(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.^{7}
C07K 14/705
C12N 15/12

(11) 공개번호	10-2005-0008699
(43) 공개일자	2005년01월21일

(21) 출원번호	10-2004-7017339		
(22) 출원일자	2004년10월28일		
번역문 제출일자	2004년10월28일		
(86) 국제출원번호	PCT/GB2003/001851	(87) 국제공개번호	WO 2003/093316
(86) 국제출원출원일자	2003년04월30일	(87) 국제공개일자	2003년11월13일

(30) 우선권주장	0209884.6	2002년04월30일	영국(GB)

(71) 출원인	아레스 트레이딩 에스.에이. 스위스, 쎄아슈-1170 오본느, 존 엥뒤스트리엘 드 루리에따
(72) 발명자	데이브스,앤드류로버츠 영국, 런던 더블유1티 2엔유 샤로테 스트리트 60, 인파마티카 파간,리차드조셉 영국, 런던 더블유1티 2엔유 샤로테 스트리트 60, 인파마티카 펠프스,크리스토퍼벤자민 영국, 런던 더블유1티 2엔유 샤로테 스트리트 60, 인파마티카 파워,크리스틴 프랑스, 티오리 에프-01710, 루에 데스 존퀼레스 10
(74) 대리인	강명구

심사청구 : 없음

(54) 면역글로불린-도메인 포함 세포 표면 인식 분자

요약

본 발명은 면역글로불린 도메인-포함 세포 표면 인식 분자로 동정된 신규한 단백질 INSP052 및 질병의 진단, 예방, 치료에서 상기 단백질과 이의 인코딩 유전자의 용도에 관한다.

명세서

기술분야

본 발명은 면역글로불린-도메인 포함 세포 표면 인식 분자로 확인된 신규한 단백질(INSP052와 INSP055) 및 질병의 진단, 예방 및 치료에서 인코딩 유전자로부터 유래된 이들 단백질과 핵산 서열의 용도에 관한다.

본원에 언급되고 있는 모든 공보, 특허 및 출원은 참고문헌으로 첨부된다.

배경기술

약물 개발 과정은 기능적 유전체학 시대가 도래하여 근본적으로 혁명 과정을 겪고 있다. '기능적 유전체학'이란 관심이 있는 단백질 서열에 기능을 부여하기 위해 생체 정보 도구를 이용하는 접근 방법에 이용된다. 이런 도구는 서열 데이터를 만드는 속도가 단백질 서열에 기능을 부여하기 위한 연구실의 능력을 너무 빨리 앞지르기 때문에 그 필요성이 점차 증가되고 있다.

생체 정보 도구는 능력 및 정확도가 크기 때문에, 생화학적 특징을 가지는 통상의 기술에 신속하게 대체되고 있다. 또한, 본 발명을 확인하는데 이용되는 진보된 생체 정보 도구는 더 높은 신뢰도를 부여할 수 있는 결과를 얻을 수 있다.

다양한 연구 기관 및 상업 조직에서는 이들이 이용할 수 있는 서열을 검사하고, 진행 단계에서 상당한 발견을 하고 있다. 그러나, 연구 및 약물 개발을 목적으 로 유전자 및 이들이 인코드하는 폴리펩티드를 확인하고 추가 특성화하는 작업이 지속적으로 요구된다.

최근에, 기능이 알려지지 않은 서열을 평가하는 주목할 만한 도구가 본원의 출원인에 개발되었다. 상기 도구는 Biopendium 검색 데이터베이스라고 하는 데이터베이스 시스템이다(WO 01/69507). 상기 데이터베이스 시스템은 특허 기술을 이용하고 설계되고 모든 가용한 단백질이나 핵산 서열로부터 발생된 정보를 보유하는 통합 데이터 자원으로 구성된다.

개별 데이터 자원으로부터 이들 서열 데이터를 통합하는 목적은 서열 자체 및 각 서열에 유관한 정보와 관련하여 가능한 많은 정보를 통합 자원으로 결합시키는 것이다. 인코드된 단백질의 3차원 구조에 관한 데이터를 비롯한 각 서열과 관련된 모든 가용한 데이터를 서로 통합하여 각 서열에서 공지된 정보를 최적으로 활용하고, 따라서 이들 서열의 비교로부터 가장 숙련된 예측을 할 수 있다. 데이터베이스에서 발생되고 각 서열 기입을 수반하는 주석은 서열 정보에 생물학적으로 유관한 관계를 부여한다.

이런 데이터 자원은 서열 단독으로부터 단백질 기능의 정확한 예측을 가능하게 한다. 전통적인 기술에서는 이런 예측이 동일 기능 집단에서 다른 단백질과 높은 서열 상동성(대략 20-30% 상동성)을 보이는 단백질에서만 가능하다. 공지된 기능의 다른 관련된 단백질에 매우 낮은 서열 상동성을 보이는 단백질에서는 정확한 예측이 불가능하다.

신호 펩티드-포함 단백질

세포외 단백질을 만들고 분비하는 세포의 능력은 많은 생물학적 과정에 중심한다. 효소, 성장 인자, 세포외 매트릭스 단백질, 신호전달 분자 모두 세포에 의해 분비된다. 이는 혈장 막과 분비 소포의 융합을 통하여 달성된다. 대부분의 경우에, 단백질은 신호 펩티드에 의해 소포체(endoplasmic reticulum)와 분비 소포(secretory vesicle)로 지향된다. 신호 펩티드는 세포질에서 막 결합된 구획, 예를 들면 분비 소포로 폴리펩티드 사슬의 수송에 영향을 주는 시스-작용 서열이다. 분비 소포로 표적되는 폴리펩티드는 세포외 매트릭스로 분비되거나 혈장 막에 유지된다. 혈장 막에 유지되는 폴리펩티드는 하나이상의 막통과 도메인을 보유한다. 세포의 기능에서 중추적인 역할을 수행하는 신호 펩티드-포함 펩티드의 예는 사이토킨, 호르몬, 세포외 매트릭스 단백질, 부착 분자, 수용체, 프로테아제, 성장 인자, 분화 인자이다.

면역글로불린 도메인-포함 세포 표면 인식 분자

면역글로불린 도메인-포함 세포 표면 인식 분자는 질병 과정과 연관된 다양한 생물학적 기능에서 중요한 역할을 하는 것으로 밝혀졌다. 이들의 활성 변화는 질병 표현형을 변화시키는 수단이고, 따라서 다양한 질환, 특히 염증 질환, 종양, 심혈관 질환에서 중요한 역할을 하는 새로운 면역글로불린 도메인-포함 세포 표면 인식 분자의 확인은 매우 적절하다.

면역글로불린 도메인-포함 세포 표면 인식 분자는 여러 생물학적 과정에 관여한다: 배형성(Martin-Bermudo, M. D. *et al,* Development. 2000 127(12): 2607-15; Chen, L. M., *et al* ., J Neurosci. 2000 20(10): 3776-84; Zweegman, S., *et al,* Exp Hematol. 2000 28(4): 401-10; Darribere, T., *et al.,* Biol Cell. 2000 92(1): 5-25), 조직 일체성의 유지(Eckes, B., *et al.,* J Cell Sci. 2000 113(Pt 13): 2455-2462; Buckwalter, J. A., *et al.,* Instr Course Lect. 2000 49: 481-9; Frenette, P. S., *et al.,* . J Exp Med. 2000 191(8): 1413-22; Delmas, V., *et al,* Dev Biol. 1999 216(2): 491-506; Humphries, M. J., *et al.,* Trends Pharmacol Sci. 2000 21(1): 29-32; Miosge, N., *et al,* Lab Invest. 1999 79(12): 1591-9; Nagaoka T, *et al* . Am J Pathol 2000 Jul 157: 1237-47; Nwariaku FE, *et al* . J Trauma 1995 39(2): 285-8; Zhu X, *et al* . Zhonghua Zheng Xing Shao Shang Wai Ke Za Zhi 1999 15(1): 53-5), 백혈구 일출(extravasation)/염증(Lim, L.H., *et al.* Am J Respir Cell Mol Biol. 2000 22(6): 693-701; Johnston, B., *et al.,* Microcirculation. 2000 7(2): 109-18; Mertens, A.V., *et al.,* Clin Exp Allergy. 1993 23(10): 868-73; Chcialowski, A., *et al.,* Pol Merkuriusz Lek. 2000 7(43): 13-7; Rojas, A.I., *et al,* Crit Rev Oral Biol Med. 19

99 10(3): 337-58; Marinova-Mutafchieva, L., *et al.,* Arthritis Rheum. 2000 43(3): 638-44; Vijayan, K.V., *et al,* J Clin Invest. 2000 105(6): 793-802; Currie, A.J., *et al,.* J Immunol. 2000 164(7): 3878-86; Rowin, M.E., *et al* ., Inflammation. 2000 24(2): 157-73; Johnston, B., *et al.,* J Immunol. 2000 164(6): 3337-44; Gerst, J.L., *et al.,* J Neurosci Res. 2000 59(5): 680-4; Kagawa, T.F., *et al.,* Proc Natl Acad Sci U S A. 2000 97(5): 2235-40; Hillan, K.J., *et al.,* Liver. 1999 19(6): 509-18; Panes, J., 1999 22(10): 514-24; Arao, T., *et al.,* J Clin Endocrinol Metab. 2000 85(1):382-9; Souza, H.S., *et al.,* Gut. 1999 45(6): 856-63; Grunstein, M.M., *et al.,* Am J Physiol Lung Cell Mol Physiol. 2000 278(6): L1154-63; Mertens, A.V., *et al.,* Clin Exp Allergy. 1993 23(10): 868-73; Berends, C., *et al.,* Clin Exp Allergy. 1993 23(11): 926-33; Fernvik, E., *et al.,* Inflammation.2000 24(1): 73-87; Bocchino, V., *et al.,* J Allergy Clin Immunol. 2000 105(1 Pt 1): 65-70; Jones SC, *et al* , Gut 1995 36(5): 724-30; Liu CM, *et al* , Ann Allergy Asthma Immunol 1998 81(2): 176-80; McMurray RW Semin Arthritis Rheum 1996 25(4): 215-33; Takahashi H, *et al* Eur J Immunol 1992 22(11): 2879-85; Carlos T, *et al* J Heart Lung Transplant 1992 11(6): 1103-8; Fabrega E, *et al* , Transplantation 2000 69(4): 569-73; Zohrens G, *et al* , Hepatology 1993 18(4): 798-802; Montefort S, *et al* . Am J Respir Crit Care Med 1994149(5): 1149-52), 종양형성(Orr, F.W., *et al.,* Cancer. 2000 88(S12): 2912-2918; Zeller, W., *et al.,* J Hematother Stem Cell Res. 1999 8(5): 539-46; Okada, T., *et al.,* Clin Exp Metastasis. 1999 17(7): 623-9; Mateo, V., *et al.,* Nat Med. 1999 5(11): 1277-84; Yamaguchi, K., *et al.,* J Exp Clin Cancer Res. 2000 19(1): 113-20; Maeshima, Y., *et al.,* J Biol Chem. 2000 275(28): 21340-8; Van Waes, C., *et al.,* Int J Oncol. 2000 16(6): 1189-95; Damiano, J.S., *et al* ., Leuk Lymphoma. 2000 38(1-2): 71-81; Seftor, R.E., *et al.,* Cancer Metastasis Rev. 1999 18(3): 359-75; Shaw, L.M., J Mammary Gland Biol Neoplasia. 1999 4(4): 367-76; Weyant, M.J., *et al.,* Clin Cancer Res. 2000 6(3): 949-56), 혈관신생(Koch AE, *et al.* Nature 1995 376 (6540): 517-9; Wagener C amp; Ergun S. Exp Cell Res 2000 261(1): 19-24; Ergun S, *et al* . Mol Cell 2000 5(2): 311-20), 골 재흡수(Hartman GD, amp; Duggan ME. Expert Opin Investig Drugs 2000 9(6): 1281-91; Tanaka Y, *et al* . J Bone Miner Res 1995 10(10): 1462-9; Lark MW, *et al* . J Pharmacol Exp Ther 1999 291(2): 612-7; Raynal C, *et al.* Endocrinology 1996 137(6):2347-54; Ilvesaro JM *, et al* . Exp Cell Res 1998 242(1): 75-83), 신경 기능 장애(Ossege LM, *et al* . Int Immunopharmacol 2001 1: 1085-100; Bitsch A, *et al* , Stroke 1998 29: 2129-35; Iadecola C amp; Alexander M. Curr Opin Neurol 2001 14:89-94; Becker K, *et al* Stroke 2001 32(1): 206-11; Relton JK, *et al* Stroke 2001 32(1): 199-205; Hamada Y, *et al* J Neurochem 1996 66: 1525-31), 혈전형성(Wang, Y.G., *et al.,* J Physiol (Lond). 2000 526(Pt 1): 57-68; Matsuno, H., *et al.,* Nippon Yakurigaku Zasshi. 2000 115(3): 143-50; Eliceiri, B.P., *et al.,* Cancer J Sci Am. 2000 6(Suppl 3): S245-9; von Beckerath, N., *et al.,* Blood. 2000 95(11):3297-301; Topol, E.J., *et al.,* Am Heart J. 2000 139(6): 927-33; Kroll, H., *et al.,* Thromb Haemost. 2000 83(3):3 92-6), 박테리아 병원균의 숙주 세포로의 침입/부착(Dersch P, *et al* . EMBO J 1999 18(5): 1199-1213).

