы определить, функционирует ли полипептид как антагонист BMP, в качестве системы для анализа может быть использован эмбрион *Xenopus*, поскольку некоторые BMP экспрессируются в эмбрионе *Xenopus* (Chang C. et al., 1999, Development 126:3347-3357, Hawley S. et al., 1995, Genes Dev. 9:2923-2935, Hemmati-Brivanlou, A., and G. H. Thomsen. 1995, Dev. Genet. 17:78-89, Jones CM. et al., 1992, 10 Development 115:639-647). Сверхэкспрессия сигналов BMP-2/4-класса или BMP-7-класса в ранней мезодерме индуцирует вентральные пути, в то время как ингибиторы этих сигналов (такие как Noggin, Xnr3, Chordin или Follistatin) индуцируют дорсальные пути. Поэтому действие полипептида на эмбриональное развитие может быть использовано для того, чтобы определить является ли данный полипептид антагонистом BMP.

Используемый термин "полипептид INSP002" включает полипептиды, содержащие полипептид экзона 1 INSP002, полипептид экзона 1 INSP002 без сигнального пептида, полипептид экзона 2 INSP002, полипептид INSP002 или полипептид INSP002 без сигнального пептида, а также полипептиды, состоящие из полипептида экзона 1 INSP002, полипептида экзона 1 INSP002 без сигнального пептида, полипептида экзона 2 INSP002, полипептида INSP002, полипептида INSP002 без сигнального пептида или варианта полипептида INSP002.

Во втором своем аспекте настоящее изобретение относится к очищенной молекуле нуклеиновой кислоты, кодирующей полипептид первого аспекта настоящего изобретения.

При этом предпочтительно, чтобы указанная очищенная молекула нуклеиновой кислоты содержала последовательность нуклеиновой кислоты, представленную в SEQ ID NO:1 (кодирующую полипептид экзона 1 INSP002), в SEQ ID NO:3 (кодирующую полипептид экзона 2 INSP002), в SEQ ID NO:5 (кодирующую полипептид INSP002) или в SEQ ID NO:13 (кодирующую вариант полипептида INSP002), или представляла собой избыточный эквивалент или фрагмент любой из этих последовательностей.

Настоящее изобретение также относится к очищенной молекуле нуклеиновой кислоты, состоящей из последовательности нуклеиновой кислоты, представленной в SEQ ID NO:1 (кодирующей полипептид экзона 1 INSP002), в SEQ ID NO:3 (кодирующей полипептид экзона 2 INSP002), в SEQ ID NO:5 (кодирующей полипептид INSP002) или в SEQ ID NO:13 (кодирующей вариант полипептида INSP002), или представляющей собой избыточный эквивалент или фрагмент любой из этих последовательностей.

В соответствии с одним из вариантов этого аспекта настоящего изобретения очищенная молекула нуклеиновой кислоты не содержит 5'-нетранслируемую область, расположенную выше последовательности нуклеиновой кислоты, кодирующей полипептид экзона 1 INSP002 и последовательности нуклеиновой кислоты, кодирующей полипептид INSP002 (нуклеотиды 1-151 SEQ ID NO:1 и SEQ ID NO:5). В соответствии с этим вариантом настоящего изобретения очищенная молекула нуклеиновой кислоты, предпочтительно, содержит нуклеотиды 152-475 SEQ ID NO:1 или нуклеотиды 152-721 SEQ ID NO:5. Настоящее изобретение, кроме того, относится к очищенной молекуле нуклеиновой кислоты, состоящей из нуклеотидов 152-475 SEQ ID NO:1 или нуклеотидов 152-721 SEQ ID NO:5. Нуклеотидная последовательность, кодирующая полипептид INSP002 без 5'-нетранслируемой области (нуклеотиды 152-721 SEQ ID NO:5), представлена в SEQ ID NO:11, а нуклеотидная последовательность, кодирующая полипептид экзона 1 INSP002 без 5'-нетранслируемой области (нуклеотиды 152-475 SEQ ID NO:1), представлена в SEQ ID NO:12.

В соответствии с другим вариантом этого аспекта настоящего изобретения очищенная молекула нуклеиновой кислоты не кодирует сигнальный пептид, расположенный в начале полипептида экзона 1 INSP002 и полипептида INSP002 (нуклеотиды 152-217 SEQ ID NO:1 и SEQ ID NO:5). В соответствии с этим вариантом настоящего изобретения, очищенная молекула нуклеиновой кислоты, предпочтительно, содержит нуклеотиды 218-475 SEQ ID NO:1 (кодирующие полипептид экзона 1 INSP002 без сигнального пептида) или нуклеотиды 218-721 SEQ ID NO:5 (кодирующие полипептид INSP002 без сигнального пептида). Настоящее изобретение, кроме того, относится к очищенной молекуле нуклеиновой кислоты, состоящей из нуклеотидов 218-475 SEQ ID NO:1 (кодирующих полипептид INSP002 без сигнального пептида) или нуклеотидов 218-721 SEQ ID NO:5 кодирующих полипептид INSP002 без сигнального пептида). Нуклеотидная последовательность, кодирующая зрелый полипептид INSP002 (SEQ ID NO:7), представлена в SEQ ID NO:9, а нуклеотидная последовательность, кодирующая зрелый полипептид экзона 1 INSP002, представлена в SEQ ID NO:10.

В соответствии с другим вариантом этого аспекта настоящего изобретения очищенная молекула нуклеиновой кислоты не содержит 5'-нетранслируемую область, расположенную выше последовательности нуклеиновой кислоты, кодирующей вариант полипептида INSP002 (нуклеотиды 1-68 SEQ ID NO:13). В соответствии с этим вариантом настоящего изобретения очищенная молекула нуклеиновой кислоты, предпочтительно, содержит нуклеотиды или состоит из нуклеотидов 69-719 SEQ ID NO:13. Нуклеотидная последовательность, кодирующая вариант полипептида INSP002 без 5'-нетранслируемой области

(нуклеотиды 69-719 SEQ ID NO:13) представлена в SEQ ID NO:15.

В третьем аспекте настоящее изобретение относится к очищенной молекуле нуклеиновой кислоты, которая гибридизуется в условиях высокой жесткости с молекулой нуклеиновой кислоты второго аспекта настоящего изобретения.

В четвертом аспекте настоящее изобретение относится к вектору, такому как экспрессирующий вектор, содержащий молекулу нуклеиновой кислоты второго или третьего аспекта настоящего изобретения.

В пятом аспекте настоящее изобретение относится к клетке-хозяину, трансформированному вектором четвертого аспекта настоящего изобретения.

В шестом аспекте настоящее изобретение относится к лиганду, который специфически связывается с полипептидом первого аспекта настоящего изобретения, обладающим активностью цитокина, имеющего цистиновые узлы в своей структуре, и который, предпочтительно, ингибирует такую активность. При этом предпочтительно, чтобы лиганд ингибировал функцию полипептида первого аспекта настоящего изобретения, который является членом подсемейства DAN цитокинов, имеющих цистиновые узлы в своей структуре. Лиганды для полипептида настоящего изобретения могут образовывать различные формы, включая природные или модифицированные субстраты, ферменты, рецепторы, небольшие органические молекулы, такие как небольшие природные или синтетические органические молекулы, имеющие молекулярную массу до 2000 Да, предпочтительно, 800 Да или менее, пептидомиметики, неорганические молекулы, пептиды, полипептиды, антитела, структурные или функциональные миметики приведенных выше соединений. В седьмом аспекте настоящее изобретение относится к соединению, которое эффективно модифицирует экспрессию природного гена, кодирующего полипептид первого аспекта настоящего изобретения, или регулирует активность полипептида первого аспекта настоящего изобретения.

Соединение седьмого аспекта настоящего изобретения может либо повышать (служить агонистом), либо снижать (служить антагонистом) уровень экспрессии гена или активности полипептида. Важно отметить, что идентификация функции полипептида INSP002 дает возможность разработать способы скрининга, позволяющие идентифицировать соединения, эффективные для лечения и/или диагностики заболеваний.

В восьмом аспекте настоящее изобретение относится к полипептиду первого аспекта настоящего изобретения, или к молекуле нуклеиновой кислоты второго или третьего аспекта настоящего изобретения, или к вектору четвертого аспекта настоящего изобретения, или к лиганду пятого аспекта настоящего изобретения, или к соединению шестого аспекта настоящего изобретения, которые могут быть использованы для лечения или диагностики. Эти молекулы могут быть также использованы для производства лекарственного средства, предназначенного для лечения клеточно-пролиферативных расстройств, аутоиммунных/воспалительных заболеваний, сердечно-сосудистых заболеваний, неврологических расстройств, нарушений развития, метаболических расстройств, инфекций и других патологических состояний.

В девятом аспекте настоящее изобретение относится к способу диагностики заболеваний у человека, включающему оценку уровня экспрессии природного гена, кодирующего полипептид первого аспекта настоящего изобретения, или уровня активности полипептида первого аспекта настоящего изобретения в ткани указанного пациента и сравнение указанного уровня экспрессии или активности с контрольным уровнем, где уровень, который отличается от указанного контрольного уровня, является показателем заболевания. Такой способ предпочтительно осуществляют *in vitro*. Аналогичные способы могут быть использованы для мониторинга терапевтического лечения заболевания у пациента, где изменение уровня экспрессии или активности полипептида или молекулы нуклеиновой кислоты в течение определенного периода времени по сравнению с контрольным уровнем служит показателем рецидива заболевания.

В соответствии с девятым и десятым аспектами настоящего изобретения расстройством или заболеванием, предпочтительно, является расстройство или заболевание, которое ассоциируется с aberrantными уровнями цитокина, имеющего цистиновые узлы в своей структуре, предпочтительно, члена подсемейства DAN. Заболеванием или расстройством может быть заболевание или расстройство, которое ассоциируется с aberrantными уровнями лиганда для цитокина, имеющего цистиновые узлы в своей структуре, предпочтительно, для члена подсемейства DAN. Так, например, таким заболеванием или расстройством может быть расстройство или заболевание, которое ассоциируется с aberrantными уровнями члена суперсемейства TGF-бета. В частности, указанным заболеванием или расстройством может быть заболевание или расстройство, которое ассоциируется с BMP, такое как невропатия, нефропатия, такая как диабетическая нефропатия, рак, заживление ран, фиброз, остеопения, остеопороз, переломы и склеростоз.

Предпочтительный способ детекции полипептидов первого аспекта настоящего изобретения включает стадии;

(a) контактирования лиганда шестого аспекта настоящего изобретения, такого как антитело, с биологическим образцом в условиях, подходящих для образования комплекса лиганд-полипептид; и

(b) детекции указанного комплекса.

Что касается девятого аспекта настоящего изобретения, то каждому специалисту известно, что для детекции аномальных уровней белка существуют различные методы, такие как методы гибридизации нуклеиновой кислоты с короткими зондами, методы анализа на точечную мутацию, метод амплификации с помощью полимеразной цепной реакции (ПЦР) и методы, в которых используются антитела. Аналогичные методы могут быть использованы в течение коротких или продолжительных промежутков времени, что позволяет наблюдать за терапевтическим лечением заболевания пациента. Настоящее изобретение также относится к набору, который может быть использован в указанных методах диагностики заболевания.

В десятом аспекте настоящее изобретение относится к использованию полипептида первого аспекта настоящего изобретения в качестве секретируемого белка. Предпочтительно, настоящее изобретение относится к применению полипептида первого аспекта настоящего изобретения в качестве цитокина, более предпочтительно, в качестве цитокина, имеющего цистиновые узлы в своей структуре, в частности, в качестве члена подсемейства DAN цитокинов, имеющих цистиновые узлы в своей структуре.

В одиннадцатом аспекте настоящее изобретение относится к фармацевтической композиции, содержащей полипептид первого аспекта настоящего изобретения или молекулу нуклеиновой кислоты второго или третьего аспекта настоящего изобретения, или вектор четвертого аспекта настоящего изобретения, или клетку-хозяина пятого аспекта настоящего изобретения, или лиганд шестого аспекта настоящего изобретения, или соединения седьмого аспекта настоящего изобретения в сочетании с фармацевтически приемлемым носителем.

В двенадцатом аспекте настоящее изобретение относится к полипептиду первого аспекта настоящего изобретения, или к молекуле нуклеиновой кислоты второго или третьего аспекта настоящего изобретения, или к вектору четвертого аспекта настоящего изобретения, или к клетке-хозяину пятого аспекта настоящего изобретения, или к лиганду шестого аспекта настоящего изобретения, или к соединению седьмого аспекта настоящего изобретения, которые могут быть использованы для производства лекарственного средства для диагностики или лечения заболеваний.

В тринадцатом аспекте настоящее изобретение относится к способу лечения заболевания у пациента, предусматривающему введение указанному пациенту полипептида первого аспекта настоящего изобретения, или молекулы нуклеиновой кислоты второго или третьего аспекта настоящего изобретения, или вектора четвертого аспекта настоящего изобретения, или лиганда шестого аспекта настоящего изобретения, или соединения седьмого аспекта настоящего изобретения.

В соответствии с двенадцатым и тринадцатым аспектами настоящего изобретения заболеванием является, предпочтительно, заболевание, которое ассоциируется с аберрантными уровнями цитокина, имеющего цистиновые узлы в своей структуре, а предпочтительно, члена подсемейства DAN. Таким заболеванием может быть заболевание, которое ассоциируется с аберрантными уровнями лиганда для цитокина, имеющего цистиновые узлы в своей структуре, а предпочтительно, лиганда для члена подсемейства DAN. Так, например, таким заболеванием может быть заболевание, которое ассоциируется с аберрантными уровнями члена суперсемейства TGF-бета. В частности, указанным заболеванием или расстройством может быть заболевание или расстройство, которое ассоциируется с BMP, такое как невропатия, нефропатия, такая как диабетическая нефропатия, рак, заживление ран, фиброз, остеопения, остеопороз, переломы и склеростеоз.

Для заболеваний, при которых у пациента уровень экспрессии природного гена, кодирующего полипептид первого аспекта настоящего изобретения, или при которых активность полипептида первого аспекта настоящего изобретения ниже, чем уровень экспрессии или активности у здорового пациента, полипептид, молекула нуклеиновой кислоты, лиганд или соединения, вводимые указанному пациенту, должны обладать агонистическими свойствами. И, наоборот, для заболеваний, при которых у данного пациента уровень экспрессии природного гена, кодирующего полипептид первого аспекта настоящего изобретения, или при которых активность полипептида первого аспекта настоящего изобретения выше, чем уровень экспрессии или активности у здорового пациента, полипептид, молекула нуклеиновой кислоты, лиганд или соединения, вводимые указанному пациенту, должны обладать антагонистическими свойствами. Примерами таких антагонистов являются молекулы антисмысловой нуклеиновой кислоты, рибозимы и лиганды, такие как антитела.

В четырнадцатом аспекте настоящее изобретение относится к трансгенным или к дефицитным по данному полипептиду животным, которые не являются человеком и которые были трансформированы так, чтобы они экспрессировали более высокие или более низкие уровни или вообще не экспрессировали полипептида первого аспекта настоящего изобретения. Такие трансгенные животные являются наиболее подходящими моделями для исследования заболеваний, и могут быть также использованы в способах скрининга для идентификации соединений, которые являются эффективными для лечения или диагностики такого заболевания.

Краткое описание стандартных способов и методик, которые могут быть использованы для осуществления настоящего изобретения, приводится ниже. Следует отметить, что настоящее изобретение не ограничивается конкретно описанными способами, протоколами, клеточными линиями, векторами и реагентами. Следует также отметить, что используемая здесь терминология приводится лишь в целях опи-

сания конкретных вариантов осуществления настоящего изобретения и не должна рассматриваться как ограничение объема настоящего изобретения. Объем настоящего изобретения ограничивается лишь прилагаемой формулой изобретения.

В данном описании используются стандартные сокращения, принятые для обозначения нуклеотидов и аминокислот.

Если не оговорено особо, то для осуществления настоящего изобретения могут быть использованы, стандартные методы молекулярной биологии и микробиологии, методы рекомбинантных ДНК и методы иммунологии, которые хорошо известны специалистам в данной области.

Более полное описание таких методов можно найти в литературе. Так, например, наиболее подходящие описания, которые могут быть использованы для консультации, приводятся в следующих работах: Sambrook Molecular Cloning; A Laboratory Manual, Second Edition (1989); DNA Cloning, Volumes I and II (D.N Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); the Methods in Enzymology series (Academic Press, Inc.), especially volumes 154 & 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds. 1987, Academic Press, London); Scopes, (1987) Protein Purification: Principles and Practice, Second Edition (Springer Verlag, N.Y.); and Handbook of Experimental Immunology, Volumes I-IV (D.M. Weir and C.C. Blackwell eds. 1986).

Используемый термин "полипептид" включает любой пептид или белок, содержащий две или более аминокислот, связанных друг с другом пептидными связями или модифицированными пептидными связями, т.е. пептидными изостерами. Этот термин относится к коротким цепям (пептидам или полипептидам) и к более длинным цепям (белкам).

Полипептид настоящего изобретения может присутствовать в форме зрелого белка, либо он может присутствовать в виде пре-, про- или препро-белка, который может быть активирован путем отщепления пре-, про- или препро-части этого белка с продуцированием активного зрелого полипептида. В таких полипептидах пре-, про- или препро-последовательность может представлять собой лидерную или секреторную последовательность, либо она может представлять собой последовательность, используемую для очистки зрелой полипептидной последовательности.

Полипептид первого аспекта настоящего изобретения может образовывать часть гибридного белка. Так, например, часто оказывается предпочтительным, чтобы данный полипептид включал одну или более дополнительных аминокислотных последовательностей, которые могут содержать секреторную или лидерную последовательности, про-последовательности, последовательности, облегчающие очистку, или последовательности, сообщающие белку более высокую стабильность, например, во время рекомбинантного продуцирования. Альтернативно или дополнительно, зрелый полипептид может быть присоединен к другому соединению, например, к соединению, увеличивающему время полужизни полипептида (например, к полиэтиленгликолю).

Полипептиды могут содержать аминокислоты, которые отличаются от 20 аминокислот, кодируемых генами, и которые являются модифицированными либо в результате природных процессов, таких как посттрансляционный процессинг, либо в результате применения химических методов модификации, хорошо известных специалистам в данной области. Примерами известных модификаций, которые могут быть осуществлены в полипептидах настоящего изобретения, являются гликозилирование, присоединение липида, сульфирование, гамма-карбоксилирование, например, остатков глутаминовой кислоты, гидроксиглирование и ADP-рибозилирование. Другими возможными модификациями являются ацетилирование, ацилирование, амидирование, ковалентное присоединение флавина, ковалентное присоединение молекулы гема, ковалентное присоединение нуклеотида или нуклеотидного производного, ковалентное присоединение липидного производного, ковалентное присоединение фосфатидилинозита, перекрестное сшивание, циклизация, образование дисульфидной связи, деметилирование, образование ковалентных поперечных связей, образование цистеина, образование пироглутамата, формилирование, образование GPI-якоря, иодирование, метилирование, миристоилирование, окисление, протеолитический процессинг, фосфорилирование, пренилирование, рацемизация, селеноилирование, присоединение аминокислот к белкам, опосредуемое переносом РНК, такое как аргинилирование и убихитинизация.

Модификации могут присутствовать в любом участке полипептида, включая пептидный остов, аминокислотные боковые цепи и амино- или карбокси-концы. Действительно, блокирование амино- или карбокси-конца в полипептиде, либо того и другого, посредством ковалентной модификации является обычным процессом, происходящим в природных и синтетических полипептидах, и такие модификации могут присутствовать в полипептидах настоящего изобретения.

Модификации, которые присутствуют в полипептиде, в большинстве случаев, будут зависеть от способа получения данного полипептида. Для полипептида, которые получены рекомбинантным методом, природа и степень модификации, по большей части, будет определяться способностью к посттрансляционной модификации конкретной клетки-хозяина и присутствием сигналов модификации в амино-

кислотной последовательности рассматриваемого полипептида. Так, например, картины гликозилирования могут варьироваться у различных типов клеток-хозяина.

Полипептиды настоящего изобретения могут быть получены любым подходящим способом. Таковыми полипептидами являются выделенные природные полипептиды (например, выделенные из клеточной культуры), полипептиды, продуцированные рекомбинантным методом (включая гибридные белки), синтетические полипептиды или полипептиды, продуцированные с использованием комбинации этих методов.

Функционально эквивалентные полипептиды первого аспекта настоящего изобретения могут представлять собой полипептиды, гомологичные полипептидам INSP002. Два полипептида могут быть определены термином "гомологичные", если последовательность одного из этих полипептидов имеет достаточно высокую степень идентичности или сходства с последовательностью другого полипептида. Термин "идентичность" означает, что в любом конкретном положении сопоставляемых последовательностей, эти две последовательности имеют идентичные аминокислотные остатки. Термин "сходство" означает, что в любом конкретном положении сопоставляемых последовательностей эти две последовательности имеют аминокислотные остатки аналогичного типа. Степень идентичности и сходства может быть легко определена. (Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing. Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part 1, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).

Поэтому гомологичными полипептидами являются природные биологические варианты (например, аллельные варианты или варианты полипептидов, образовавшиеся в результате географических изменений видов, от которых они происходят), и мутанты (такие как мутанты, содержащие аминокислотные замены, инсерции или делеции) полипептидов INSP002. Такими мутантами могут быть полипептиды, в которых один или более аминокислотных остатков заменены консервативным или неконсервативным аминокислотным остатком (предпочтительно, консервативным аминокислотным остатком), и такой замененный аминокислотный остаток может быть, а может и не быть остатком, кодируемым генетическим кодом. Обычно такие замены происходят между Ala, Val, Leu и Ile; между Ser и Thr; между кислотными остатками Asp и Glu; между Asn и Gln; между основными остатками Lys и Arg; или между ароматическими остатками Phe и Tyr. Особенно предпочтительными являются варианты, в которых несколько, т.е. от 5 до 10, от 1 до 5, от 1 до 3, от 1 до 2 аминокислот или только одна аминокислота являются замененными, делетированными или добавленными в любой комбинации. Особенно предпочтительными являются "молчащие" замены, добавления и делеции, которые не влияют на свойства и активность указанного белка. Также предпочтительными являются консервативные замены. Такими мутантами также являются полипептиды, в которых один или более аминокислотных остатков включают группу-заместитель.

Обычно считается, что два полипептида, имеющие более чем 30% идентичность, являются функционально эквивалентными. Предпочтительно, чтобы функционально эквивалентные полипептиды первого аспекта настоящего изобретения имели степень идентичности последовательности полипептида INSP002 или его активных фрагментов, составляющую более 35%. Более предпочтительными являются полипептиды, имеющие степень идентичности более чем 35, 40, 50, 60, 70, 80, 90, 95, 98, 99% или более соответственно.

Функционально эквивалентными полипептидами первого аспекта настоящего изобретения могут быть также полипептиды, которые были идентифицированы с применением одного или более методов сопоставления структур. Так, например, для идентификации полипептидов с неизвестными функциями, в отношении которых, несмотря на их низкую степень идентичности полипептидам INSP002, делается предположение, что они обладают активностью секретированной молекулы, где указанное предположение основано на их значительной структурной гомологии с последовательностями полипептида INSP002, может быть использована технология информационной обработки потока данных для генома (Inpharmatica Genome Threader), которая составляет один из аспектов методов поиска, используемых для генерирования базы данных поиска Biopendium (см. совместно рассматриваемую заявку на патент Великобритании, PCT/GB01/01105). Термин "значительная структурная гомология" означает, что технология Inpharmatica Genome Threader позволяет предсказать, с достоверностью 10% и выше, что два белка имеют структурную гомологию.

Полипептиды первого аспекта настоящего изобретения также включают фрагменты полипептидов INSP002 и фрагменты функциональных эквивалентов полипептидов INSP002, при условии, что эти фрагменты сохраняют активность цитокина, имеющего цистиновые узлы в своей структуре, а предпочтительно, активность члена подсемейства DAN, имеющего цистиновые узлы в своей структуре, либо они имеют антигенную детерминанту, общую для полипептидов INSP002.

Используемый здесь термин "фрагмент" означает полипептид, имеющий аминокислотную последовательность, которая идентична части, но не всей, аминокислотной последовательности полипептидов INSP002, либо одному из их функциональных эквивалентов. Эти фрагменты должны содержать, по крайней мере n смежных аминокислот последовательности, и в зависимости от конкретной последова-

тельности, n , предпочтительно, равно 7 или более (например, 8, 10, 12, 14, 16, 18, 20 или более). Небольшие фрагменты могут образовывать антигенную детерминанту.

Указанные фрагменты могут присутствовать в "свободной форме", т.е. не являться частью другого полипептида или не быть присоединенными к другим аминокислотам или полипептидам, либо они могут входить в состав более крупного полипептида, часть или область которого они образуют. Если фрагмент настоящего изобретения входит в состав более крупного полипептида, то наиболее предпочтительно, чтобы такой фрагмент образовывал одну непрерывную область. Так, например, некоторые предпочтительные варианты настоящего изобретения относятся к фрагменту, имеющему пре- и/или про-полипептидную область, присоединенную к амино-концу указанного фрагмента, и/или дополнительную область, присоединенную к карбоксильному концу этого фрагмента. При этом в состав одного более крупного полипептида могут входить несколько фрагментов.

Полипептиды настоящего изобретения или их иммуногенные фрагменты (содержащие, по крайней мере одну антигенную детерминанту) могут быть использованы для генерирования лигандов, таких как поликлональные или моноклональные антитела, которые обладают иммуноспецифичностью к полипептидам. Указанные антитела могут быть использованы для выделения или для идентификации клонов, экспрессирующих полипептиды настоящего изобретения, или для очистки данных полипептидов аффинной хроматографией. Как совершенно очевидно для специалиста, такие антитела, помимо других применений, могут быть также использованы в качестве диагностических или терапевтических средств.

Термин "иммуноспецифический" означает, что антитела имеют, в основном, более высокую аффинность к полипептидам настоящего изобретения, чем к другим родственным полипептидам известного уровня. Используемый здесь термин "антитело" означает интактные молекулы, а также их фрагменты, такие как Fab, $F(ab')_2$ и Fv, которые способны связываться с рассматриваемой антигенной детерминантой. Такие антитела связываются с полипептидами первого аспекта настоящего изобретения.

Под понятием "значительно более высокая аффинность" подразумевается заметно повышенная аффинность к полипептиду настоящего изобретения по сравнению с аффинностью известных секретируемых белков.

При этом, предпочтительно, чтобы аффинность к полипептиду настоящего изобретения, по крайней мере, в $1,5$, 2 , 5 , 10 , 100 , 10^3 , 10^4 , 10^5 или 10^6 раз превышала аффинность к известным секретируемым белкам, таким как цитокины, имеющие цистиновые узлы в своей структуре, в частности, таким как члены подсемейства DAN.

Если желательно использовать поликлональные антитела, то выбранное млекопитающее, такое как мышь, кролик, коза или лошадь, может быть иммунизовано полипептидом первого аспекта настоящего изобретения. Полипептид, используемый для иммунизации животного, может быть получен методами рекомбинантных ДНК, либо он может быть синтезирован методом химического синтеза. Если необходимо, то данный полипептид может быть конъюгирован с белком-носителем. Обычно используемыми носителями, к которым могут быть химически присоединены данные полипептиды, являются альбумин бычьей сыворотки, тироглобулин и гемоцианин лимфы улитки. Затем связанный полипептид может быть использован для иммунизации животного. Сыворотку, взятую у иммунизованного животного, собирают и обрабатывают известными методами, например, иммуноаффинной хроматографией.

Моноклональные антитела против полипептидов первого аспекта настоящего изобретения могут быть легко продуцированы любым специалистом. Общая методика получения моноклональных антител с использованием гибридомной техники хорошо известна в данной области (см. например, Kohler G. & Milstein, C. *Nature* 256:495-497 (1975); Kozbor et al. *Immunology Today* 4:72(1983); Cole et al., 77-96, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985)).

Панели моноклональных антител, продуцированных против полипептидов первого аспекта настоящего изобретения, могут быть скринированы на различные свойства, т.е. на изотип, эпитоп, аффинность и т.п. Моноклональные антитела являются особенно подходящими для очистки отдельных полипептидов, против которых они направлены. Альтернативно, гены, кодирующие нужные моноклональные антитела, могут быть выделены из гибридом, например, методами ПЦР, известными в данной области, а также они могут быть клонированы и экспрессированы в соответствующих векторах.

Могут быть также использованы химерные антитела, в которых не-человеческие переменные области соединены или лигированы с человеческими константными областями (см. например, Liu et al., *Proc. Natl. Acad. Sci. USA*, 84, 3439 (1987)).

Эти антитела могут быть модифицированы так, чтобы они были менее иммуногенными для индивидуума, например, путем их "гуманизации", см. Jones et al. *Nature*, 321, 522 (1986); Verhoeyen et al. *Science*, 239, 1534 (1988); Kabat et al. *J. Immunol.*, 147, 1709 (1991); Queen et al. *Proc. Natl. Acad. Sci. USA*, 86, 10029 (1989); Gorman et al., *Proc. Natl. Acad. Sci. USA*, 88, 34181 (1991); and Hodgson et al., *Bio/Technology*, 9, 421 (1991)). Используемый термин "гуманизованное антитело" означает молекулы антитела, в которых аминокислоты CDR и другие выбранные аминокислоты в переменных доменах тяжелой и/или легкой цепей нечеловеческого донорного антитела были заменены эквивалентными аминокислотами человеческого антитела. Таким образом, гуманизованное антитело имеет близкое сходство с человеческим антителом, но, при этом, оно обладает способностью связываться с донорным антителом.

В другом альтернативном варианте изобретения, таким антителом может быть "биспецифическое" антитело, т.е. антитело, имеющее два различных антиген-связывающих домена, каждый из которых обладает специфичностью к различным эпитопам.

Для отбора генов, кодирующих антитела, обладающие способностью связываться с полипептидами настоящего изобретения, либо из набора ПЦР-амплифицированных V-генов лимфоцитов человека, скринированных на способность вырабатывать соответствующие антитела, либо из библиотеки "необученных" лимфоцитов, может быть использована техника фагового представления (McCafferty J. et al. (1990), Nature 348, 552-554; Marks J. et al. (1992), Biotechnology 10, 779-783). Аффинность этих антител может быть также увеличена путем перестановки цепей (Clackson T. et al. (1991), Nature 352, 624-628).

Антитела, генерированные описанными выше методами, независимо от того, являются ли они поликлональными или моноклональными, обладают и другими ценными свойствами, т.е. они могут быть использованы в качестве реагентов в иммуноанализах, радиоиммуноанализах (РИА) или твердофазных иммуноферментных анализах (ELISA). Для этих целей антитела могут быть помечены аналитически детектируемым реагентом, таким как радиоизотоп, флуоресцентная молекула или фермент.