여러 면역글로불린 도메인-포함 세포 표면 인식 분자 집단의 구조와 기능의 상세한 특성화는 염증, 종양, 신경병, 면역, 심혈관 기능을 수반하는 질환의 치료에 사용되는 조절인자를 개발하려는 많은 제약 기업의 적극적인 프로그램을 유도하 였다. 면역글로불린 도메인-포함 세포 표면 인식 분자는 배형성에서 아폽토시스까지 생물학의 거의 모든 측면에 관여한다. 이들은 조직의 구조적 일체성과 항상성 기능에 필수적이다. 따라서, 면역글로불린 도메인-포함 세포 표면 인식 분자에서 결함은 질병을 초래하고, 여러 질병은 면역글로불린 도메인-포함 세포 표면 인식 분자 기능의 조절을 수반한다. 이런 집단의 구성원은 표 1에 기술한다.

면역글로불린 도메인-포함 세포 표면 인식 분자 집단은 여러 별개의 집단으로 구성된다. 이들 집단중에서, 일부는 소형 분자 추적성(tractibility)으로 인하여 제약분야에서 특히 주목을 받고 있다. 이들은 아래와 같다:

1. 면역글로불린 부착 분자는 인테그린(integrin)에 대한 반대 수용체(counter receptor)를 대표하고, 세포내 부착 분자(ICAM)와 혈관 세포 부착 분자(VCAM)를 포괄한다. 막은 가변적 숫자의 원형 면역글로불린-유사 세포외 도메인으로 구성된다. 이런 집단의 일부 구성원은 예로써 PECAM-1(CD31)과 NCAM이고, 동형 부착(homotypic adhesion)을 매개한다. 이런 집단의 다른 구성원은 예로써 ICAM-1과 VCAM-1이고, 인테그린과의 상호작용으로 부착을 매개한다.

2. 세포 표면 성장 인자 수용체. 성장 인자는 세포외에 기원하고, 생물학적 효과를 발휘하기 위하여 표적 세포의 혈장막에 위치하는 특정한 고친화성 수용체와 상호작용한다. 다양한 성장 인자 수용체의 분자 특성화에서 이들 수용체는 정의된 집단: 티로신 키나아제 수용체, G-단백질 결합된 7개 막통과 수용체, 세린/트레오닌 키나아제 수용체에 속하는 것으로 밝혀졌다. 티로신 키나아제 수용체는 티로신 키나아제 활성을 보유하는 세포외 도메인, 막통과 도메인, 세포내 도메인으로 특성화된다. VEGFR, PDGFR, FGFR, CSF-1R, c-KIT는 세포외 영역에 면역글로불린 도메인을 포함하는 티로신 키나아제 성장 인자 수용체의 전형이다. 성장 인자 기능의 조절 장애는 종양(Bartucci M *et al* , (2001) Cancer Res. Sep 15;61(18): 6747-54, Dias S *et al* ., (2001) Proc Natl Acad Sci U S A. Sep 11;98(19): 10857-62, Djavan B *et al.* , (2001) World J Urol. 19(4): 225-33), 염증(Fiocchi C. (2001) J Clin Invest. Aug;108(4): 523-6, Hodge S *et al.,* (2001) Respirology. Sep;6(3): 205-211, Fenwick SA *et al.,* (2001) J Anat. Sep; 199(Pt 3):231-40), 신경병(Cooper JD *et al.* , (2001) Proc Natl Acad Sci U S A 98(18): 10439-44, Fahnestock M *et al* , (2001) Mol Cell Neurosci 18(2): 210-20), 대사(Vickers MH *et al.* , (2001) Endocrinology. 142(9):3964-73)

를 비롯한 다양한 질병 표현형을 초래한다.

표 1

면역글로불린 도메인-포함 세포 표면 인식 분자

수용체	*리간드*	*분포*
ICAM-1 5 Ig 도메인	LFA-1 (CD11a/CD18) Mac-1 (CD11b/CD18), CD43	광범위, 내피 세포, 섬유아세포, 상피, 단핵구, 림프구, 수상돌기 세포, 연골세포.
ICAM-2 2 Ig 도메인	LFA-1 (CD11b)	내피 세포(多): 림프구, 단핵구, 호염기구, 혈소판(少).
ICAM-3 5 Ig 도메인	LFA-1 (αd/CD18)	림프구, 단핵구, 중성구, 호산구, 호염기구.
VCAM-1 6 또는 7 Ig 도메인	α4β1, α4β7	내피 세포, 단핵구, 섬유아세포, 수상돌기 세포, 골수 간질 세포, 근육모세포.
LFA-3 6 Ig 도메인	CD2	내피 세포, 백혈구, 상피 세포
PECAM-1 (CD31)	CD31, 헤파린	내피 세포 (EC-EC 접합에서), T 세포 하위집단, 혈소판, 중성구, 호산구, 단핵구, 평활근 세포, 골수 줄기 세포.
NCAM	NCAM, 헤파린 SO_4	신경 세포, 근육
MAdCAM-1 4 Ig 도메인	α4β7, L-셀렉틴	파이어 소절(Peyer's patch), 장간막 림프절, 점막 내피 세포, 비장.
CD2	CD58, CD59, CD48	T 림프구
VEGFR	VEGF	광범위, 망막, 배꼽 정맥, 부신, NT2 신경 선구 세포
FGFR	FGF	광범위, 뇌, 결장, 난소
KIT	줄기 세포 인자, MGF	광범위, 태아, 멜라닌세포, 방광, 소뇌, 위 상피(少)
PDGFR	PDGF	광범위, 유방, 태반, 섬유아세포, 폐, 난소, 피부, 심장
CSF-1R	CSF	광범위, 태반, 간, 다발성 경화증 병소, 비장, 폐, 유방.

이런 연유로, 면역글로불린 도메인-포함 세포 표면 인식 분자는 질병 과정과 연관된 다양한 생물학적 기능에서 중요한 역할을 하는 것으로 밝혀졌다. 이들의 활성 변화는 질병 표현형을 변화시키는 수단이고, 따라서 다양한 질환, 특히 염증 질환, 종양, 심혈관 질환, 중추 신경계 질환, 감염에서 중요한 역할을 하는 새로운 면역글로불린 도메인-포함 세포 표면 인식 분자의 확인은 매우 적절하다.

발명의 상세한 설명

본 발명은 INSP052와 INSP055 단백질이 면역글로불린 도메인-포함 세포 표면 인식 분자로써 기능한다는 발견에 기초한다. 면역글로불린 도메인-포함 세포 표면 인식 분자의 예는 표 1에 기재한다.

본 발명의 제 1 특징의 한 구체예에서, 아래의 폴리펩티드를 제공한다:

(i) SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 또는 SEQ ID NO:16에 상술된 아미노산 서열, 또는 INSP052의 세포외 도메인의 아미노산 서열을 포함하거나 이들 서열로 구성되는 폴리펩티드;

(ii) (i)에 따른 폴리펩티드의 활성을 보유하거나 (i)에 따른 폴리펩티드와 공통의 항원 결정부위를 보유하는 단편;

(iii) (i) 또는 (ii)의 기능적 등가체.

본원에서 '(i)에 따른 폴리펩티드의 활성'은 면역글로불린 도메인-포함 세포 표면 인식 분자 활성을 의미한다. 면역글로불린 도메인-포함 세포 표면 인식 분자는 면역글로불린 도메인-포함 세포 표면 인식 분자 집단에서 보존된 특징으로 확인될 수 있는 아미노산 서열 또는 구조적 특징을 포함하는 폴리펩티드를 의미한다.

SEQ ID NO:2에 상술된 서열을 보유하는 폴리펩티드는 'INSP052 엑손 1 폴리펩티드'라고 한다. SEQ ID NO:4에 상술된 서열을 보유하는 폴리펩티드는 'INSP052 엑손 2 폴리펩티드'라고 한다. SEQ ID NO:6에 상술된 서열을 보유하는 폴리펩티드는 'INSP052 엑손 3 폴리펩티드'라고 한다. SEQ ID NO:8에 상술된 서열을 보유하는 폴리펩티드는 'INSP052 엑손 4 폴리펩티드'라고 한다. SEQ ID NO:10에 상술된 서열 을 보유하는 폴리펩티드는 'INSP052 엑손 5 폴리펩티드'라고 한다. SEQ ID NO:12에 상술된 서열을 보유하는 폴리펩티드는 'INSP052 엑손 6 폴리펩티드'라고 한다. SEQ ID NO:14에 상술된 서열을 보유하는 폴리펩티드는 'INSP052 엑손 7 폴리펩티드'라고 한다. SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14를 통합하면 SEQ ID NO:16에 상술된 서열이 만들어진다. SEQ ID NO:16에 상술된 서열을 보유하는 폴리펩티드는 INSP052 폴리펩티드라고 한다.

본원에서 'INSP052 엑손 폴리펩티드'는 INSP052 엑손 1 폴리펩티드, INSP052 엑손 2 폴리펩티드, INSP052 엑손 3 폴리펩티드, INSP052 엑손 4 폴리펩티드, INSP052 엑손 5 폴리펩티드, INSP052 엑손 6 폴리펩티드, INSP052 엑손 7 폴리펩티드, INSP052 폴리펩티드, INSP052의 세포외 도메인으로 구성되거나 이들을 포함하는 폴리펩티드를 포괄한다.

한 구체예에서, 상기 구체예에 따른 폴리펩티드 SEQ ID NO:16에 상술된 아미노산 서열, 이의 단편 또는 기능적 등가체로 구성된다. 다른 구체예에서, 폴리펩티드는 SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12 또는 SEQ ID NO:14에 상술된 아미노산 서열, 또는 이들의 변이체로 구성된다.

본 발명의 제 1 특징의 한 구체예에서, 아래의 폴리펩티드를 제공한다:

(i) INSP052의 세포외 도메인의 아미노산 서열을 포함하거나 이들 서열로 구성되는 폴리펩티드;

(ii) (i)에 따른 폴리펩티드의 활성을 보유하거나 (i)에 따른 폴리펩티드와 공통의 항원 결정부위를 보유하는 단편;

(iii) (i) 또는 (ii)의 기능적 등가체.

INSP052의 세포외 도메인은 아미노산 1-240에 해당한다(실시예 참조). 또한, INSP052의 세포외 도메인과 관련하여 도 7을 참조한다.

본 발명의 제 1 특징의 2번째 구체예에서, 아래의 폴리펩티드를 제공한다:

(i) SEQ ID NO:18에 상술된 아미노산 서열을 포함하거나 이들 서열로 구성되는 폴리펩티드;

(ii) (i)에 따른 폴리펩티드의 활성을 보유하거나 (i)에 따른 폴리펩티드와 공통의 항원 결정부위를 보유하는 단편;

(iii) (i) 또는 (ii)의 기능적 등가체.

본원에서 '(i)에 따른 폴리펩티드의 활성'은 면역글로불린 도메인-포함 세포 표면 인식 분자 활성을 의미한다.

적절하게는, 상기 구체예에 따른 펩티드는 SEQ ID NO:18에 상술된 아미노산 서열, 이의 단편 또는 기능적 등가체로 구성된다. SEQ ID NO:18에 상술된 아미노산 서열을 보유하는 폴리펩티드는 이후 'INSP055 폴리펩티드'라고 한다.

본 발명의 제 2 특징에서, 본 발명의 제 1 특징의 폴리펩티드를 인코드하는 정제된 핵산 분자를 제공한다.

적절하게는, 정제된 핵산 분자는 SEQ ID NO:1에 상술된 핵산 서열(INSP052 A 엑손 1 폴리펩티드를 인코딩), SEQ ID NO:3에 상술된 핵산 서열(INSP052 엑손 2 폴 리펩티드를 인코딩), SEQ ID NO:5에 상술된 핵산 서열(INSP052 엑손 3 폴리펩티드를 인코딩), SEQ ID NO:7에 상술된 핵산 서열(INSP052 엑손 4 폴리펩티드를 인코딩), SEQ ID NO:9에 상술된 핵산 서열(INSP052 엑손 5 폴리펩티드를 인코딩), SEQ ID NO:11에 상술된 핵산 서열(INSP052 엑손 6 폴리펩티드를 인코딩), SEQ ID NO:13에 상술된 핵산 서열(INSP052 엑손 7 폴리펩티드를 인코딩), SEQ ID NO:15에 상술된 핵산 서열(INSP052 폴리펩티드를 인코딩) 또는 SEQ ID NO:17에 상술된 핵산 서열(INSP055 폴리펩티드를 인코딩)을 포함하거나 이들 서열, 이들의 다양한 등가체 또는 단편으로 구성된다.

SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13을 통합하면 SEQ ID NO:15에 상술된 서열이 만들어진다.

한 구체예에서, INSP052의 세포외 도메인을 하거나 이런 도메인으로 구성되는 폴리펩티드를 인코드하는 핵산 분자를 제공한다. 적절하게는, 핵산 분자는 도 7에 제시된 핵산 서열 또는 이의 코딩 영역을 포함하거나 이런 서열로 구성된다.

본 발명의 제 3 특징에서, 매우 높은 엄밀도 조건하에서 본 발명의 제 2 특징의 핵산 분자와 혼성화할 수 있는 정제된 핵산 분자를 제공한다.

본 발명의 제 4 특징에서, 본 발명의 제 2 또는 제 3 특징의 핵산 분자를 포함하는 벡터, 예를 들면 발현벡터를 제공한다.

본 발명의 제 5 특징에서, 본 발명의 제 4 특징의 벡터로 형질전환된 숙주 세포를 제공한다.

본 발명의 제 6 특징에서, 본 발명의 제 1 특징의 폴리펩티드에 특이적으로 결합하여 이의 활성을 저해하는 리간드를 제공한다.

본원에서 '본 발명에 따른 폴리펩티드의 활성'및 유사 표현은 면역글로불린 도메인-포함 세포 표면 인식 분자의 특징적인 활성을 의미한다.

본 발명의 제 7 특징에서, 본 발명의 제 1 특징의 폴리펩티드를 인코드하는 자연 유전자 발현을 변경하거나 본 발명의 제 1 특징의 폴리펩티드 활성을 조절하는데 효과적인 화합물을 제공한다.

본 발명의 제 7 특징의 화합물은 폴리펩티드의 활성 또는 유전자의 발현 수준을 증가(항진) 또는 감소(길항)시킬 수 있다.

중요하게는, INSP052와 INSP055 폴리펩티드의 기능 확인은 질병의 치료 또는 진단에 효과적인 화합물을 확인하는 스크리닝 방법의 설계를 가능하게 한다. 본 발명의 제 6과 제 7 특징에 따른 리간드와 화합물을 이런 방법으로 확인할 수 있다. 이들 방법은 본 발명의 특징에 포섭된다.

본 발명의 제 8 특징에서, 치료 또는 진단에 사용되는 본 발명의 제 1 특징의 폴리펩티드; 본 발명의 제 2 또는 제 3 특징의 핵산 분자; 본 발명의 제 4 특징의 벡터; 제 5 특징의 숙주 세포; 제 6 특징의 리간드; 제 7 특징의 화합물을 제공한다. 이들 분자는 세포 증식성 질환, 자가면역/염증 질환, 심혈관 질환, 신경과 정신 질환, 발달 장애, 유전 질환, 대사 질환, 감염 및 다른 병인성 질환을 포함하지만 이들에 국한되지 않는 질병의 치료 약물을 제조하는 데에도 사용할 수 있다. 적절하게는, 이들 질환은 신생물, 암, 뇌종양, 신경교종, 골암, 폐암, 유암암, 전립선암, 결장암, 혈관종, 골수증식성 질환, 백혈병, 혈액학적 질환, 호중구 감소 증, 혈소판 감소증, 혈관신생 질환, 피부 질환, 노화, 상처, 화상, 섬유증, 심혈관 질환, 재협착증, 심장 질환, 말초 혈관 질환, 관상 동맥 질환, 부종, 혈전색전증, 월경통, 자궁내막증, 전자간증, 폐 질환, COPD, 천식 골 질환, 신장 질환, 사구체신염, 간 질환, 크론병, 위염, 궤양성 대장염, 궤양, 면역 질환, 자가면역 질환, 관절염, 류머티스 관절염, 건선, 수포성 표피박리증, 전신 홍반성 루퍼스, 강직성 척추염, 라임병, 다발성 경화증, 신경퇴행, 뇌졸중, 뇌/척수 손상, 알츠하이머병, 파킨슨병, 운동 신경 질환, 신경근육 질환, HIV, AIDS, 시토메갈로바이러스 감염, 진균 감염, 안구 질환, 망막황반 이상, 녹내장, 당뇨성 망막병증, 안구 고혈압 및 면역글로불린 도메인-포함 세포 표면 인식 분자가 관여하는 다른 이상 등이다. 본 발명의 제 1, 제 2, 제 3, 제 4, 제 6 또는 제 7 특징의 이들 부분은 이런 질환의 치료를 위한 약물의 제조에도 사용될 수 있다.

제 9 특징에서, 환자에서 질병을 진단하는 방법을 제공하는데, 상기 방법은 환자의 조직에서 본 발명의 제 1 특징의 폴리펩티드를 인코드하는 자연 유전자의 발현 수준을 평가하거나 제 1 특징의 폴리펩티드 활성을 평가하고, 발현이나 활성 수준을 대조 수준과 비교하는 단계로 구성되고, 대조 수준과 상이한 수준은 질병을 암시한다. 이런 방법은 *in vitro* 에서 적절하게 실시할 수 있다. 환자에서 질병 치료 효과를 모니터하는데 유사한 방법을 이용할 수 있는데, 시간이 경과함에 따라 폴리펩티드 또는 핵산 분자의 활성이나 발현 수준이 대조 수치에 근접하는 것은 질병으로부터 회복을

암시한다.

본 발명의 제 1 특징의 폴리펩티드를 감지하는 바람직한 방법은 (a) 리간드- 폴리펩티드 복합체의 형성에 적합한 조건하에 리간드, 예를 들면 본 발명의 제 6 특징의 항체를 생물학적 시료와 접촉시키고; (b) 상기 복합체를 감지하는 단계로 구성된다.

당업자가 인지하는 바와 같이, 본 발명의 제 9 특징에 따른 다양한 방법, 예를 들면 짧은 프로브와의 핵산 혼성화 반응, 점 돌연변이 분석; 중합효소 연쇄 반응(PCR) 증폭 및 단백질 수준의 이상을 감지하는 항체를 이용하는 방법이 존재한다. 단기적 또는 장기적 계획하에 환자에서 모니터되는 질병을 치료하는데 유사한 방법을 이용할 수 있다. 또한, 본 발명은 이들 방법에서 질병을 진단하는데 유용한 키트를 제공한다.

적절하게는, 본 발명의 제 9 특징의 방법으로 진단된 질병은 면역글로불린 도메인-포함 세포 표면 인식 분자가 관여하는 상기한 질병이다.

본 발명의 제 10 특징에서, 면역글로불린 도메인-포함 세포 표면 인식 분자로써 본 발명의 제 1 특징의 폴리펩티드의 용도를 제공한다. 세포 표면 수용체에서 Ig 도메인의 중요성은 Lokker NA et al., 'Functional importance of platelet-derived growth factor (PDGF) receptor extracellular immunoglobulin-like domains. Identification of PDGF binding site and neutralizing monoclonal antibodies,' *J Biol Chem* 1997 Dec 26;272(52):33037-44.

또한, 본 발명은 면역글로불린 도메인-포함 세포 표면 인식 분자 활성을 보유하는 단백질을 발현하는 본 발명의 제 2 또는 제 3 특징에 따른 핵산 분자의 용도를 제공한다. 더 나아가, 본 발명은 면역글로불린 도메인-포함 세포 표면 인식 분자 활성을 달성하는 방법을 제공하는데, 상기 방법은 본 발명의 제 1 특징의 폴리펩티드를 사용하는 단계로 구성된다.

본 발명의 제 11 특징에서, 제약학적으로 수용가능한 담체와 함께, 본 발명의 제 1 특징의 폴리펩티드, 본 발명의 제 2 또는 제 3 특징의 핵산 분자, 본 발명의 제 4 특징의 벡터, 본 발명의 제 5 특징의 숙주 세포, 본 발명의 제 6 특징의 리간드, 또는 본 발명의 제 7 특징의 화합물을 함유하는 제약학적 조성물을 제공한다.

본 발명의 제 12 특징에서, 치료 또는 진단에 사용되는 본 발명의 제 1 특징의 폴리펩티드, 본 발명의 제 2 또는 제 3 특징의 핵산 분자, 본 발명의 제 4 특징의 벡터, 본 발명의 제 5 특징의 숙주 세포, 본 발명의 제 6 특징의 리간드, 또는 본 발명의 제 7 특징의 화합물을 제공한다. 이들 분자들은 신생물, 암, 뇌종양, 신경교종, 골암, 폐암, 유암암, 전립선암, 결장암, 혈관종, 골수증식성 질환, 백혈병, 혈액학적 질환, 호중구 감소증, 혈소판 감소증, 혈관신생 질환, 피부 질환, 노화, 상처, 화상, 섬유증, 심혈관 질환, 재협착증, 심장 질환, 말초 혈관 질환, 관상 동맥 질환, 부종, 혈전색전증, 월경통, 자궁내막증, 전자간증, 폐 질환, COPD, 천식 골 질환, 신장 질환, 사구체신염, 간 질환, 크론병, 위염, 궤양성 대장염, 궤양, 면역 질환, 자가면역 질환, 관절염, 류머티스 관절염, 건선, 수포성 표피박리증, 전신 홍반성 루퍼스, 강직성 척추염, 라임병, 다발성 경화증, 신경퇴행, 뇌졸중, 뇌/척수 손상, 알츠하이머병, 파킨슨병, 운동 신경 질환, 신경근육 질환, HIV, AIDS, 시토메갈로바이러스 감염, 진균 감염, 안구 질환, 망막황반 이상, 녹내장, 당뇨성 망막병증, 안구 고혈압 및 면역글로불린 도메인-포함 세포 표면 인식 분자가 관여하는 다른 이상을 포함하지만 이들에 국한되지 않는 질환의 치료를 위한 약물의 제조에도 사용될 수 있다.

본 발명의 제 13 특징에서, 제 1 특징의 폴리펩티드, 제 2 또는 제 3 특징의 핵산 분자, 제 4 특징의 벡터, 제 5 특징의 숙주 세포, 제 6 특징의 리간드 또는 제 7 특징의 화합물을 환자에게 투여하는 단계로 구성되는 환자의 질병을 치료하는 방법을 제공한다.

본 발명의 제 1 특징의 폴리펩티드를 인코드하는 자연 유전자의 발현이나 제 1 특징의 폴리펩티드의 활성이 건강한 사람의 발현이나 활성 수준과 비교하여 병든 환자에서 감소되는 질병의 경우, 환자에 투여되는 폴리펩티드, 핵산분자, 리간드 화합물은 항진제가 된다. 역으로, 제 1 특징의 폴리펩티드를 인코드하는 자연 유전자 발현이나 제 1 특징의 폴리펩티드의 활성이 건강한 사람의 발현이나 활성 수준과 비교하여 병든 환자에서 증가되는 질병의 경우, 환자에 투여되는 폴리펩티드, 핵산분자, 리간드 화합물은 길항제가 된다. 이런 길항제로는 안티센스 핵산 분자, 리보자임, 항체와 같은 리간드가 포함된다.