Предпочтительными молекулами нуклеиновой кислоты второго и третьего аспектов настоящего изобретения являются молекулы, которые кодируют полипептидные последовательности, представленные в SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:7, SEQ ID NO:6, SEQ ID NO:8 и SEQ ID NO:14, и функционально эквивалентные полипептиды. Эти молекулы нуклеиновой кислоты могут быть использованы в способах и для целей, описанных в настоящей заявке. Молекулы нуклеиновой кислоты настоящего изобретения, предпочтительно, содержат, по крайней мере n смежных нуклеотидов в описываемых последовательностях, и в зависимости от конкретной последовательности, n предпочтительно, равно 10 или более (например, 12, 14, 15, 18, 20, 25, 30, 35, 40 или более).

Молекулы нуклеиновой кислоты настоящего изобретения также включают последовательности, комплементарные последовательностям молекул нуклеиновой кислоты, описанных выше (например, для использования в качестве антисмысловых последовательностей или в качестве зондов).

Молекулы нуклеиновой кислоты настоящего изобретения могут присутствовать в форме РНК, такой как мРНК, или в форме ДНК, включая, например, кДНК, синтетическую ДНК или геномную ДНК. Такие молекулы нуклеиновой кислоты могут быть получены методами клонирования, химического синтеза или их комбинацией. Молекулы нуклеиновой кислоты могут быть получены, например, методом химического синтеза, таким как твердофазный фосфорамидитный химический синтез, выделение из геномных или кДНК-библиотек или выделение из микроорганизма. РНК-молекулы могут быть, в основном, генерированы путем *in vitro*- или *in vivo*-транскрипции ДНК-последовательностей.

Молекулы нуклеиновой кислоты могут быть двухцепочечными или одноцепочечными. Одноцепочечной ДНК может быть кодирующая цепь, также известная как смысловая цепь, либо некодирующая цепь, также называемая антисмысловой цепью.

Термин "молекула нуклеиновой кислоты" также охватывает аналоги ДНК и РНК, такие как аналоги, содержащие модифицированные остовы и связанные с пептидом нуклеиновые кислоты (PNA). Используемый термин "PNA" означает антисмысловую молекулу или антиген, который содержит олигонуклеотид, состоящий, по крайней мере, из пяти нуклеотидов, и присоединенный к пептидному остову из аминокислотных остатков, которые, предпочтительно, заканчиваются лизином. Концевой лизин придает данной композиции растворимость. PNA могут быть ПЭГилированы для продления их времени жизни в клетке, где они предпочтительно, связываются с комплементарной одноцепочечной ДНК и РНК и приостанавливают элонгацию транскрипта (Nielsen P.E. et al. (1993), Anticancer Drug Des. 8:53-63).

Молекула нуклеиновой кислоты, которая кодирует полипептид SEQ ID NO:2, может быть идентична кодирующей последовательности молекулы нуклеиновой кислоты, показанной в SEQ ID NO:1 (нуклеотиды 152-475), и представлена в SEQ ID NO:12. Молекула нуклеиновой кислоты, которая кодирует полипептид SEQ ID NO:7, может быть идентична кодирующей последовательности молекулы нуклеиновой кислоты, показанной в SEQ ID NO:1 (нуклеотиды 218-475), и представлена в SEQ ID NO:10. Молекула нуклеиновой кислоты, которая кодирует полипептид SEQ ID NO:4, может быть идентична кодирующей последовательности молекулы нуклеиновой кислоты, показанной в SEQ ID NO:3. Молекула нуклеиновой кислоты, которая кодирует полипептид SEQ ID NO:6, может быть идентична кодирующей последовательности молекулы нуклеиновой кислоты, показанной в SEQ ID NO:5 (нуклеотиды 152-721), и представлена в SEQ ID NO:11. Молекула нуклеиновой кислоты, которая кодирует полипептид SEQ ID NO:8, может быть идентична кодирующей последовательности молекулы нуклеиновой кислоты, показанной в SEQ ID NO:5 (нуклеотиды 218-721), и представлена в SEQ ID NO:9. Молекула нуклеиновой кислоты, которая кодирует полипептид SEQ ID NO:14, может быть идентична кодирующей последовательности молекулы нуклеиновой кислоты, показанной в SEQ ID NO:13 (нуклеотиды 69-719), и представлена в SEQ ID NO:15.

Эти молекулы могут также иметь различные последовательности, которые, вследствие вырожденности генетического кода, кодируют полипептид SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:7, SEQ ID NO:6, SEQ ID NO:8 или SEQ ID NO:14. Такими молекулами нуклеиновой кислоты, которые кодируют полипептид SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:7, SEQ ID NO:6, SEQ ID NO:8 или SEQ ID NO:14,

могут быть, но не ограничиваются ими, кодирующая последовательность для отдельного зрелого полипептида; кодирующая последовательность для зрелого полипептида вместе с другими кодирующими последовательностями, такими как последовательности, кодирующие лидерную или секреторную последовательность, такую как про-, пре- или препро-полипептидную последовательность; кодирующая последовательность зрелого полипептида с приведенными выше дополнительными кодирующими последовательностями или без них, либо вместе с дополнительными не кодирующими последовательностями, включая не кодирующие 5'- и 3'-последовательности, такие как транскрибированные нетранслированные последовательности, которые играют определенную роль в транскрипции (включая сигналы терминации) и в связывании с рибосомой и обеспечивают стабильность мРНК. Молекулами нуклеиновой кислоты могут также включать вспомогательные последовательности, кодирующие дополнительные аминокислоты, такие как аминокислоты, обладающие дополнительными функциональными свойствами.

Молекулы нуклеиновой кислоты второго и третьего аспектов настоящего изобретения могут также кодировать фрагменты или функциональные эквиваленты полипептидов и фрагментов первого аспекта настоящего изобретения. Такая молекула нуклеиновой кислоты может представлять собой природный вариант, такой как природный аллельный вариант, либо указанной молекулой может быть вариант, не встречающийся в природе. Указанные неприродные варианты молекулы нуклеиновой кислоты могут быть получены методами мутагенеза, включая методы, применяемые к молекулам нуклеиновой кислоты, к клеткам или к целым организмам.

Из рассматриваемых вариантов можно отметить варианты, которые отличаются от приведенных выше молекул нуклеиновой кислоты тем, что они имеют нуклеотидные замены, делеции или инсерции. Замены, делеции или инсерции могут быть осуществлены по отношению к одному или более нуклеотидам. Варианты могут быть модифицированы в кодирующей или в не кодирующей области или в той и другой области. Альтерации в кодирующих областях могут приводить к консервативным или неконсервативным аминокислотным заменам, делециям или инсерциям.

Молекулы нуклеиновой кислоты настоящего изобретения могут быть также сконструированы методами, в основном, известными в данной области, включая, в зависимости от различных целей, модификацию клонирования, процессинга и/или экспрессии генного продукта (полипептида). Перестановка ДНК путем рандомизированной фрагментации и повторной сборки с помощью ПЦР генных фрагментов и синтетических олигонуклеотидов представляет собой технологию, которая может быть использована для конструирования нуклеотидных последовательностей. Сайт-направленный мутагенез может быть использован для введения новых рестрикционных сайтов, изменения характера гликозилирования, изменения предпочтительности кодонов, продуцирования вариантов сплайсинга, введения мутаций и т.п.

Молекулы нуклеиновой кислоты, которые кодируют полипептид первого аспекта настоящего изобретения, могут быть лигированы с гетерологичной последовательностью так, чтобы комбинированная молекула нуклеиновой кислоты кодировала гибридный белок. Такие комбинированные молекулы нуклеиновой кислоты входят во второй или третий аспект настоящего изобретения. Так, например, для скрининга пептидных библиотек на ингибиторы активности указанного полипептида с использованием такой комбинированной молекулы нуклеиновой кислоты может оказаться полезным экспрессировать гибридный белок, который будет распознаваться коммерчески доступным антителом. Гибридный белок может быть также сконструирован таким образом, чтобы он содержал сайт расщепления, расположенный между последовательностью полипептида настоящего изобретения и последовательностью гетерологичного белка, и чтобы такой полипептид мог быть отщеплен и выделен из указанного гетерологичного белка.

Молекулами нуклеиновой кислоты настоящего изобретения могут быть также антисмысловые молекулы, которые являются частично комплементарными молекулам нуклеиновой кислоты, кодирующим полипептиды настоящего изобретения, и поэтому, они гибридизуются с кодирующими молекулами нуклеиновой кислоты (в результате гибридизации). Антисмысловые молекулы, такие как олигонуклеотиды, могут быть сконструированы так, чтобы они распознавали нужную нуклеиновую кислоту, кодирующую полипептид настоящего изобретения, специфически связывались с этой нуклеиновой кислотой и предупреждали ее транскрипцию, в соответствии с методами, хорошо известными специалистам (см. например, Cohen J.S., *Trends in Pharm. Sci.*, 10, 435 (1989), Okano, J. *Neurochem.* 56, 560 (1991); O'Connor, J. *Neurochem* 56, 560 (1991); Lee et al. *Nucleic Acids Res* 6, 3073 (1979); Cooney et al., *Science* 241, 456 (1988); Dervan et al. *Science* 251, 1360 (1991)).

Используемый термин "гибридизация" означает присоединение двух молекул нуклеиновой кислоты друг к другу посредством водородных связей. Обычно одна молекула может быть фиксирована на твердом носителе, а другая молекула может находиться в растворе в свободном состоянии. Затем эти две молекулы могут контактировать друг с другом в условиях, благоприятствующих образованию водородной связи. Факторами, влияющими на образование таких связей, являются: тип и объем растворителя; температура реакции; время гибридизации; перемешивание; присутствие агентов, блокирующих неспецифическое связывание молекулы в жидкой фазе с твердым носителем (реагент Денхардта или BLOTTO); концентрация молекул; использование соединений, повышающих скорость ассоциации молекул (сульфата декстрана или полиэтиленгликоля); и жесткость условий промывки после гибридизации (Sambrook et al.

[см. выше]).

Ингибирование гибридизации полностью комплементарной молекулы с молекулой-мишенью может быть оценено с использованием гибридизационного анализа, известного в данной области (Sambrook et al. [см. выше]). В основном, гомологичная молекула будет затем конкурировать с полностью гомологичной молекулой за связывание с молекулой-мишенью и ингибировать это связывание в различных условиях жесткости, как описано у Wahl G.M. и S.L. Berger (1987; *Methods Enzymol.* 152:399-407) и Kimmel A.R. (1987; *Methods Enzymol.* 152:507-511).

Термин "жесткость" означает условия реакции гибридизации, которые способствуют ассоциации наиболее сходных молекул, но не ассоциации молекул, не обладающих таким сходством. Условия гибридизации высокой жесткости определены как условия инкубирования в течение ночи при 42°C в растворе, содержащем 50% формамид, 5X SSC (150мМ NaCl, 15мМ тринатрийцитрат), 50мМ фосфат натрия (рН 7,6), 5х раствор Денхардта, 10% сульфат декстрана и 20 мкг/мл денатурированной фрагментированной ДНК спермы лосося, с последующей промывкой фильтров в 0,1X SSC приблизительно при 65°C. Условия низкой жесткости предусматривают реакцию гибридизации, осуществляемую при 35°C (Sambrook et al. [см. выше]). Предпочтительными условиями гибридизации являются условия гибридизации высокой жесткости.

В предпочтительных вариантах этого аспекта, настоящее изобретение относится к молекулам нуклеиновой кислоты, которые по всей своей длине по крайней мере на 70% идентичны молекуле нуклеиновой кислоты, кодирующей полипептиды INSP002 (SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:7, SEQ ID NO:6, SEQ ID NO:8 или SEQ ID NO:14), и к молекулам нуклеиновой кислоты, которые, в основном, комплементарны указанным молекулам нуклеиновой кислоты. В соответствии с этим аспектом настоящего изобретения, предпочтительно, чтобы молекула нуклеиновой кислоты содержала область, которая, по всей своей длине, по крайней мере, на 80% идентична кодирующим последовательностям для SEQ ID NO:2 и SEQ ID NO:7, представленным в SEQ ID NO:1, SEQ ID NO:10 и SEQ ID NO:12; кодирующей последовательности для SEQ ID NO:4, представленной в SEQ ID NO:3; кодирующим последовательностям для SEQ ID NO:6 и SEQ ID NO:8, представленным в SEQ ID NO:5, SEQ ID NO:9 и SEQ ID NO:11; или кодирующим последовательностям для SEQ ID NO:14, представленным в SEQ ID NO:13 и SEQ ID NO:15; либо, чтобы она представляла собой комплементарную им молекулу нуклеиновой кислоты. При этом особенно предпочтительно, чтобы молекулы нуклеиновой кислоты по всей своей длине были по крайней мере на 90%, более предпочтительно по крайней мере на 95%, особенно предпочтительно по крайней мере на 98% или 99% идентичны указанным последовательностям. Предпочтительными вариантами этого аспекта являются молекулы нуклеиновой кислоты, кодирующие полипептиды, которые обладают, в основном, такой же биологической функцией или активностью, что и полипептиды INSP002.

Настоящее изобретение относится к способу детекции молекулы нуклеиновой кислоты настоящего изобретения, включающему стадии:

(а) контактирования нуклеинового зонда настоящего изобретения с биологическим образцом в условиях гибридизации, способствующих образованию дуплексов; и

(b) детекции любого такого образованного дуплекса.

Как будет дополнительно обсуждаться ниже в связи с анализами, которые могут быть использованы в соответствии с настоящим изобретением, молекула нуклеиновой кислоты, описанная выше, может быть использована в качестве гибридизационного зонда для РНК, кДНК или геномной ДНК в целях выделения полноразмерных кДНК и геномных клонов, кодирующих полипептиды INSP002, и в целях выделения кДНК и геномных клонов гомологичных и ортологичных генов, имеющих высокую степень сходства с генами, кодирующими этот полипептид.

В этой связи, наряду с другими известными методами, могут быть использованы описанные и обсуждаемые методы, которые приводятся ниже в качестве иллюстрации. Методы секвенирования и анализа ДНК хорошо известны и, в основном, доступны в данной области, и могут быть реально использованы для осуществления многих вариантов настоящего изобретения, обсуждаемых в настоящей заявке. В указанных методах могут использоваться ферменты, такие как фрагмент Кленова ДНК-полимеразы I, секвенназа (US Biochemical Corp., Cleveland, OH), полимеразы Taq (Perkin Elmer), термостабильная полимеразы T7 (Amersham, Chicago, IL) или комбинация таких полимераз и корректирующие экзонуклеазы, такие как экзонуклеазы, присутствующие в ELONGASE Amplification System, и поставляемые Gibco/BRL (Gaithersburg, MD). Способ секвенирования может быть, предпочтительно, автоматизирован с использованием устройств, таких как аппарат Hamilton Micro Lab 2200 (Hamilton, Reno, NV), термоячейка Peltier Thermal Cycler (PTC200; MJ Research, Watertown, MA), катализатор ABI и ДНК-секвенаторы 373 и 377 DNA Sequencers (Perkin Elmer).

Одним из методов выделения молекулы нуклеиновой кислоты, кодирующей полипептид с функцией, эквивалентной функции полипептида INSP002, является зондирование библиотеки геномных ДНК или кДНК природным или искусственно сконструированным зондом с использованием стандартных методов, известных в данной области (см. например, "Current Protocols in Molecular Biology", Ausubel et al. (eds). Greene Publishing Associates and Wiley Interscience, New York, 1989, 1992). Особенно подходящими являются зонды, содержащие 15, предпочтительно по крайней мере 30, более предпочтительно по край-

ней мере 50 непрерывно следующих друг за другом оснований, которые соответствуют или комплементарны последовательностям нуклеиновой кислоты, происходящей от соответствующего кодирующего гена (SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13 или SEQ ID NO:15). Для облегчения идентификации такие зонды могут быть помечены аналитически детектируемым реагентом. Подходящими реагентами являются, но не ограничиваются ими, радиоизотопы, флуоресцентные красители и ферменты, способные катализировать образование детектируемого продукта. С использованием этих зондов специалист может самостоятельно выделить комплементарные копии полинуклеотидов геномной ДНК, кДНК или РНК, кодирующих представляющие интерес белки, происходящие от различных источников, например человека, млекопитающего или других животных, и скринировать эти источники на присутствие родственных последовательностей, например отдельных членов семейства, типа и/или подтипа.

Во многих случаях выделенные кДНК-последовательности будут неполными, т.е. в этих последовательностях область, кодирующая полипептид, будет обрезана обычно у 5'-конца. Для получения полноразмерных кДНК или для удлинения коротких кДНК существует несколько методов. Такие последовательности можно удлинить с использованием неполной нуклеотидной последовательности и с применением различных известных в данной области методов детекции расположенных выше последовательностей, таких как промоторы и регуляторные элементы. Так, например, одним из методов, который может быть использован в данном случае, является метод быстрой амплификации кДНК-концов (RACE; см., например, Frohman et al., PNAS USA 85, 8998-9002, 1998). Недавно разработанные модификации этой технологии, проиллюстрированные Marathon TM (Clontech Laboratories Inc.), значительно упростили поиск более длинных кДНК. Для поиска неизвестной последовательности нуклеиновой кислоты, которая является смежной с известным локусом, может быть использован слегка модифицированный метод, который называется "сайт-рестрикционной" ПЦР и предусматривает использование универсальных праймеров (Sarkar G. (1993) PCR Methods Applic. 2:318-322). Для амплификации или для удлинения последовательностей с использованием различных праймеров, полученных на основе известной области, может быть также использована обратная ПЦР (Triglia T. et al. (1988) Nucleic Acids Res. 16:8186). Другим методом, который может быть использован в данном случае, является ПЦР "с захватом", которая представляет собой ПЦР-амплификацию ДНК-фрагментов, смежных с известной последовательностью в искусственной хромосомной ДНК человека и дрожжей (Lagerstrom M. et al. (1991) PCR Methods Applic. 1, 111-119). Другим методом, который может быть использован для поиска неизвестных последовательностей, является метод Parker J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Кроме того, можно использовать ПЦР, "гнездовые" праймеры и библиотеки PromoterFinderTM для "прогулки" по геномной ДНК (Clontech, Palo Alto, CA). Этот способ позволяет избежать необходимости скрининга библиотек и может быть использован для обнаружения участков стыка интрон/экзон.

При скрининге на полноразмерные ДНК предпочтительно использовать библиотеки, которые были отобраны по размерам и содержат более крупные кДНК. Предпочтительными также являются библиотеки рандомизированных праймеров, которые могут включать дополнительные последовательности, содержащие 5'-области генов. Использование библиотеки рандомизированных праймеров может оказаться особенно предпочтительным в том случае, если библиотека oligo-d(T) не дает полноразмерной кДНК. Геномные библиотеки могут быть использованы для удлинения последовательности с получением 5'-нетранскрибированных регуляторных областей.

В одном из вариантов осуществления изобретения, молекулы нуклеиновой кислоты настоящего изобретения могут быть использованы для определения локализации в хромосоме. В этом методе молекула нуклеиновой кислоты может быть нацелена на конкретный участок, либо она может быть гибридизована с конкретным участком на отдельной хромосоме человека. В соответствии с настоящим изобретением, картирование релевантных последовательностей в хромосомах является важным звеном в подтверждении корреляции этих последовательностей с геноассоциированным заболеванием. После картирования последовательности с определением ее точной локализации, физическое положение последовательности на хромосоме может быть сопоставлено с данными генетической карты. Эти данные можно найти, например, в работе V. McKusick, Mendelian Inheritance in Man (имеющейся в настоящее время в John Hopkins University Welch Medical Library). Взаимосвязь между генами, которые могут быть картированы на одной и той же области хромосомы, и заболеваниями, ассоциированными с ними, затем идентифицируют с помощью анализа на сцепление генов (совместное наследование физически смежных генов). Это позволяет исследователям получить ценную информацию, которая может быть использована для поиска генов, ассоциированных с данным заболеванием, с применением метода позиционного клонирования или других методов обнаружения генов. После предварительного определения локализации генов, ассоциированных с данным заболеванием или синдромом, путем анализа на сцепление генов с конкретной областью генома, любое картирование последовательностей на данном участке может дать информацию об ассоциированных или регуляторных генах, которая может быть использована для последующих исследований. Молекула нуклеиновой кислоты может быть также использована для детекции различий в хромосомной локализации, обусловленных транслокацией, инверсией и т.п., у нормальных индивидуумов, индивидуумов-носителей и у индивидуумов с заболеванием.

Молекулы нуклеиновой кислоты настоящего изобретения также являются ценным материалом для определения локализации в тканях. Такие методы позволяют определять характер экспрессии полипептида в тканях путем детекции мРНК, которая их кодирует. Такими методами являются методы гибридизации *in situ* и методы амплификации нуклеотидов, такие как ПЦР. Результаты этих исследований позволили получить определенные сведения относительно нормальных функций данного полипептида в организме. Кроме того, в этих исследованиях сравнение нормального характера экспрессии мРНК с экспрессией мРНК, кодируемой мутантным геном, позволяет получить важную информацию о роли мутантных полипептидов в заболевании. Такая anomальная экспрессия может иметь кратковременную, пространственную или количественную природу.

Для ингибирования эндогенной экспрессии гена, кодирующего полипептид настоящего изобретения, могут быть также использованы методы "отключения" гена. Одним из методов, который может быть использован для "отключения" последовательность-специфического посттрансляционного гена, является интерференция РНК (RNAi) (Elbashir S.M. et al. *Nature* 2001, 411, 494-498). Короткие дцРНК-олигонуклеотиды синтезируют *in vitro* и вводят в клетку. Последовательность-специфическое связывание этих дцРНК-олигонуклеотидов запускает процесс деградации мРНК-мишени, приводит к снижению уровня или к отмене экспрессии целевого белка.

Эффективность методов "отключения" гена, описанных выше, может быть оценена путем определения уровня экспрессии полипептидов (например, с помощью Вестерн-блоттинга), и на РНК-уровне - с помощью методики, основанной на TaqMan.

Векторы настоящего изобретения содержат молекулы нуклеиновой кислоты настоящего изобретения и могут представлять собой клонирующие или экспрессирующие векторы. Клетки-хозяева настоящего изобретения, которые могут быть трансформированы, трансфецированы или трансдуцированы векторами настоящего изобретения, могут представлять собой прокариотические или эукариотические клетки-хозяева.

Полипептиды настоящего изобретения могут быть получены в рекомбинантной форме посредством экспрессии кодирующих эти полипептиды молекул нуклеиновой кислоты в векторах, содержащихся в клетке-хозяине. Такие методы экспрессии хорошо известны специалистам, и многие из них подробно описаны у Sambrook et al. (см. выше) и Fernandez & Hoeffler (1998, eds. "Gene expression systems. Using nature for the art of expression". Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto).

Для продуцирования полипептида в нужном хозяине могут быть использованы, в основном, любые системы или любые векторы, подходящие для поддержания, амплификации или экспрессии молекул нуклеиновой кислоты. Подходящая нуклеотидная последовательность может быть встроена в экспрессионную систему любым из хорошо известных и рутинных методов, таких как методы, описанные у Sambrook et al. (см. выше). В общих чертах, кодирующий ген может быть помещен под контроль регуляторного элемента, такого как промотор, сайт связывания с рибосомой (для экспрессии в бактериях) и необязательно оператор, так, чтобы ДНК-последовательность, кодирующая нужный полипептид, транскрибировалась в РНК в трансформированной клетке-хозяине.

Примерами подходящих систем экспрессии являются, например, хромосомная, эпизомная и вирусные системы, включая, например, векторы, происходящие от бактериальных плазмид, бактериофага, транспозонов, дрожжевых эпизом, инсерционных элементов, дрожжевых хромосомных элементов, вирусов, таких как бакуловирус, паповавирусы, такие как SV40, вирусы коровьей оспы, аденовирусы, вирусы оспы домашней птицы, псевдорабивирусы и ретровирусы, или их комбинации, а также векторы, происходящие от плазмидных и бактериофаговых генетических элементов, включая космиды и фагмиды. Для доставки более крупных фрагментов ДНК, чем те, которые могут содержаться и экспрессироваться в плазмиде, могут быть также использованы искусственные человеческие хромосомы (НАС).

Особенно подходящими экспрессионными системами являются микроорганизмы, такие как бактерии, трансформированные рекомбинантным бактериофагом, плазмидными или космидными ДНК-экспрессирующими векторами; дрожжи, трансформированные векторами для экспрессии в дрожжах; клеточные системы насекомых, инфицированные вирусными экспрессирующими векторами (например, бакуловирусом); клеточные системы растений, трансформированные вирусными экспрессирующими векторами (например, вирусом мозаики цветной капусты, CaMV; вирусом мозаики табака, TMV) или векторами для экспрессии в бактериях (например, плазмидами Ti или pBR322); или клеточные системы животных. Для продуцирования полипептидов настоящего изобретения могут быть также использованы бесклеточные системы трансляции.

Введение молекул нуклеиновой кислоты, кодирующих полипептид настоящего изобретения, в клетки-хозяева может быть осуществлено методами, описанными во многих известных лабораторных руководствах, таких как руководство Davis et al. *I Basic Methods in Molecular Biology* (1986) и Sambrook et al. [см. выше]. Особенно подходящими методами являются трансфекция с использованием фосфата кальция; трансфекция, опосредованная DEAE-декстраном; трансфекция; микроинъекция; трансфекция, опосредованная катионным липидом; электропорация; трансдукция; загрузка путем соскоба; введение методом биобаллистики или инфицирование (см. Sambrook et al., 1989, [см. выше], Ausubel et al. [см. выше],

Spector, Goldman & Leinwald, 1998). В эукариотических клетках экспрессионные системы, в зависимости от целей их использования, могут быть либо временными (например, эпизодическими), либо перманентными (хромосомная интеграция).

Кодирующая молекула нуклеиновой кислоты может включать, а может и не включать последовательность, кодирующую регуляторную последовательность, такую как сигнальный пептид, или лидерную последовательность, если это необходимо, например, для секреции транслируемого полипептида в просвет эндоплазматического ретикулума, в периплазматическое пространство или во внеклеточное пространство. Эти сигналы могут быть эндогенными по отношению к полипептиду, либо они могут быть гетерологичными. Лидерные последовательности могут быть удалены бактериальным хозяином при посттрансляционном процессинге.

Помимо регуляторных последовательностей может оказаться желательным введение регуляторных последовательностей, которые обеспечивают регуляцию экспрессии полипептида в зависимости от роста клетки-хозяина. Примерами регуляторных последовательностей являются последовательности, которые обеспечивают увеличение или снижение уровня экспрессии гена в ответ на химическую или физическую стимуляцию, включая присутствие регуляторного соединения, или в ответ на различные температурные или метаболические условия. Регуляторными последовательностями являются нетранслируемые области вектора, такие как энхансеры, промоторы и 5'- и 3'-нетранслируемые области. Эти последовательности взаимодействуют с клеточными белками хозяина, в результате чего осуществляется транскрипция и трансляция. Такие регуляторные последовательности могут варьироваться по своей длине и специфичности. В зависимости от выбранной векторной системы и выбранного хозяина могут быть использованы любые подходящие транскрипционные и трансляционные элементы, включая конститутивные и индуцибельные промоторы. Так, например, для клонирования в бактериальных системах могут быть использованы индуцибельные промоторы, такие как гибридный промотор *lacZ* фагмиды BlueScript (Stratagene, LaJolla, CA) или плазмиды pSport1™ (Gibco BRL) и т.п. В клетках насекомых может использоваться бакуловирусный промотор полиэдрина. Промоторы или энхансеры, происходящие от геномов растительных клеток (например, гена белка теплового шока, RUBISCO и гена запасных белков) или от вирусов растений (например, вирусных промоторов или лидерных последовательностей), могут быть клонированы в указанный вектор. В клеточных системах млекопитающих предпочтительными являются промоторы, происходящие от генов млекопитающих или от вирусов млекопитающих. Если необходимо генерировать клеточную линию, содержащую множество копий данной последовательности, то могут быть использованы векторы на основе SV40 или EBV с соответствующим селективным маркером.

Экспрессирующий вектор конструируют так, чтобы конкретная кодирующая последовательность нуклеиновой кислоты присутствовала в векторе вместе с соответствующими регуляторными последовательностями, и чтобы положение и ориентация кодирующей последовательности по отношению к регуляторным последовательностям позволяли такой кодирующей последовательности транскрибироваться под "контролем" регуляторных последовательностей, т.е. так, чтобы РНК-полимераза, которая связывается с ДНК-молекулой у регуляторных последовательностей, осуществляла транскрипцию кодирующей последовательности. В некоторых случаях может оказаться необходимым модифицировать последовательность так, чтобы она могла присоединяться к регуляторным последовательностям в соответствующей ориентации, т.е. с сохранением рамки считывания.

Регуляторные последовательности и другие контрольные последовательности могут быть лигированы с кодирующей последовательностью нуклеиновой кислоты перед встраиванием в вектор. Альтернативно, такая кодирующая последовательность может быть клонирована непосредственно в экспрессирующий вектор, который уже содержит регуляторные последовательности и соответствующий рестрикционный сайт.

Для длительного высокоэффективного продуцирования рекомбинантного полипептида, предпочтительной является стабильная экспрессия. Так, например, клеточные линии, которые стабильно экспрессируют нужный полипептид, могут быть трансформированы с использованием экспрессирующих векторов, которые могут содержать вирусные сайты инициации репликации и/или эндогенные экспрессионные элементы и выбранный маркерный ген, присутствующий на том же самом или на отдельном векторе. После введения этого вектора, клетки можно оставить на 1-2 дня для роста в обогащенной среде, которую затем заменяют селективной средой. Селективный маркер необходим для сообщения резистентности, используемой в целях отбора, и его присутствие позволяет культивировать и выделять клетки, успешно экспрессирующие введенные последовательности. Резистентные клоны стабильно трансформированных клеток могут быть подвергнуты пролиферации методами культивирования ткани, подходящими для клеток такого типа.

Клеточные линии млекопитающих, подходящие в качестве хозяев для экспрессии, хорошо известны в данной области, и такими клеточными линиями являются многие иммортализованные клеточные линии, имеющиеся в Американской коллекции типовых культур (ATCC), включая, но не ограничиваясь ими, клетки яичника китайского хомячка (CHO), клетки HeLa, клетки почек детёнышей хомячка (BHK), клетки почек обезьяны (COS), клетки C127, клетки 3T3, клетки BHK, клетки HEK 293, клетки меланомы Боуэса, клетки гепатоцеллюлярной карциномы человека (например, Hep G2) и ряд других клеточных

линий.

В бакуловирусной системе материалы для получения экспрессионных систем бакуловирус/клетка насекомого являются коммерчески доступными и поставляются в наборах, *inter alia*, от Invitrogen, San Diego CA (набор "MaxBac"). В общих чертах, эти методы известны специалистам и подробно описаны у Summers & Smith, Texas Agricultural Experiment Station Bulletin № 1555 (1987). Клетками-хозяевами, особенно подходящими для использования в данной системе, являются клетки насекомых, такие как клетки *Drosophila* S2 и клетки *Spodoptera* Sf9.

В данной области известно множество систем экспрессии генов в клеточных культурах растений и в целых растениях. Примерами подходящих систем экспрессии генов в клетках растений являются системы, описанные в US 5693506, US 5659122 и US 5608143. Другие примеры экспрессии генов в клеточных культурах растений описаны Zenk, *Phytochemistry* 30, 3861-3863 (1991).

В частности, можно использовать любые растения, из которых могут быть выделены и культивированы протопласты с последующим генерированием из них целых регенерированных растений, которые содержали бы перенесенный ген. В частности, из культивируемых клеток или тканей могут быть регенерированы все растения, включая, но, не ограничиваясь ими, все основные виды сахарного тростника, сахарной свеклы, хлопчатника, плодовых и других деревьев, бобовых и овощных растений.

Примерами особенно предпочтительных бактериальных клеток-хозяев являются клетки стрептококков, стафилококков, *E.coli*, *Streptomyces* и *Bacillus subtilis*.

Примерами особенно подходящих клеток-хозяев для экспрессии в грибах являются дрожжевые клетки (например, *S. cerevisiae*) и клетки *Aspergillus*.

Любые системы отбора, которые могут быть использованы для выделения трансформированных клеточных линий, известны в данной области. Примерами являются гены тимидинкиназы (Wigler M. et al. (1977) *Cell* 11:223-32) и аденин-фосфорибозилтрансферазы вируса простого герпеса (Lowy I. et al. (1980) *Cell* 22:817-23), которые могут быть использованы в tk- или apt⁺-клетках соответственно.

Кроме того, в качестве основы для отбора могут быть использованы гены резистентности к антиметаболиту, антибиотику или к гербициду, например, ген дигидрофолатредуктазы (DHFR), который придает резистентность к метотрексату (Wigler M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); ген npt, который придает резистентность к аминогликозидам, неомицину и к G-418 (Colbere-Garapin F. et al. (1981) *J. Mol. Biol.* 150:1-14), и гены als или pat, которые придают резистентность к хлорсульфурон- и фосфинотицинацетилтрансферазе соответственно. Были описаны и другие селективные гены, примеры которых хорошо известны специалистам.

Хотя присутствие или отсутствие экспрессии маркерного гена позволяет предположить, что присутствует также и нужный ген, однако его присутствие и экспрессия еще нуждаются в подтверждении. Так, например, если соответствующая последовательность была встроена в последовательность маркерного гена, то трансформированные клетки, содержащие соответствующие последовательности, могут быть идентифицированы по отсутствию функции маркерного гена. Альтернативно, маркерный ген может находиться в тандеме с последовательностью, кодирующей полипептид настоящего изобретения, находящийся под контролем одного промотора. Экспрессия маркерного гена в ответ на индуцирование или отбор обычно указывает также на экспрессию тандемного гена.

Альтернативно, клетки-хозяева, которые содержат последовательность нуклеиновой кислоты, кодирующую полипептид настоящего изобретения, и которые экспрессируют указанный полипептид, могут быть идентифицированы различными методами, известными специалистам. Такими методами являются, но не ограничиваются ими, ДНК-ДНК или ДНК-РНК-гибридизация и биоанализы на присутствие белка, например, сортировка клеток по интенсивности флуоресценции (FACS) или иммуноанализы (таких как твердофазный иммуноферментный анализ [ELISA] и радиоиммуноанализ [РИА]), которые предусматривают использование мембран, растворов или чипов для детекции и/или количественной оценки нуклеиновой кислоты или белка (см. Hampton R. et al. (1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul, M.N. и Maddox D.E. et al., (1983) *J. Exp. Med.* 158, 1211-1216).

Существует широкий ряд методов мечения и конъюгирования, известных специалистам, и эти методы могут быть использованы в различных анализах на присутствие нуклеиновых кислот и аминокислот. Методами, используемыми в целях продуцирования меченых зондов для гибридизации или ПЦР-зондов для детекции последовательностей, родственных молекулам нуклеиновой кислоты, кодирующим полипептиды настоящего изобретения, являются мечение олигонуклеотидами, ник-трансляция, мечение по концам или ПЦР-амплификация с использованием меченного полинуклеотида. Альтернативно, последовательности, кодирующие полипептид настоящего изобретения, могут быть клонированы в вектор для продуцирования мРНК-зонда. Такие векторы известны в данной области, являются коммерчески доступными и могут быть использованы для синтеза РНК-зондов *in vitro* путем добавления соответствующей РНК-полимеразы, такой как T7, T3 или SP6, и меченых нуклеотидов. Эти методики могут быть осуществлены с использованием различных коммерчески доступных наборов (Pharmacia & Upjohn (Kalamazoo, MI); Promega (Madison WI) и U.S. Biochemical Corp., Cleveland, OH).

Подходящими репортерными молекулами или метками, которые могут быть использованы для облегчения детекции, являются радионуклиды, ферменты и флуоресцентные, хемилюминисцентные или

хромогенные агенты, а также субстраты, кофакторы, ингибиторы, магнитные частицы и т.п.

Молекулы нуклеиновой кислоты настоящего изобретения могут быть также использованы для генерирования трансгенных животных, в частности, грызунов. Такие трансгенные животные входят в дополнительный аспект настоящего изобретения. Такое генерирование может быть осуществлено путем локальной модификации соматических клеток или манипуляций с зародышевыми линиями: для введения наследуемых модификаций. Такие трансгенные животные могут быть, в частности, использованы для генерирования животных-моделей в целях поиска молекул лекарственных средств, которые являются эффективными в качестве модуляторов полипептидов настоящего изобретения.

Полипептид может быть выделен и очищен из рекомбинантных клеточных культур хорошо известными методами, включая преципитацию сульфатом аммония или этанолом, экстракцию кислотой, анионообменную или катионообменную хроматографию, хроматографию на фосфоцеллюлозе, гидрофобную хроматографию, аффинную хроматографию, хроматографию на гидроксипатитах и хроматографию на лектине. Для очистки наиболее подходящей является высокоэффективная жидкостная хроматография. В случае если данный полипептид был денатурирован в процессе выделения или очистки, то для восстановления активной конформации может быть использована хорошо известная техника рефолдинга белков.

Если необходимо, то для облегчения очистки белков могут быть также использованы специальные векторные конструкции, полученные путем присоединения последовательностей, кодирующих полипептиды настоящего изобретения, к нуклеотидной последовательности, кодирующей полипептидный домен, который облегчает очистку растворимых белков. Примерами таких доменов, облегчающих очистку белков, являются пептиды, образующие хелатные комплексы с металлами, такие как: гистидин-триптофановые модули, которые позволяют проводить очистку на иммобилизованных металлах; домены белка А, которые позволяют проводить очистку на иммобилизованном иммуноглобулине; и домен, используемый в системе FLAGS удлинение/аффинная очистка (Immunex Corp., Seattle, WA). Для облегчения очистки, между доменом для очистки и полипептидом настоящего изобретения могут быть включены расщепляемые линкерные последовательности, такие как последовательности, специфичные к фактору ХА или энтерокиназе (Invitrogen, San Diego, CA). Один из таких экспрессирующих векторов обеспечивает экспрессию гибридного белка, содержащего полипептид настоящего изобретения, присоединенный к нескольким гистидиновым остаткам, за которыми следуют тиоредоксин или рестрикционный сайт энтерокиназы. Гистидиновые остатки облегчают проведение очистки с помощью IMAC (аффинной хроматографии на иммобилизованном ионе металла, описанной Porath J. et al. (1992), *Prot. Exp. Purif.* 3:263-281), тогда как тиоредоксин или рестрикционный сайт энтерокиназы позволяют выделять полипептид из гибридного белка. Обсуждение векторов, содержащих гибридные белки, можно найти у Kroll D.J. et al. (1993; *DNA Cell Biol.* 12:441-453).

Если экспрессируемый полипептид используется в аналитическом скрининге, то обычно предпочтительно, чтобы он был продуцирован на поверхности клетки-хозяина, в которой он экспрессируется. В этом случае, клетки-хозяева могут быть собраны до их использования в скрининг-анализе, например, таким методом, как сортировка клеток с возбуждением флуоресценции (FACS) или иммуноаффинным методом. Если полипептид секретируется в среду, то такая среда может быть восстановлена для выделения и очистки экспрессированного полипептида. Если полипептид продуцируется внутри клетки, то перед выделением полипептида клетки должны быть сначала подвергнуты лизису.

Полипептид настоящего изобретения может быть использован для скрининга библиотек соединений в любом из известных методов, применяемых для скрининга лекарственных средств. Такие соединения могут стимулировать (служить агонистами) или ингибировать (служить антагонистами) уровень экспрессии гена или активность полипептида настоящего изобретения, и поэтому они составляют дополнительный аспект настоящего изобретения. Предпочтительными соединениями являются соединения, способные влиять на экспрессию природного гена, кодирующего полипептид первого аспекта настоящего изобретения, или регулировать активность полипептида первого аспекта настоящего изобретения.

Соединения-агонисты или соединения-антагонисты могут быть выделены, например, из клеток, бесклеточных препаратов, химических библиотек или смесей природных продуктов. Такими агонистами или антагонистами могут быть природные или модифицированные субстраты, лиганды, ферменты, рецепторы, либо структурные или функциональные миметики. Подходящее описание таких методов скрининга можно найти в работе Coligan et al., *Current Protocols in Immunology* I(2):Chapter 5 (1991).

Соединения, которые могут рассматриваться как наиболее вероятные кандидаты на хорошие антагонисты, представляют собой молекулы, которые связываются с полипептидом настоящего изобретения, но не индуцируют каких-либо биологических эффектов полипептида после связывания с ним. Потенциальными антагонистами являются небольшие органические молекулы, пептиды, полипептиды и антитела, которые связываются с полипептидом настоящего изобретения и, тем самым, ингибируют или подавляют его активность. Таким образом, связывание полипептида с нормальными связывающимися клеточными молекулами может ингибироваться так, чтобы подавлялась нормальная биологическая активность полипептида.

Полипептид настоящего изобретения, который используется в таком методе скрининга, может на-

ходиться в растворе в свободном состоянии, может быть иммобилизован на твердом носителе, может присутствовать на клеточной поверхности, либо он может быть локализован внутри клетки. В основном такие методы скрининга предусматривают использование соответствующих клеток или клеточных мембран, экспрессирующих полипептид, который контактирует с тестируемым соединением, что приводит к связыванию, либо к стимуляции или ингибированию функционального ответа. Функциональный ответ клеток, контактируемых с тестируемым соединением, сравнивают с контрольными клетками, которые не контактировали с тестируемым соединением. С помощью подходящей системы детекции, такой анализ позволяет определить, может ли тестируемое соединение давать сигнал, генерируемый активацией указанного полипептида. Ингибиторы активации обычно анализируют в присутствии известного агониста, и оценивают влияние этого агониста на активацию в присутствии тестируемого соединения.

Предпочтительный способ идентификации соединения-агониста или соединения-антагониста по отношению к полипептиду настоящего изобретения включает:

(а) контактирование клетки, экспрессирующей на клеточной поверхности полипептид первого аспекта настоящего изобретения, который ассоциирован со вторым компонентом, способным давать детектируемый сигнал в ответ на связывание соединения с указанным полипептидом, и при этом соединение скринируют в условиях, стимулирующих связывание с полипептидом; и

(b) определение события связывания указанного соединения с указанным полипептидом, либо его активации или ингибирования путем измерения уровня сигнала, генерируемого в результате взаимодействия указанного соединения с указанным полипептидом.

Еще более предпочтительный способ идентификации агониста или антагониста по отношению к полипептиду настоящего изобретения включает:

(а) контактирование клетки, экспрессирующей на клеточной поверхности полипептид, который ассоциирован со вторым компонентом, способным давать детектируемый сигнал в ответ на связывание соединения с полипептидом, и при этом соединение скринируют в условиях, стимулирующих связывание с полипептидом; и

(b) определение события связывания соединения с полипептидом, либо его активации или ингибирования путем сравнения уровня сигнала, генерируемого в результате взаимодействия соединения с полипептидом, с уровнем сигнала, полученного в отсутствие указанного соединения.

В других предпочтительных вариантах изобретения общие методы, описанные выше, могут дополнительно включать осуществление идентификации агониста или антагониста в присутствии меченного или немеченного лиганда для полипептида.

В другом варианте изобретения способ идентификации агониста или антагониста полипептида настоящего изобретения включает: определение факта ингибирования связывания лиганда с клетками, которые содержат полипептид настоящего изобретения на своей поверхности, или с клеточными мембранами, содержащими такой полипептид, в присутствии соединения-кандидата в условиях, стимулирующих связывание с полипептидом, и определение количества лиганда, связанного с полипептидом. Считается, что соединение, способное приводить к снижению уровня связывания с лигандом, является агонистом или антагонистом. При этом предпочтительно, чтобы указанный лиганд был меченым.

Более конкретно, способ скрининга на соединение, являющееся агонистом или антагонистом к полипептиду, включает стадии:

(а) инкубирования меченного лиганда с целой клеткой, экспрессирующей полипептид настоящего изобретения на клеточной поверхности, или с клеточной мембраной, содержащей полипептид настоящего изобретения;

(b) измерения количества меченного лиганда, связанного с целой клеткой или с клеточной мембраной;

(с) добавления соединения-кандидата к смеси меченого лиганда и целой клетки или клеточной мембраны стадии (а) и доведение этой смеси до состояния равновесия;

(d) измерения количества меченого лиганда, связанного с целой клеткой или с клеточной мембраной, после проведения стадии (с); и

(е) сравнения различия в количествах связанного меченого лиганда стадий (b) и (d), где соединение, вызывающее снижение уровня связывания в стадии (d), считается агонистом или антагонистом.

В вышеописанных анализах может быть установлено, что указанные полипептиды модулируют различные физиологические и патологические процессы дозозависимым способом. Так, например, "функциональными эквивалентами" полипептидов настоящего изобретения являются полипептиды, которые в приведенных выше анализах обладают той же самой дозозависимой модулирующей активностью. Хотя уровень дозозависимой активности необязательно является идентичным уровню полипептидов настоящего изобретения, однако, предпочтительно, чтобы в данном анализе на активность "функциональные эквиваленты" обнаруживали в основном такую же аналогичную зависимость от дозы, как и полипептиды настоящего изобретения. В некоторых описанных выше вариантах изобретения могут быть использованы простые анализы на связывание, в которых адгезию тестируемого соединения к поверхности, несущей данный полипептид, детектируют с помощью метки, непосредственно или опосредованно ассоциированной с тестируемым соединением, или с помощью анализа на конкуренцию с меченым со-

единением-конкурентом. В другом варианте осуществления изобретения могут быть использованы конкурентные анализы на скрининг лекарственных средств, в которых нейтрализующие антитела, способные связываться с данным полипептидом, специфически конкурируют за связывание с тестируемым соединением. Таким образом, указанные антитела могут быть использованы для детекции на присутствие любого тестируемого соединения, обладающего специфической аффинностью связывания с указанным полипептидом.

Могут быть также разработаны анализы для детекции влияния добавленных тестируемых соединений на продуцирование мРНК, кодирующей указанный полипептид, в клетках. Так, например, анализ ELISA может быть разработан так, чтобы можно было измерить секретируемые или клеточно-ассоциированные уровни полипептида с использованием моноклональных или поликлональных антител стандартными методами, известными в данной области, и этот анализ может быть использован для поиска соединений, которые могут ингибировать или усиливать продуцирование полипептида из соответствующим образом модифицированных клеток или тканей. Затем может быть определен уровень образования связывающих комплексов между полипептидом и тестируемым соединением.

Аналитическими способами, которые также входят в объем настоящего изобретения, являются способы, предусматривающие использование генов и полипептидов настоящего изобретения в анализах на сверхэкспрессию или в анализах на исключение. Такие анализы предусматривают манипуляцию с уровнями указанных генов/полипептидов в клетках и оценку влияния этой манипуляции на физиологию клеток, подвергнутых манипуляции. Так, например, такие эксперименты позволяют точно выявить пути передачи сигнала и метаболические пути, в которых участвуют конкретные гены/полипептиды, а также получить информацию относительно идентичности полипептидов, с которыми взаимодействуют исследуемые полипептиды, и найти ключ к разгадке механизмов, с помощью которых регулируются родственные гены и белки.

Другой метод, который может быть использован для скрининга на лекарственные средства, дает возможность осуществлять высокоэффективный скрининг соединений, обладающих подходящей аффинностью связывания с нужным полипептидом (см., Международную патентную заявку WO 84/03564). В этом методе, большое число различных небольших тестируемых соединений синтезируют на твердом субстрате, который затем может быть подвергнут реакции с полипептидом настоящего изобретения и промывке. Одним из способов иммобилизации полипептида является использование не-нейтрализующих антител. Связанный полипептид может быть затем детектирован методами, хорошо известными в данной области. Очищенный полипептид также может быть непосредственно нанесен на планшеты для последующего использования в приведенных выше способах скрининга на лекарственные средства.

Полипептид настоящего изобретения может быть использован для идентификации мембрано-ассоциированных или растворимых рецепторов с помощью стандартной техники связывания с рецептором, известной в данной области, такой как анализы на связывание и перекрестное сшивание с лигандом, в которых полипептид метят радиоактивным изотопом, химически модифицируют или присоединяют к пептидной последовательности, которая облегчает его детекцию или очистку, и инкубируют с источником предполагаемого рецептора (например, с композицией клеток, клеточными мембранами, клеточными супернатантами, тканевыми экстрактами или с физиологическими жидкостями). Эффективность связывания может быть определена биофизическими методами, такими как поверхностный плазменный резонанс и спектроскопия. Анализы на связывание могут быть использованы для очистки и клонирования рецептора, но они могут быть также использованы для идентификации агонистов и антагонистов указанного полипептида, которые конкурируют с указанным полипептидом за связывание с рецептором. Стандартные методы проведения скрининг-анализов хорошо известны в данной области.

Настоящее изобретение также относится к набору для скрининга, используемому в методах идентификации агонистов, антагонистов, лигандов, рецепторов, субстратов и ферментов, описанных выше.

Настоящее изобретение также относится к агонистам, антагонистам, лигандам, рецепторам, субстратам, ферментам и к другим соединениям, модулирующим активность или антигенность полипептида настоящего изобретения в соответствии с описанными выше механизмами.

Настоящее изобретение также относится к фармацевтическим композициям, содержащим полипептид, нуклеиновую кислоту, лиганд или соединение настоящего изобретения в комбинации с подходящим фармацевтическим носителем. Эти композиции могут быть использованы в качестве терапевтических или диагностических реагентов, вакцин или в качестве других иммуногенных композиций, подробно описанных ниже.

В соответствии с используемой здесь терминологией, композиция, содержащая полипептид, нуклеиновую кислоту, лиганд или соединение [X], считается "в основном не содержащей" примесей [здесь, Y], если по крайней мере 85 мас.% от всего количества X+Y в данной композиции составляет X. Предпочтительно X составляет по крайней мере примерно 90% по общей массе X+Y в данной композиции, более предпочтительно по крайней мере примерно 95, 98 или даже 99 мас.%.

Фармацевтические композиции предпочтительно должны содержать терапевтически эффективное количество полипептида, молекулы нуклеиновой кислоты, лиганда или соединения настоящего изобретения. Используемый здесь термин "терапевтически эффективное количество" означает количество тера-

терапевтического агента, необходимого для лечения, ослабления или профилактики рассматриваемого заболевания или состояния, или для достижения детектируемого терапевтического или профилактического эффекта. Для каждого соединения терапевтически эффективная доза может быть сначала оценена либо в анализе клеточной культуры, например, опухолевых клеток, либо на животных-моделях, обычно на мышах, кроликах, собаках или свиньях. Животное-модель может быть также использовано для определения соответствующего интервала концентраций и способа введения. Такая информация может быть затем использована для определения подходящих доз и способов их введения человеку.

Точное эффективное количество соединения для введения индивидууму будет зависеть от тяжести патологического состояния, общего состояния здоровья индивидуума, возраста, массы и пола индивидуума, его режима питания и частоты введения лекарственного средства, а также от комбинации(й) лекарственных средств, чувствительности данного индивидуума на реакцию и его переносимость/восприимчивость к данной терапии. Такое количество может быть определено путем рутинного экспериментирования и может быть установлено врачом-клиницистом. В общих чертах, эффективная доза может составлять от 0,01 до 50 мг/кг, предпочтительно от 0,05 до 10 мг/кг. Композиции могут быть введены пациенту либо отдельно, либо в комбинации с другими агентами, лекарственными средствами или гормонами.

Фармацевтическая композиция может также содержать фармацевтически приемлемый носитель, подходящий для введения терапевтического агента. Такими носителями являются антитела и другие полипептиды, гены и другие терапевтические агенты, такие как липосомы, при условии, что этот носитель сам по себе не индуцирует выработку антител, оказывающих негативное воздействие на индивидуума, которому вводят указанную композицию, и что данный вводимый носитель не является чрезмерно токсичным. Подходящими носителями могут быть крупные макромолекулы с замедленным метаболизмом, такие как белки, полисахариды, полимолочные кислоты, полигликолевые кислоты, полимерные аминокислоты, сополимеры аминокислот и неактивные вирусные частицы.

Используемыми фармацевтически приемлемыми солями могут быть, например, соли минеральных кислот, такие как гидрохлориды, гидробромиды, фосфаты, сульфаты и т.п., и соли органических кислот, такие как ацетаты, пропионаты, малонаты, бензоаты и т.п. Подробное обсуждение фармацевтически приемлемых носителей можно найти в Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Фармацевтически приемлемые носители в терапевтических композициях могут, кроме того, содержать жидкости, такие как вода, физиологический раствор, глицерин и этанол. Кроме того, в указанных композициях могут присутствовать и вспомогательные вещества, такие как смачивающие или эмульгирующие агенты, pH-забуферизующие вещества и т.п. С помощью таких носителей могут быть получены фармацевтические композиции в виде таблеток, пилюль, драже, капсул, жидкостей, гелей, сиропов, взвесей, суспензий и т.п., для перорального введения пациенту.

Композиции настоящего изобретения, после их приготовления, могут быть непосредственно введены индивидууму. Индивидуумами, подвергаемыми лечению, могут быть животные, в частности человек.

Фармацевтические композиции, используемые в настоящем изобретении, могут быть введены различными способами, включая, но не ограничиваясь ими, пероральное, внутривенное, внутримышечное, внутриартериальное, интрамедулярное, интратекальное, интравентрикулярное, трансдермальное или чрескожное введение (см. например, WO 98/10734), а также подкожное, внутривентрикулярное, интраназальное, кишечное, местное, подязычное, интравагинальное или ректальное введение. Для введения фармацевтических композиций настоящего изобретения могут быть также использованы «генные ружья» или безигольные шприцы. В основном, терапевтические композиции могут быть приготовлены в виде растворов для инъекций, либо в виде жидких растворов или суспензий, либо они могут быть получены в виде твердых форм, подходящих для получения растворов или суспензий в жидких носителях перед их введением путем инъекции.

Указанные композиции могут быть непосредственно доставлены, в основном, путем подкожной, внутривентрикулярной, внутривенной или внутримышечной инъекции, либо они могут быть доставлены в интерстициальное пространство ткани. Композиции могут быть также введены в пораженный участок. При этом лечение может быть проведено по схеме введения разовой дозы или дробных доз.

Если активность полипептида настоящего изобретения превышает активность, требуемую для лечения конкретного патологического состояния, то в данном случае может быть рассмотрено несколько подходов. Один из таких подходов включает введение индивидууму вышеописанного соединения-ингибитора (антагониста) вместе с фармацевтически приемлемым носителем, где указанный ингибитор вводят в количестве, эффективном для ингибирования функции полипептида, например, блокирования связывания с лигандами, субстратами, ферментами или рецепторами, либо ингибирования вторичного сигнала, и, тем самым, эффективным для ослабления симптомов аномального состояния. Такими антагонистами, предпочтительно, являются антитела. Для минимизации иммуногенности антител, описанных выше, наиболее предпочтительно, чтобы такие антитела были химерными и/или гуманизованными.

В соответствии с другим подходом могут быть введены растворимые формы полипептидов, которые сохраняют аффинность связывания с рассматриваемым лигандом, субстратом, ферментом или рецептором. В основном, такой полипептид может быть введен в виде фрагментов, в которых сохраняются

соответствующие части.

В альтернативном подходе, экспрессия гена, кодирующего полипептид, может быть ингибирована методами блокирования экспрессии, такими как использование молекул антисмысловой нуклеиновой кислоты (описанных выше), которые могут быть либо эндогенно генерированы, либо введены отдельно. Модификации экспрессии генов могут быть получены путем конструирования комплементарных последовательностей или антисмысловых молекул (ДНК, РНК или PNA) для контролируемых элементов, 5'-последовательностей или регуляторных последовательностей (сигнальной последовательности, промоторов, энхансеров и интронов) гена, кодирующего полипептид. Аналогичным образом, ингибирование может быть осуществлено с использованием методики спаривания оснований с образованием "тройной спирали". Спаривание с образованием тройной спирали используется потому, что оно приводит к нарушению способности двойной спирали раскрываться так, чтобы это оказалось достаточным для связывания с полимеразой, факторами транскрипции или регуляторными молекулами. Успехи в клинической терапии, достигнутые за последнее время благодаря использованию ДНК-триплекса, были описаны в литературе (Gee J.E. et al. (1994): Huber B.E. & B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing Co., Mt. Kisco, NY). Комплементарная последовательность или антисмысловая молекула могут быть также сконструированы в целях блокирования трансляции мРНК посредством предотвращения связывания транскрипта с рибосомами. Такие олигонуклеотиды могут быть введены или генерированы *in situ* в результате экспрессии *in vivo*.

Кроме того, экспрессия полипептида настоящего изобретения может быть предотвращена с использованием рибозимов, специфичных к мРНК-последовательности, кодирующей этот полипептид. Рибозимы представляют собой каталитически активные РНК, которые могут быть природными или синтетическими (см. например, Usman N. et al., *Curr. Opin. Struct. Biol.* (1996) 6(4), 527-33). Синтетические рибозимы могут быть сконструированы так, чтобы они специфически расщепляли мРНК в выбранных положениях и, тем самым, предотвращали трансляцию мРНК в функциональный полипептид. Рибозимы могут быть синтезированы с использованием природного рибозофосфатного остова и природных оснований, обычно присутствующих в РНК-молекулах. Альтернативно, рибозимы могут быть синтезированы с использованием неприродных остовов, например, 2'-О-метил-РНК, в целях защиты от расщепления рибонуклеазой, и эти рибозимы могут содержать модифицированные основания.

РНК-молекулы могут быть модифицированы в целях увеличения внутриклеточной стабильности и времени полужизни. Возможными модификациями являются, но не ограничиваются ими, присоединение фланкирующих последовательностей у 5'- и/или у 3'-концов молекулы или использование в указанном остове молекулы фосфоритоата или 2'-О-метила вместо фосфодиэстеразных связей. Эта концепция может рассматриваться и при продуцировании PNA и может быть применена ко всем указанным молекулам посредством включения нетрадиционных оснований, таких как инозин, квеозин и бутозин, а также ацетил-, метил-, тио- и аналогичные модифицированные формы аденина, цитидина, гуанина, тимина и уридина, которые не могут легко распознаваться эндогенными эндонуклеазами.

Для лечения аномальных состояний, ассоциированных с недостаточной экспрессией полипептида настоящего изобретения и его активностью, имеется несколько способов. Один из таких способов предусматривает введение индивидууму терапевтически эффективного количества соединения, которое активирует полипептид, т.е. соединение-агонист, описанное выше, в целях ослабления симптомов патологического состояния. Альтернативно, терапевтическое количество указанного полипептида в комбинации с подходящим фармацевтическим носителем может быть введено в целях сохранения соответствующего физиологического равновесия данного полипептида.

Для осуществления эндогенного продуцирования у индивидуума данного полипептида соответствующими клетками может быть применена генотерапия. Генотерапию используют для перманентного лечения заболеваний, связанных с аномальным продуцированием полипептида, путем замены дефектного гена "правильным" терапевтическим геном.

Генотерапия настоящего изобретения может быть осуществлена *in vivo* или *ex vivo*. Для генотерапии *ex vivo* требуется выделение и очистка клеток, взятых у пациента, введение терапевтического гена и введение генетически модифицированных клеток обратно пациенту. В противоположность этому, генотерапия *in vivo* не требует выделения и очистки клеток пациента.

Для введения пациенту, терапевтический ген обычно является "упакованным". Носители для доставки генов могут быть невирусными, такими как липосомы, либо они могут представлять собой дефектные по репликации вирусы, такие как аденовирус, описанный Berkner K.L., *Curr. Top. Microbiol. Immunol.*, 158, 39-66 (1992) или векторы на основе адено-ассоциированного вируса (AAV), описанные Muzyczka N. *Curr. Top. Microbiol. Immunol.*, 158, 97-129 (1992) и в патенте США № 5252479. Так, например, молекула нуклеиновой кислоты, кодирующая полипептид настоящего изобретения, может быть сконструирована для экспрессии в дефектном по репликации ретровирусном векторе. Затем эта экспрессионная конструкция может быть выделена и введена в упаковывающую клетку, трансдуцированную ретровирусным плазмидным вектором, содержащим РНК, кодирующую указанный полипептид, так, чтобы указанная упаковывающая клетка продуцировала инфекционные вирусные частицы, содержащие нужный ген. Эти клетки-продуценты могут быть введены индивидууму для конструирования клеток *in*

vivo и экспрессии полипептид in vivo (см. Chapter 20 Gene Therapy and other Molecular Genetic-based Therapeutic Approaches (и цитируемые там ссылки), Human Molecular Genetics (1996), T. Strachan & A.P.Read, BIOS Scientific Publishers Ltd).

Другим способом является введение "оголенной ДНК", где терапевтический ген непосредственно инъецируют в кровотоки или в мышечную ткань.

В случае, когда полипептидами или молекулами нуклеиновой кислоты настоящего изобретения являются агенты, вызывающие заболевание, настоящее изобретение относится к их возможному применению в вакцинах для вырабатывания антител против указанного агента, вызывающего заболевание. Если приведенный выше полипептид или молекула нуклеиновой кислоты должны быть активированы, то разрабатываемая вакцина должна вырабатывать антитела или Т-клетки против указанных агентов (как описано в WO 00/29428).

Вакцины настоящего изобретения могут быть либо профилактическими (т.е. предназначенными для предупреждения инфекции), либо терапевтическими (т.е. предназначенными для лечения заболеваний после инфицирования). Такие вакцины содержат иммунизирующий(е) антиген(ы), иммуноген(ы), полипептид(ы), белок(ки) или нуклеиновую кислоту, обычно в комбинации с описанными выше фармацевтически приемлемыми носителями, которыми может быть любой носитель, который сам по себе не индуцирует вырабатывание антител, оказывающих негативное воздействие на индивидуума, которому вводят указанную композицию. Кроме того, эти носители могут действовать как иммуностимуляторы ("адъюванты"). Кроме того, антиген или иммуноген может быть конъюгирован с бактериальным токсиндом, таким как токсин дифтерии, столбняка, холеры, *H. pylori* и другие патогены.