본 발명의 제 14 특징에서 제 1 특징의 폴리펩티드의 발현 수준을 더 높게, 더 낮게 또는 발현이 없도록 형질전환된 유전자도입 또는 유전자적중(knockout)된 사람-아닌 동물을 제공한다. 이런 유전자도입 동물은 질병 연구에 매우 유용한 모델이며, 질병의 치료 또는 진단에 효과적인 화합물을 확인하는 스크리닝 방법에도 이용될 수 있다.

적절하게는, 질병은 면역글로불린 도메인-포함 세포 표면 인식 분자가 관여하는 상기한 질병이다.

본 발명의 폴리펩티드와 핵산에서 추구되는 보호 범위는 천연 자원에 존재하는 핵산 또는 폴리펩티드에 적용되지 않

는다. 오히려, 본 발명에서 청구되는 폴리펩티드와 핵산은 '분리된' 또는 '정제된'폴리펩티드와 핵산으로 간주한다. 본 명세서에서 '분리된'과 '정제된'은 천연 자연의 핵산 또는 폴리펩티드에 존재하는 적어도 한가지 다른 성분(예, 핵산 또는 폴리펩티드)로부터 분리된 핵산 또는 폴리펩티드를 의미한다. 가령, 조직 추출물에 포함된 폴리펩티드는 합성이나 재조합 방식으로 만들어지는 경우에 '분리된'또는 '정제된' 폴리펩티드를 구성한다. 한 구체예에서, 핵산 또는 폴리펩티드는 용매, 완충액, 이온, 또는 동일 용액에 정상적으로 존재하는 다른 성분의 존재하에서만 발견된다.

'분리된'과 '정제된'은 폴리펩티드 또는 핵산이 수득되는 방법 또는 제조의 순도 수준을 의미하지 않는다. 따라서, 이런 분리된 또는 정제된 종류는 재조합되거나, 관심있는 세포 또는 조직으로부터 직접 분리되거나 또는 결정된 서열에 기초하여 합성될 수 있다.

본 발명을 실시하는데 이용할 수 있는 표준 기술 및 과정은 아래에 요약하였다. 인지하는 바와 같이, 본 발명에서 설명하고 있는 방법, 프로토콜, 세포 주, 벡터 및 시약에 국한되지 않는다. 본 명세서에 이용되는 용어는 본 발명의 특정 구체예를 설명하기 위한 것으로, 본 발명의 범위를 한정하지 않는다.

본 명세서에서는 뉴클레오티드 및 핵산에 대하여 표준 약어를 이용한다.

본 발명은 달리 명시하지 않는 경우에, 당분야에 공지된 분자 생물학, 미생물학, 재조합 DNA 기술, 면역학의 통상적인 기술을 이용하여 실시한다.

이런 기술은 기존 문헌에서 충분히 설명되고 있다. 적절한 참고 문헌에는 Sambrook Molecular Cloning; A Laboratory Manual, Second Edition (1989); DNA Cloning, Volumes I and 11 (D.N Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames amp; S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames amp; 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 amp; 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)등이 포함된다 .

본 명세서에서 '폴리펩티드'는 2개 이상의 아미노산이 펩티드 결합 또는 변형된 펩티드 결합(가령, peptide isosteres)에 의해 서로 연결된 임의의 펩티드 또는 단백질을 포괄한다. 폴리펩티드 용어에는 짧은 사슬(펩티드 및 올리고펩티드)과 긴 사슬(단백질)이 모두 포함된다.

본 발명의 폴리펩티드는 성숙한 단백질 형태이거나 또는 pre-, pro-, prepro- 부분의 절단으로 활성화되면 활성 성숙 폴리펩티드를 생산할 수 있는 pre-, pro- , prepro- 단백질 형태이다. 이런 폴리펩티드에서, pre-, pro-, prepro- 서열은 리더, 분비 서열이거나 또는 성숙한 폴리펩티드 서열의 정제에 이용되는 서열이다.

본 발명의 제 1 특징의 폴리펩티드는 융합 단백질 형태의 일부를 형성할 수도 있다. 가령, 분비 또는 리더 서열, pro- 서열, 정제에 도움이 되는 서열, 또는 예로써 재조합 생산동안 좀더 높은 단백질 안정성을 공여하는 서열을 비롯한 한가지 이상의 추가 아미노산을 포함하는 것이 유리하다. 대안으로, 성숙한 폴리펩티드를 다른 화합물, 예를 들면 폴리펩티드의 반감기를 증가시키는 화합물(가령, 폴리에틸렌 글리콜)과 융합시킬 수도 있다.

폴리펩티드에는 20개 유전자를 인코드하는 아미노산 이외의 아미노산도 포함될 수 있는데, 이들 아미노산은 자연적인 공정, 예를 들면 번역후 가공 또는 당분야에 공지된 화학적 변형 기에 의해 변형될 수 있다. 본 발명의 폴리펩티드에 통상적으로 존재할 수 있는 이런 공지의 변형에는 글리코실화반응, 지질 부착, 설페이션, 예로써 글루타민산 잔기의 감마-카르복실화반응, 수산화반응, ADP-리보실화반응 등이 포함된다. 다른 가능한 변형에는 아세틸화반응, 아실화반응, 아미데이션, 플라빈의 공유 부착, 헤미(haeme) 잔기의 공유 부착, 뉴클레오티드 또는 뉴클레오티드 유도체의 공유 부착, 지질 유도체의 공유 부착, 포스파티딜이노시톨의 공유 부착, 교차-연결, 고리화반응, 이황화결합 형성, 디메틸화반응, 공유 교차 연결의 형성, 시스테인 형성, 피로글루타메이트 형성, 포밀화반응, GPI 고정기 형성, 요오드화반응, 메틸화반응, 미리스톨화반응, 산화, 단백분해 공정, 포스포릴화반응, 프레닐화반응, 라셈화반응, 세리노일화반응, 아르기닐화와 같은 t-RNA 매개된 아미노산의 단백질에 추가 및 유비퀴닌화반응 등이 포함된다.

폴리펩티드의 모든 위치, 예를 들면 펩티드 골격, 아미노산 측쇄 및 아미노 또는 카르복시 말단에도 변형이 일어날 수 있다. 사실, 폴리펩티드에서 아미노 또는 카르복시 말단, 또는 양 말단 모두의 공유 변형에 의한 차단은 자연 발생 폴리펩티드와 합성 폴리펩티드에서 공통한데, 이런 변형은 본 발명의 폴리펩티드에도 존재한다.

폴리펩티드에서 일어날 수 있는 변형은 폴리펩티드가 어떻게 만들어지는 가에 따라 달라진다. 재조합에 의해 만들어진 폴리펩티드의 경우에, 상당 부분에서 변형의 성질과 정도는 특정 숙주 세포의 번역후 변형 능력 및 관심있는 폴리펩티드 아미노산 서열에 존재하는 변형 신호에 의해 결정된다. 가령, 상이한 유형의 숙주 세포간에 글리고실화 반응 패턴이 변한다.

본 발명의 폴리펩티드는 임의의 적절한 방법으로 만들 수 있다. 이런 폴리펩티드에는 자연 발생 폴리펩티드(예를 들면, 세포 배양물로부터 정제된 분리 폴리펩티드), 재조합에 의해 만들어진 폴리펩티드(융합 단백질 포함), 합성에 의해 만들어진 폴리펩티드 또는 이들 방법의 복합으로 만든 폴리펩티드 등이 포함된다.

본 발명의 제 1 특징의 기능적으로 등가인 폴리펩티드는 INSP052와 INS055 폴리펩티드에 상동성인 폴리펩티드이다. 폴리펩티드의 한 서열이 다른 폴리펩티드 의 서열과 동일하거나 유사성이 충분한 경우에, 이들 두 폴리펩티드는 '상동성'이라고 한다. '동일성'이란 배열된 서열의 임의 지점에서 서열간에 아미노산 잔기가 동일함을 의미한다. '유사성'이란 배열된 서열의 임의 지점에서 아미노산 잔기가 유사한 유형임을 의미한다. 서열의 동일성 및 유사성 정도는 용이하게 계산할 수 있다(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; Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991).

따라서, 상동성 폴리펩티드에는 천연 생물학적 변이체(가령, 폴리펩티드가 유도된 종내에 대립형질 변이체 또는 지형적 변이체), INSP052와 INSP055 폴리펩티드의 돌연변이체(아미노산 치환, 결손 및 삽입을 포함하는 돌연변이체) 등이 포함된다. 이런 돌연변이체에는 보존형 또는 비-보존형 아미노산(바람직하게는 보존형 아미노산)에 의해 하나이상의 아미노산이 치환된 폴리펩티드가 포함되는데, 이런 치환된 아미노산 잔기는 유전자 코드에 의해 인코드되거나 인코드되지 않는다.

전형적으로 이런 치환은 다음의 군내에서 이루어진다; Ala, Val, Leu, Ile; Ser, Thr; 산성 아미노산 잔기 Asp, Glu; Asn, Gln; 염기성 잔기 Lys, Arg; 또는 방향족 잔기 Phe, Tyr. 5 내지 1O개, 1 내지 5개, 1 내지 3개, 1 내지 2개 또는 1 개 아미노산이 치환, 결손 또는 임의의 조합으로 부가된 변이체가 바람직하다. 특히, 단백질의 성질 및 활성에 변화를 주지 않는 침묵 치환, 추가 및 결실이 바람직하다. 또한, 보존성 치환이 바람직하다. 이런 돌연변이체에는 한가지 이상의 아미노산 잔기가 치환기를 보유하는 폴리펩티드가 포함된다.

일반적으로, 2개 폴리펩티드간 30%이상의 상동성은 기능적 등가로 간주된다. 적절하게는, 본 발명의 제 1 특징의 기능적 등가 폴리펩티드는 INSP052와 INSP055 폴리펩티드, 또는 이의 활성 단편과 80%이상의 서열 상동성을 가진다. 좀더 적절하게는, 폴리펩티드는 85%, 90%, 95%, 98% 또는 99%이상의 상동성을 가진다.

본 발명의 제 1 특징의 기능적 등가 폴리펩티드는 한 가지 이상의 구조적 배열 기술을 이용하여 확인된 폴리펩티드일 수도 있다. 가령, Biopendium 검색 데이터베이스를 만드는데 이용되는 검색 도구의 한 측면을 형성하는 Inpharmatica Genome Threader(PCT patent application PCT/GBO1/01105, WO 01/69507) 기술을 이용하여 현재 기능이 알려지지 않고 INSP052와 INSP055 폴리펩티드와 비교하여 서열 상동성이 낮으며 면역글로불린 도메인-포함 세포 표면 인식 분자로 예측되는 폴리펩티드를 확인할 수 있는데, 상기 방법에서는 INSP052와 INSP055 폴리펩티드 서열과의 현저한 구조적 상동성을 공유하는 본 발명의 제 1 특징의 폴리펩티드를 사용한다. '현저한 구조적 상동성'은 Inpharmatica Genome Threader에 의해 두 단백질의 구조적 상동성이 적어도 10%, 적절하게는 적어도 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% 및 그 이상의 확실성으로 예측됨을 의미한다.

본 발명의 제 1 특징의 폴리펩티드에는 INSP052와 INSP055 폴리펩티드의 단 편 및 INSP052와 INSP055 폴리펩티드의 기능적 등가체의 단편이 포함되는데, 여기서 이들 단편은 면역글로불린 도메인-포함 세포 표면 인식 분자 활성을 보유하거나 INSP052와 INSP055 폴리펩티드와 공통되는 항원 결정부위를 갖는다.

본 명세서에서 '단편'은 INSP052와 INSP055 폴리펩티드의 아미노산 서열과 일부분이 동일한 아미노산 서열을 보유하는 폴리펩티드 또는 이의 기능적 등가체를 의미한다. 단편은 서열로부터 유래된 적어도 n개의 연속 아미노산을 포함하고, 특정 서열에 따라 n은 7 또는 그 이상(예, 8, 10, 12, 14, 16, 18, 20 또는 그 이상)이다. 소단편은 항원 결정부위를 형성할 수 있다.

이런 단편은 'free-standing' 즉, 다른 아미노산 또는 폴리펩티드와 융합되거나 일부분으로 남아있지 않을 수 있고, 또는 이런 단편은 좀더 큰 폴리펩티드의 일부분을 구성할 수도 있다. 좀더 큰 폴리펩티드의 일부분을 구성하는 경우에, 본 발명의 단편은 가장 적절하게는 단일 연속 부분을 형성하게 된다. 가령, 바람직한 구체예는 단편의 카르보닐 단부에 융합된 또는 단편의 아미노 단부에 융합된 Pre- 또는 pro-폴리펩티드를 가지는 단편에 관한다.

본 발명의 폴리펩티드 또는 이의 면역원성 단편(적어도 한 개의 항원 결정부위 포함)을 이용하여, 폴리펩티드에 면역특이적인 리간드, 예를 들면 다클론 또는 단클론 항체를 만들 수 있다. 이런 항체를 이용하여, 본 발명의 폴리펩티드를 발현하는 클론을 확인하거나 친화성 크로마토그래피로 본 발명의 폴리펩티드를 정제할 수 있다. 항체는 치료 또는 진단 목적으로 이용할 수도 있는데, 이는 당업자에게 명백하다.

'면역특이적'이란 선행기술에서 다른 관련 폴리펩티드에 대한 친화력보다 본 발명의 폴리펩티드에 대하여 좀더 큰 친화성을 항체가 갖는다는 것을 의미한다. 본 명세서에서 '항체'란 고유 분자 및 이의 단편, 예를 들면 문제의 항원 결정부위에 결합할 수 있는 Fab, $F(ab')_2$, Fv를 의미한다. 이런 항체는 본 발명의 제 1 특징인 폴리펩티드에 결합한다.

'실질적으로 높은 친화성'이란 공지된 면역글로불린 도메인-포함 세포 표면 인식 분자에 대한 친화성과 비교하여 본 발명의 펩티드에 대한 친화성에서 측정가능한 증가가 존재한다는 것을 의미한다.

적절하게는, 친화성은 공지된 면역글로불린 도메인-포함 세포 표면 인식 분자와 비교하여 본 발명의 펩티드에 대하여 적어도 1.5배, 2배, 5배, 10배, 100배, 10^3 배, 10^4 배, 10^5 배 또는 10^6 배 높다.

다클론 항체가 바람직한 경우에, 생쥐, 토끼, 염소 또는 말과 같은 선택된 동물에 본 발명의 폴리펩티드를 면역한다. 동물을 면역하는데 이용되는 폴리펩티드는 재조합 DNA 기술을 이용하여 유도하거나 또는 화학적으로 합성할 수 있다. 필요에 따라서 폴리펩티드에 운반체 단백질을 연결시킬 수 있다. 폴리펩티드에 화학적으로 결합시키는데 흔히 이용되는 운반체는 소혈청 알부민, 티로글로블린, 키홀 림펫 헤모시아닌 등이다. 결합된 폴리펩티드를 이용하여 동물을 면역할 수 있다. 면역주사를 맞은 동물의 혈청을 취하여 공지의 과정, 예를 들면 면역 친화성 크로마토그래피로 처리한다.

본 발명의 제 1 특징의 폴리펩티드에 대한 단클론 항체는 당업자가 용이하게 만들 수 있다. 하이브리도마 기술을 이용하여 단클론항체를 만드는 일반적인 방법은 널리 알려져 있다(Kohler, G. and Milstein, C., Nature 256: 495-497 (1975); Kozbor et al., Immunology Today 4: 72 (1983); Cole et al., 77-96 in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. (1985).

본 발명의 폴리펩티드에 대한 단클론 항체 패널은 다양한 특성, 예를 들면, 이소타입, 에피토프, 친화력 등으로 선별할 수 있다. 단클론 항체는 이들의 개별 폴리펩티드를 정제하는데 특히 유용하다. 대안으로, 관심있는 단클론 항체를 인코드하는 유전자는 당분야에 공지의 기술인 PCR을 이용하여 하이브리도마에서 분리하고 클론하며 적절한 벡터에서 발현시킬 수 있다.

사람이외의 동물에서 취한 가변 영역이 사람의 불변 영역에 결합되거나 융합된 키메라 항체를 이용할 수도 있다(Liu et al., Proc. Natl. Acad. Sci. USA, 84, 3439(1987)).

항체는 예로써 인화(humanisation)반응을 통하여 개체에서 항원성을 적게 나타내도록 변형시킬 수 있다(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); Hodgson et al., Bio/Technology, 9, 421 (1991)). 본 명세서에서 '인화 항체'는 사람아닌 공여자 항체의 중쇄 또는 경쇄의 가변 도메인의 선별된 아미노산과 CDR 아미노산을 사람 항체에서 등가의 아 미노산 잔기로 대체시킨 항체 분자를 말한다. 따라서, 인화 항체는 사람의 항체와는 상당히 유사하지만, 공여자 항체의 결합 능력을 가진다.

다른 대안에서, 항체는 '이중-특이성' 항체가 될 수도 있는데, 이는 2가지 상이한 항원 결합 도메인을 보유하는 항체를 말하는데, 각 도메인은 상이한 에피토프에 관련된다.

파아지 디스플레이 기술을 이용하여, 유관한 항체를 보유하는지에 대해 스크리닝되는 사람 림프구의 PCR 증폭된 V-유전자 레퍼토리, 또는 고유 라이브러리로부터 본 발명의 폴리펩티드에 대한 결합 활성을 가지는 항체를 인코드하는 유전자를 선별할 수 있다(McCafferty, J. et al., (1990), Nature 348, 552-554; Marks, J. et al., (1992) Biotechnology 10, 779-783). 이들 항체의 친화력은 체인 셔플링(chain shuffling)에 의해 개선될 수 있다(Clackson, T. et al., (1991) Nature 352, 624-628).

상기 기술에 의해 만들어지는 단클론 또는 다클론 항체는 면역검사, 방사능면역검사(RIA), 효소-연결된 면역 흡착 검사(ELISA)에서 시약으로 이용되는 추가 용도를 가진다. 이런 분야에서, 항체는 방사능동위원소, 형광 분자 또는 효소와 같은 분석적으로 감지가능한 시약으로 라벨될 수 있다.

본 발명의 제 2와 제 3 특징의 바람직한 핵산 분자는 SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16 또는 SEQ ID NO:18에 상술된 폴리펩티드 서열 또는 이의 기능적 등

가 폴리펩티드를 인코드한다. 이들 핵산 분자들은 본 명세서에서 기술된 분야에 이용될 수 있다. 본 발명의 핵산 분자는 본 명세서에 개시된 서열로부터 적어도 n개 연속 뉴클레오티드로 구성되는데, 이때 n은 특정 서열에 따라 10개 또는 그 이상(예, 12개, 14개, 15개, 18개, 20개, 25개, 30개, 35개, 40개 또는 그 이상)이 된다.

본 발명의 핵산에는 상기한 핵산에 상보적인 서열(예, 안티센스 또는 프로브 목적)도 포함된다.

본 발명의 핵산 분자는 mRNA와 같은 RNA, 또는 cDNA, 합성 DNA, 게놈 DNA와 같은 DNA형이 될 수 있다. 클로닝, 화학적 합성 기술, 또는 이의 복합 기술을 이용하여 이런 핵산 분자를 얻을 수 있다. 유기체로부터 분리하거나 또는 게놈이나 cDNA 라이브러리로부터 고체 상 포스포아미디트 화학적 합성과 같은 화학적 합성 방법을 이용하여 핵산 분자를 만들 수 있다. 일반적으로, RNA 분자는 DNA 서열의 *in vitro* 또는 *in vivo* 전사에 의해 만들 수 있다.

핵산 분자는 이중 가닥 또는 단일 가닥으로 되어 있다. 단일 가닥 DNA는 센스 가닥으로 알려진 코딩 가닥이거나 안티-센스 가닥으로 알려진 비-코딩 가닥일 수도 있다.

'핵산 분자'에는 DNA와 RNA 유사체, 예를 들면 변형된 골격구조를 포함하는 유사체 및 펩티드 핵산(PNA) 역시 포함된다. 본 명세서에서 'PNA'는 리신으로 종료되는 아미노산 잔기의 펩티드 골격에 연결된 적어도 5개 뉴클레오티드로 구성되는 안티센스 분자 또는 항-유전자 물질을 의미한다. 말단의 리신은 조성물에 안정성을 부여한다. PNA는 세포에서 수명을 연장시키기 위해 페길화(pegylation)되는데, 여기서 이들은 상보적인 단일 가닥의 DNA 및 RNA에 선호적으로 결합하고 전사 연장을 종료시킨다(Nielsen, P.E. et al. (1 993) Anticancer Drug Des. 8:53-63).

이들 분자는 유전자 축중의 결과로써 SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16 또는 SEQ ID NO:18의 폴리펩티드를 인코드하는 서로 다른 서열을 보유할 수도 있다. 이런 핵산 분자에는 성숙한 폴리펩티드의 코딩 서열 자체; 성숙한 폴리펩티드와 추가 코딩 서열, 예를 들면 pro-, pre-, prepro- 폴리펩티드 서열과 같은 리더 또는 분비 서열을 인코드하는 코딩 서열; 상기한 추가 코딩 서열의 존부하에 전사, 리보솜 결합, mRNA 안정성에 중요한 역할을 수행하는 전사된 비-번역 서열과 같은 비-코딩 5' amp; 3' 서열(종료 신호 포함)을 비롯한 다른 비-코딩 서열과 함께 성숙 폴리펩티드의 코딩 서열이 포함되지만 이들에 국한되지는 않는다. 핵산 분자에는 다른 기능을 제공하는 부가적인 아미노산을 인코드하는 추가 서열이 포함될 수도 있다.

본 발명의 제 2와 제 3 특징의 핵산 분자는 제 1 특징의 폴리펩티드의 단편 또는 기능적 등가체를 인코드할 수도 있다. 이런 핵산 분자는 자연-발생 대립형질 변이체와 같은 자연 발생 변이체이거나 자연 상태에서는 발생하지 않은 변이체이다. 이런 비-자연적 핵산 변이체는 핵산 분자, 세포 또는 유기체에 이용될 수 있는 돌연변이유발 기술로 만들 수 있다.

이런 변이체에는 치환, 결손 또는 삽입에 의해 상기 핵산 분자와는 상이한 변이체가 포함된다. 치환, 결손 또는 삽입은 한 가지 이상의 뉴클레오티드가 관계한다. 변이체는 코딩이나 비-코딩 또는 둘 모두에서 변화될 수 있다. 코딩 서열에서 변화는 보존성 또는 비-보존성 아미노산 치환, 결손 또는 삽입을 발생시킬 수 있다.

본 발명의 핵산 분자는 당분야에 공지된 일반적인 방법을 이용하여, 유전자 생성물(폴리펩티드)의 클로닝, 프로세싱 및/또는 발현을 변화시키는 것을 비롯한 다양한 이유로 조작할 수 있다. 뉴클레오티드 서열을 조작하는데 이용될 수 있는 기술에는 무작위 절단에 의한 DNA 셔플링 및 유전자 단편과 합성 올리고뉴클레오티드의 PCR 재조합 등이 포함된다. 부위-직접 돌연변이유발은 새로운 제한효소 부위를 추가하고 글리코실화 패턴을 변경하며 코돈 선호도를 변화시키고 절단 변이체를 만들고 돌연변이를 도입하는데 이용할 수 있다.

본 발명의 제 1 폴리펩티드를 인코드하는 핵산을 이형 서열에 결찰시키면 복합된 핵산 분자는 융합 단백질을 인코드하게 된다. 이런 복합된 핵산 분자 역시 본 발명의 제 2 또는 제 3 특징에 포함된다. 가령, 폴리펩티드 활성 저해물질에 대한 폴리펩티드 라이브러리를 스크리닝하기 위하여, 복합된 핵산 분자를 이용하여 상업적으로 이용가능한 항체에 의해 인지될 수 있는 융합단백질을 발현하는 것이 유용할 수 있다. 융합 단백질에는 본 발명의 폴리펩티드 서열과 이형 단백질의 서열사이에 절단 부위가 위치하도록 조작하여, 폴리펩티드가 이형단백질로부터 절단 및 정제될 수 있도록 한다.

본 발명의 핵산 분자에는 본 발명의 폴리펩티드를 인코드하는 핵산 분자에 부분적으로 상보적이고, 따라서 인코딩 핵산 분자에 혼성화되는 안티센스 분자 역시 포함된다. 올리고뉴클레오티드와 같은 안티센스 분자는 본 발명의 폴리펩티드를 인코드하는 표적 핵산을 인지하고 특이적으로 결합하여 핵산의 전사를 방해해도록 설계될 수 있다(참조: 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).