Поскольку полипептиды могут разрушаться в желудке, то вакцины, содержащие полипептиды, предпочтительно, вводят парентерально (например, путем подкожной, внутримышечной, внутривенной или чрескожной инъекции). Композициями, подходящими для парентерального введения, являются водные и безводные стерильные растворы для инъекций, которые могут содержать антиоксиданты, буферы, бактериостаты и растворенные вещества, сообщающие данной композиции изотоничность с кровью реципиента, а также водные и безводные стерильные суспензии, которые могут содержать суспендирующие агенты или загустители.

Вакцинные композиции настоящего изобретения могут быть помещены в упаковки для разовых лекарственных форм или для дробных лекарственных форм. Так, например, эти лекарственные формы могут быть помещены в запаиваемые ампулы и сосуды и могут храниться в замороженном виде, в которые, непосредственно перед их использованием, необходимо лишь добавить стерильный жидкий носитель. Конкретная доза будет зависеть от удельной активности данной вакцины и может быть легко определена путем рутинного экспериментирования.

Доставка антител, связывающихся с полипептидами настоящего изобретения, методами генотерапии может быть осуществлена, например, как описано в Международной патентной заявке WO 98/55607.

Для получения вакцинных композиций может быть также использована технология безигольного впрыскивания (см., например, www.powderject.com). Ряд подходящих методов вакцинации и систем для доставки вакцин описаны в Международной патентной заявке WO 00/29428.

Настоящее изобретение также относится к применению молекул нуклеиновой кислоты настоящего изобретения в качестве диагностических реагентов. Детекция мутированной формы гена, характеризующейся существованием молекул нуклеиновой кислоты настоящего изобретения, которые ассоциируются с дисфункцией, используется в качестве диагностического средства, которое может облегчать установление диагноза заболевания или установления восприимчивости к заболеванию, возникающему в результате пониженной экспрессии, сверхэкспрессии или измененной пространственной или временной экспрессии данного гена. Индивидуумы, несущие мутации в данном гене, могут быть идентифицированы на уровне ДНК рядом методов.

Молекулы нуклеиновой кислоты, используемые для диагностики, могут быть получены из клеток данного индивидуума, таких как клетки крови, мочи, слюны, биоптата ткани или материала, полученного после аутопсии. Геномные ДНК могут быть использованы непосредственно для детекции либо перед проведением анализа они могут быть амплифицированы ферментативно с использованием ПЦР, лигазной цепной реакции (ЛЦР), амплификации с замещением цепи (SDA) или другими методами амплификации (см. Saiki et al., Nature, 324, 163-166 (1986); Bej et al. Crit. Rev. Biochem. Molec. Biol., 26, 301-334 (1991); Birkenmeyer et al. J. Virol. Meth., 35, 117-126 (1991), Van Brunt, J. Bio/Technology, 8, 291-294 (1990)).

В одном из аспектов настоящее изобретение относится к способу диагностики заболевания у пациента, включающему оценку уровня экспрессии природного гена, кодирующего полипептид настоящего изобретения, и сравнение полученного уровня экспрессии с контрольным уровнем, где уровень, отличающийся от указанного контрольного уровня, является признаком наличия данного заболевания. Этот способ может включать следующие стадии:

а) контактирование образца ткани, взятой у пациента, с нуклеиновокислотным зондом в жестких условиях, в результате чего образуется гибридный комплекс между молекулой нуклеиновой кислоты настоящего изобретения и указанным зондом;

b) контактирование контрольного образца с указанным зондом в условиях, аналогичных условиям стадии (a); и

c) детекцию присутствия гибридных комплексов в указанных образцах, где детекция уровней гибридного комплекса в образце данного пациента, которые отличаются от уровней гибридного комплекса в контрольном образце, является признаком наличия данного заболевания.

В другом аспекте настоящее изобретение относится к способу диагностики, включающему следующие стадии:

a) получения образцов тканей от пациента, исследуемого на наличие заболевания;

b) выделения молекулы нуклеиновой кислоты настоящего изобретения из указанного образца ткани; и

c) установления диагноза заболевания у данного пациента путем обнаружения присутствия мутации молекулы нуклеиновой кислоты, которая ассоциируется с данным заболеванием.

Для облегчения детекции молекулы нуклеиновой кислоты в описанных выше способах, может быть проведена стадия амплификации, например, с помощью ПЦР.

Делеции и инсерции могут быть детектированы по изменению размера амплифицированного продукта по сравнению с нормальным генотипом. Точечная мутация может быть идентифицирована путем гибридизации амплифицированной ДНК с меченой РНК настоящего изобретения, или альтернативно, с меченой антисмысловой ДНК-последовательностью настоящего изобретения. Полностью соответствующие последовательности могут быть дифференцированы от дуплексов с ошибочным спариванием путем гидролиза РНКазой или путем оценки различий в температурах плавления. Присутствие или отсутствие мутации у пациента может быть определено посредством контактирования ДНК с нуклеиновокислотным зондом, который гибридизуется с ДНК в жестких условиях, в результате чего образуется гибридная двухцепочечная молекула, которая имеет негибридизованную часть цепи нуклеиновокислотного зонда в любой области, соответствующей мутации, ассоциированной с заболеванием; и установления присутствия или отсутствия негибридизованной части цепи нуклеиновокислотного зонда как показателя наличия или отсутствия ассоциированной с заболеванием мутации в соответствующей области ДНК-цепи.

Такие способы диагностики являются особенно ценными для проведения пренатальных и даже неонатальных тестов.

Точечные мутации и другие отличия между последовательностями эталонного гена и "мутантных" генов могут быть идентифицированы другими хорошо известными методами, такими как прямое секвенирование ДНК или определение конформационного полиморфизма одноцепочечных последовательностей (см. Orita et al., *Genomics*, 5, 874-879 (1989)). Так, например, секвенирующий праймер может быть использован вместе двухцепочечным ПЦР-продуктом или с одноцепочечной матричной молекулой, генерированной с помощью модифицированной ПЦР. Определение последовательности осуществляют в соответствии со стандартными методиками с использованием радиоактивно меченых нуклеотидов или в соответствии с методами автоматического секвенирования с использованием флуоресцентных меток. Клонированные ДНК-сегменты могут быть также использованы в качестве зондов для детекции специфических ДНК-сегментов. Чувствительность этого метода значительно повышается при использовании в комбинации с ПЦР. Кроме того, точечные мутации и другие модификации последовательностей, такие как полиморфизм, могут быть детектированы, как описано выше, например, с использованием аллель-специфических олигонуклеотидов для ПЦР-амплификации последовательностей, отличающихся одним нуклеотидом.

Различия в ДНК-последовательностях могут быть также детектированы по изменениям электрофоретической подвижности ДНК-фрагментов в гелях, в присутствии или в отсутствии денатурирующих агентов, либо путем прямого секвенирования ДНК (например, Myers et al. *Science* (1985) 230:1242). Изменения в конкретных положениях данных последовательностей могут быть также выявлены либо путем проведения анализов на защиту от нуклеазы, таких как анализ на защиту от РНКазы или S1, либо путем химического расщепления (см. Cotton et al., *Proc. Natl. Acad. Sci. USA* (1985) 85:4397-4401).

Помимо стандартного гель-электрофореза и секвенирования ДНК, мутации, такие как микроделеции, анеуплоидии, транслокации и инверсии, могут быть также детектированы путем проведения анализа *in situ* (см. например, Keller et al. *DNA Probes*, 2nd Ed., Stockton Press, New York, N.Y., USA (1993)), т.е. ДНК или РНК-последовательности в клетках могут быть проанализированы на мутации, не прибегая к их выделению и/или иммобилизации на мембране. Гибридизация *in situ* с флуоресценцией (FISH) является в настоящее время наиболее широко используемым методом, и множество его описаний уже появилось в литературе (см. например, Trachuck et al. *Science* 250, 559-562 (1990) и Trask et al. *Trends, Genet.*, 7, 149-154 (1991)).

В другом варианте осуществления изобретения для проведения эффективного скрининга генетических вариантов, мутаций и полиморфизма может быть создан массив олигонуклеотидных зондов, содержащий молекулу нуклеиновой кислоты настоящего изобретения. Технология создания массивов хорошо известна в данной области, находит широкое применение и может быть использована для решения ряда проблем молекулярной генетики, касающихся экспрессии генов, сцепления генов и генетической измен-

чивости (см. например, M. Chee et al., Science (1996). Vol. 274, pp.610-613).

В одном из вариантов осуществления изобретения, такой массив может быть получен и использован в соответствии с методами, описанными в заявке PCT WO 95/11995 (Chee et al.); Lockhart D.J. et al., (1996). Nat. Biotech. 14:1675-1680; и Schema M. et al. (1996). Proc. Natl. Acad. Sci. 93:10614-10619. Олигонуклеотидные пары могут быть использованы в количестве от двух пар до одного миллиона пар. Олигомеры синтезируют в соответствующих участках на субстрате с использованием оптически направленного химического синтеза. Таким субстратом может быть бумага, нейлон или мембрана другого типа, фильтр, чип, покровное стекло или любой другой твердый носитель. В другом аспекте настоящего изобретения олигонуклеотид может быть синтезирован на поверхности субстрата с использованием процедуры химического связывания и струйного аппарата для нанесения путем разбрызгивания, как описано в заявке PCT WO 95/251116 (Baldeschweiler et al.). В другом своем аспекте, для определения порядка расположения кДНК-фрагментов или олигонуклеотидов на поверхности субстрата и связывания этих фрагментов или олигонуклеотидов с поверхностью субстрата могут быть использованы "сетчатые" массивы, аналогичные массивам для дот (или слот)-блотов, с использованием вакуумной системы и методов термического, УФ-, механического или химического связывания. Массив, такой как массивы, описанные выше, может быть создан вручную или с использованием подходящих устройств (слот-блот или дот-блот-аппарата), материалов (любого твердого носителя) и оборудования (включая роботизированную аппаратуру), и этот массив может содержать 8, 24, 96, 384, 1536 или 6144 олигонуклеотидов или любое другое число от двух и выше одного миллиона, которое является подходящим для эффективного использования коммерчески доступного оборудования.

Помимо методов, описанных выше, заболевания могут быть диагностированы методами, включающими определение в образце, взятом у индивидуума, аномально низкого или аномально высокого уровня полипептида или мРНК. Для количественной оценки полипептидов, пониженный или повышенный уровень экспрессии может быть измерен на РНК-уровне любыми методами, хорошо известными в данной области, например, путем амплификации нуклеиновых кислот, а именно, с помощью ПЦР или ОТ-ПЦР, путем проведения анализов на защиту от РНКазы, с помощью Нозерн-блот-анализов и другими методами гибридизации.

Аналитические методы, которые могут быть использованы для определения уровней полипептида настоящего изобретения в образце, взятом у хозяина, хорошо известны специалистам, и подробно обсуждаются выше (включая радиоиммуноанализы, анализы на конкурентное связывание, Вестерн-блот-анализ и ELISA-анализы). В этом аспекте настоящее изобретение относится к способу диагностики, который включает стадии:

(а) контактирования лиганда, описанного выше, с биологическим образцом в условиях, подходящих для образования комплекса лиганд-полипептид; и

(b) детекции указанного комплекса.

Протоколы, такие как ELISA, РИА и FACS, разработанные для измерения уровней полипептидов, могут быть, кроме того, взяты за основу при определении измененных или аномальных уровней экспрессии полипептида. Нормальные или стандартные значения для экспрессии полипептида получают путем объединения физиологических жидкостей или клеточных экстрактов, взятых у нормальных млекопитающих, предпочтительно, у человека, с антителом против полипептида в условиях, подходящих для образования комплекса. Количество образованного стандартного комплекса может быть оценено различными методами, такими как фотометрические методы.

Антитела, которые специфически связываются с полипептидом настоящего изобретения, могут быть использованы для диагностики состояний или заболеваний, характеризующихся экспрессией полипептида, или в анализах для наблюдения за пациентами, которым было проведено лечение полипептидами, молекулами нуклеиновой кислоты, лигандами и другими соединениями настоящего изобретения. Антитела, подходящие для использования в диагностических целях, могут быть получены способом, аналогичным способу, описанному выше для терапии.

Диагностические анализы на присутствие полипептида осуществляют методами, в которых используют антитело и метку для детекции полипептида в физиологических жидкостях или экстрактах клеток или тканей человека. При этом могут быть использованы модифицированные или немодифицированные антитела, и эти антитела могут быть помечены путем их ковалентного или нековалентного связывания с репортерной молекулой. При этом могут быть использованы репортерные молекулы широкого ряда, известные в данной области, и некоторые из них описаны выше.

Количества полипептида, экспрессируемого у индивидуума, в контрольном и зараженном образцах и образце ткани, взятой при биопсии, сравнивают со стандартными величинами. Отклонение величин, полученных для данного индивидуума, от стандартных величин позволяет определить параметры для диагностики заболевания. Диагностический анализ может быть использован для выявления отсутствия, присутствия и избыточной экспрессии полипептида и для мониторинга регуляции уровней полипептида во время терапевтического лечения. Такие анализы могут быть также использованы для оценки эффективности конкретного курса терапевтического лечения при исследовании животных в клинических испытаниях или для наблюдения за лечением отдельного пациента.

Диагностический набор настоящего изобретения может содержать:

- (a) молекулу нуклеиновой кислоты настоящего изобретения;
- (b) полипептид настоящего изобретения; или
- (c) лиганд настоящего изобретения.

В одном из аспектов настоящего изобретения, диагностический набор может включать первый контейнер, содержащий нуклеиновокислотный зонд, который гибридизуется в жестких условиях с молекулой нуклеиновой кислоты настоящего изобретения; второй контейнер, содержащий праймеры, подходящие для амплификации молекулы нуклеиновой кислоты; и инструкции по использованию зонда и праймеров для облегчения диагностики заболевания. Этот набор может, кроме того, включать третий контейнер, содержащий агент для расщепления негибридизованной РНК.

В альтернативном аспекте настоящего изобретения, диагностический набор может содержать массив молекул нуклеиновой кислоты, по крайней мере, одной из которых является молекула нуклеиновой кислоты настоящего изобретения.

Для детекции полипептида настоящего изобретения диагностический набор может содержать одно или более антител, связывающихся с полипептидом настоящего изобретения, и реагент, используемый для детекции реакции связывания между антителом и полипептидом.

Такие наборы могут быть использованы для диагностики заболевания или восприимчивости к заболеванию, в частности, для диагностики пролиферативных расстройств, аутоиммунных/воспалительных заболеваний, сердечно-сосудистых заболеваний, неврологических расстройств, нарушений развития, метаболических расстройств, инфекций и других патологических состояний. Такими заболеванием или расстройством являются, предпочтительно, заболевание или расстройство, ассоциированное с аберрантными уровнями цитокина, имеющего цистиновые узлы в своей структуре, предпочтительно, члена подсемейства DAN. Таким заболеванием или расстройством может быть также расстройство или заболевание, которое ассоциируется с аберрантными уровнями лиганда для цитокина, имеющего цистиновые узлы в своей структуре, предпочтительно, для члена подсемейства DAN. Так, например, таким заболеванием или расстройством может быть заболевание или расстройство, которое ассоциируется с аберрантными уровнями члена суперсемейства TGF-бета. В частности, указанным заболеванием или расстройством может быть заболевание или расстройство, которое ассоциируется с BMP, такое как невропатия, нефропатия, такая как диабетическая нефропатия, рак, заживление ран, фиброз, остеопения, остеопороз, переломы и склеростеоз. Различные аспекты и варианты настоящего изобретения будут более подробно проиллюстрированы на примерах, в частности, для полипептидов INSP002.

При этом следует отметить, что в настоящее изобретение могут быть внесены конкретные модификации, не выходящие за объем изобретения.

Краткое описание графического материала

Фиг. 1 - результаты, полученные с помощью BLAST в неизбыточной базе данных NCBI с использованием объединенной полипептидной последовательности SEQ ID NO:2 и SEQ ID NO:4;

фиг. 2 - сопоставление, осуществляемое с помощью программы BLAST, объединенной полипептидной последовательности SEQ ID NO:2 и SEQ ID NO:4 с близкородственной последовательностью сегбегус-родственного белка 1 Homo sapiens;

фиг. 3 - четыре наилучших соответствия в базах данных NCBI-nr и NCBI-nt BLAST, полученных для INSP002 26 ноября 2002;

фиг. 4 - сопоставление INSP002 с AK095926.1;

фиг. 5 - сопоставление INSP002 с IMAGE: 4558384;

фиг. 6 - нуклеотидная последовательность INSP002 с трансляцией;

фиг. 7 - неполная клонированная последовательность INSP002 с трансляцией;

фиг. 8 - карта для неполного PCRII-ТОПО-INSP002;

фиг. 9 - сопоставление предсказанной последовательности INSP002 (верхняя строка) с неполной клонированной последовательностью (нижняя строка);

фиг. 10 - нуклеотидная последовательность и трансляция кДНК-вставки в Image: 4558384;

фиг. 11 - сопоставление последовательностей предсказанного INSP002 (верхняя строка) с IMAGE: 4558384 (BC025333.1)(нижняя строка);

фиг. 12 - нуклеотидная последовательность и трансляция INSP002V, генерированного с помощью ПЦР, в Image: 4558384;

фиг. 13 - карта pCR4blunt-ТОПО-INSP002V;

фиг. 14 - сравнение предсказанного INSP002 (верхняя строка) с вариантом последовательности INSP002V (нижняя строка);

фиг. 15 - карта экспрессирующего вектора pEAK12d;

фиг. 16 - карта вектора pDONR201 Gateway;

фиг. 17 - карта pEAK12d-INSP002-V-6HIS;

фиг. 18 - последовательность полноразмерного INSP002, клонированного из сердца;

фиг. 19 - карта плазмиды pCR4blunt-ТОПО-INSP002FL, кодирующей полноразмерный INSP002, клонированного из сердца.

Примеры

Пример 1. Сравнение белка INSP002 с белками в базе данных последовательностей.

Полипептидную последовательность, происходящую от объединенных последовательностей SEQ ID NO:2 и SEQ ID NO:4, которая представляет собой транслированную последовательность следующих друг за другом экзонов INSP002, использовали в качестве запрашиваемой последовательности в программе BLAST для поиска в базе данных неизбыточных последовательностей NCBI. Наилучшие десять соответствий дают последовательности, аннотированные как белки cerberus или cerberus-родственные белки, которые являются членами семейства цитокинов с цистиновыми узлами, и которые сопоставляли с запрашиваемой последовательностью с высокосignифицируемыми величинами E ($2E^{-10}$ - $2E^{-06}$) (фиг. 1). На фиг. 2 показано сопоставление запрашиваемой последовательности INSP002 с последовательностью cerberus-родственного белка 1 Homo sapiens (Feng et al. 2001).

Полипептидную последовательность, происходящую от объединенных последовательностей SEQ ID NO:2 и SEQ ID NO:4, которая представляет собой транслированную последовательность следующих друг за другом экзонов INSP002, вводили в программу SignalP V2.0.b2 (Nielsen et al. 1997 Protein Eng. 1:1-6). С помощью этой программы было предсказано, что указанная полипептидная последовательность имеет сигнальный пептид. Наиболее вероятно, что сайт отщепления сигнального пептида находится между остатками 22 и 23 указанной полипептидной последовательности INSP002, происходящей от объединенных последовательностей SEQ ID NO:2 и SEQ ID NO:4.

Нуклеотидная последовательность SEQ ID NO:1, кодирующая экзон 1 полипептида SEQ ID NO:2, содержит 5'-нетранслируемую область (5'UTR) и белок-кодирующую последовательность (CDS). CDS начинается с нуклеотида 152.

Пример 2. Повторение BLAST-поиска.

BLAST-поиск в базах данных NCBI-пг и NCBI-нт проводили 26 ноября 2002 с использованием полипептидной последовательности SEQ ID NO:6, происходящей от объединенных последовательностей SEQ ID NO:2 и SEQ ID NO:4. Наилучшие четыре соответствия, идентифицированные в результате проведенного поиска, показаны на фиг. 3.

Этот поиск выявил, что полипептид INSP002 идентичен гипотетическому белку FLJ38067 на аминокислотном уровне и соответствующей нуклеотидной последовательности AK095926, клонированной из сердца и депонированной 16 июля 2002 г. На фиг. 4 показано сопоставление запрашиваемой последовательности INSP002 с белком, происходящем от кДНК-клона AK095926.

Граница сплайсинга экзона 1 и экзона 2, предсказанная для INSP002, была экспериментально подтверждена присутствием AK095926.

```

450      . : . : . : . : . :
451  GATGTGTAAGGCTGTGCCCTTCGTT CAG      GTGTTCTCCCGGC
      |||||
10225  GATGTGTAAGGCTGTGCCCTTCGTT CAGGTG...CAGGTGTTCTCCCGGC
      Экзон1      Экзон2
500      . : . : . : . : . :
492  CCGGCTGCTCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTGC
      |||||
13670  CCGGCTGCTCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTGC

```

Исследования также показали, что части INSP002 идентичны IMAGE-клону 4558384 (BC025333.1), депонированному 8 марта 2002 г. На фиг. 5 показано сопоставление частей запрашиваемой последовательности INSP002 с IMAGE-клоном 4558384.

Пример 3. Частичное клонирование кДНК INSP002.

(i) Библиотеки кДНК.

Библиотеки человеческих кДНК (в векторах бактериофага лямбда (λ)) закупали у Stratagene или Clontech или получали в Научно-исследовательском институте фармакологии Сероно в векторах λ ZAP или λ GT10 в соответствии с протоколом производителей (Stratagene). ДНК бактериофага λ получали из лабораторных культур инфицированного штамма-хозяина E.coli с использованием системы очистки ДНК Wizard Lambda Preps в соответствии с инструкциями производителей (Promega Corporation, Madison WI). Список используемых библиотек и штаммов-хозяев приведен в табл. 1.

(ii) ПЦР предполагаемых кДНК, полученных из фаговой библиотеки ДНК.

Неполную кДНК, кодирующую INSP002 (фиг. 6), получали в виде продукта ПЦР-амплификации, состоящего из 159 п.н. (фиг. 7), с использованием геноспецифических клонирующих праймеров (INSP002-CP1 и INSP002-CP2, фиг. 6 и табл. II). ПЦР осуществляли в конечном объеме 50 мкл, содержащем IX буфер AmpliTaq™, 200 мкМ dNTP, 50 пмоль каждого клонирующего праймера, 2,5 единиц AmpliTaq™ (Perkin-Elmer) и 100 нг каждой фаговой библиотеки ДНК с использованием аппарата MJ Research DNA Engine с установленной программой, работающей в следующем режиме; 94°C, 1 мин; 40

циклов - 94°C, 1 мин, x °C и y мин и 72°C (где x означает наименьшую температуру Tm=5°C, и y=1 мин на т.п.н. продукта); затем 1 цикл - 72°C, 7 мин и цикл выдерживания при 4°C.

Продукты амплификации визуализировали на 0,8% агарозном геле в 1X буфере TAE (Life Technologies) и ПЦР-продукты, мигрирующие в предполагаемой молекулярной массе, выделяли из геля с использованием системы очистки ДНК Wizard PCR Preps DNA Purification System (Promega). ПЦР-продукты, элюированные в 50 мкл стерильной воды, либо непосредственно субклонировали, либо хранили при -20°C.

(iii) Геноспецифические клонирующие праймеры для ПЦР.

Пары ПЦР-праймеров, имеющих длину от 18 до 25 нуклеотидов, конструировали для амплификации полноразмерной последовательности предполагаемой кДНК с использованием программы Primer Designer Software (Scientific & Educational Software, PO Box 72045, Durham, NC 27722-2045, USA). ПЦР-праймеры оптимизировали так, чтобы Tm была близка к 55±10°C, и содержание GC составляло 40-60%. Отбирали праймеры, которые обнаруживали высокую селективность для последовательности-мишени INSP002 (низкий уровень или отсутствие неспецифического праймирования с другими матрицами).

(iv) Субклонирование ПЦР-продуктов.

ПЦР-продукты субклонировали в клонирующий вектор, модифицированный топоизомеразой I (pCRII-TOPO) с использованием набора для клонирования TOPO TA, закупленного у Invitrogen Corporation (cat. № K4600-01 и K3575-01 соответственно), и с использованием условий, указанных производителем. Вкратце, 4 мкл выделенного из геля ПЦР-продукта, полученного в результате амплификации библиотеки последовательностей почки плода человека (библиотека номер 12), инкубировали в течение 15 мин при комнатной температуре с 1 мкл вектора TOPO и 1 мкл солевого раствора. Затем реакционную смесь трансформировали в штамм TOP10 E.coli (Invitrogen) следующим образом: 50 мкл-аликвоту клеток One Shot TOP10 оттаивали на льду и добавляли 2 мкл реакционной смеси TOPO. Смесь инкубировали в течение 15 ч на льду и затем подвергали тепловому шоку путем инкубирования при 42°C в течение точно 30 с. Затем образцы снова помещали на лед и добавляли 250 мкл теплой среды SOC (при комнатной температуре). Образцы инкубировали со встряхиванием (220 об/мин) в течение 1 ч при 37°C. Затем смесь для трансформации высевали на чашки с L-бульоном (LB), содержащие ампициллин (100 мкг/мл), и инкубировали в течение ночи при 37°C. Резистентные к ампициллину колонии, содержащие кДНК-вставки, идентифицировали с помощью ПЦР колоний.

(v) ПЦР колоний.

Колонии инокулировали в 50 мкл стерильной воды с использованием стерильной зубочистки. Затем 10 мкл-аликвоту инокулята подвергали ПЦР в общем реакционном объеме 20 мкл, как описано выше, за исключением того, что в качестве пар праймеров использовали SP6 и T7. Циклы проводили в следующих условиях: 94°C, 2 мин, 30 циклов при 94°C, 30 с.; 47°C, 30 с. и 72°C, 1 мин; 1 цикл, 72°C, 7 мин. Затем перед проведением анализа образцы выдерживали при 4°C (цикл выдерживания).

Продукты ПЦР-реакции анализировали на 1% агарозных гелях в буфере 1 X TAE. Колонии, которые давали ПЦР-продукт предполагаемого размера (159 п.н. -кДНК + 187 п.н., что обусловлено присутствием сайта множественного клонирования или MCS), культивировали в течение ночи при 37°C в 5 мл L-бульона (LB), содержащего ампициллин (100 мкг/мл), со встряхиванием при 220 об/мин при 37°C.

(vi) Получение и секвенирование плазмидной ДНК.

Минипрепарат плазмидной ДНК получали из 5 мл культур с использованием роботизированной системы Qiaprep Turbo 9600 (Qiagen) или набора Wizard Plus SV Miniprep (Promega cat. № 1460) в соответствии с инструкциями производителей. Плазмидную ДНК элюировали в 100 мкл стерильной воды. Концентрацию ДНК измеряли с использованием фотометра Эппендорфа BO. Затем плазмидную ДНК (200-500 нг) подвергали ДНК-секвенированию с использованием праймера 17 и праймера SP6 и с использованием системы BigDye Terminator (Applied Biosystems cat. № 4390246) в соответствии с инструкциями производителей. Реакционные смеси для секвенирования очищали на колонках с Dye-Ex (Qiagen) или на планшетах для очистки Montage SEQ 96 (Millipore cat. № LSKS09624) и затем анализировали на секвенаторе Applied Biosystems 3700.

(vii) Идентификация кДНК-библиотек, содержащих INSP002.

ПЦР-продукты, полученные с использованием INSP002-CP1 и INSP002-CP2 и мигрирующие при "правильном" размере (159 п.н.), идентифицировали в библиотеках кДНК коры головного мозга, толстой кишки, легких плода и почек плода (библиотеки 8, 9, 11 и 12). Последовательность ПЦР-продукта, клонированного в вектор pCRII-TOPO, показана на фиг. 7 и карта плазмиды (плазмида ID 13422) - на фиг. 8. Неполная клонированная кДНК представляла собой часть экзона 2 INSP002, как было показано путем сопоставления предсказанной нуклеотидной последовательности INSP002 и клонированной неполной нуклеотидной последовательности на фиг. 9а, и сопоставления предсказанной последовательности белка INSP002 и клонированной неполной последовательности белка на фиг. 9б.

Пример 4. Генерирование ORF INSP002 из Image: 4558384.

Image-клон 4558384 (в плазмиде pOTB7), выделенный из ретинобластомы, закупали у Resgen (Invitrogen Corp.). Уколоточную культуру E.coli 4558384 засевали путем разбрасывания на LB-чашки, содержащие ампициллин (100 мкг/мл) и культивировали в течение ночи при 37°C. Отдельные резистентные к

ампициллину моноколонии инокулировали в 5 мл среды LB, содержащей ампициллин (100 мкг/мл), и инкубировали со встряхиванием при 220 об/мин в течение ночи при 37°C. Получали минипрепарат плазмидной ДНК, который секвенировали с использованием праймеров SP6, T7, M13F, INSP002-CP1 и INSP002-CP2, как описано в примере 3(vi).

Последовательность вставки показана на фиг. 10. Сопоставление нуклеотидной последовательности и предполагаемой аминокислотной последовательности, предсказанной, исходя из кДНК Image 4558384, с INSP002, показано на фиг. 11. Очевидно, что кДНК Image 4558384 представляет собой вариант сплайсинга INSP002. Этот вариант, по сравнению с предсказанной последовательностью INSP002, содержит 87 п.н.-вставку, которая вводит сдвиг рамки считывания и ранний стоп-кодон. Помимо 3'-нетранслируемой последовательности, он также содержит Alu-повтор, что указывает на контаминирующее включение геномной ДНК в кДНК. Экзоны 2 и 4 Image 4558384 эквивалентны экзонам 1 и 2 предсказанного INSP002. Однако Image 4558384 вводит дополнительный экзон между экзонами 1 и 2 предсказанного INSP002. Дополнительный экзон кодирует ранний стоп-кодон, что предотвращает трансляцию домена "цистиновый узел". Границы сплайсинга в Image 4558384 показаны ниже:

```

500      .   :   .   :   .   :   .   :   .   :   .   :
492 GGCTGTGCCCTTCGTTTCAG          ACACGGGAGTCTCGCTATGTTG
      |||||>>>...>>>|||||
10234 GGCTGTGCCCTTCGTTTCAGGTG...TAGACACGGGAGTCTCGCTATGTTG
      Экзон2                      Экзон3
550      .   :   .   :   .   :   .   :   .   :   .   :
533 CCCAAGCTAGTCTTGAGCTTCTGGCCTCAAGCAATCCTCCACCTCAGCC
      |||||>>>...>>>|||||
10741 CCCAAGCTAGTCTTGAGCTTCTGGCCTCAAGCAATCCTCCACCTCAGCC
600      .   :   .   :   .   :   .   :   .   :   .   :
583 TCC[ ]GTCTAG          GTGTTCTCCCGGCCCGGCTGCTCAGCCA
      |||>>>...>>>|||||
10791 TCC[ ]GTCTAGGTG...CAGGTGTTCTCCCGGCCCGGCTGCTCAGCCA
  
```

Для удаления геномной ДНК в целях генерирования полноразмерной кДНК, кодирующей ORF INSP002, рекомендуется использовать ПЦР-стратегию. ПЦР-праймеры были сконструированы для амплификации 5'-конца (выше) и 3'-конца (ниже) последовательности INSP002, которая фланкирует 87 п.н.-вставку в клоне Image. Обратный праймер для расположенной выше последовательности и прямой праймер для расположенной ниже последовательности содержали комплементарные последовательности у своих 3'- и 5'-концов соответственно для создания перекрывающихся концов так, чтобы ПЦР-продукты от каждой реакции могли смешиваться и подвергаться отжигу, что обеспечивало бы амплификацию полноразмерной кДНК в третьей ПЦР-реакции, проводимой с использованием присоединяемого выше "гнездового" прямого праймера и присоединяемого ниже "гнездового" обратного праймера.