본 명세서에서 '혼성화 반응'은 수소 결합에 의해 2개 핵산의 상호 연합을 의미한다. 일반적으로, 한 분자는 고형 서포

트에 고정되고, 다른 분자는 용액에 자유로운 형태로 존재한다. 이후, 수소 결합이 일어나기 좋은 조건하에서 두 분자를 서로 접촉시킨다.

이런 결합에 영향을 주는 인자는 아래와 같다; 용매의 형태와 용적, 반응 온도; 혼성화 반응 시간; 교반; 고형 서포트에 액상 분자의 비-특이적 결합을 차단하는 물질(Denhardt's reagent 또는 BLOTTO); 분자의 농도; 분자의 연합 속도를 증가시키는 화합물의 이용(덱스트란 설페이트 또는 폴리에틸렌 글리콜); 혼성화 반응이후 세척 조건의 엄밀도(Sambrook et al. [supra]).

표적 분자에 완전하게 상보적인 분자의 혼성화 반응 저해는 당분야에 공지된 혼성화 반응 분석법으로 검사할 수 있다(Sambrook et al .,supral). 실질적으로 상동한 분자는 다양한 엄밀도 조건하에 표적 분자에 완전 상동한 분자의 결합을 경쟁하고 저해하게 된다(Wahl, G.M. and S.L. Berger 1987; Methods Enzymol. 152:399-407; Kimmel, A.R. 1987; Methods Enzymol. 152:507-511).

'엄밀도'란 상이한 분자의 연합보다 매우 유사한 분자들의 연합을 유리하게 하는 혼성화 반응 조건을 말한다. 높은 엄밀도 조건은 아래와 같이 정의할 수 있 다;

50% 포름아미드, 5XSSC(150 mM NaCl, 15 mM 구연산3나트륨), 5O mM 인산나트륨 (pH7.6), 5x Denhardts 용액, 10% 덱스트란 설페이트, 20 ㎍/㎖ 변성되고 전단된 연어 정자 DNA로 구성된 용액에 42℃에서 하룻밤동안 배양, 이후 65℃에서 0.1X SSC에서 필터를 세척. 낮은 엄밀도 조건은 35℃에서 실행되는 혼성화 반응이다(Sambrook et al. [supra]). 적절하게는, 혼성화 반응에 이용되는 조건은 높은 엄밀도 조건이다.

본 발명의 적절한 구체예는 INSP052 또는 INS055 폴리펩티드(SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15 또는 SEQ ID NO:17)를 인코드하는 핵산 분자에 전체 길이의 적어도 70% 동일한 핵산 분자 및 이런 핵산 분자에 실제적으로 상보적인 핵산 분자가 된다. 적절하게는, SEQ ID NO:1에 상술된 SEQ ID NO:2, SEQ ID NO:3에 상술된 SEQ ID NO:4, SEQ ID NO:5에 상술된 SEQ ID NO:6, SEQ ID NO:7에 상술된 SEQ ID NO:8, SEQ ID NO:9에 상술된 SEQ ID NO:10, SEQ ID NO:11에 상술된 SEQ ID NO:12, SEQ ID NO:13에 상술된 SEQ ID NO:14, SEQ ID NO:15에 상술된 SEQ ID NO:16 또는 SEQ ID NO:17에 상술된 SEQ ID NO:18에 대한 코딩 서열에 전체 길이의 적어도 80% 동일한 핵산 분자 및 이런 핵산 분자에 실제적으로 상보적인 핵산 분자이다. 이와 관련하여, 전체 길이의 적어도 90%, 바람직하게는 적어도 95%, 좀더 바람직하게는 적어도 98% 또는 99% 동일한 것이 특히 선호된다. 이와 관련하여 바람직한 구체예는 INSP052와 INSP055 폴리펩티드와 실질적으로 동일한 기능이나 활성을 보유하는 폴리펩티드를 인코드하는 핵산 분자이다.

본 발명은 또한 핵산 분자를 감지하는 방법을 제공하는데, 상기 방법은 아래의 단계로 구성된다;

(a) 이중나선을 형성하는 혼성화 조건하에서 생물학적 시료와 본 발명의 핵산 프로브를 접촉시키고;

(b) 형성된 이중나선을 감지한다.

본 발명에 이용될 수 있는 분석법과 연관하여 아래에 부언된 같이, 상기한 핵산 분자는 INSP052와 INSP055 폴리펩티드를 인코드하는 전장 cDNA와 게놈 클론을 분리하고 이런 폴리펩티드를 인코드하는 유전자와 높은 서열 유사성을 가지는 상동성 유전자의 cDNA와 게놈 클론을 분리하기 위한 RNA, cDNA, 게놈 DNA의 혼성화 반응 프로브로 이용될 수 있다.

이와 관련하여, 당분야에 공지된 기술 중에서 아래의 기술을 이용하며 하기에 설명 목적으로 제시한다. DNA 서열화와 분석 방법은 당분야에 널리 공지되어 있고, 일반적으로 이용할 수 있으며, 본 명세서에 기술된 본 발명의 많은 구체예를 실행하는데 실제로 이용된다. 이런 방법은 DNA 중합효소 I의 클리노우 단편(Klenow fragment), Sequenase(US Biochemical Corp, Cleveland, OH), Taq 중합효소(Perkin Elmer), 열안정성 T7 중합효소(Amersham, Chicago, IL), ELONGASE Amplification System(Gibco/BRL, Gaithersburg, MD)에서 볼 수 있는 중합효소와 프루프-리딩 엑소뉴클레아제를 복합과 같은 효소 등을 이용할 수 있다. 적절하게는, 서열화 공정은 Hamilton Micro Lab 2200(Hamilton, Reno, NV), Peltier Thermal Cycler(PTC200; MJ Research, Watertown, MA), ABI Catalyst, 373과 377 DNA Sequencers(Perkin Elmer)와 같은 기계를 이용하여 자동화된다.

INSP052와 INSP055 폴리펩티드와 등가의 기능을 가지는 폴리펩티드를 인코드하는 핵산 분자를 분리하는 한 가지 방법은 당분야에서 인정되는 표준 과정을 이용하여 천연 또는 인공-설계된 프로브로 게놈이나 cDNA 라이브러리를 탐색하는 것이다('Current Protocols in Molecular Biology', Ausubel et al. (eds). Greene Publishing Association and John Wiley Interscience, New York, 1989,1992). 적절한 인코딩 유전자의 핵산 서열(SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, SEQ ID NO:15 또는 SEQ ID NO:17)에 상응하거나 이들에 상보적인 적어도 15개, 바람하게는 30개, 좀더 바람하게는 적어도 50개 연속 염기로 구

성된 프로브가 특히 유용한 프로브이다. 이런 프로브는 확인을 용이하게 하는 분석적으로 감지가능한 물질로 라벨된다. 유용한 반응물에는 방사능동위원소, 형광 염료 및 감지가능한 산물 형성을 촉진할 수 있는 효소 등이 포함되지만 이들에 국한되지 않는다. 이런 프로브를 이용하면 당업자는 사람, 포유류 또는 다른 동물 종으로부터 관심있는 단백질을 인코드하는 게놈 DNA, cDNA, RNA 폴리뉴클레오티드의 상보적인 사본을 분리하고, 이런 자원에서 족, 타입, 서브타입의 다른 구성원과 관련된 서열을 추가 스크리닝할 수 있다.

많은 경우에, 분리된 cDNA 서열은 폴리펩티드를 인코드하는 부분이 보통 5' 말단에서 짧게 절단되기 때문에 불완전하다. 여러 방법을 이용하여 전장 cDNA를 수득하거나 짧은 cDNA를 연장할 수 있다. 이들 서열은 부분 뉴클레오티드 서열을 이 용하고 프로모터와 조절 요소와 같은 상류 서열을 감지하는 당분야에 공지된 다양한 방법을 활용하여 연장할 수 있다. 가령, 이용가능한 한 가지 방법을 cDNA 말단의 신속 증폭 방법에 기초한다(RACE; Frohman et al., PNAS USA 85, 8998-9002, 1988). MarathonTM 기술(Clontech Laboratories Inc.)로 예시되는 이런 기술의 최근 변형 기술로 좀더 긴 cDNA에 대한 검색이 상당히 단순화되었다. 약간 다른 기술인 '제한-부위' PCR에서는 범용 프라이머를 이용하여 공지된 위치에 인접하는 미지의 핵산 서열을 찾는다(Sarkar, G. (1993) PCR Methods Applic. 2:318-322). 또한, 역 PCR을 이용하여 공지된 부분에 기초한 다변 프라이머를 이용하여 서열을 연장 또는 증폭할 수 있다(Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). 이용가능한 다른 방법은 사람과 이스트 인공 염색체 DNA에서 공지 서열에 인접한 DNA 단편의 PCR 증폭을 수반하는 캡쳐 PCR이다(Lagerstrom, M. et al. (1991) PCR Methods Applic., 1, 111-119). 미지 서열을 확인하는데 이용할 수 있는 또 다른 방법은 Parker, J.D. et al.,((1991); Nucleic Acids Res. 19:30553060)의 방법이다. 이에 더하여, PCR, nested 프라이머, 게놈 DNA를 보행하는 PromoterFinderTM 라이브러리를 이용할 수 있다(Clontech, Palo Alto, CA). 이런 공정은 라이브러리 스크리닝이 필요하지 않고, 인트론/엑손 결합 부분을 찾는데 유용하다.

전장 cDNA를 스크리닝하는 경우에, 좀더 큰 cDNA를 포함하도록 크기-선별된 라이브러리를 이용하는 것이 바람직하다. 또한, 유전자의 5' 부분을 포함하는 서열을 좀더 많이 보유하는 무작위-프라임된 라이브러리이 바람직하다. 무작위-프라임된 라이브러리는 올리고 d(T)가 전장 cDNA를 만들지 못하는 경우에 특히 유용하다. 게놈 라이브러리는 서열을 5' 비-전사된 조절 부분으로 연장하는데 이용할 수 있다.

본 발명의 한 구체예에서, 본 발명의 핵산은 염색체 위치 확인에 이용할 수 있다. 이런 기술에서, 핵산 분자는 개별 사람 염색체상의 특정 위치로 특이적으로 표적화되고 혼성화할 수 있다. 본 발명에 따른 염색체에 관련된 서열의 매핑은 유전자-연관된 질환과 이들 서열의 상관성을 확인하는데 중요한 단계이다. 서열이 정확한 염색체 위치로 매핑되면, 염색체상에서 서열의 물리적인 위치는 유전자 지도 자료와 연계될 수 있다. 이런 데이터는 예로써 V. McKusick, Mendelian Inheritance in Man(available on-line through Johns Hopkins University Welch Medical Library)에서 찾아볼 수 있다. 이후, 연관 분석(물리적으로 인접한 유전자의 공동 유전)을 통하여 동일한 염색체 부분에 매핑된 유전자와 질병의 상관관계를 확인한다. 위치 클로닝 또는 다른 유전자 발견 기술을 이용하면 질병 유전자를 조사하는데 귀중한 정보를 얻을 수 있다. 질병 또는 증후군이 특정 게놈 부분에 유전적으로 연관되어 위치하는 경우에, 상기 부분의 매핑 서열은 추가 조사를 위한 연관된 유전자 또는 조절 유전자를 나타낸다. 핵산 분자는 정상 개체, 보균자 또는 병든 개체간에 전위, 역전 등으로 인한 염색체상 위치에서 차이를 감지하는 데에도 이용할 수 있다.

또한, 본 발명의 핵산 분자는 조직 위치 측정(localisation)에도 유망하다. 이런 기술을 이용하면 폴리펩티드를 인코드하는 mRNA를 감지하여 조직에서 폴리펩티드의 발현 패턴을 결정할 수 있다. 이런 기술에는 *in situ* 혼성화 반응 기술 및 PCR과 같은 뉴클레오티드 증폭 기술 등이 포함된다. 이런 연구 결과로부터 유기체내 폴리펩티드의 정상적인 기능을 제시할 수 있다. 이에 더하여, mRNA의 정상 발현 패턴과 돌연변이 유전자에 의해 인코드된 mRNA의 발현 패턴을 비교 연구는 질병에서 돌연변이 폴리펩티드의 역할에 대한 중요한 정보를 제공한다. 이런 부적절한 발현은 일시적이거나, 지엽적이거나 정성적인 특징을 가질 수 있다.