Первая ПЦР-реакционная смесь, используемая для амплификации 5'-конца INSP002 (выше от 87 п.н.-вставки), содержала, в конечном объеме 50 мкл, 5 мкл 10X буфера Pfx Platinum, 1,5 мкл dNTP (10 мМ), 1 мкл MgSO₄ (50 мМ), 1,5 мкл INSP002V-5'-F (10 мкМ), 1,5 мкл INSP002V-5'-R (10 мкМ), 0,75 мкл Pfx Platinum и 135 нг плазмидной кДНК IMAGE:4558384. Амплификацию осуществляли в следующих условиях: 1 цикл при 94°C, 2 мин, 30 циклов при 94°C, 15 с и 68°C, 1 мин. и 1 цикл при 68°C, 7 мин. Вторую ПЦР-реакцию амплификации 3'-конца INSP002 (ниже от 87 п.н.-вставки) осуществляли в тех же самых условиях, за исключением того, что в качестве праймеров использовали INSP002V-3'-F и INSP002V-3'-R.

Продукты амплификации визуализировали на 0,8% агарозном геле в 1X буфере TAE (Invitrogen). ПЦР-продукты, мигрирующие в предполагаемой молекулярной массе (520 п.н. и 448 п.н. для ПЦР 1 и 2 соответственно) выделяли из геля с использованием системы очистки ДНК Wizard PCR Preps DNA Purification System (Promega). ПЦР-продукты элюировали в 50 мкл стерильной воды, и концентрацию ДНК измеряли с использованием фотометра Эппендорфа ВО. Затем 50 нг каждого очищенного ПЦР-продукта использовали в качестве матрицы для "гнездовой" ПЦР в 50 мкл реакционной смеси, содержащей 5 мкл 10X буфера Pfx Platinum, 1,5 мкл dNTP (10 мМ), 1 мкл MgSO₄ (50 мМ), 1,5 мкл INSP002V-5'-nest-F (10 мкМ), 1,5 мкл INSP002V-3'-nest-R (10 мкМ). Реакционную смесь нагревали при 95°C в течение 3 мин и добавляли 0,75 мкл полимеразы Pfx Platinum. Амплификацию осуществляли в следующих условиях: 1 цикл при 94°C, 2 мин, 30 циклов при 94°C, 15 с, 61°C, 30 с и 68°C 1 мин и 1 цикл 68°C, 7 мин. ПЦР-продукты, мигрирующие в предполагаемой молекулярной массе 719 п.н., выделяли из геля с использованием системы очистки ДНК Wizard PCR Preps DNA Purification System и элюировали в 50 мкл стерильной воды. Затем 4 мкл очищенного ПЦР-продукта лигировали в вектор pCR4blunt-TOPO, как описано в

разделе 1.4. Резистентные к ампициллину колонии тестировали на вставки с помощью ПЦР колоний с использованием праймеров T3 и T7, как описано в разделе 1.5. Колонии, которые давали ПЦР-продукт ожидаемого размера (719 п.н.+106 п.н., что обусловлено присутствием сайта множественного клонирования или MCS), культивировали в 5 мл LB-бульона, содержащего ампициллин (100 мкг/мл), со встряхиванием при 220 об/мин при 37°C. Минипрепарат плазмидной ДНК получали из 5 мл культур и секвенировали с использованием праймеров T3 и T7, как описано в разделе 1.6. Последовательность одного из полученных клонов и соответствующая плазмидная карта (pCR4blunt-TOPO-INSP002V) показаны на фиг. 12 и 13 соответственно. Трансляция клонированной последовательности показала, что INSP002V содержит 2 аминокислотных делеции ($\Delta V107$ и $\Delta Q108$) и одну аминокислотную замену (F110L) по сравнению с предсказанной последовательностью INSP002. Сопоставление нуклеотидной и аминокислотной последовательностей предсказанного INSP002 и INSP002V проиллюстрировано на фиг. 14.

При сравнении с предсказанным INSP002 было выявлено, что на границе сплайсинга между экзонами 1 и 2 используется расположенный выше донорный сайт длиной 6 п.н. Это приводит к делеции двух аминокислот (ValGlu). Используемый акцептор сплайсинга аналогичен акцептору, используемому в предсказанном INSP002, однако за этим акцептором имеются две ошибки, обнаруженные при секвенировании. Это приводит к аминокислотной замене (Phe → Leu).

Пример 5. Конструирование плазмиды для экспрессии INSP002V в клетках HEK293/EBNA.

Клон pCR4blunt-TOPO, содержащий полноразмерную кодирующую последовательность (ORF) INSP002V, идентифицированную путем секвенирования ДНК (фиг. 13), затем использовали для субклонирования вставки в вектор для экспрессии pEAK12d в клетках млекопитающего (фиг. 15) с использованием методики клонирования Gateway™ (Invitrogen).

(i) Генерирование Gateway-совместимой ORF INSP002, присоединенной к последовательности-метке 6HIS, с сохранением рамки считывания.

Первая стадия способа клонирования Gateway включает двухстадийную ПЦР-реакцию, которая генерирует ORF INSP002V, фланкированную у 5'-конца сайтом рекомбинации attB1 и последовательностью Козака, и у 3'-конца - последовательностью, кодирующей, с сохранением рамки считывания, 6-гистидиновую (6HIS) метку стоп-кодоном и сайтом рекомбинации attB2 (Gateway-совместимой кДНК). Первая ПЦР-реакционная смесь (в конечном объеме 50 мкл) содержала: 25 нг pCR4blunt-TOPO-INSP002V (плазмида 13075 и фиг. 13), 1,5 мкл dNTP (10 мМ), 5 мкл 10X полимеразного буфера Pfx, 1 мкл MgSO₄ (50 мМ), 0,5 мкл каждого геноспецифического праймера (100 мкМ) (INSP002V EX1 и INSP002V-EX2) и 0,5 мкл ДНК-полимеразы Pfx Platinum (Invitrogen). ПЦР-реакцию осуществляли с использованием начальной стадии денатурации при 95°C, 2 мин, с последующим проведением 12 циклов при 94°C, 15 с, и 68°C, 30 с. ПЦР-продукты очищали непосредственно из реакционной смеси с использованием системы очистки ДНК Wizard PCR Prep (Promega) в соответствии с инструкциями производителя. Вторая ПЦР-реакционная смесь (в конечном объеме 50 мкл) содержала: 10 мкл очищенного ПЦР-продукта, 1,5 мкл dNTP (5мМ), 1 мкл MgSO₄ (50 мМ), 5 мкл 10X полимеразного буфера Pfx Platinum, 0,5 мкл каждого праймера для конверсии Gateway (100 мкМ) (прямого праймера GCP и обратного праймера GCP) и 0,5 мкл ДНК-полимеразы Pfx Platinum. Вторую ПЦР-реакцию осуществляли при следующих условиях: 95°C, 1 мин., 4 цикла при 94°C, 15 с; 45°C, 30 с и 68°C, 3,5 мин; 25 циклов при 94°C, 15 с; 55°C, 30 с и 68°C, 3,5 мин. ПЦР-продукты очищали, как описано выше.

(ii) Субклонирование Gateway-совместимой ORF INSP002V во встраивающий вектор pDONR201 Gateway и экспрессирующий вектор pEAK12d.

Вторая стадия способа клонирования Gateway включает субклонирование Gateway-модифицированного ПЦР-продукта во встраивающий вектор pDONR201 Gateway (Invitrogen, фиг. 16), которое осуществляли следующим образом: 5 мкл очищенного ПЦР-продукта инкубировали с 1,5 мкл вектора pDONR201 (0,1 мкг/мкл), 2 мкл буфера BP и 1,5 мкл клоназной ферментной смеси BP (Invitrogen) при комнатной температуре в течение 1 ч. Реакцию прекращали добавлением протеиназы К (2 мкг) и инкубировали при 37°C еще 10 мин. Аликвоту реакционной смеси (2 мкл) трансформировали в клетки DH10B E.coli путем электропорации с использованием импульсного устройства Biorad Gene Pulser. Трансформанты высевали на LB-планшеты с канамицином. Плазмидный минипрепарат ДНК (Mini-prep) приготавливали из 1-4 полученных колоний с использованием набора Wizard Plus SV Minipreps (Promega), и затем 1,5 мкл плазмидного элюата использовали в реакционной смеси для рекомбинации, содержащей 1,5 мкл вектора pEAK12d (фиг. 9) (0,1 мкг/мл), 2 мкл буфера LR и 1,5 мкл клоназы LR (Invitrogen) в конечном объеме 10 мкл. Смесь инкубировали при комнатной температуре в течение 1 ч, а затем реакцию прекращали добавлением протеиназы К (2 мкг) и смесь инкубировали при 37°C в течение еще 10 мин. Аликвоту этой реакционной смеси (1 мкл) использовали для трансформации клеток DH10B E.coli путем электропорации.

Клоны, содержащие "правильную" вставку, идентифицировали путем проведения ПЦР колоний, как описано выше, за исключением того, что для ПЦР использовали праймеры pEAK12d (pEAK12d F и pEAK12d R). Плазмидный минипрепарат ДНК выделяли из клонов, содержащих "правильную" вставку, с использованием роботизированной системы Qiagrep Turbo 9600 (Qiagen) или вручную с использованием набора Wizard Plus SV Miniprep (Promega), и последовательность подтверждали с использованием прай-

меров pEAK12d F и pEAK12d R.

Очищенный в градиенте CsCl максипрепарат ДНК плазмиды pEAK12d-INSP002V-6HIS (плазмида ID номер 13227, фиг. 17) получали из 500 мл культуры клонов с подтвержденной последовательностью (Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press), ресуспендировали при концентрации 1 мкг/мл в стерильной воде и хранили при -20°C.

(iii) Конструирование экспрессирующего вектора pEAK12d.

Вектор pEAK12d представляет собой совместимый с системой клонирования Gateway вариант вектора pEAK12 для экспрессии в клетках млекопитающих (закупленного у Edge Biosystems), в котором нужная кДНК экспрессируется под контролем человеческого промотора EF1 α . Вектор pEAK12d генерировали, как описано ниже.

pEAK12 расщепляли рестриктирующими ферментами HindIII и NotI, затупляли по концам фрагментом Кленова (New England Biolabs) и дефосфорилировали с использованием щелочной фосфатазы кишечника теленка (Roche). После дефосфорилирования вектор лигировали с сохранением рамки считывания с затупленным по концам кластером C Gateway (система конверсии вектора Gateway, Invitrogen, cat № 11828-019), который содержит сайты рекомбинации AttR, фланкирующие ген ccdB и ген резистентности к хлорамфениколу и трансформировали в клетки DB3.1 E.coli (которые обеспечивают размножение векторов, содержащих ген ccdB). Минипрепарат ДНК выделяли из нескольких полученных колоний с использованием набора Wizard Plus SV Miniprep (Promega) и расщепляли ферментами AseI/EcoRI для идентификации клонов с получением фрагмента длиной 670 п.н., что указывало на то, что этот кластер был встроен в "правильной" ориентации. Полученная плазида была обозначена pEAK12d (фиг. 15).

Пример 6. Экспрессия INSP002-SV-6HIS-V1 в клетках млекопитающих (плазмида № 13227) и очистка.

Клетки почек эмбриона человека 293, экспрессирующие ядерный антиген вируса Эпштейна Барра (HEK293-EBNA, Invitrogen), поддерживали в суспензии бессывороточной среды Ex-клеток VPRO (посевной материал, поддерживающая среда, JRH). За 16-20 ч до трансфекции (день 1), клетки высевали во флаконы с 2х T225 (50 мл на флакон в среде DMEM/F12 (1:1)), содержащие посевную среду с 2% FBS (JRH) при плотности 2х10⁵ клеток/мл. На следующий день (день трансфекции 0) осуществляли трансфекцию с использованием реагента JetPEITM (2 мкл/мкг плазмидной ДНК, PolyPlus-трансфекция). Для каждого флакона, 113 мкг ДНК (№ 13227) ко-трансфекциировали 2,3 мкг GFP (флуоресцентного репортерного гена). Затем смесь для трансфекции добавляли во флаконы с 2х T225 и инкубировали при 37°C (5% CO₂) в течение 6 дней. Для повышения шансов получить большее количество материала, эту процедуру повторяли в еще в двух колбах, в результате чего получали всего 200 мл продукта. Подтверждение позитивной трансфекции проводили путем качественной оценки флуоресценции на день 1 и на день 6 (Axiovert 10 Zeiss).

На 6-й день (день сбора), супернатанты (200 мл) из четырех флаконов объединяли и центрифугировали (4°C, 400 г), затем помещали в сосуд, содержащий уникальный идентификатор.

Одну аликвоту (500 мкл) оставляли для проведения количественной оценки (QC) 6xHis-меченного белка (QC внутреннего биопроцессинга).

Способ очистки

200 мл образца культуральной среды, содержащего рекомбинантный белок с C-концевой меткой 6His, разводили холодным буфером А (50 mM NaH₂PO₄; 600мM NaCl; 8,7 % (мас/об) глицерина, pH 7,5) до конечного объема 400 мл. Образец фильтровали через стерильный фильтр 0,22 мкм (Millipore, 500 мл-фильтровальное устройство) и выдерживали при 4°C в стерильном квадратном 500 мл-сосуде со средой (Nalgene).

Очистку осуществляли при 4°C на рабочей станции VISION (Applied Biosystems), подсоединенной к автоматическому устройству для загрузки образцов (Labomatic). Процедура очистки состояла из двух последовательных стадий, т.е., сначала проводили металлоаффинную хроматографию на колонке с Poros 20 MC (Applied Biosystems), загруженной ионами Ni (4,6х50 мм, 0,83 мл), затем проводили гелефильтрацию на колонке (1,0х10 см) со средой, содержащей сефадекс G-25 (Amersham Pharmacia).

Для проведения первой стадии хроматографии, металлоаффинную колонку регенерировали 30 колоночными объемами раствора EDTA (100 mM EDTA; 1M NaCl; pH 8,0), снова загружали ионы металла путем промывки 15 колоночными объемами 100 mM раствора NiSO₄, после чего промывали 10 колоночными объемами буфера А, затем промывали 7 колоночными объемами буфера В (50 mM NaH₂PO₄; 600 mM NaCl; 8,7 % (мас/об) глицерин, 400 mM имидазол; pH 7,5) и, наконец, уравнивали 15 колоночными объемами буфера А, содержащего 15 mM имидазола. Образец загружали на аффинную колонку с металлическим Ni партиями по 200 мл. 200 мл образца переносили с помощью устройства для загрузки образцов Labomatic в 200 мл-петлю для образца и затем загружали на аффинную колонку с металлическим Ni при скорости потока 10 мл/мин. Для загрузки 400 мл образца на колонку процедуру переноса и загрузки повторяли еще раз. По окончании процедуры загрузки колонку промывали 12 колоночными объемами буфера А, затем 28 колоночными объемами буфера А, содержащего 20 mM имидазола. Во время промывки 20 mM имидазолом, высвобождаемые примесные белки элюировались с колонки. И нако-

нец, рекомбинантный His-меченый белок элюировали 10 колоночными объемами буфера В при скорости потока 2 мл/мин и элюированный белок собирали в 1,6 мл-фракции.

Для проведения второй стадии хроматографии, гель-фильтрационную колонку с сефадексом G-25 регенерировали 2 мл буфера D (1,137 М NaCl; 2,7 мМ KCl; 1,5 мМ KH_2PO_4 ; 8 мМ Na_2HPO_4 ; pH 7,2) и затем уравнивали 4 колонками буфера С (137 мМ NaCl, 2,7 мМ KCl; 1,5 мМ KH_2PO_4 ; 8 мМ Na_2HPO_4 ; 20% (мас/об) глицерина; pH 7,4). Фракции пиков, элюированные с Ni-колонки, автоматически пропускали через многофункциональное устройство для загрузки образцов, подсоединенное к рабочей станции VISION с колонки, загружали в колонку с сефадексом G-25 и затем белок элюировали буфером С при скорости потока 2 мл/мин. Обессоленный образец собирали в 2,2 мл-фракцию. Эту фракцию фильтровали через стерильный центрифужный фильтр 0,22 мкм (Millipore), разделяли на аликвоты, замораживали и хранили при -80°C . Аликвоту данного образца анализировали с помощью электрофореза в ДСН-ПААГ (в 4-12% геле NuPAGE; Novex) путем проведения Вестерн-блоттинга с использованием анти-His антител.

После электрофореза белки подвергали электрофоретическому переносу из геля на нитроцеллюлозную мембрану при 290 мА в течение 1 ч при 4°C . Мембрану блокировали 5% сухим молоком в буфере Е (137 мМ NaCl; 2,7 мМ KCl; 1,5 мМ KH_2PO_4 ; 8 мМ Na_2HPO_4 ; 0,1 % твина 20, pH 7,4) в течение 1 ч при комнатной температуре и затем инкубировали со смесью 2 кроличьих поликлональных анти-His антител (G-18 и H-15, 0,2 мкг/мл каждого; Santa Cruz) в 2,5% сухом молоке в буфере Е в течение ночи при 4°C . После инкубирования в течение еще 1 ч при комнатной температуре, мембрану промывали буфером Е (3x10 мин) и затем инкубировали со "вторым" ПХ-конъюгированным антикроличьим антителом (ДАКО, HRP 0399), разведенным 1/3000 в буфере Е, содержащем 2,5% сухое молоко, в течение 2 ч при комнатной температуре. После промывки буфером Е (3x10 мин), мембрану проявляли в течение 1 мин с помощью ECL-набора (Amersham Pharmacia). Затем мембрану экспонировали с пленкой Hyperfilm (Amersham Pharmacia), пленку проявляли и снимок вестерн-блота визуально анализировали.

Пример 7. Клонирование полноразмерного INSP002 из сердца.

Полноразмерную кодирующую последовательность INSP002 клонировали из кДНК сердца, как описано ниже.

(i) Генерирование кДНК-матрицы сердца.

Полноразмерную РНК сердца человека закупали у Clontech. Качество и концентрацию РНК анализировали с использованием биоанализатора Agilent 2100.

Реакционная смесь для синтеза кДНК содержала 1 мкл oligo(dT)₁₅ праймера (500 мкг/мл, Promega cat. № C1101), 2 мкг полноразмерной РНК и 1 мкл dNTP (10 мМ) в объеме 12 мкл. Полученную смесь нагревали при 65°C в течение 5 мин и затем охлаждали на льду. Затем добавляли следующие реагенты: 4 мкл 5X буфера для первой цепи, 2 мкл DTT (0,1 М), 1 мкл ингибитора рекомбинантной рибонуклеазы RNaseOut (40 единиц/мкл, Promega, cat. № 2511) и смесь инкубировали при 42°C в течение 2 ч, после чего добавляли 1 мкл (200 единиц) Superscript II (Invitrogen cat. № 18064-014). Смесь инкубировали при 42°C в течение 50 мин и затем нагревали при 70°C в течение 15 мин. Для удаления матричной РНК добавляли 1 мкл (2 единицы) РНКазы Н E.coli (Invitrogen, cat. № 18021-014) и реакционную смесь инкубировали еще 20 мин при 37°C . Конечную реакционную смесь разводили 200 мкл стерильной воды и хранили при -80°C .

(ii) Клонирование полноразмерной кодирующей последовательности INSP002 с помощью ПЦР.

Полноразмерную кодирующую последовательность INSP002 клонировали из кДНК сердца человека с помощью ПЦР в 50 мкл реакционной ПЦР-смеси, содержащей 5 мкл кДНК сердца, 5 мкл 10X буфера Pfx, 1,5 мкл dNTP (10 мМ), 1 мкл MgSO_4 (50 мМ), 1,5 мкл геноспецифического прямого праймера INSP002-FL-F (10 мкл), 1,5 мкл геноспецифического обратного праймера INSP002-FL-R (10 мкл) и 0,5 мкл ДНК-полимеразы Pfx Platinum (Invitrogen). Циклы проводили в следующих условиях: 1 цикл при 94°C , 4 мин, 35 циклов при 94°C , 15 с, при 55°C , 30 с, при 68°C , 1 мин и 1 цикл при 68°C , 10 мин, затем цикл выдерживания при 4°C .

Продукты амплификации визуализировали на 0,8% агарозном геле в 1X буфере TAE (Invitrogen) и ПЦР-продукты, мигрирующие в предполагаемой молекулярной массе (589 п.н.), выделяли из геля с использованием набора для экстракции в геле Qiagen MinElute (Qiagen). ПЦР-продукты элюировали в 50 мкл 10 мМ Трис-НCl, pH 8,5 и субклонировали в вектор pCR4blunt-TOPO, как описано выше (раздел 1.4). Несколько резистентных к ампициллину колоний подвергали ПЦР-реакции, как описано в разделе 1.5. Колонии, содержащие вставку "правильного" размера (589 п.н.+106 п.н., что обусловлено присутствием MCS), культивировали в течение ночи при 37°C в 5 мл L-бульона (LB), содержащего ампициллин (100 мкг/мл), со встряхиванием при 220 об/мин при 37°C . Минипрепарат плазмидной ДНК получали из 5 мл культур с использованием роботизированной системы Qiaprep Turbo 9600 (Qiagen) или с использованием набора Wizard Plus SV Miniprep (Promega cat. № 1460) в соответствии с инструкциями производителей, и затем 200-500 нг минипрепарата ДНК секвенировали, как описано в разделе 1.6 с использованием праймеров T3 и T7 (табл. III). Клонированная последовательность представлена на фиг. 18. Карта полученной плазмиды, pCR4blunt-TOPO-INSP002FL (плазмида ID № 13514) показана на фиг. 19.

Таблица I

Библиотеки человеческих кДНК

Библиотек а	Источник ткани/клетки	Вектор	Штамм- хозяин	Поставщи к	№ в каталоге
1	Головной мозг плода человека	Zap II	XL1- Blue MRF'	Stratage n	936206
2	Яичник человека	GT10	LE392	Clontech	HL1098a
3	Гипофиз человека	GT10	LE392	Clontech	HL1097a
4	Плацента человека	GT11	LE392	Clontech	HL1075b
5	Яички человека	GT11	LE392	Clontech	HL1010b
6	Черное вещество человека	GT10	LE392	Получено авторами	
7	Головной мозг плода человека	GT10	LE392	Получено авторами	
8	Кора головного мозга человека	GT10	LE392	Получено авторами	
9	Толстая кишка человека	GT10	LE392	Clontech	HL1034a
10	Головной мозг плода человека	GT10	LE392	Clontech	HL1065a
11	Легкие плода человека	GT10	LE392	Clontech	HL1072a
12	Почки плода человека	GT10	LE392	Clontech	HL1071a
13	Печень плода человека	GT10	LE392	Clontech	HL1064a
14	Костный мозг человека	GT10	LE392	Clontech	HL1058a
15	Моноциты периферической крови человека	GT10	LE392	Clontech	HL1050a
16	Плацента человека	GT10	LE392	Получено авторами	
17	SHSY5Y человека	GT10	LE392	Получено авторами	
18	Человеческая клеточная линия U373	GT10	LE392	Получено авторами	
19	Человеческая клеточная линия CFRoc-1	Uni Zap	XL1- Blue MRF'	Stratage n	936206
20	Сетчатка глаза человека	GT10	LE392	Clontech	HL1132a
21	Мочевой пузырь человека	GT10	LE392	Получено авторами	
22	Человеческие тромбоциты	Uni Zap	XL1- Blue MRF'	Получено авторами	
23	Человеческая нейробластома Kan+TS	GT10	LE392	Получено авторами	
24	Гладкая мышца бронхов человека	GT10	LE392	Получено авторами	
25	Гладкая мышца бронхов человека	GT10	LE392	Получено авторами	
26	Тимус человека	GT10	LE392	Clontech	HL1127a
27	5'-фрагмент селезенки человека	GT11	LE392	Clontech	HL1134b
28	Моноциты периферической крови человека	GT10	LE392	Clontech	HL1050a
29	Яички человека	GT10	LE392	Clontech	HL1065a
30	Головной мозг плода человека	GT10	LE392	Clontech	HL1065a
31	Черное вещество человека	GT10	LE392	Clontech	HL1093a
32	Плацента человека #11	GT11	LE392	Clontech	HL1075a
33	Головной мозг плода человека	GT10	LE392	Clontech	Сделано по заказу
34	Плацента человека #59	GT10	LE392	Clontech	HL5014a
35	Гипофиз человека	GT10	LE392	Clontech	HL1097a
36	Поджелудочная железа человека #63	Uni Zap XR	XL1- Blue MRF'	Stratage n	937208
37	Плацента человека #19	GT11	LE392	Clontech	HL1008
38	5'-фрагмент печени человека	GT11	LE392	Получено авторами	HL1115b

39	Матка человека	Zap-CMV XR	XL1-Blue MRF'	Stratagen	980207
40	кДНК-библиотека крупных вставок почки человека	TriplEx2	XL1-Blue	Получено авторами	HL5507u

Таблица II

Праймеры для клонирования INSP002

Праймер	Последовательность (5'→3')
INSP002-CP1	CTC AGC CAT ACG CCT CCG AA
INSP002-CP2	GCT GAG CTG CCA GTG AGA CA
INSP002V-5'-F	ACC TGG AAG GAA GCG ACT GCA CTG A
INSP002V-5'-R	GCA GCC GGG CCG GGA GAG AAC GAA GGG CAC AGC CTT A
INSP002V-3'-F	AGG CTG TGC CCT TCG TTC TCT CCC GGC CCG GCT GCT C
INSP002V-3'-R	ACT CCA GGA CGG GCA CTG TGT CTA C
INSP002V-5'-nest-F	GTC GAC TGC TAG TGA CCT TGA G
INSP002V-3'-nest-R	ACA TCA TCC AGG TCC ACG TCT T

Таблица III

Праймеры для субклонирования и секвенирования INSP002

Праймер	Последовательность (5'→3')
SP6	ATT TAG GTG ACA CTA TAG
T7	TAA TAC GAC TCA CTA TAG GG
T3	ATT AAC CCT CAC TAA AGG GA
M13F	TGT AAA ACG ACG GCC AGT
pEAK12-F	GCC AGC TTG GCA CTT GAT GT
pEAK12-R	GAT GGA GGT GGA CGT GTC AG
INSP002V-EX1	AA GCA GGC TTC <u>GCC ACC</u> ATG CTC CTT GGC CAG CTA TC
INSP002VEX2	<i>GTG ATG GTG ATG GTG</i> TGC TTT TGG GCT GCA GTG AC
GCP Прямой	G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTC <u>GCC ACC</u>
GCP Обратный	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTT TCA ATG <i>GTG ATG GTG ATG GTG</i>

Таблица IV

ПЦР-праймеры для клонирования полноразмерного INSP002

Праймер	Последовательность (5'→3')
INSP002V-FL-F	GAT GCT CCT TGG CCA GCT AT
INSP002V-FL-R	CCA TCC ACG ATG CTC AGT TC

Подчеркнутая последовательность = последовательность Козака

Последовательность, обозначенная жирным шрифтом = стоп-кодон

Последовательность, обозначенная курсивом = метка His

Список последовательностей

Примечание: для аминокислот, кодируемых участком стыка "экзон-экзон", аминокислота будет дополнительно приписана 5'-экзону

SEQ ID NO:1 (INSP002-нуклеотидная последовательность, экзон 1)

```

1 AAATGCCTCC CAGGCTATCC AGGAGGGGCC AAGAGATTAA AAGCAGGTT
51 AGAAGGCTCA GATGCCACTC ACCAGACAGC AGGGTCGACT GCTAGTGACC
101 TTGAGCCAG TCCGGACAGA CAGACAGGCA GACAGACGCA CGGACAAGCA
151 GATGCTCCTT GGCCAGCTAT CCACTCTTCT GTGCCTGCTT AGCGGGGCC
201 TGCTACAGG CTCAGGGAGG CCTGAACCCC AGTCTCCTCG ACCTCAGTCC
251 TGGGCTGCAG CCAATCAGAC CTGGGCTCTG GGCCAGGGG CCTGCCCCC
301 ACTGGTGCCA GCTTCTGCCC TTGGGAGCTG GAAGGCCTTC TTGGGCTGC
351 AGAAAGCCAG GCAGCTGGGG ATGGGCAGGC TGCAGCGTGG GCAGACGAG
401 GTGGCTGCTG TGA CTCTGCC GCTGAACCT CAGGAAGTGA TCCAGGGGAT
451 GTGAAGGCT GTGCCCTCG TTCAG

```

SEQ ID NO:2 (INSP002-последовательности белка, экзон 1)

```

1 MLLGQLSTLL CLLSGALPTG SGRPEQSPR PQSWAANQT WALPGALPP
51 LVPASALGSW KAFLGLQKAR QLGMRRLQRG QDEVAAVTLP LNPQEVIGM
101 CRAVPFVQ

```

SEQ ID NO:3 (INSP002-нуклеотидная последовательность, экзон 2)

```

1 GTGTTCTCCC GGGCGGGCTG CTCAGCCATA CGCTCCGAA ATCATCTGTG
51 CTTTGGTCAT TGCTCCTCTC TCTACATCCC TGGCTCGAC CCCACCCAC
101 TAGTCTGTG CAACAGCTGT ATGCTGCTC GCAAGCGTTG GGCACCCGTG
151 GTCCTGTGGT GTCTCACTGG CAGCTCAGCC TCCCGTCGAC GGGTGAAGAT
201 ATCCACCATG CTGATCGAGG GGTGTCACTG CAGCCCAAAA GCATGA

```

SEQ ID NO:4 (INSP002-последовательности белка, экзон 2)

```

1 VFSRPGCSAI RLNNHLCFGH CSSLYIPGSD PTPLVLCNSC MPARKRWAPV
51 VLWCLTGSSA SRRRVKISTM LIEGCHCSPK A

```

SEQ ID NO:5 (INSP002-нуклеотидная последовательность)

```

1 AAATGCCTCC CAGGCTATCC AGGAGGGGCC AAGAGATTAA AAGCAGGTT
51 AGAAGGCTCA GATGCCACTC ACCAGACAGC AGGGTCGACT GCTAGTGACC
101 TTGAGCCAG TCCGGACAGA CAGACAGGCA GACAGACGCA CGGACAAGCA
151 GATGCTCCTT GGCCAGCTAT CCACTCTTCT GTGCCTGCTT AGCGGGGCC
201 TGCTACAGG CTCAGGGAGG CCTGAACCCC AGTCTCCTCG ACCTCAGTCC
251 TGGGCTGCAG CCAATCAGAC CTGGGCTCTG GGCCAGGGG CCTGCCCCC
301 ACTGGTGCCA GCTTCTGCCC TTGGGAGCTG GAAGGCCTTC TTGGGCTGC
351 AGAAAGCCAG GCAGCTGGGG ATGGGCAGGC TGCAGCGTGG GCAAGACGAG
401 GTGGCTGCTG TGA CTCTGCC GCTGAACCT CAGGAAGTGA TCCAGGGGAT
451 GTGAAGGCT GTGCCCTTCG TTCAGGTGTT CTCCGGGCC GGCTGCTCAG
501 CCATACGCCT CCGAAATCAT CTGTGCTTTG GTCATTGCTC CTCTCTTAC
551 ATCCCTGGCT CGGACCCAC CCCACTAGTC CTGTGCAACA GCTGTATGCC
601 TGCTCGCAAG CGTTGGGCAC CCGTGGTCTT GTGGTGTCTC ACTGGCAGCT
651 CAGCTCCCG TCGACGGGTG AAGATATCCA CCATGCTGAT CGAGGGGTGT
701 CACTGCAGCC CAAAAGCATG A

```

SEQ ID NO:6 (INSP002-последовательность белка)

```

1 MLLGQLSTLL CLLSGALPTG SGRPEQSPR PQSWAANQT WALPGALPP
51 LVPASALGSW KAFLGLQKAR QLGMRRLQRG QDEVAAVTLP LNPQEVIGM

```

101 CKAVPFVQVF SRPGCSAIRL RNHLFCGHCS SLYIPGSDPT PLVLCNSCMP
151 ARKRWAPVVL WCLTGSSASR RRVKISTMLI EGCHCSPKA

SEQ ID NO:7 (INSP002-последовательности белка без сигнального пептида, экзон 1)

1 RPEQSPRPQ SWAAANQWA LGPGALPPLV PASALGSWKA FLGLQKARQL
51 GMGRLOQGD EVAAVTLPLN PQEVIQGMCK AVPFVQ

SEQ ID NO:8 (INSP002-последовательность белка без сигнального пептида)

1 RPEQSPRPQ SWAAANQWA LGPGALPPLV PASALGSWKA FLGLQKARQL
51 GMGRLOQGD EVAAVTLPLN PQEVIQGMCK AVPFVQVFSR PGCSAIRLRN
101 HLCFGHCSL YIPGSDPTPL VLCNSCHPAR KRWAPVVLWC LTGSSASRRR
151 VKISTMLIEG CHCSPKA

SEQ ID NO:9 (Нуклеотидная последовательность, кодирующая последовательность белка INSP002 без сигнального пептида-зрелый полипептид INSP002)

1 AGGCCCTGAAC CCCAGTCTCC TCGACCTCAG TCCTGGGCTG CAGCCAATCA GACCTGGGCT
61 CTGGGCCAG GGGCCCTGCC CCCACTGGTG CCAGCTTCTG CCCTTGGGAG CTGGAAGGCC
121 TTCTTGGGCC TGCAGAAAGC CAGGCAGCTG GGGATGGGCA GGCTGCAGCG TGGGCAAGAC
181 GAGGTGGCTG CTGTGACTCT GCCGCTGAAC CCTCAGGAAG TGATCCAGGG GATGTGTAAG
241 GCTGTGCCCT TCGTTCAGGT GTTCTCCCGG CCCGGCTGCT CAGCCATACG CCTCCGAAAT
301 CATCTGTGCT TTGGTCATTG CTCCTCTCTC TACATCCCTG GCTCGGACCC CACCCCACTA
361 GTCTGTGCA ACAGCTGTAT GCCTGCTCGC AAGCGTTGGG CACCCGTGGT CCTGTGGTGT
421 CTCACCTGCA GCTCAGCCTC CCGTCGACGG GTGAAGATAT CCACCATGCT GATCGAGGGG
461 TGTCACTGCA GCCCAAAGC ATGA

SEQ ID NO:10 (Нуклеотидная последовательность экзона 1, кодирующего белок INSP002 без сигнального пептида-экзон 1 зрелого полипептида INSP002)

1 AGGCCCTGAAC CCCAGTCTCC TCGACCTCAG TCCTGGGCTG CAGCCAATCA GACCTGGGCT
61 CTGGGCCAG GGGCCCTGCC CCCACTGGTG CCAGCTTCTG CCCTTGGGAG CTGGAAGGCC
121 TTCTTGGGCC TGCAGAAAGC CAGGCAGCTG GGGATGGGCA GGCTGCAGCG TGGGCAAGAC
181 GAGGTGGCTG CTGTGACTCT GCCGCTGAAC CCTCAGGAAG TGATCCAGGG GATGTGTAAG
241 GCTGTGCCCT TCGTTCAG

SEQ ID NO:11 (Нуклеотидная последовательность, кодирующая белок INSP002 без 5'-нетранслируемой области)

1 ATGCTCCTT GGCCAGCTAT CCACTCTTCT GTGCTGCTT AGCGGGGCCCT
51 GCCTACAGG CTCAGGGAGG CCTGAACCCC AGTCTCCTCG ACCTCAGTCT
101 GGGCTGCAG CCAATCAGAC CTGGGCTCTG GGGCCAGGGG CCCTGCCCCCA
151 CTGGTGCCA GCTTCTGCCC TTGGGAGCTG GAAGGCCTTC TTGGGCTGCA
201 GAAAGCCAG GCAGCTGGGG ATGGGCAGGC TGCAGCGTGG GCAAGACGAGG
251 TGCTGCTG TGACTCTGCC GCTGAACCCCT CAGGAAGTGA TCCAGGGGATG
301 TGTAAAGCT GTGCCCTTCG TTCAGGTGTT CTCCTGGGCC GGCTGCTCAGC
351 CATACGCT CCGAAATCAT CTGTCTTTG GTCATTTGCT CTCTCTTACA
401 TCCTGGCT CGGACCCAC CCCACTAGTC CTGTGCAACA GCTGTATGCT
451 GCTCGCAAG CGTTGGGCAC CCGTGGTCTT GTGGTGTCTC ACTGGCAGTC
501 AGCCTCCG TCGACGGGTG AAGATATCCA CCATGCTGAT CGAGGGGTGTC
551 ACTGCAGCC CAAAGCATG A

SEQ ID NO:12 (Нуклеотидная последовательность экзона 1, кодирующего белок INSP002 без 5'-нетранслируемой области)

1 ATGCTCCTT GGCCAGCTAT CCACTCTTCT GTGCTGCTT AGCGGGGCCCT
51 GCCTACAGG CTCAGGGAGG CCTGAACCCC AGTCTCCTCG ACCTCAGTCT
101 GGGCTGCAG CCAATCAGAC CTGGGCTCTG GGGCCAGGGG CCCTGCCCCCA
151 CTGGTGCCA GCTTCTGCCC TTGGGAGCTG GAAGGCCTTC TTGGGCTGCA

201 GAAAGCCAG GCAGCTGGGG ATGGGCAGGC TGCAGCGTGG GCAAGACGAGG
 251 TGGCTGCTG TGA CTCTGCC GCTGAACCCT CAGGAAGTGA TCCAGGGGATG
 301 TGTAAAGCT GTGCCCTTCG TTCAG

SEQ ID NO:13 (Нуклеотидная последовательность, кодирующая
 вариант полипептида INSP002)

1 GTCGACTGCT AGTGACCTTG AGCCCACTCC GGACAGACAG ACAGGCAGAC AGACGCACGG
 61 ACAAGCAGAT GCTCCTTGGC CAGCTATCCA CTCTTCTGTG CCTGCTTAGC GGGGCCCTGC
 121 CTACAGGCTC AGGGAGGCCT GAACCCCACT CTCTCGACC TCAGTCTTGG GCTGCAGCCA
 181 ATCAGACCTG GGCTCTGGGC CCAGGGGCC TCCCCCACT GGTGCCAGCT TCTGCCCTTG
 241 GGAGCTGGAA GGCTTCTTG GGCTGCAGA AAGCCAGGCA GCTGGGGATG GGCAGGCTGC
 301 AGCGTGGGCA AGACGAGGTG GCTGCTGTGA CTCTGCCGCT GAACCTCAG GAAGTGATCC
 361 AGGGGATGTG TAAGGCTGTG CCCTTGTTC TCTCCCGGC CGGCTGTCA GCCATAGCC
 421 TCCGAAATCA TCTGTCTTT GGTCTGTCT CCTCTCTTA CATCCCTGGC TCGGACCCCA
 481 CCCCCTAGT CCTGTGCAAC AGCTGTATGC CTGCTGCAA GCGTTGGGCA CCCGTGGTCC
 541 TGTGCTGTCT CACTGGCAGC TCAGCCTCCC GTCGACGGGT GAAGATATCC ACCATGCTGA
 601 TCGAGGGGTG TCACTGCAGC CCAAAAGCAT GAACTGAGCA TCTGGATGGG TGCACGGAGA
 661 CACGCACCTT GGAGAAATGA GGGGAGATGG ACCAAGAAAG ACGTGGACCT GGATGATGT

SEQ ID NO:14 (Вариант полипептида INSP002)

1 MLLGQLSTLL CLLSGALPTG SGRPEPQSPR PQSWAANOT WALGPGALPP LVPASALGSW
 61 KAFGLQKAR QLGMRLLRG QDEVAAVTLP LNPQEVIOGH CRAVPFVLSR PGCSAIRLRN
 121 HLCFGHCSSL YIPGSDPTPL VLCNSCMPAR KRWAPVVLWC LTGSSASRRR VKISTMLIEG
 161 CRCSPKA

SEQ ID NO:15 (Нуклеотидная последовательность, кодирующая
 вариант полипептида INSP002 без 5'-нетранслируемой области)

1 ATGCTCCTTG GCCAGTATC CACTCTTCTG TGCTGCTTA GCGGGGCCCT GCTACAGGC
 61 TCAGGGAGGC CTGAACCCCA GTCTCCTCGA CTCAGTCTCT GGGCTGCAGC CAATCAGACC
 121 TGGGCTCTGG GCCCAGGGGC CTGCCCCCA CTGGTGCCAG CTCTGCCCT TGGGAGCTGG
 181 AAGGCCCTTCT TGGGCTGCA GAAAGCCAGG CAGCTGGGGA TGGCAGGCT GCAGCGTGGG
 241 CAAGACGAGG TGGCTGCTGT GACTTGCCG CTGAACCTC AGGAAGTGAT CCAGGGGATG
 301 TGTAAAGCTG TGCCCTTCGT TCTCTCCCG CCCGGCTGCT CAGCCATAGC CCTCCGAAT
 361 CATCTGTGCT TTGGTCATTG CTCCTCTCTC TACATCCCTG GCTCGGACCC CACCCCACTA
 421 GTCTGTGCA ACAGCTGTAT GCTGTCTCGC AAGCGTTGGG CACCGTGGT CCTGTGCTGT
 481 CTCACTGCA GCTCAGCCTC CCGTGCAGG GTGAAGATAT CCACCATGCT GATCGAGGGG
 541 TGCACTGCA GCCCAAAGC A

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Полипептид с функцией секреторного белка, являющегося членом суперсемейства цитокинов, имеющих в своей структуре цистеиновые узлы, предпочтительно с функцией члена подсемейства DAN, где полипептид:

(i) состоит из или представляет собой аминокислотную последовательность SEQ ID NO:6, SEQ ID NO:8 или SEQ ID NO:14;

(ii) представляет собой фрагмент последовательности (i); или

(iii) идентичен полипептидам (i) или (ii) более чем на 70%.

2. Полипептид по п.1, отличающийся тем, что состоит из 7 или более аминокислотных остатков (например, из 8, 10, 12, 14, 16, 18, 20 или более) аминокислотной последовательности SEQ ID NO:6, SEQ ID NO:8 или SEQ ID NO:14.

3. Очищенная молекула нуклеиновой кислоты, которая кодирует полипептид по п.1 или 2.

4. Очищенная молекула нуклеиновой кислоты по п.3, отличающаяся тем, что состоит из или представляет собой последовательность нуклеиновой кислоты SEQ ID NO:5, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13 или SEQ ID NO:15.

5. Очищенная молекула нуклеиновой кислоты, которая гибридизуется с молекулой нуклеиновой кислоты по п.3 или 4 в условиях высокой жесткости.

6. Вектор, содержащий молекулу нуклеиновой кислоты по любому из пп.3-5.

7. Клетка-хозяин, трансформированная вектором по п.6.

8. Антитело, которое специфически связывается с полипептидом по п.1 или 2 и которое предпочтительно ингибирует полипептид с активностью цитокина, имеющего структуру с цистиновыми узлами, предпочтительно полипептид с активностью подсемейства DAN по п.1 или 2.

9. Применение полипептида по п.1 или 2 для лечения или диагностики заболевания.

10. Применение молекулы нуклеиновой кислоты по любому из пп.3-5 для лечения или диагностики заболевания.

11. Применение вектора по п.6 для лечения или диагностики заболевания.

12. Применение антитела по п.8 для лечения или диагностики заболевания.

13. Способ диагностики заболевания у пациента, предусматривающий оценку уровня экспрессии

природного гена, кодирующего полипептид по п.1 или 2, или оценку активности полипептида по п.1 или 2 в ткани указанного пациента и сравнение указанного уровня экспрессии или активности с контрольным уровнем, где уровень, отличный от указанного контрольного уровня, свидетельствует о наличии заболевания.

14. Способ по п.13, осуществляемый *in vitro*.

15. Способ по п.13 или 14, который предусматривает стадии:

(а) контактирования антитела по п.8 с биологическим образцом в условиях, подходящих для образования комплекса антитело-полипептид; и

(b) детекции указанного комплекса.

16. Способ по п.13 или 14, предусматривающий стадии:

а) контактирования образца ткани, взятой у пациента, с нуклеотидным зондом в жестких условиях, при которых возможно образование гибридного комплекса между молекулой нуклеиновой кислоты по любому из пп.3-5 и зондом;

б) контактирования контрольного образца с указанным зондом в условиях, аналогичных условиям стадии (а); и

с) обнаружения гибридных комплексов в указанных образцах; причем обнаружение уровней гибридного комплекса в образце пациента, которые отличаются от уровней гибридного комплекса в контрольном образце, свидетельствует о наличии заболевания.

17. Способ по п.13 или 14, предусматривающий:

а) контактирование образца нуклеиновой кислоты из ткани, взятой у пациента, с нуклеотидным праймером в жестких условиях, при которых возможно образование гибридного комплекса между молекулой нуклеиновой кислоты по любому из пп.3-5 и праймером;

б) контактирование контрольного образца с указанным праймером в условиях, аналогичных условиям стадии (а);

с) амплификацию образца нуклеиновой кислоты и

д) обнаружение уровня амплифицированной нуклеиновой кислоты в образцах, взятых у пациента, и в контрольных образцах, причем обнаружение уровней амплифицированной нуклеиновой кислоты в образце пациента, которые существенно отличаются от уровней амплифицированной нуклеиновой кислоты в контрольном образце, свидетельствует о наличии заболевания.

18. Способ по п.13 или 14, предусматривающий:

а) получение образца тканей пациента, исследуемого на наличие заболевания;

б) выделение молекулы нуклеиновой кислоты по любому из пп.3-5 из указанного образца ткани и

с) диагностирование заболевания у пациента путем обнаружения мутации в молекуле нуклеиновой кислоты, связанной заболеванием, которая свидетельствует о наличии заболевания.

19. Способ по п.18, дополнительно предусматривающий амплификацию молекулы нуклеиновой кислоты с образованием продукта амплификации и обнаружение наличия или отсутствия мутации в продукте амплификации.

20. Способ по п.18 или 19, где наличие или отсутствие мутации у пациента определяют путем осуществления контакта указанной молекулы нуклеиновой кислоты с нуклеотидным зондом, который гибридизуется с указанной молекулой нуклеиновой кислоты в жестких условиях с образованием гибридной двухцепочечной молекулы, где гибридная двухцепочечная молекула имеет негибридируемую часть цепи нуклеотидного зонда в любой области, соответствующей мутации, связанной с заболеванием; и обнаружение наличия или отсутствия негибридируемой части цепи зонда, что свидетельствует о наличии или отсутствии связанной с заболеванием мутации.

21. Способ по любому из пп.13-20, где указанное заболевание выбрано из пролиферативного заболевания, аутоиммунного/воспалительного заболевания, сердечно-сосудистого заболевания, неврологического расстройства, нарушения развития, метаболического расстройства, инфекции или другого патологического состояния.

22. Применение полипептида по п.1 или 2 в качестве секретируемого белка.

23. Применение полипептида по п.1 или 2 в качестве цитокина, предпочтительно в качестве цитокина с цистиновыми узлами, предпочтительно в качестве цитокина подсемейства DAN с цистиновыми узлами.

24. Фармацевтическая композиция, содержащая полипептид по п.1 или 2.

25. Фармацевтическая композиция, содержащая молекулу нуклеиновой кислоты по любому из пп.3-5.

26. Фармацевтическая композиция, содержащая вектор по п.6.

27. Фармацевтическая композиция, содержащая антитело по п.8.

28. Композиция вакцины, содержащая полипептид по п.1 или 2.

29. Композиция вакцины, содержащая молекулу нуклеиновой кислоты по любому из пп.3-5.

30. Применение полипептида по п.1 или 2 для получения лекарственного средства для лечения пролиферативных расстройств, аутоиммунных/воспалительных заболеваний, сердечно-сосудистых заболеваний, неврологических расстройств, нарушений развития, метаболических расстройств, инфекций и других патологических состояний.

риода времени, при этом изменение указанного уровня экспрессии или активности за определенный период времени по сравнению с контрольным уровнем свидетельствует о регрессии указанного заболевания.

51. Способ мониторинга терапевтического лечения заболевания у пациента, предусматривающий мониторинг уровня экспрессии молекулы нуклеиновой кислоты по любому из пп.3-5 в течение определенного периода времени, при этом изменение указанного уровня экспрессии или активности за определенный период времени по сравнению с контрольным уровнем свидетельствует о регрессии указанного заболевания.

52. Способ идентификации соединения, которое является эффективным для лечения и/или диагностики заболевания, предусматривающий контактирование полипептида по п.1 или 2 с одним или несколькими соединениями, предположительно обладающими аффинностью связывания с указанным полипептидом, и отбор соединения, которое специфически связывается с указанным полипептидом.

53. Способ идентификации соединения, которое является эффективным для лечения и/или диагностики заболевания, предусматривающий контактирование молекулы нуклеиновой кислоты по любому из пп.3-5 с одним или несколькими соединениями, предположительно, обладающими аффинностью связывания с указанной молекулой нуклеиновой кислоты, и отбор соединения, которое специфически связывается с указанной молекулой нуклеиновой кислоты.

54. Набор, применяемый для диагностики заболевания, содержащий первый контейнер с нуклеотидным зондом, который гибридизуется в жестких условиях с молекулой нуклеиновой кислоты по любому из пп.3-5; второй контейнер с подходящими для амплификации указанными молекулами нуклеиновой кислоты; и инструкции по применению зонда и праймеров для осуществления диагностики заболевания.

55. Набор по п.54, отличающийся тем, что дополнительно содержит третий контейнер с агентом для расщепления негибридизованной РНК.

56. Набор, содержащий массив молекул нуклеиновой кислоты, по крайней мере одна из которых является молекулой нуклеиновой кислоты по любому из пп.3-5.

57. Набор, содержащий одно или более антител, который связываются с полипептидом по п.1 или 2; и реагент, используемый для обнаружения реакции связывания между указанным антителом и указанным полипептидом.

58. Трансгенное животное или животное с «нокаут», не являющееся человеком, трансформированное таким образом, чтобы происходила повышенная или пониженная экспрессия полипептида по п.1 или 2, или экспрессия не происходила вообще.

59. Способ скрининга соединения, эффективного для лечения заболевания, предусматривающий контактирование трансгенного животного, не являющегося человеком, по п.58 с соединением-кандидатом и определение эффективности указанного соединения для лечения заболевания указанного животного.

Список последовательностей

Gln Asp Glu Val Ala Ala Val Thr Leu Pro Leu Asn Pro Gln Glu Val
 85 90 95

Ile Gln Gly Met Cys Lys Ala Val Pro Phe Val Gln
 100 105

<210> 3
 <211> 246
 <212> ДНК
 <213> Homo sapiens

<220>
 <221> Нуклеотидная последовательность INSP002 - экзон 2

<400> 3
 gtgttctccc ggcccggctg ctcagccata cgctccgaa atcatctgtg ctttggtcat 60
 tgctcctctc tctacatccc tggctcggac cccaccccac tagtctgtg caacagctgt 120
 atgcctgctc gcaagcgtg ggcacccgtg gtctgtggt gtctactgg cagctcagcc 180
 tcccgctgac ggggtgaagat atccaccatg ctgatcgagg ggtgtcactg cagcccaaaa 240
 gcatga 246

<210> 4
 <211> 81
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Последовательность белка INSP002 - экзон 2

<400> 4
 Val Phe Ser Arg Pro Gly Cys Ser Ala Ile Arg Leu Arg Asn His Leu
 1 5 10 15

Cys Phe Gly His Cys Ser Ser Leu Tyr Ile Pro Gly Ser Asp Pro Thr
 20 25 30

Pro Leu Val Leu Cys Asn Ser Cys Met Pro Ala Arg Lys Arg Trp Ala
 35 40 45

Pro Val Val Leu Trp Cys Leu Thr Gly Ser Ser Ala Ser Arg Arg Arg
 50 55 60

Val Lys Ile Ser Thr Met Leu Ile Glu Gly Cys His Cys Ser Pro Lys
 65 70 75 80

Ala

<210> 5
 <211> 721
 <212> ДНК
 <213> Homo sapiens

<220>
 <221> Нуклеотидная последовательность INSP002

<400> 5
 aaatgcctcc caggctatcc aggaggggccc aagagattaa aagcagggtc agaaggctca 60

gatgccactc accagacagc agggctcact gctagtacc ttgagcccag tccggacaga 120
 cagacaggca gacagacgca cggacaagca gatgctcctt ggccagctat ccactcttct 180
 gtgcctgctt agcggggccc tgcctacagg ctacgggagg cctgaacccc agtctcctcg 240
 acctcagtc tgggctgcag ccaatcagac ctgggctctg ggcccagggg ccctgcccc 300
 actggtgcca gctctgccc ttgggagctg gaaggcctc ttgggcctgc agaaagccag 360
 gcagctgggg atgggcaggc tgcagcgtgg gcaagacgag gtggctgctg tgactctgcc 420
 gctgaaccct caggaagtga tccaggggat gtgtaaggct gtgcccttcg ttcagggtgt 480
 ctcccgccc ggctgctcag ccatacgctt ccgaaatcat ctgtgcttg gtcattgctc 540
 ctctcttac atccctggct cggacccccc cccactagtc ctgtgcaaca gctgtatgcc 600
 tgctcgcaag cgttgggcac ccgtggtcct gtggtgtctc actggcagct cagcctccc 660
 tcgacgggtg aagatatcca ccatgctgat cgaggggtgt cactgcagcc caaaagcatg 720
 a 721

<210> 6
 <211> 189
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Последовательность белка INSP002

<400> 6
 Met Leu Leu Gly Gln Leu Ser Thr Leu Leu Cys Leu Leu Ser Gly Ala
 1 5 10 15
 Leu Pro Thr Gly Ser Gly Arg Pro Glu Pro Gln Ser Pro Arg Pro Gln
 20 25 30
 Ser Trp Ala Ala Ala Asn Gln Thr Trp Ala Leu Gly Pro Gly Ala Leu
 35 40 45
 Pro Pro Leu Val Pro Ala Ser Ala Leu Gly Ser Trp Lys Ala Phe Leu
 50 55 60
 Gly Leu Gln Lys Ala Arg Gln Leu Gly Met Gly Arg Leu Gln Arg Gly
 65 70 75 80
 Gln Asp Glu Val Ala Ala Val Thr Leu Pro Leu Asn Pro Gln Glu Val
 85 90 95
 Ile Gln Gly Met Cys Lys Ala Val Pro Phe Val Gln Val Phe Ser Arg
 100 105 110
 Pro Gly Cys Ser Ala Ile Arg Leu Arg Asn His Leu Cys Phe Gly His
 115 120 125
 Cys Ser Ser Leu Tyr Ile Pro Gly Ser Asp Pro Thr Pro Leu Val Leu
 130 135 140
 Cys Asn Ser Cys Met Pro Ala Arg Lys Arg Trp Ala Pro Val Val Leu
 145 150 155 160
 Trp Cys Leu Thr Gly Ser Ser Ala Ser Arg Arg Arg Val Lys Ile Ser
 165 170 175
 Thr Met Leu Ile Glu Gly Cys His Cys Ser Pro Lys Ala
 180 185

<210> 7

<211> 86
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Последовательность белка INSP002 - экзон 1 без
 сигнального пептида

<400> 7
 Arg Pro Glu Pro Gln Ser Pro Arg Pro Gln Ser Trp Ala Ala Ala Asn
 1 5 10 15

Gln Thr Trp Ala Leu Gly Pro Gly Ala Leu Pro Pro Leu Val Pro Ala
 20 25 30

Ser Ala Leu Gly Ser Trp Lys Ala Phe Leu Gly Leu Gln Lys Ala Arg
 35 40 45

Gln Leu Gly Met Gly Arg Leu Gln Arg Gly Gln Asp Glu Val Ala Ala
 50 55 60

Val Thr Leu Pro Leu Asn Pro Gln Glu Val Ile Gln Gly Met Cys Lys
 65 70 75 80

Ala Val Pro Phe Val Gln
 85

<210> 8
 <211> 167
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Последовательность белка INSP002 без
 сигнального пептида

<400> 8
 Arg Pro Glu Pro Gln Ser Pro Arg Pro Gln Ser Trp Ala Ala Ala Asn
 1 5 10 15

Gln Thr Trp Ala Leu Gly Pro Gly Ala Leu Pro Pro Leu Val Pro Ala
 20 25 30

Ser Ala Leu Gly Ser Trp Lys Ala Phe Leu Gly Leu Gln Lys Ala Arg
 35 40 45

Gln Leu Gly Met Gly Arg Leu Gln Arg Gly Gln Asp Glu Val Ala Ala
 50 55 60

Val Thr Leu Pro Leu Asn Pro Gln Glu Val Ile Gln Gly Met Cys Lys
 65 70 75 80

Ala Val Pro Phe Val Gln Val Phe Ser Arg Pro Gly Cys Ser Ala Ile
 85 90 95

Arg Leu Arg Asn His Leu Cys Phe Gly His Cys Ser Ser Leu Tyr Ile
 100 105 110

Pro Gly Ser Asp Pro Thr Pro Leu Val Leu Cys Asn Ser Cys Met Pro
 115 120 125

Ala Arg Lys Arg Trp Ala Pro Val Val Leu Trp Cys Leu Thr Gly Ser
 130 135 140

Ser Ala Ser Arg Arg Arg Val Lys Ile Ser Thr Met Leu Ile Glu Gly
 145 150 155 160

Cys His Cys Ser Pro Lys Ala
 165

<210> 9

<211> 504

<212> ДНК

<213> Homo sapiens

<220>

<221> Нуклеотидная последовательность INSP002 без сигнального пептида

<400> 9

```
aggcctgaac cccagtctcc tcgacctcag tcctgggctg cagccaatca gacctgggct 60
ctgggcccag gggccctgcc cccactggcg ccagctctcg ccctgggag ctggaaggcc 120
ttctgggcc tgcagaaagc caggcagctg gggatgggca ggctgcagcg tgggcaagac 180
gagggtggctg ctgtgactct gccgctgaac cctcaggaag tgatccaggg gatgtgtaag 240
gctgtgccct tcgttcaggt gtctcccg cccggctgct cagccatacg cctccgaaat 300
catctgtgct ttggctcattg ctctctctc tacatccctg gctcggacc caccacctacta 360
gtcctgtgca acagctgtat gcctgctgc aagcgttggg caccctgggt cctgtggtgt 420
ctcactggca gctcagcctc ccgtcgacgg gtgaagatat ccaccatgct gatcgagggg 480
gtcactgca gcccaaaagc atga 504
```

<210> 10

<211> 258

<212> ДНК

<213> Homo sapiens

<220>

<221> Нуклеотидная последовательность INSP002 - экзон 1 без сигнального пептида

<400> 10

```
aggcctgaac cccagtctcc tcgacctcag tcctgggctg cagccaatca gacctgggct 60
ctgggcccag gggccctgcc cccactggcg ccagctctcg ccctgggag ctggaaggcc 120
ttctgggcc tgcagaaagc caggcagctg gggatgggca ggctgcagcg tgggcaagac 180
gagggtggctg ctgtgactct gccgctgaac cctcaggaag tgatccaggg gatgtgtaag 240
gctgtgccct tcgttcag 258
```

<210> 11

<211> 570

<212> ДНК

<213> Homo sapiens

<220>

<221> Нуклеотидная последовательность INSP002 без 5' UTR

<400> 11

```
atgctccttg gccagctatc cactctctg tgctgctta gggggccct gcctacaggc 60
tcaggaggc ctgaaccca gtctctcga cctcagctc gggctgcagc caatcagacc 120
tgggctctgg gccaggggc cctgccccca ctggtgccag ctctgccct tgggagctgg 180
aaggcctct tgggcctgca gaaagccagg cagctgggga tgggcaggct gcagcgtggg 240
caagacgagg tggctgctgt gactctgcc ctgaaccctc aggaagtgat ccaggggatg 300
tgtaaggctg tgccctctgt tcaggtgttc tccggcccg gctgctcagc catacgcctc 360
cgaaatcatc tgtctttgg tcattgtcc tctctctaca tcctggctc ggacccacc 420
```

ccactagtcc tgtgcaacag ctgtatgcct gctcgcaagc gttgggcacc cgtggctctg 480
 tgggtgtca ctggcagctc agcctcccgt cgacgggtga agatatccac catgctgac 540
 gaggggtgtc actgcagccc aaaagcatga 570