본 발명의 폴리펩티드를 인코드하는 유전자의 내생 발현을 하향 조절하기 위한 유전자 침묵 접근 방법을 실시할 수도 있다. RNA 간섭(RNAi)(Elbashir, SM et al.,Nature 2001, 411, 494-498)은 이용될 수 있는 서열 특이적 전사후 유전자 침묵화 방법중 하나이다. 짧은 dsRNA 올리고뉴클레오티드를 *in vitro* 에서 합성하여 세포내로 도입한다. 이들 dsRNA 올리고뉴클레오티드의 서열 특이적인 결합은 표적 mRNA 분해를 자극하여, 표적 단백질 발현을 감소시키거나 발현되지 않도록 한다.

상기에서 평가한 유전자 침묵화 방법의 효과는 폴리펩티드 발현(웨스턴 블랏팅)으로 측정하여 평가하고, RNA 수준에서는 TaqMan-기초한 방법을 이용하여 평가할 수 있다.

본 발명의 벡터는 본 발명의 핵산 분자를 포함하며, 클로닝 또는 발현벡터가 될 수 있다. 본 발명의 숙주 세포는 벡터에 의해 형질전환, 형질감염 또는 형질도입된 진핵 또는 원핵 세포이다.

본 발명의 폴리펩티드는 숙주내에 포함된 벡터에 존재하는 인코딩된 핵산 분자의 발현으로 재조합 형태로 만들어질 수 있다. 이런 발현 방법은 당분야에 공지되어 있다(Sambrook et al (supra); Fernandez amp; Hoeffler (1998, eds.

'Gene expression systems. Using nature for the art of expression'. Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto).

일반적으로, 필요 숙주에서 폴리펩티드를 만들기 위하여 핵산 분자를 유지, 증폭 또는 발현하는데 적절한 임의의 시스템 또는 벡터를 이용한다. 적절한 뉴클레오티드 서열은 당분야에 공지된 일반적인 기술, 예를 들면 Sambrook et al., (supra)에서 기술하는 다양한 기술을 이용하여 발현 시스템내로 도입한다. 일반적으로, 인코딩 유전자는 프로모터, 리보솜 결합 부분(박테리아 발현의 경우), 선택적으로 오퍼레이트와 같은 조절 요소의 통제하에 위치시켜, 원하는 폴리펩티드를 인코드하는 DNA 서열이 형질도입된 숙주 세포내에서 RNA로 전사되도록 한다.

적절한 발현 시스템에는 크로모좀, 에피솜, 바이러스-유도된 시스템 등이 포함되는데, 예를 들면 세균성 플라스미드, 박테리오파아지, 트랜스포존, 이스트 에피좀, 삽입 요소, 이스트 염색체 원소; 배큘로바이러스와 같은 바이러스, SV40과 같은 파포바 바이러스, 우두 바이러스, 아데노바이러스 , 가금류 폭스 바이러스, 가사공수병 바이러스, 레트로바이러스, 또는 이들의 조합, 예를 들면 플라스미드 및 코스미드와 파아지미드를 비롯한 박테리오파아지 유전자 요소로부터 유래된 조합 등이 포함된다. 또한, 사람 인공 염색체(HAC)를 이용하여 플라스미드에서 유지되고 발현될 수 있는 것보다 훨씬 큰 DNA 단편을 운반할 수 있다.

특히 적절한 발현 시스템에는 재조합 박테리오파아지, 플라스미드, 코스미드 발현 벡터로 형질전환된 박테리아; 이스트 발현벡터에 의해 형질전환된 이스트; 바이러스 발현 벡터(예, 배큘로바이러스)에 의해 감염된 곤충 세포 시스템; 바이러스 발현 벡터(가령, 꽃양배추 모자이크 바이러스, CaMV; 담배 모자이크 바이러스, TMV) 또는 박테리아 발현 벡터(Ti 또는 pBR322 플라스미드)에 의해 형질도입된 식물 세포 시스템; 또는 동물 세포 시스템이 포함된다. 무-세포 번역 시스템을 이용하여 본 발명의 폴리펩티드를 만들 수도 있다.

본 발명의 폴리펩티드를 인코드하는 핵산 분자의 숙주 세포로의 도입은 많은 실험 매뉴얼에 기술된 방법으로 달성할 수 있다; Davis et al., Basic Methods in Molecular Biology(1986) and Sambrook et al.,[supra]. 특히 적절한 방법에는 인산칼슘 트랜스펙션, DEAE-덱스트란 매개된 트랜스펙션, 트랜스벡션, 미세주사, 양이온 지질-매개된 트랜스펙션, 전기천공, 형질도입, 스크레이프 로딩, 발리스틱 도입(ballistic introduction) 또는 감염 등이 포함된다(Sambrook et al., 1989 [supra]; Ausubel et al., 1991 [supra]; Spector, Goldman amp; Leinwald 1998). 진핵 세포에서 발현 시스템은 필요에 따라 달라 일시적(에피솜)이거나 영구적(염색체 통합)일 수 있다.

인코딩 핵산 분자에는 소포체의 내강, 세포질 공간 또는 세포외 환경으로 번역된 폴리펩티드의 분비를 조절 서열, 예를 들면 신호 펩티드 또는 리더 서열을 인코드하는 서열이 포함되거나 포함되지 않을 수 있다. 이런 신호는 폴리펩티드에 내생이거나 이형 신호일 수 있다. 리더 서열은 번역후 가공에서 세균 숙주에 의해 제거될 수 있다.

조절 서열 이외에, 숙주 세포 성장과 관련된 폴리펩티드의 발현 조절을 가능하게 하는 조절 서열을 추가하는 것도 바람직하다. 조절 서열의 예는 조절 화합물 의 존재, 다양한 온도 또는 대사 조건을 비롯한 화학적 또는 물리적인 자극에 반응하여 유전자 발현을 증가 또는 감소시키는 서열이다. 조절 서열은 벡터의 비-번역 부분인데, 예를 들면 인헨서, 프로모터, 5'와 3' 비-번역 부분이다. 이들은 숙수 세포 단백질과 상호작용하여 전사 및 번역을 실행한다. 이런 조절 서열은 강도 및 특이성에서 가변적이다. 이용되는 벡터 시스템 및 숙주에 따라, 본질성과 유도성 프로모터를 비롯한 많은 임의의 적절한 전사 및 번역 요소를 이용할 수 있다. 가령, 세균 시스템에 클로닝되는 경우에, Bluescript 파아지미드의 하이브리드 lacZ 프로모터(Stratagene, LaJolla, CA), pSportITM 플라스미드(Gibco BRL) 및 이와 유사한 유도성 프로모터를 이용할 수 있다. 곤충 세포에는 배큘로바이러스 폴리헤드린 프로모터를 이용할 수 있다. 식물 세포의 게놈(예, 열 쇼크, RUBISCO 및 저장 단백질 유전자) 또는 식물 바이러스(예, 바이러스성 프로모터 또는 리더 서열)로부터 유래된 프로모터 또는 인헨서는 벡터내로 클론될 수 있다. 포유류 세포 시스템에서는 포유류 유전자 또는 포유류 바이러스로부터 유래된 프로모터가 적절하다. 서열의 다중 사본을 포함하는 세포주를 만들 필요가 있는 경우에, 적절한 선택성 표식과 함께 SV40 또는 EBV에 기초한 벡터를 이용할 수 있다.

특정 핵산 코딩 서열이 적절한 조절 서열을 가지는 벡터에 위치하도록 발현벡터를 작제하는데, 조절 서열에 대한 코딩 서열의 위치와 방향은 코딩 서열이 조절 서열의 '통제'하에 전사되도록 한다, 다시 말하면 조절 서열에서 DNA 분자에 결합하는 RAN 중합효소가 코딩 서열을 전사한다. 일부 경우에, 리딩 프레임을 유지하는 적절한 방향으로 조절 서열에 부착되도록 서열의 변화가 필요할 수도 있다.

벡터로 코딩 서열을 도입하기에 앞서 조절 서열 및 다른 조절 서열을 핵산 코딩 서열에 결찰시킬 수도 있다. 대안으로, 코딩 서열은 조절 서열 및 적절한 제한 효소 부위를 이미 포함하고 있는 발현벡터에 직접 클론될 수 있다.

재조합 폴리펩티드를 장기간 다량 생산하고자 하는 경우에, 발현을 안정화시키는 것이 바람직하다. 가령, 관심있는 폴리펩티드를 안정적으로 발현하는 세포주는 동일 또는 별개의 벡터에 바이러스 복제 원점 또는 선택적 표식 유전자 발현 요소를 포함하는 발현벡터로 형질전환될 수 있다. 벡터의 도입이후, 세포는 선택성 배지로 옮기기 이전에 풍부한 배지에서 1-2일간 생장시킨다. 선택적 표식의 목적은 선별에 대한 저항성을 부여하고, 이의 존재가 도입된 서열을 성

공적으로 발현하는 세포의 생장 및 회수를 가능하도록 하는 것이다.

안정적으로 형질도입된 세포의 저항성 클론은 세포 유형에 적합한 조직 배양 기술을 이용하여 증식시킬 수 있다.

발현을 위한 숙주로 이용가능한 포유류 세포주는 당분야에 공지되어 있는데, 중국 햄스터 난소(CHO), HeLa, 아기 햄스터 신장(BHK), 원숭이 신장(COS), C127, 3T3, BHK, HEK 293, Bowes 흑색종, 사람 간세포 육종(예, Hep G2) 세포, 다른 다양한 세포주를 포함하나 이에 국한되지 않은 American Type Culture Collection(ATCC)으로부터 입수가능한 많은 영속화 세포주 등이 포함된다.

배큘로바이러스 시스템의 경우에, 배큘로바이러스/곤충 세포 발현 시스템의 재료는 키트 형태로 이용가능하다(Invitrogen, San Diego CA(the 'MaxBac' kit)). 이런 기술은 당업자에 공지된 것으로, Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987)에서 상세하게 기술한다. 이런 시스템에 사용하기 적합한 숙주 세포는 초파리(Drosophila) S2와 Spodoptera Sf9 세포와 같은 곤충 세포이다.

많은 식물 세포 배양물 및 전체 식물 유전자 발현 시스템이 당분야에 공지되어 있다. 적절한 식물 세포 유전자 발현 시스템의 예로는 US 5,693,506; US 5,659,122; US 5,608,143에 기술된 것 등이 포함된다. 식물 세포 배양물에서 유전자 발현의 다른 예는 Zenk, Phytochemistry 30, 3861-3863 (1991)에서 설명한다.

특히, 원형질체를 분리할 수 있고 온전한 재생 식물로 배양할 수 있는 모든 식물을 이용하여 전이된 유전자를 보유하는 식물을 회수할 수 있다. 실질적으로 모든 식물이 사탕수수, 사탕무, 목화, 과일, 다른 나무, 콩과식물, 채소 등을 포함하지만 이들에 국한되지 않는 배양된 세포 또는 조직으로부터 재생될 수 있다.

특히 적절한 세균성 숙주 세포로는 연쇄상구균(*streptococci*), 포도상구균(*staphylococci*), 대장균(*E. coli*), 스트렙토마이세스(*Streptomyces*), 바실러스 서브틸리스 세포(*Bacillus subtilis*) 등이 포함된다.

곰팡이 발현을 위한 적절한 숙주 세포로는 효모(*S. cerevisiae*) 및 아스퍼질러스(*Aspergillus*) 세포 등이 포함된다.

당분야에 공지된 선별 시스템을 이용하여 형질전환된 세포주를 회수할 수 있다. 전형적으로, tk- 세포 또는 aprt□ 세포에서 각각 이용될 수 있는 허피스 심플렉스 바이러스 티미딘 카이나제 유전자(Wigler, M. et al.,(1977) Cell 11:223-32) 및 아데닌 포스포리보실트란스퍼라제(Lowy, 1. et al. (1980) Cell 22:817-23) 유전자이다.