<210> 12
 <211> 324
 <212> ДНК
 <213> Homo sapiens

<220>

<221> Нуклеотидная последовательность INSP002 - экзон 1 без 5'UTR

<400> 12
 atgtctcttg gccagctatc cactctctg tgcctgctta gcggggccct gcctacaggc 60
 tcaggagggc ctgaacccca gtctctcga cctcagtcct gggctgcagc caatcagacc 120
 tgggctctgg gccagggggc cctgccccca ctggtgccag ctctgccct tggagctgg 180
 aaggccttct tgggcctgca gaaagccagg cagctgggga tgggcaggct gcagcgtggg 240
 caagacgagg tggctgctgt gactctgccg ctgaaccctc aggaagtgat ccagggggatg 300
 tgtaaggctg tgccctctgt tcag 324

<210> 13
 <211> 719
 <212> ДНК
 <213> Homo sapiens

<220>

<221> Нуклеотидная последовательность INSP002 - вариант полипептида INSP002

<400> 13
 gtcgactgt agtgacctg agcccagtc ggacagacag acaggcagac agacgcacgg 60
 acaagcagat gtctctggc cagctatcca ctctctgtg cctgcttagc ggggccctgc 120
 ctacaggctc agggaggcct gaaccccagt ctctcgacc tcagtctgg gctgcagcca 180
 atcagacctg ggctctggc ccaggggccc tgccccact ggtgccagct tctgccctg 240
 ggagctggaa ggccttctg ggctgcaga aagccaggca gctggggatg ggcaggctgc 300
 agcgtgggca agacgagggt gctgctgtga ctctgccgt gaacctcag gaagtatcc 360
 aggggatgtg taaggctgt cctctgttc tctccggcc cggctgctca gccatagcc 420
 tccgaaatca tctgtctt ggctatgtc cctctctta catccctggc tcggacccca 480
 cccactagt cctgtgcaac agctgtatgc ctgctcgca gcgtgggca cccgtggtcc 540
 tgtgtgtct cactggcagc tcagcctccc gtcgacgggt gaagatatcc accatgtga 600
 tcgaggggtg tcactgcagc caaaagcat gaactgagca tctggatggg tgcacggaga 660
 cagcacctt ggagaaatga ggggagatgg accaagaaag acgtggacct ggatgatgt 719

<210> 14
 <211> 187
 <212> PRT
 <213> Homo sapiens

<220>

<221> Нуклеотидная последовательность INSP002 - вариант полипептида INSP002

<400> 14

Met Leu Leu Gly Gln Leu Ser Thr Leu Leu Cys Leu Leu Ser Gly Ala

1 5 10 15

Leu Pro Thr Gly Ser Gly Arg Pro Glu Pro Gln Ser Pro Arg Pro Gln

20 25 30

Ser Trp Ala Ala Ala Asn Gln Thr Trp Ala Leu Gly Pro Gly Ala Leu

35 40 45

Pro Pro Leu Val Pro Ala Ser Ala Leu Gly Ser Trp Lys Ala Phe Leu
50 55 60

Gly Leu Gln Lys Ala Arg Gln Leu Gly Met Gly Arg Leu Gln Arg Gly
65 70 75 80

Gln Asp Glu Val Ala Ala Val Thr Leu Pro Leu Asn Pro Gln Glu Val
85 90 95

Ile Gln Gly Met Cys Lys Ala Val Pro Phe Val Leu Ser Arg Pro Gly
100 105 110

Cys Ser Ala Ile Arg Leu Arg Asn His Leu Cys Phe Gly His Cys Ser
115 120 125

Ser Leu Tyr Ile Pro Gly Ser Asp Pro Thr Pro Leu Val Leu Cys Asn
130 135 140

Ser Cys Met Pro Ala Arg Lys Arg Trp Ala Pro Val Val Leu Trp Cys
145 150 155 160

Leu Thr Gly Ser Ser Ala Ser Arg Arg Arg Val Lys Ile Ser Thr Met
165 170 175

Leu Ile Glu Gly Cys His Cys Ser Pro Lys Ala
180 185

<210> 15

<211> 561

<212> ДНК

<213> Homo sapiens

<220>

<221> Нуклеотидная последовательность INSP002 - вариант полипептида INSP002 без 5' UTR

<400> 15

atgctccttg gccagctatc cactctctg tgctgctta gcggggccct gcctacaggc 60
tcaggagggc ctgaacccca gtctcctga cctcagtcct gggctgcagc caatcagacc 120
tggtgctctg gccaggggc cctgccccca ctggtgccag cttctgccct tgggagctgg 180
aaggccttct tgggcctgca gaaagccagg cagctgggga tgggcaggct gcagcgtggg 240
caagacgagg tggctgctgt gactctgccg ctgaaccctc aggaagtgtat ccagggggatg 300
tgtaaggctg tgcccttctg tctctcccg cccggctgct cagccatacg cctccgaaat 360
catctgtgct ttgtcattg ctctctctc tacatccctg gctcggacc caccacctacta 420
gtcctgtgca acagctgtat gctgctcgc aagcgttggg caccctgtgt cctgtggtgt 480
ctcactggca gtcagcctc ccgtgacgg gtgaagatat ccaccatgct gatcgagggg 540
tgtcactgca gcccaaaagc a 561

<210> 16

<211> 20

<212> ДНК

<213> Синтезированная последовательность

<220>

<223> ПЦР - праймер

<400> 16

ctcagccata cgctccgaa

20

<210> 17
 <211> 20
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 17
 gctgagctgc cagtgagaca 20

 <210> 18
 <211> 25
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 18
 acctggaagg aagcgactgc actga 25

 <210> 19
 <211> 37
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 19
 gcagccgggc cgggagagaa cgaagggcac agcctta 37

 <210> 20
 <211> 37
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 20
 aggtgtgcc ctctgtctc tcccgcccg gctgctc 37

 <210> 21
 <211> 25
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 21
 actccaggac gggcactgtg tctac 25

 <210> 22
 <211> 22
 <212> ДНК
 <213> Синтезированная последовательность

<220>
 <223> ПЦР - праймер

 <400> 22
 gtcgactgct agtgaccttg ag 22

 <210> 23
 <211> 22
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 23
 acatcatcca ggtccacgtc tt 22

 <210> 24
 <211> 18
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 24
 attagtgga cactatag 18

 <210> 25
 <211> 20
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 25
 taatacgact cactataggg 20

 <210> 26
 <211> 20
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

 <400> 26
 attaaccctc actaaaggga 20

 <210> 27
 <211> 18
 <212> ДНК
 <213> Синтезированная последовательность

 <220>
 <223> ПЦР - праймер

<400> 27	
tgtaaaacga cggccagt	18
<210> 28	
<211> 20	
<212> ДНК	
<213> Синтезированная последовательность	
<220>	
<223> ПЦР - праймер	
<400> 28	
gccagcttgg cacttgatgt	20
<210> 29	
<211> 20	
<212> ДНК	
<213> Синтезированная последовательность	
<220>	
<223> ПЦР - праймер	
<400> 29	
gatggagggtg gacgtgtcag	20
<210> 30	
<211> 37	
<212> ДНК	
<213> Синтезированная последовательность	
<220>	
<223> ПЦР - праймер	
<400> 30	
aagcaggctt cgccaccatg ctccttggcc agctatc	37
<210> 31	
<211> 35	
<212> ДНК	
<213> Синтезированная последовательность	
<220>	
<223> ПЦР - праймер	
<400> 31	
gtgatggtga tgggtgtgctt ttgggctgca gtgac	35
<210> 32	
<211> 37	
<212> ДНК	
<213> Синтезированная последовательность	
<220>	
<223> ПЦР - праймер	
<400> 32	
ggggacaagt ttgtacaaaa aagcaggctt cgccacc	37
<210> 33	

<211> 51
 <212> ДНК
 <213> Синтезированная последовательность

<220>
 <223> ПЦР - праймер

<400> 33
 ggggaccact ttgtacaaga aagctgggtt tcaatggtga tggatggt g 51

<210> 34
 <211> 20
 <212> ДНК
 <213> Синтезированная последовательность

<220>
 <223> ПЦР - праймер

<400> 34
 gatgctcctt ggccagctat 20

<210> 35
 <211> 20
 <212> ДНК
 <213> Синтезированная последовательность

<220>
 <223> ПЦР - праймер

<400> 35
 ccatccacga tgctcagttc 20

<210> 36
 <211> 1756
 <212> ДНК
 <213> Синтезированная последовательность

<220>
 <223> Векторная вставка

<220>
 <221> Последовательность векторной вставки, содержащая Image 4558384
 (фигура 10)

<400> 36
 cggcacgagg gggagacctg gaaggaagcg actgcactga tccagaaatg cctcccaggc 60
 tatccaggag gggccaagag attaaaagca gggtcagaag gctcagatgc cactcaccag 120
 acagcagggg cgactgctag tgacctgag cccagtccgg acagacagac aggcagacag 180
 acgcacggac aagcagatgc tcctggcca gctatccact cttctgtgcc tgcttagcgg 240
 ggccctgcct acaggctcag ggaggcctga accccagtct cctcgacctc agtcctgggc 300
 tgcagccaat cagacctggg ctctgggcc aggggccctg cccccactgg tgccagcttc 360
 tgccctggg agctggaagg cctcttggg cctgcagaaa gccaggcagc tggggatggg 420
 caggctgcag cgtgggcaag acgaggtggc tgctgtgact ctgccgtga accctcagga 480
 agtgcaccag gggatgtgta aggctgtgcc ctctgtcag acacgggagt ctgcctatgt 540
 tgccaagct agtctgagc ttctggcctc aagcaatcct cccacctcag cctcctgagt 600
 tctagggtgt ctccggccc ggctgctcag ccatacgct cggaaatcat ctgtgcttg 660
 gtcattgtc ctctctac atccctggct cggacccac cccactagtc ctgtgcaaca 720
 gctgtatgcc tgctcgcaag cgtggggcac ccgtggtcct gtggtgtctc actggcagct 780
 cagcctccg tcgacgggtg aagatatcca ccatgctgat cgaggggtgt cactgcagcc 840
 caaaagcatg aactgagcat ctggatgggt gcacggagac acgcacctg gagaaatgag 900

gggagatgga ccaagaaaga cgtggacctg gatgatgtac tctgggtcaa gagaccaggg 960
 atgcagggtt aggcagacag gtccccagag tctcacctt gctccccaga cagtagacac 1020
 agtgcccgtc ctggagtgc accactgata gtcacagcac acaatgattg acaactcact 1080
 tttttttt ttttgagat ggagtctcgc tctgtcgccc aggtggagt gcagtggcgc 1140
 aatctcagct cactgcaagc tcacacctcc gggtttatgc cattctctg tctcagcctc 1200
 ccgagtagct gggactacag gcacccgcca acacgcccg staattttt gtatttttag 1260
 taaagacagg gttcacctg gtagccagg atggtctcta tctcctgacc tctgtatctg 1320
 cctgccttg ccttattatt tttttttaa ggacagagtc tctctctgc acccaggctg 1380
 gagtgaatg gcgcgatctt ggctcactgt aactccact tgcaggctc aagcagttct 1440
 cctgcctcag cctcctgagt agctgggact acaggcacc gccaccatgc ccagctaatt 1500
 ttgtatttt tagtagagac agagttcac catattagcc tggctgtct caaactcctg 1560
 gcctcagggt atctgccac ctgcgcctcc caaagtctg ggatcaaat cactgtaat 1620
 cattaggctg aactgtctt tatagaatga ggtcaaagac actccaggt gcaggagggg 1680
 tagatggccc caccagacc gagagacaca gtgatgacct cagcctaggg acmccaaaaa 1740
 aaaaaaaaaa aaaaaa 1756

<210> 37

<211> 133

<212> PRT

<213> Синтезированная последовательность

<220>

<223> Homo sapiens

<220>

<221> Предсказанный продукт трансляции SEQ ID No:36

<400> 37

Met Leu Leu Gly Gln Leu Ser Thr Leu Leu Cys Leu Leu Ser Gly Ala

1 5 10 15

Leu Pro Thr Gly Ser Gly Arg Pro Glu Pro Gln Ser Pro Arg Pro Gln

20 25 30

Ser Trp Ala Ala Ala Asn Gln Thr Trp Ala Leu Gly Pro Gly Ala Leu

35 40 45

Pro Pro Leu Val Pro Ala Ser Ala Leu Gly Ser Trp Lys Ala Phe Leu

50 55 60

Gly Leu Gln Lys Ala Arg Gln Leu Gly Met Gly Arg Leu Gln Arg Gly

65 70 75 80

Gln Asp Glu Val Ala Ala Val Thr Leu Pro Leu Asn Pro Gln Glu Val

85 90 95

Ile Gln Gly Met Cys Lys Ala Val Pro Phe Val Gln Thr Arg Glu Ser

100 105 110

Arg Tyr Val Ala Gln Ala Ser Leu Glu Leu Leu Ala Ser Ser Asn Pro

115 120 125

Pro Thr Ser Ala Ser

130

<210> 38

<211> 589

<212> ДНК

<213> Homo sapiens

<220>

<221> Полноразмерный INSP002, клонированный из кДНК человеческого сердца

<400> 38

```

gatgctcctt ggccagctat ccaactcttct gtgcctgctt agcggggccc tgcctacagg 60
ctcaggaggagg cctgaacccc agtctcctcg acctcagtc tgggctgcag ccaatcagac 120
ctgggctctg ggcccagggg cctgcccccc actggtgccca gcttctgccc ttgggagctg 180
gaaggccttc ttgggcctgc agaaagccag gcagctgggg atgggcaggc tgcagcgtgg 240
gcaagacgag gtggctgctg tgactctgcc gctgaaccct caggaagtga tccaggggat 300
gtgtaaggct gtgcccttcg ttcaggtgtt ctccggccc ggctgctcag ccatacgctt 360
ccgaaatcat ctgtgctttg gtcattgctc ctctctctac atccctggct cggaccccac 420
cccactagtc ctgtgcaaca gctgtatgcc tgctcgcaag cgttgggcac ccgtggctct 480
gtgggtgtct actggcagct cagcctcccg tcgacgggtg aagatatcca ccatgctgat 540
cgaggggtgt cactgcagcc caaaagcatg aactgagcat cgtggatgg 589

```

Query= INSP002

(189 letters)

Database: ncbi-nt

1,021,646 sequences; 4,527,726,597 total letters

Searching.....done

	Score	E
Sequences producing significant alignments:	(bits)	Value
gi 15217064 gb AF400435.1 AF400435 Homo sapiens cerberus-related...	69	2e-10
gi 11427432 ref XM_005320.1 Homo sapiens cerberus 1 (Xenopus la...	69	2e-10
gi 4885134 ref NM_005454.1 Homo sapiens cerberus 1 (Xenopus lae...	69	2e-10
gi 5802091 gb AF139721.1 AF139721 Gallus gallus cerberus homolog...	68	3e-10
gi 5902633 gb AF179484.1 AF179484 Gallus gallus caronte (CAR) mR...	67	5e-10
gi 2853615 gb AF035579.1 AF035579 Mus musculus cerberus homolog ...	61	5e-08
gi 6753409 ref NM_009887.1 Mus musculus cerberus 1 homolog (Xen...	61	5e-08
gi 2654026 gb AF031896.1 AF031896 Mus musculus cerberus-related ...	61	5e-08
gi 1513087 gb U64831.1 XLU64831 Xenopus laevis head-inducing fac...	57	9e-07
gi 6679017 ref NM_008675.1 Mus musculus neuroblastoma, suppress...	55	3e-06

Фиг. 1. Наилучшие десять соответствий для INSP002, полученные с помощью BLAST в неизбыточной базе данных NCBI

```

>gi|15217064|gb|AF400435.1|AF400435 Homo sapiens cerberus-related 1
      (CER1) mRNA, complete cds
      Length = 804

      Score = 68.9 bits (167), Expect = 2e-10
      Identities = 36/102 (35%), Positives = 53/102 (51%)
      Frame = +1

Query: 87  VTLPLNPQEVIIQGMCKAVPFVQVFSRPGCSAIRLRNHLFCGHCSSLYIPGSDPTPLVLCN 146
          V LP+  EV   C+ VPF Q  +  GC  + ++N+LCFG C S++ PG+      C+
Sbjct: 442 VILPIKSHEVHWETCRTVPFSQTITHEGCEKVVVQNNLCFGKCGSVHFPGAAQHSHTSCS 621

Query: 147 SCMPARKRWAPVVLWCLTGSSASRRRVKISTMLIEGCHCSPK 188
          C+PA+      + L C   SS      +K+  ML+E C C  K
Sbjct: 622 HCLPAKFTTMHLPLNCTELSSV----IKV-VMLVEECQCKVK 732

```

Фиг. 2. Сопоставление с использованием наилучших соответствий для cerberus-родственного белка Homo sapiens

i) ncbi-nr

Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
1,242,768 sequences; 395,571,179 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
ref NP_689867.1 hypothetical protein FLJ38607 [Homo sapiens] >g...	394	e-109
ref XP_113990.1 hypothetical protein XP_113990 [Homo sapiens]	221	5e-57
ref XP_164658.1 similar to hypothetical protein FLJ38607 [Homo ...	70	1e-11
ref NP_005445.1 cerberus 1; cerberus-related 1; cerberus 1 (Xen...	69	3e-11
gb AAD51610.1 AF139721_1 cerberus homolog [Gallus gallus]	68	6e-11

ii) ncbi-nt

Database: All GenBank+EMBL+DDBJ+PDB sequences (but no EST, STS, GSS,
or phase 0, 1 or 2 HTGS sequences)
1,469,831 sequences; 7,238,625,236 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
ref NM_152654.1 Homo sapiens hypothetical protein FLJ38607 (FLJ...	394	e-108
dbj AK095926.1 Homo sapiens cDNA FLJ38607 fis, clone HEART2004941	394	e-108
gb BC025333.1 Homo sapiens, clone IMAGE:4558384, mRNA, partial cds	221	e-103
ref XM_113990.1 Homo sapiens LOC199699 (LOC199699), mRNA	221	e-103
gb AC092069.2 Homo sapiens chromosome 19 clone CTC-425F1, compl...	222	1e-56

Фиг. 3. Наилучшие четыре соответствия для INSP002, полученные с помощью BLAST в базах данных NCBI-nr и NCBI-nt

```
>dbj|AK095926.1| Homo sapiens cDNA FLJ38607 fis, clone HEART2004941
Length = 1731

Score = 394 bits (1013), Expect = e-108
Identities = 189/189 (100%), Positives = 189/189 (100%)
Frame = +2

Query: 1 MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW 60
      MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW
Sbjct: 44 MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW 223

Query: 61 KAFLGLQKARQLGMGRLLQRGQDEVAAVTLPLNPQEVIQGMCKAVPFVQVFSRPGCSAIRL 120
      KAFLGLQKARQLGMGRLLQRGQDEVAAVTLPLNPQEVIQGMCKAVPFVQVFSRPGCSAIRL
Sbjct: 224 KAFLGLQKARQLGMGRLLQRGQDEVAAVTLPLNPQEVIQGMCKAVPFVQVFSRPGCSAIRL 403

Query: 121 RNHLFCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSASRRRVKISTMLI 180
      RNHLFCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSASRRRVKISTMLI
Sbjct: 404 RNHLFCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSASRRRVKISTMLI 583

Query: 181 EGCHCSPKA 189
      EGCHCSPKA
Sbjct: 584 EGCHCSPKA 610
```

Фиг. 4. Сопоставление INSP002 с AK095926.1

```
>gb|BC025333.1| Homo sapiens, clone IMAGE:4558384, mRNA, partial cds
Length = 1746

Score = 221 bits (562), Expect(2) = e-103
Identities = 108/108 (100%), Positives = 108/108 (100%)
Frame = +1

Query: 1 MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW 60
      MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW
Sbjct: 187 MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW 366

Query: 61 KAFLGLQKARQLGMGRLLQRGQDEVAAVTLPLNPQEVIQGMCKAVPFVQ 108
      KAFLGLQKARQLGMGRLLQRGQDEVAAVTLPLNPQEVIQGMCKAVPFVQ
Sbjct: 367 KAFLGLQKARQLGMGRLLQRGQDEVAAVTLPLNPQEVIQGMCKAVPFVQ 510

Score = 178 bits (451), Expect(2) = e-103
Identities = 81/81 (100%), Positives = 81/81 (100%)
Frame = +2

Query: 109 VFSRPGCSAIRLRNHLFCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSA 168
      VFSRPGCSAIRLRNHLFCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSA
Sbjct: 596 VFSRPGCSAIRLRNHLFCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSA 775

Query: 169 SRRRVKISTMLIEGCHCSPKA 189
      SRRRVKISTMLIEGCHCSPKA
Sbjct: 776 SRRRVKISTMLIEGCHCSPKA 838
```

Фиг. 5. Сопоставление INSP002 с IMAGE: 4558384

```
1 AAATGCCTCC CAGGCTATCC AGGAGGGGCC AAGAGATTAA AAGCAGGTTT AGAAGGCTCA
61 GATGCCACTC ACCAGACAGC AGGCTCGACT GCTAGTGACC TTGAGCCGAG TCCGGACAGA
121 CAGACAGGCA GACAGACGCA CGGACAAGCA GATGCTCTT GGCAGCTAT CCACTCTTCT
      m l l g q l s t l
181 GTGCTGCTT AGCGGGGCC TGCCTACAGG CTCAGGAGG CCGTGAACCC AGTCTCTCG
      l c l l s g a l p t g s g r p e p q s p
241 ACCTCAGTCC TGGGCTGCAG CCAATCAGAC CTGGGCTCTG GGCCAGGGG CCGTGCCTCC
      r p q s w a a a n q t w a l g p g a l p
301 ACTGGTGCCA GCTTCTGCCC TTGGGAGCTG GAAGGCTTC TTGGGCTGC AGAAGCCAG
      p l v p a s a l g s w k a f l g l q k a
361 GCAGCTGGG ATGGGAGGC TGCAGCGTGG GCAAGACGAG GTGGCTGCTG TGACTCTGCC
      r q l g m g r l q r g q d e v a a v t l
421 GCTGAACCTC CAGGAAGTGA TCCAGGGGAT GTGTAAGGCT GTGCCCTTCG TTCAGGTGTT
      p l n p q e v i q g m c k a v p f v q v
481 CTCCCGGCC GGCTGCTCAG CCATACGCTT CCGAATCAT CTGTGCTTTG GTCATTGCTC
      f s r p g c s a l r l f n h l c f g h c
      INSP002-CP1
541 CTCTCTCTAC ATCCCTGGCT CGGACCCAC CCCACTAGTC CTGTGCAACA GCTGTATGCC
      s s l y i p g s d p t p l v l c n s c m
601 TGCTCGCAAG CGTTGGGCAC CGTGGTCTCT GTGGTGTCTC ACTGGCAGCT CAGCCTCCCG
      p a r k r w a p v v l w c l t g s s a s
      INSP002-CP2
661 TCGACGGGTG AAGATATCCA CCATGCTGAT CGAGGGGTGT CACTGCAGCC CAAAAGCATG
      r r r v k i s t m l i e g c h c s p k a
721 AACTGAGCAT CGTGGATGGG TGCACGGAGA CACGCACCTT GGAGAAATGA GGGGAGAT
```

Фиг. 6. Нуклеотидная последовательность INSP002 с трансляцией

```

1  CTCAGCCATA CGCCTCCGAA ATCATCTGTG CTTTGGTCAT TGCTCCTCTC TCTACATCCC
   s a i r l r n h l c f g h c s s l y i

61  TGGCTCGGAC CCCACCCAC TAGTCCTGTG CAACAGCTGT ATGCCTGCTC GCAAGCGTTG
   p g s d p t p l v l c n s c m p a r k r

121 GGCACCCGTG GTCCTGTGGT GTCTCACTGG CAGCTCAGC
   w a p v v l w c l t g s s

```

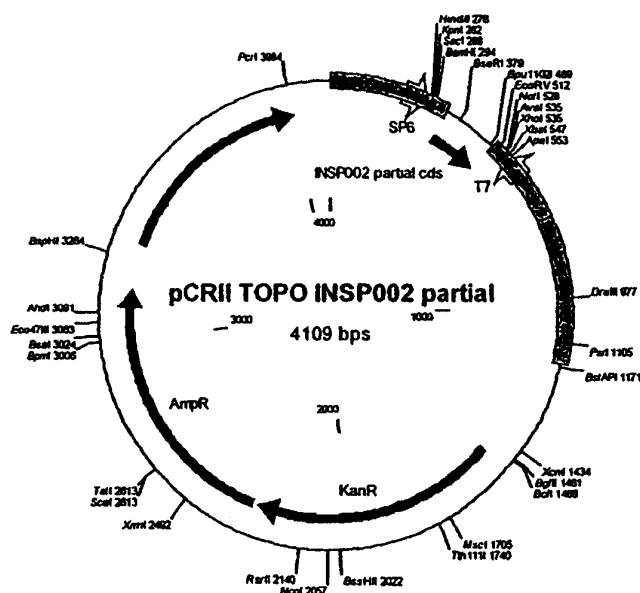
Фиг. 7. Неполная клонированная последовательность INSP002 с трансляцией

Молекула: product1, 4109 bps DNA Circular
 Название файла: ~7879480.cm5

Описание: Ligation of NoName into PCRII-TOPO-open

Особенности молекулы:

Тип	Начало	Конец	Название	Описание
Область	1	336		LacZa'
Маркер	239		SP6	
Ген	337	495	INSP002 partial cds	Inserted PCR product
Область	496	747		'LacZa
Маркер	584		C T7	
Область	749	1163		f1 ori
Ген	1497	2291	KanR	Канамицин-резистентный ген
Ген	2309	3169	AmpR	Ампициллин-резистентный ген
Ген	3314	3987		pUC ori



Фиг. 8. Карта неполной pCRII-TOPO-INSP002

Фиг. 9. Сопоставление предсказанной INSP002 (верхняя строка) с неполной клонированной последовательностью (нижняя строка)

```

      10      20      30      40      50      60
INSPpr CAGAAATGCCTCCAGGCTATCCAGGAGGGCCAGAGATTAAAGCAGGTTCAGAAGGC
:
Serono C-----

      70      80      90     100     110     120
INSPpr TCAGATGCCACTCACCAGACAGCAGGTCGACTGCTAGTGACCTTGAGCCAGTCCGGAC
Serono -----

      130     140     150     160     170     180
INSPpr AGACAGACAGGCAGACAGACGACAGCAGATGCTCTTGGCCAGCTATCCACTCT
Serono -----

      190     200     210     220     230     240
INSPpr TCTGTGCTGCTTAGCGGGGCCCTGCCTACAGGCTCAGGGAGGCTGAACCCAGTCTCC
Serono -----

      250     260     270     280     290     300
INSPpr TCGACCTCAGTCTGGGCTGCAGCCAATCAGACCTGGGCTCTGGGCCAGGGGCCCTGCC
Serono -----

      310     320     330     340     350     360
INSPpr CCCACTGGTGCCAGCTTCTGCCCTTGGGAGCTGGAAGGCCTTCTTGGGCTGCAGAAAGC
Serono -----

      370     380     390     400     410     420
INSPpr CAGGCAGCTGGGGATGGGCAGGCTGCAGCGTGGGCAAGACGAGTGGTGTGTGACTCT
Serono -----

      430     440     450     460     470     480
INSPpr GCCGCTGAACCTCAGGAAGTGATCCAGGGGATGTGTAAGGCTGTGCCCTTCGTTCAGGT
Serono -----

      490     500     510     520     530     540
INSPpr GTTCTCCCGCCCGGCTGCTCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTG
:
Serono -----TCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTG
              10      20      30      40

      550     560     570     580     590     600
INSPpr CTCCTCTCTTACATCCCTGGCTCGGACCCACCCCACTAGTCTGTGCAACAGCTGTAT
:
Serono CTCCTCTCTTACATCCCTGGCTCGGACCCACCCCACTAGTCTGTGCAACAGCTGTAT
      50      60      70      80      90     100

      610     620     630     640     650     660
INSPpr GCCTGCTCGCAAGCGTTGGGCACCCGTCCTGTGCTGTCTCACTGGCAGCTCAGCCTC
:
Serono GCCTGCTCGCAAGCGTTGGGCACCCGTCCTGTGCTGTCTCACTGGCAGCTCAGC---
      110     120     130     140     150

      670     680     690     700     710     720
INSPpr CCGTCGACGGGTGAAGATATCCACATGCTGATCGAGGGGTGCTCACTGCAGCCCAAAAGC
Serono -----

```

Фиг. 9а. Сопоставление нуклеотидных последовательностей.

```

Serono -----
INSP00 MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW
      10      20      30      40      50      60

Serono -----SAIRL
INSP00 KAFLGLQKARQLGMGRLOQCODEVAAVTLPLNPQEVIGMCKAVPFVQVFSRPGCSAIRL
      70      80      90     100     110     120

      10      20      30      40      50
Serono RNHLCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSA-----
:
INSP00 RNHLCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSASRRRVKISTMLI
      130     140     150     160     170     180

Serono -----
INSP00 EGCHCSFKA

```

Фиг. 9b. Сопоставление последовательностей белка

```

1  CGGCACGAGG GGGAGACCTG GAAGGAAGCG ACTGCACTGA TCCAGAAATG CCTCCCAGGC
      INSP002V-5'-F
61  TATCCAGGAG GGGCCAAGAG ATTTAAAGCA GGTTCAGAAG GCTCAGATGC CACTCACCAG
121 ACAGCAGGGT CGACTGCTAG TGACCTTGAG CCCAGTCCGG ACAGACAGAC AGGCAGACAG
      INSP002V-5'nest-F
181 ACGCACGGAC AAGCAGATGC TCCTTGGCCA GCTATCCACT CTTCTGTGCC TGCTTAGCGG
      m l l g q l s t l l c l l s
241 GGCCCTGCCT ACAGGCTCAG GGAGGCCTGA ACCCCAGTCT CCTCGACCTC AGTCCTGGGC
      g a l p t g s g r p e p q s p r p q s w
301 TGCAGCCAAT CAGACCTGGG CTCTGGGCCC AGGGGCCCTG CCCCCACTGG TGCCAGCTTC
      a a a n q t w a l g p g a l p p l v p a
361 TGCCCTTGGG AGCTGGAAGG CCTTCTTGGG CCTGCAGAAA GCCAGGCAGC TGGGGATGGG
      s a l g s w k a f l g l q k a r q l g m
421 CAGGCTGCAG CGTGGGCAAG ACGAGGTGGC TGCTGTGACT CTGCCGCTGA ACCCTCAGGA
      g r l q r g q d e v a a v t l p l n p q
481 AGTGATCCAG GGGATGTGTA AGGCTGTGCC CTTCGTTTCTG ACAGCGACT CAGCAGATG
      e v i q g m c k a v p f v q t r e s r y
      INSP002V-5'-R
541 TCCTTGGGCTT AGTCTTGGGCTT TCTTGGGCTT AGGCAATGCT CCAAGCCTG CTCTTGGGCT
      v a q a s l e l l a s s n p p t s a s
601 CTCTTGGGCTT CTCCCGGCCG GGCTGCTCAG CCATACGCCT CCGAAATCAT CTGTGCTTTG
      INSP002V-3'-F
661 GTCATTGCTC CTCTCTCTAC ATCCCTGGCT CGGACCCAC CCCACTAGTC CTGTGCAACA
721 GCTGTATGCC TGCTCGCAAG CGTTGGGCAC CCGTGGTCCT GTGGTGTCTC ACTGGCAGCT
781 CAGCCTCCCG TCGACGGGTG AAGATATCCA CCATGCTGAT CGAGGGGTGT CACTGCAGCC
841 CAAAAGCATG AACTGAGCAT CTGGATGGGT GCACGGAGAC ACGCACCTTG GAGAAATGAG
901 GGGAGATGGA CCAAGAAAGA CGTGACCTG GATGATGTAC TCTGGGTCAA GAGACCAGGG
      INSP002V-3'nest-R
961 ATGCAGGGTT AGGCAGACAG GTCCCCAGAG TCCTCACCCCT GCTCCCCAGA CAGTAGACAC
1021 AGTGCCCGTC CTGGAGTTGC ACCACTGATA GTCACAGCAC ACAATGATTG ACAACTCACT
      INSP002V-3'-R
1081 TTTTTTTTTT TTTTGTAGAT GGAGTCTCGC TCTGTGCGCC AGGCTGGAGT GCAGTGGCGC
1141 AATCTCAGCT CACTGCAAGC TCCACCTCCC GGGTTTATGC CATTCTCTCG TCTCAGCCTC
1201 CCGAGTAGCT GGGACTACAG GCACCCGCCA ACACGCCCGG STAATTTTTT GTATTTTTAG
1261 TAAAGACAGG GTTTCACCGT GTTAGCCAGG ATGGTCTCTA TCTCCTGACC TCGTGATCTG
1321 CCTGCCTTGG CTTATTATT TTTTTTTTAA GGACAGAGTC TCTCTCTGTC ACCCAGGCTG
1381 GAGTGCAATG GCGCGATCTT GGCTCACTGT AACTTCCACT TGCCAGGCTC AAGCAGTTCT

1441 CCTGCCTCAG CCTCCTGAGT AGCTGGGACT ACAGGCACCC GCCACCATGC CCAGCTAATT
1501 TTTGTATTTT TAGTAGAGAC AGAGTTTCAC CATATTAGCC TGGCTGGTCT CAAACTCCTG
1561 GCCTCAGGTG ATCTGCCCAC CTCGGCCTCC CAAAGTGCTG GGATCAAATC CACTGTTAAT
1621 CATTAGGCTG AACTGTCTCT TATAGAATGA GGTCAAAGAC ACTCCAGTT GCAGGGAGGG
1681 TAGATGGCCC CACCCAGACC GAGAGACACA GTGATGACCT CAGCCTAGGG ACMCCAAAAA
1741 AAAAAAAAAA AAAAAA

```

Фиг. 10. Нуклеотидная последовательность и трансляция кДНК-вставки в Image 4558384. Положение 87 п.н-вставки по сравнению с предсказанным INSP002 показано в затененной рамке. Последовательность Alu в 3'-некодирующей последовательности показана курсивом.

Фиг. 11. Сопоставление последовательностей предсказанного EVSP002 (верхняя строка)
с Image: 4558384 (BC025333.1) (нижняя строка)

```

INSP00 -----CAGAAATGCCTCCCAGGCTATCCAGGAG
BC0253 GGGAGACCTGGAAGGAAGCGACTGCACTGATCCAGAAATGCCTCCCAGGCTATCCAGGAG
      10      20      30      40      50      60

      30      40      50      60      70      80
INSP00 GGGCCAAGAGATTAAAGCAGGTTCAGAAGGCTCAGATGCCACTCACCAGACAGCAGGGT
BC0253 GGGCCAAGAGATTAAAGCAGGTTCAGAAGGCTCAGATGCCACTCACCAGACAGCAGGGT
      70      80      90     100     110     120

      90     100     110     120     130     140
INSP00 CGACTGCTAGTGACCTTGAGCCCAGTCCGGACAGACAGAGCAGGCAGACAGCACGGAC
BC0253 CGACTGCTAGTGACCTTGAGCCCAGTCCGGACAGACAGAGCAGGCAGACAGCACGGAC
      130     140     150     160     170     180

      150     160     170     180     190     200
INSP00 AAGCAGATGCTCCTTGCCAGCTATCCAATCTTCTGTGCTGCTTAGCGGGGCCCTGCCT
BC0253 AAGCAGATGCTCCTTGCCAGCTATCCAATCTTCTGTGCTGCTTAGCGGGGCCCTGCCT
      190     200     210     220     230     240

      210     220     230     240     250     260
INSP00 ACAGGCTCAGGGAGGCTGAACCCAGTCTCCTCGACCTCAGTCTGGGCTGCAGCCAAT
BC0253 ACAGGCTCAGGGAGGCTGAACCCAGTCTCCTCGACCTCAGTCTGGGCTGCAGCCAAT
      250     260     270     280     290     300

      270     280     290     300     310     320
INSP00 CAGACCTGGGCTCTGGGCCAGGGGCCCTGCCCCACTGGTGCCAGCTTCTGCCCTTGGG
BC0253 CAGACCTGGGCTCTGGGCCAGGGGCCCTGCCCCACTGGTGCCAGCTTCTGCCCTTGGG
      310     320     330     340     350     360

      330     340     350     360     370     380
INSP00 AGCTGGAAGGCCTTCTTGGGCTGCAGAAAGCCAGGCAGCTGGGGATGGGCAGGCTGCAG
BC0253 AGCTGGAAGGCCTTCTTGGGCTGCAGAAAGCCAGGCAGCTGGGGATGGGCAGGCTGCAG
      370     380     390     400     410     420

      390     400     410     420     430     440
INSP00 CGTGGGCAAGACGAGGTGGCTGCTGTGACTCTGCCGCTGAACCCCTCAGGAAGTGATCCAG
BC0253 CGTGGGCAAGACGAGGTGGCTGCTGTGACTCTGCCGCTGAACCCCTCAGGAAGTGATCCAG
      430     440     450     460     470     480

      450     460     470
INSP00 GGGATGTGTAAGGCTGTGCCCTTCGTTTCAG-----
BC0253 GGGATGTGTAAGGCTGTGCCCTTCGTTTCAGACACGGGAGTCTCGCTATGTTGCCCAAGCT
      490     500     510     520     530     540

      480
INSP00 -----GTGTT
BC0253 AGTCTTGAGCTTCTGGCCTCAAGCAATCCTCCACCTCAGCCTCCTGAGTTCTAGGTGTT
      550     560     570     580     590     600

      490     500     510     520     530     540
INSP00 CTCCCGGCCCGGCTGCTCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTGCTC
BC0253 CTCCCGGCCCGGCTGCTCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTGCTC
      610     620     630     640     650     660

```

```

      550      560      570      580      590      600
INSP00 CTCTCTCTACATCCCTGGCTCGGACCCACCCCACTAGTCCTGTGCAACAGCTGTATGCC
      :
BC0253 CTCTCTCTACATCCCTGGCTCGGACCCACCCCACTAGTCCTGTGCAACAGCTGTATGCC
      670      680      690      700      710      720

      610      620      630      640      650      660
INSP00 TGCTCGCAAGCGTTGGGCACCCGTGGTCCTGTGGTGTCTCACTGGCAGCTCAGCCTCCCG
      :
BC0253 TGCTCGCAAGCGTTGGGCACCCGTGGTCCTGTGGTGTCTCACTGGCAGCTCAGCCTCCCG
      730      740      750      760      770      780

      670      680      690      700      710      720
INSP00 TCGACGGGTGAAGATATCCACCATGCTGATCGAGGGGTGTCACTGCAGCCCAAAAGCATG
      :
BC0253 TCGACGGGTGAAGATATCCACCATGCTGATCGAGGGGTGTCACTGCAGCCCAAAAGCATG
      790      800      810      820      830      840

INSP00 A-----
      :
BC0253 AACTGAGCATCTGGATGGGTGCACGGAGACACGCACCTTGGAGAAATGAGGGGAGATGGA
      850      860      870      880      890      900

INSP00 -----
BC0253 CCAAGAAAGACGTGGACCTGGATGATGTACTCTGGGTCAAGAGACCAGGGATGCAGGGTT
      910      920      930      940      950      960

INSP00 -----
BC0253 AGGCAGACAGGTCCCCAGAGTCTCACCCTGCTCCCCAGACAGTAGACACAGTGCCCGTC
      970      980      990      1000      1010      1020

INSP00 -----
BC0253 CTGGAGTTGCACCACTGATAGTCACAGCACACAATGATTGACAACTCACTTTTTTTTTT
      1030      1040      1050      1060      1070      1080

INSP00 -----
BC0253 TTTTGGAGATGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCAATCTCAGCT
      1090      1100      1110      1120      1130      1140

INSP00 -----
BC0253 CACTGCAAGCTCCACCTCCCGGGTTTATGCCATTCTCCTGTCTCAGCCTCCCGAGTAGCT
      1150      1160      1170      1180      1190      1200

INSP00 -----
BC0253 GGGACTACAGGCACCCGCCAACACGCCCGGCTAATTTTTTGTATTTTAGTAAAGACAGG
      1210      1220      1230      1240      1250      1260

INSP00 -----
BC0253 GTTTCACCGTGTTAGCCAGGATGGTCTCTATCTCCTGACCTCGTGATCTGCCTGCCTTGG
      1270      1280      1290      1300      1310      1320

INSP00 -----
BC0253 CCTTATTATTTTTTTTTTAAGGACAGAGTCTCTCTGTGTCACCCAGGCTGGAGTGCAATG
      1330      1340      1350      1360      1370      1380

```

```

INSP00 -----
BC0253 GCGCGATCTTGGCTCACTGTAACCTTCCACTTGCCAGGCTCAAGCAGTTCTCCTGCCTCAG
      1390      1400      1410      1420      1430      1440

INSP00 -----
BC0253 CCTCCTGAGTAGCTGGGACTACAGGCACCCGCCACCATGCCAGCTAATTTTGTATTTT
      1450      1460      1470      1480      1490      1500

INSP00 -----
BC0253 TAGTAGAGACAGAGTTTCACCATATTAGCCTGGCTGGTCTCAAACCTCCTGGCCTCAGGTG
      1510      1520      1530      1540      1550      1560

INSP00 -----
BC0253 ATCTGCCACCTCGGCCTCCCAAAGTGCTGGGATCAAATCCACTGTTAATCATTAGGCTG
      1570      1580      1590      1600      1610      1620

INSP00 -----
BC0253 AACTGTCTCTTATAGAATGAGGTCAAAGACACTCCAGTTGCAGGGAGGGTAGATGGCCC
      1630      1640      1650      1660      1670      1680

INSP00 -----
BC0253 CACCCAGACCGAGAGACACAGTGATGACCTCAGCCTAGGGACACCAAAAAAAAAAAAAAA
      1690      1700      1710      1720      1730      1740

INSP00 -----
BC0253 AAAAAA

```

Фиг. 11a. Нуклеотидные последовательности

```

4558384_OR      MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW
INSP002_PR      MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW
                *****

4558384_OR      KAFLGLQKARQLGMGRLQRQDEVAAVTLPLNPQEVIQGMCKAVPFVQTRE-----
INSP002_PR      KAFLGLQKARQLGMGRLQRQDEVAAVTLPLNPQEVIQGMCKAVPFVQVFSRPGCSAIRL
                *****

4558384_OR      -----SRYVAQAS---LELLASSNPP-----TSAS-----
INSP002_PR      RNHLCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSASRRRVKISTMLI
                * *           * * * * *           ***

4558384_OR      -----
INSP002_PR      EGCHCSPKA

```

Фиг. 11b. Последовательности белка

```

1  GTCGACTGCT AGTGACCTTG AGCCCACTCC GGACAGACAG ACAGGCAGAC AGACGCACGG
61  ACAAGCAGAT GTCCTTGGC CAGCTATCCA CTCTTCTGTG CTGCTTAGC GGGGCCCTGC
    м l l g q l s t l l c l l s g a l

121 CTACAGGCTC AGGGAGGCCT GAACCCAGT CTCCTCGACC TCAGTCCTGG GCTGCAGCCA
    p t g s g r p e p q s p r p q s w a a a

181 ATCAGACCTG GGCTCTGGGC CCAGGGGCCC TGCCCCACT GGTGCCAGCT TCTGCCCTTG
    n q t w a l g p g a l p p l v p a s a l

241 GGAGCTGGAA GGCCTTCTTG GGCCTGCAGA AAGCCAGGCA GCTGGGGATG GGCAGGCTGC
    g s w k a f l g l q k a r q l g m g r l

301 AGCGTGGGCA AGACGAGGTG GCTGCTGTGA CTCTGCCGCT GAACCCCTCAG GAAGTGATCC
    q r g q d e v a a v t l p l n p q e v i

361 AGGGGATGTG TAAGGCTGTG CCCTTCGTTT TCTCCCGGCC CGGCTGTCTA GCCATACGCC
    q g m c k a v p f v l s r p g c s a i r

421 TCCGAAATCA TCTGTGCTTT GGTCAATTGCT CCTCTCTCTA CATCCCTGGC TCGACCCCA
    l r n h l c f g h c s s l y i p g s d p

481 CCCCACTAGT CCTGTGCAAC AGCTGTATGC CTGCTCGCAA GCGTTGGGCA CCCGTGGTCC
    t p l v l c n s c m p a r k r w a p v v

541 TGTGGTGTCT CACTGGCAGC TCAGCCTCCC GTCGACGGGT GAAGATATCC ACCATGCTGA
    l w c l t g s s a s r r r v k i s t m l

601 TCGAGGGGTG TCACTGCAGC CAAAAGCAT GAACTGAGCA TCTGGATGGG TGCACGGAGA
    i e g c h c s p k a

661 CACGCACCTT GGAGAAATGA GGGGAGATGG ACCAAGAAAG ACGTGGACCT GGATGATGT

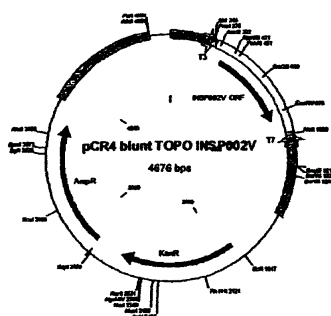
```

Фиг. 12. Нуклеотидная последовательность и трансляция последовательности INSP002V, генерированной с помощью ПЦР из Image 4558384

Молекула: pCR4 blunt TOPO INSP002V, кольцевая ДНК 4676 п.н.
 Название файла: pCR4 blunt TOPO INSP002V.cm5
 Описание: Лигирование продукта INSP002-nestFR в pCR4blunt-Topo
 Примечания: Плазмида ID 13075

Особенности молекулы:

Тип	Начало	Конец	Название	Описание
Область	2	216		область lac-промотора
Область	205	221		сайт обратного праймирования M13
Область	217	294		гибридный ген LacZa-ccdB
Маркер	243		T3	
Область	262	294		полилинкер
Область	294	294		сайт клонирования TOPO
Область	295	1013		ПЦР-продукт INSP002
Ген	363	923	INSP002V ORF	
Область	1014	1529		'гибридный ген LacZa-ccdB
Область	1014	1031		'полилинкер
Область	1014	1014		'сайт клонирования TOPO
Маркер	1066		C T7	
Область	1089	1074	C	- Сайт прямого праймирования 20M13
Ген	1878	2672	KanR	
Область	2864	2868		Сайт связывания с рибосомой
Ген	2876	3736	AmpR	
Область	3881	4554		Сайт инициации pUC



Фиг. 13. Карта pCR4blunt-Topo-INSP002V

Фиг. 14. Сопоставление предсказанной последовательности INSP002 (верхняя строка) с последовательностью варианта INSP002V (нижняя строка)

```

INSPpr CAGAAATGCCTCCCAGGCTATCCAGGAGGGCCAAGAGATTAAGCAGGTTTCAGAAGGC
INSPen -----

          70      80      90      100     110     120
INSPpr TCAGATGCCACTCACCAGACAGCAGGGTCGACTGCTAGTGACCTTGAGCCCAGTCCGGAC
          :
INSPen -----GTCGACTGCTAGTGACCTTGAGCCCAGTCCGGAC
                   10      20      30

          130     140     150     160     170     180
INSPpr AGACAGACAGGCAGACAGACGACGGACAAGCAGATGCTCCTTGGCCAGCTATCCACTCT
          :
INSPen AGACAGACAGGCAGACAGACGACGGACAAGCAGATGCTCCTTGGCCAGCTATCCACTCT
          40      50      60      70      80      90

          190     200     210     220     230     240
INSPpr TCTGTGCCTGCTTAGCGGGGCCCTGCCTACAGGCTCAGGGAGGCCTGAACCCAGTCTCC
          :
INSPen TCTGTGCCTGCTTAGCGGGGCCCTGCCTACAGGCTCAGGGAGGCCTGAACCCAGTCTCC
          100     110     120     130     140     150

          250     260     270     280     290     300
INSPpr TCGACCTCAGTCCTGGGCTGCAGCCAATCAGACCTGGGCTCTGGGCCCAGGGGCCCTGCC
          :
INSPen TCGACCTCAGTCCTGGGCTGCAGCCAATCAGACCTGGGCTCTGGGCCCAGGGGCCCTGCC
          160     170     180     190     200     210

          310     320     330     340     350     360
INSPpr CCCACTGGTGCCAGCTTCTGCCCTTGGGAGCTGGAAGGCCTTCTTGGGCCTGCAGAAAGC
          :
INSPen CCCACTGGTGCCAGCTTCTGCCCTTGGGAGCTGGAAGGCCTTCTTGGGCCTGCAGAAAGC
          220     230     240     250     260     270

          370     380     390     400     410     420
INSPpr CAGGCAGCTGGGGATGGGCAGGCTGCAGCGTGGGCAAGACGAGGTGGCTGCTGACTCT
          :
INSPen CAGGCAGCTGGGGATGGGCAGGCTGCAGCGTGGGCAAGACGAGGTGGCTGCTGACTCT
          280     290     300     310     320     330

```

Фиг. 14а. Нуклеотидные последовательности

```

      430      440      450      460      470      480
INSPpr GCCGCTGAACCCTCAGGAAGTGATCCAGGGGATGTGTAAGGCTGTGCCCTTCGTTACAGGT
      :
INSPen GCCGCTGAACCCTCAGGAAGTGATCCAGGGGATGTGTAAGGCTGTGCCCTTCGTTTC----
      340      350      360      370      380      390

      490      500      510      520      530      540
INSPpr GTTCTCCCGGCCCGGCTGCTCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTG
      :
INSPen --TCTCCCGGCCCGGCTGCTCAGCCATACGCCTCCGAAATCATCTGTGCTTTGGTCATTG
      400      410      420      430      440

      550      560      570      580      590      600
INSPpr CTCCTCTCTCTACATCCCTGGCTCGGACCCACCCCACTAGTCCTGTGCAACAGCTGTAT
      :
INSPen CTCCTCTCTCTACATCCCTGGCTCGGACCCACCCCACTAGTCCTGTGCAACAGCTGTAT
      450      460      470      480      490      500

      610      620      630      640      650      660
INSPpr GCCTGCTCGCAAGCGTTGGGCACCCGTGGTCCTGTGGTGTCTCACTGGCAGCTCAGCCTC
      :
INSPen GCCTGCTCGCAAGCGTTGGGCACCCGTGGTCCTGTGGTGTCTCACTGGCAGCTCAGCCTC
      510      520      530      540      550      560

      670      680      690      700      710      720
INSPpr CCGTCGACGGGTGAAGATATCCACCATGCTGATCGAGGGGTGTCAGTGCAGCCCAAAAGC
      :
INSPen CCGTCGACGGGTGAAGATATCCACCATGCTGATCGAGGGGTGTCAGTGCAGCCCAAAAGC
      570      580      590      600      610      620

INSPpr ATGA-----
      :
INSPen ATGAACTGAGCATCTGGATGGGTGCACGGAGACACGCACCTTGGAGAAATGAGGGGAGAT
      630      640      650      660      670      680

INSPpr -----

INSPen GGACCAAGAAAGACGTGGACCTGGATGATGT
      690      700      710

INSP002V 23 MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW
INSP002  1 MLLGQLSTLLCLLSGALPTGSGRPEPQSPRPQSWAAANQTWALGPGALPPLVPASALGSW
      *****

INSP002V 83 KAFLGLQKARQLGMRLQRGQDEVAATLPLNPQEV IQGMCKAVPF--VLSRPGCSAIRL
INSP002 61 KAFLGLQKARQLGMRLQRGQDEVAATLPLNPQEV IQGMCKAVPFVQVFSRPGCSAIRL
      *****

INSP002V 141 RNHLCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSASRRRVKISTMLI
INSP002 121 RNHLCFGHCSSLYIPGSDPTPLVLCNSCMPARKRWAPVVLWCLTGSSASRRRVKISTMLI
      *****

INSP002V 201 EGCHCSPKA
INSP002 181 EGCHCSPKA
      *****

```

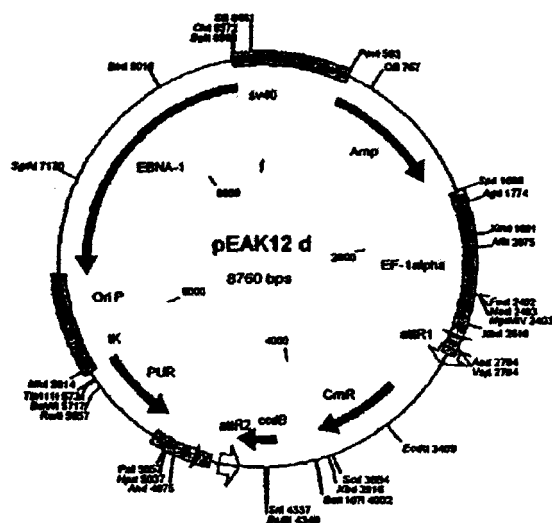
Фиг. 14b. Последовательности белка

Молекула: pEAK12 d, кольцевая ДНК 8760 п.н.
 Название файла: pEAK12DEST.cm5

Описание: Вектор для экспрессии в клетках
 млекопитающего (плазмиды ID 11345)

Особенности молекулы:

Тип	Начало	Конец	Название	Описание
Область	2	595		pmh-ori
Ген	596	1519	Amp	
Область	1690	2795	EF-1alpha	
Область	2703	2722		положение праймера pEAK12F MCS
Область	2796	2845		
Маркер	2855		attR1	
Ген	3256	3915	CmR	
Ген	4257	4562	ccdB	
Маркер	4603		C attR2	
Область	4733	4733		MCS
Область	4734	5162		poly A/сплайсинг
Область	4819	4848	C	положение праймера pEAK12R
Ген	5781	5163	C PUR	PUROMYCIN
Область	6005	5782	C tK	промотор tK
Область	6500	6006	C Ori P	
Ген	8552	6500	C EBNA-1	
Область	8553	8752	sv40	



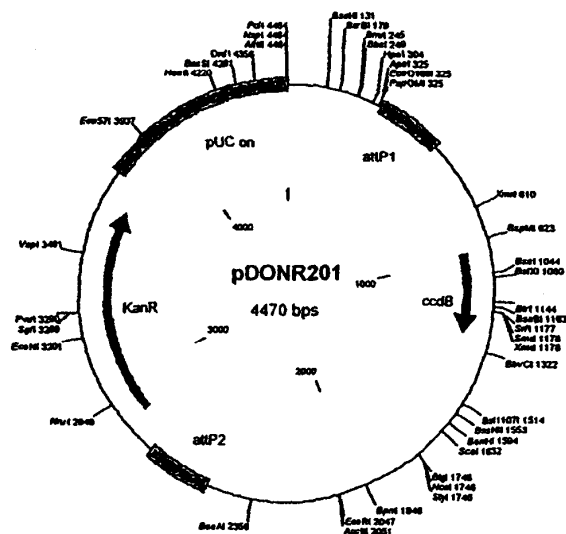
Фиг. 15. Карта экспрессирующего вектора pEAK12d

Молекула: pDONR201, кольцевая ДНК 4470 п.н.
 Название файла: pDONR201.cm5

Описание: вводный вектор Gateway (Invitrogen)-плазмиды ID 13309

Особенности молекулы:

Тип	Начало	Конец	Название
Область	332	563	attP1
Ген	959	1264	ccdB
Область	2513	2744	attP2
Ген	2868	3677	KanR
Область	3794	4467	pUC ori

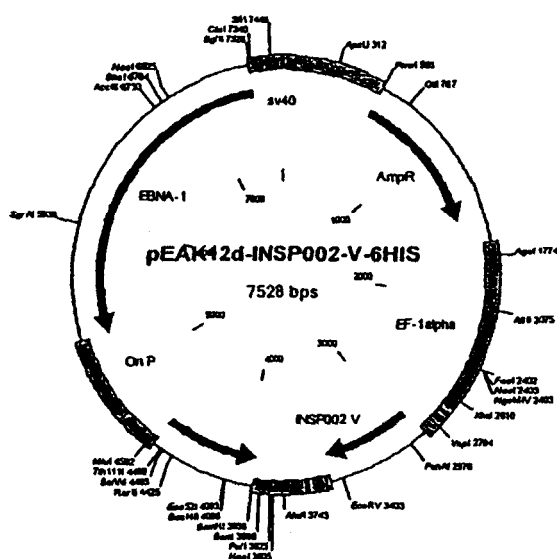


Фиг. 16. Карта вектора pDONR201 Gateway

Молекула: pEAK12d-INSP002-V-6HIS, кольцевая ДНК 7528 п.н.
 Название файла: pEAK12d-INSP002V-6HIS.cm5
 Описание: кДНК встроена между сайтами attR1 и attR2
 Примечание: плазмиды ID 13227

Особенности молекулы:

Тип	Начало	Конец	Название	Описание
Область	2	595		pmb-ori
Ген	596	1519	AmpR	
Область	1690	2795	EF-1alpha	промотор
Область	2703	2722		pEAK12F
Область	2796	2845		MCS''
Область	2855	2874		attB1
Ген	2888	3469	INSP002-V	
Область	3474	3495		attB2
Область	3501	3501		'MCS
Область	3502	3930		poly A/ сплайсинг
Область	3616	3597 C		pEAK12R
Ген	4549	3931 C		Пуромицин R
Область	4773	4550 C		промотор tK
Область	5268	4774 C	Ori P	
Ген	7320	5268 C	EBNA-1	
Область	7321	7520	sv40	



Фиг. 17. Карта pEAK12d-INSP002-V-6HIS

```

1  GATGCTCCTT GGCCAGCTAT CCACTCTTCT GTGCCTGCTT AGCGGGGCCC TGCCTACAGG
   m l l g q l s t l l c l l s g a l p t
   INSP002-FL-F
61  CTCAGGGAGG CCTGAACCCC AGTCTCCTCG ACCTCAGTCC TGGGCTGCAG CCAATCAGAC
   g s g r p e p q s p r p q s w a a a n q
121 CTGGGCTCTG GGCCAGGGG CCCTGCCCCC ACTGGTGCCA GCTTCTGCCC TTGGGAGCTG
   t w a l g p g a l p p l v p a s a l g s
181 GAAGGCCTTC TTGGGCCTGC AGAAAGCCAG GCAGCTGGGG ATGGGCAGGC TGCAGCGTGG
   w k a f l g l q k a r q l g m g r l q r
241 GCAAGACGAG GTGGCTGCTG TGA CTCTGCC GCTGAACCCT CAGGAAGTGA TCCAGGGGAT
   g q d e v a a v t l p l n p q e v i q g
301 GTGTAAGGCT GTGCCCTTCG TTCAGGTGTT CTCCCGGCCC GGCTGCTCAG CCATACGCCT
   m c k a v p f v q v f s r p g c s a i r
361 CCGAAATCAT CTGTGCTTTG GTCAFTGCTC CTCTCTCTAC ATCCCTGGCT CGGACCCAC
   l r n h l c f g h c s s l y i p g s d p
421 CCCACTAGTC CTGTGCAACA GCTGTATGCC TGCTCGCAAG CGTTGGGCAC CCGTGGTCCT
   t p l v l c n s c m p a r k r w a p v v
481 GTGGTGTCTC ACTGGCAGCT CAGCCTCCCG TCGACGGGTG AAGATATCCA CCATGCTGAT
   l w c l t g s s a s r r r v k i s t m l
541 CGAGGGGTGT CACTGCAGCC CAAAAGCATG AACTGAGCAT CGTGGATGG
   i e g c h c s p k a
                        INSP002-FL-R

```

Фиг. 18. Полноразмерный INSP002, клонированный из кДНК человеческого сердца

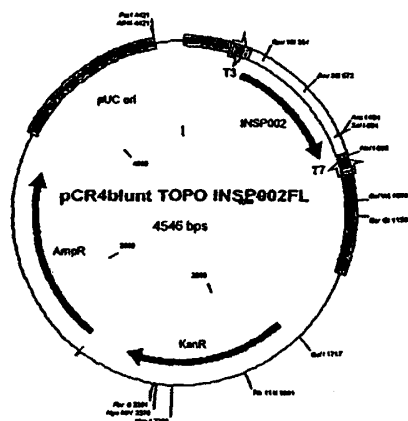
Молекула: pCR4blunt TOPO INSP002FL, кольцевая ДНК 4546 п.н.

Название файла: pCR4 blunt TOPO INSP002 FL.cm5,

Описание: Лигирование Noname в открытый pCR4blunt-TOPO*

Особенности молекулы:

Тип	Начало	Конец	Название	Описание
Область	2	216		область lac-промотора
Область	205	221		сайт обратного праймирования M13
Область	217	294		гибридный ген LacZa-ccdB
Маркер	243		T3	
Область	262	294		полилинкер
Область	294	294		сайт клонирования TOPO
Область	295	883		ПЦР-продукт INSP002-F3R3
Ген	296	865	INSP002	ОРС INSP002
Область	884	1399		'гибридный ген LacZa-ccdB
Область	884	901		полилинкер
Область	884	884		'сайт клонирования TOPO
Маркер	936		C T7	промотор T7
Область	959	944 C		Сайт прямого праймирования - 20M13
Ген	1748	2542	KanR	ОРС гена резистентности к канамицину
Область	2734	2738		Сайт связывания с рибосомой
Ген	2746	3606	AmpR	ОРС гена резистентности к ампициллину
Область	3751	4424	pUC ori	сайт инициации pUC



Фиг. 19



Евразийская патентная организация, ЕАПВ

Россия, 109012, Москва, Малый Черкасский пер., 2/6