또한, 선별을 위한 기초로써 항-대사물질, 항생제 또는 제초제 저항성을 이용할 수 있는데, 예를 들면, 메토트렉세이트에 저항성을 부여하는 디하이드로폴레이트 환원효소(DHFR)(Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77: 3567-70); 아미노글리코시드 네오마이신 및 G418에 저항성을 부여하는 npt (Colbere-Garapin, F. et al (1981) J. Mol. Biol. 150:1-14); 클로로설푸론 및 포스포나트로신 아세틸트란스퍼라제에 각각 저항성을 부여하는 als 또는 pat 등이 포함된다. 부가적인 선별 유전자에 대해서도 설명하고 있는데, 이들의 예는 당업자에게 자명하다.

표식 유전자의 발현 유무가 관심있는 유전자의 존재를 암시하긴 하지만, 이의 존재와 발현은 확증을 필요로 한다. 가령, 관련 서열이 표식 유전자내로 삽입되는 경우에, 표식 유전자 기능의 부재로 적절한 서열을 포함하는 형질도입된 서열을 확증할 수 있다. 대안으로, 표식 유전자는 단일 프로모터 조절하에 본 발명의 폴리펩티드를 인코드하는 서열과 나란하게 위치시킬 수 있다. 일반적으로, 유도 또는 선별에 대한 반응으로 표식 유전자의 발현은 나란히 위치하는 유전자의 발현을 암시한다.

대안으로, 본 발명의 폴리펩티드를 인코드하는 핵산 서열을 포함하고 상기 폴리펩티드를 발현하는 숙주는 당분야에 공지된 다양한 기술을 이용하여 확인할 수 있다. 이런 과정에는 DNA-DNA 또는 DNA-RNA 하이브리드반응 및 핵산이나 단백질의 감지 또는 정량화를 위한 막, 용액 또는 칩 기초한 기술을 수반하는 단백질 생물분 석, 예를 들면 형광 활성화된 세포 소팅(FACS) 또는 면역검사 기술(효소-결합된 면역흡착 검사[ELISA] 및 방사능면역검사[RIA]) 등이 포함된다(Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN); Maddox, D.E. et al. (I 983) J. Exp. Med, 158, 1211-1216).

다양한 라벨 및 접합 기술은 당분야에 공지되어 있고, 다양한 핵산 및 아미노산 검정에 이용될 수 있다. 본 발명의 폴리펩티드를 인코드하는 핵산 분자와 관련된 서열을 감지하기 위한 라벨된 혼성화 또는 PCR 프로브의 생산 수단에는 라벨된 폴리뉴클레오티드를 이용한 올리고라벨링, 니크 번역, 말단-라벨링 또는 PCR 증폭등이 포함된다.

대안으로, 본 발명의 폴리펩티드를 인코드하는 서열은 mRNA 프로브를 만들기 위해 벡터내로 클론될 수 있다. 이런 벡터는 당분야에 공지되어 있고 상업적으로 가용하며 T7, T3 또는 SP6과 같은 적절한 RNA 중합효소 및 라벨된 뉴클레오티드를 추가하여 *in vitro* 에서 RNA 프로브를 합성하는데 이용할 수 있다. 이런 과정은 사용 키트를 이용하여

실행할 수 있다(Pharmacia amp; Upjohn, (Kalamazoo, MI); Promega (Madison WI); U.S. Biochemical Corp., Cleveland, OH)).

용이한 감지에 이용하는데 적절한 리포터 분자에는 방사능핵종, 효소, 형광, 화학발광이나 발색 물질, 기질, 보조인자, 저해물질, 자성 입자 등이 포함된다.

본 발명에 따른 핵산 분자를 이용하여 유전자도입 동물, 특히 설치류 동물을 만들 수 있다. 이런 유전자도입 동물 역시 본 발명의 특징에 속한다. 체세포를 변형하거나 유전가능한 변형을 통합하는 생식세포 요법을 이용하여 유전자도입 동물 을 만들 수 있다. 이런 유전자도입 동물은 특히 본 발명의 폴리펩티드의 조절물질로써 효과적인 약물 분자에 대한 동물 모델을 만드는데 이용될 수 있다.

공지의 방법으로 재조합 세포 배양물로부터 폴리펩티드를 회수하고 정제할 수 있는데, 이용할 수 있는 방법으로는 암모늄 설페이트이나 에탄올 침전, 산 추출, 음이온 또는 양이온 교환 크로마토그래피, 포스포셀룰로오즈 크로마토그래피, 소수성 상호작용 크로마토그래피, 친화력 크로마토그래피, 하이드록실아파타이트 크로마토그래피, 렉틴 크로마토그래피 등이 포함된다. 정제에는 고 성능 액체 크로마토그래피(HPLC)가 특히 유용하다. 단백질을 재폴팅하는데 공지된 기술을 이용하여, 정제 또는 분리 과정동안 폴리펩티드가 변성되는 경우에 활성 형으로 재생할 수 있다.

원하는 경우에 특화된 벡터 작제를 이용하여, 가용성 단백질의 정제를 조장하는 폴리펩티드 도메인을 인코드하는 뉴클레오티드 서열을 본 발명의 폴리펩티드를 인코드하는 서열에 결합시켜, 단백질 정제를 촉진할 수 있다. 이런 정제-촉진 도메인에는 고정된 금속에서 정제를 허용하는 히스티딘-트립토판 모듈과 같은 금속-킬레이트화 도메인, 고정된 면역글로불린에서 정제를 가능하게 하는 단백질 A 도메인; FLAGS 연장/친화력 정제 시스템에 이용되는 도메인 등이 포함된다(Immunex Corp., Seattle, WA). 정제 도메인과 본 발명의 폴리펩티드 사이에 인자 XA 또는 엔테로키나제(Invitrogen, San Diego, CA)와 같은 절단가능 링커 서열을 포함시켜 정제를 촉진할 수 있다. 이런 발현 벡터는 티오레독신 또는 엔테로카이나제 절단 부위를 선행하는 여러 히스티딘 잔기에 융합된 본 발명의 폴리펩티드를 보유하는 융합 단백질의 발현을 제공한다. 히스티딘 잔기는 IMAC(고정된 금속 이온 친화력 크로마토그래피 Porath, J. et al. (1992), Prot. Exp. Purif. 3: 263-281)에 의한 정제를 촉진하고, 티오레독신 또는 엔테로카이나제 절단 부위는 융합 단백질로부터 폴리펩티드를 정제하는 수단을 제공한다. 융합 단백질을 포함하는 벡터는 Kroll, D.J. et al. (1993; DNA Cell Biol. 12.441-453)에서 설명한다.

스크리닝 검정에 이용되는 폴리펩티드를 발현하는 경우에, 일반적으로 폴리펩티드는 이를 발현하는 숙주 세포의 표면에서 생산되도록 하는 것이 바람직하다. 이런 경우에, 스크리닝 검정에 앞서 형광 활성화된 세포 소트(FACS) 또는 면역친화력 기술 등을 이용하여 숙주 세포를 수득한다. 폴리펩티드가 배지내로 분비되는 경우에, 발현된 폴리펩티드를 회수하고 정제하기 위하여 배지를 회수할 수 있다. 폴리펩티드가 세포내에서 만들어지는 경우에, 세포를 먼저 용해시키고, 폴리펩티드를 회수한다.

본 발명의 폴리펩티드는 약물 스크리닝 기술을 이용하여 화합물 라이브러리를 스크리닝하는데 이용할 수 있다. 이런 화합물은 본 발명의 유전자의 발현 수준 또는 본 발명의 폴리펩티드의 활성 수준을 증가(항진제) 또는 저해(길항제)할 수 있는데, 이 또한 본 발명의 특징이 된다. 바람직한 화합물은 본 발명의 폴리펩티드를 인코드하는 자연 유전자의 발현을 변화시키거나 본 발명의 제 1 특징의 폴리펩티드 활성을 조절하는데 유효하다.

항진제 또는 길항제 화합물은 세포, 무-세포 침전물, 화학 라이브러리 또는 천연 생성물 혼합물로부터 분리할 수 있다. 이런 항진제 또는 길항제는 천연이나 변형된 기질, 리간드, 효소, 리포터, 구조적 또는 기능적 모방체 등이 될 수 있다. 이런 스크리닝 기술에 대하여 Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991)를 참조한다.

길항제로써 유망한 화합물은 결합시에 폴리펩티드의 생물학적 효과를 유도하지 않으면서 본 발명의 폴리펩티드에 결합할 수 있는 분자이다. 유망한 항진제로는 본 발명의 폴리펩티드에 결합하여 활성을 저해 또는 제거할 수 있는 작은 유기 분자, 펩티드, 폴리펩티드, 항체 등이 포함된다. 이런 방식으로, 정상적인 세포 결합 분자에 폴리펩티드의 결합을 방해하여, 폴리펩티드의 정상적인 생물학적 활성을 저해한다.

이런 스크리닝 기술에 이용되는 본 발명의 폴리펩티드는 용액에서 유리된 형태, 고형 서포트에 고착된 형태, 세포 표면에서 생성된 형태 또는 세포내에 위치하는 형태로 존재할 수 있다. 일반적으로, 이런 스크리닝 과정은 시험 화합물과 접촉하는 폴리펩티드를 발현하는 적절한 세포 또는 세포막을 이용하여 결합, 또는 기능적 반응의 촉진이나 저해를 관찰하는 단계를 수반한다. 시험 화합물과 접촉된 세포의 기능적 반응은 시험 화합물과 접촉하지 않은 대조 세포의 기능적 반응과 비교한다. 이런 검정에서는 적절한 감지 시스템을 이용하여 시험 화합물이 폴리펩티드 활성화에 의해 발생된 신호를 유발하는 지를 평가한다. 일반적으로, 활성화의 저해물질은 공지된 항진제의 존재하에 분석하고, 항진제에 의한 활성화에 대한 효과는 시험 화합물 존재하에 관찰한다.

본 발명의 폴리펩티드의 항진제 또는 길항제를 확인하는 바람직한 방법은 아 래와 같다;

(a) 폴리펩티드에 결합을 허용하는 조건하에서 본 발명의 제 1 특징의 폴리펩티드를 세포 표면에서 발현하는 세포와 시험 화합물을 접촉시키고, 이때 폴리펩티드는 폴리펩티드에 화합물의 결합에 반응하여 감지가능한 신호를 제공할 수 있는 제 2 성분과 연합되고,

(b) 폴리펩티드와 화합물의 상호작용으로 발생되는 신호의 수준을 측정함으로써 화합물이 폴리펩티드에 결합하여 이를 활성화시키는지 또는 저해하는지를 결정한다.

본 발명의 폴리펩티드의 항진제 또는 길항제를 확인하는 다른 바람직한 방법은 아래와 같다;

(a) 폴리펩티드에 결합을 허용하는 조건하에서 본 발명의 제 1 특징의 폴리펩티드를 세포 표면에서 발현하는 세포와 시험 화합물을 접촉시키고, 이때 폴리펩티드는 폴리펩티드에 화합물의 결합에 반응하여 감지가능한 신호를 제공할 수 있는 제 2 성분과 연합되고,

(b) 폴리펩티드와 화합물의 상호작용으로 발생되는 신호 수준을 화합물이 없을 경우의 신호 수준과 비교함으로써 화합물이 폴리펩티드에 결합하여 이를 활성화시키는지 또는 저해하는지를 결정한다.

다른 바람직한 구체예에서, 상기한 일반적인 방법은 폴리펩티드에 대한 라벨되거나 라벨되지 않은 리간드의 존재하에 항진제 또는 길항제를 확인하는 단계가 추가로 포함된다.

본 발명의 폴리펩티드의 항진제 또는 길항제를 확인하는 방법의 또 다른 구체예는 아래와 같다;

폴리펩티드에 결합을 허용하는 조건하에서 후보 화합물 존재하에 세포 표면에 본 발명의 폴리펩티드를 포함하는 세포 또는 이런 폴리펩티드를 포함하는 세포막에 리간드의 결합 저해를 결정하고;

폴리펩티드에 결합된 리간드 양을 결정한다. 리간드 결합을 감소시키는 화합물은 항진제 또는 길항제로 간주된다. 적절하게는, 리간드는 라벨된다.

좀더 구체적으로, 폴리펩티드 길항제 또는 항진제 화합물을 스크리닝하는 방법은 아래와 같은 단계로 구성된다;

(a) 본 발명의 폴리펩티드를 발현하는 전체 세포 또는 본 발명의 폴리펩티드를 포함하는 세포 표면에 함께 라벨된 리간드를 배양하고;

(b) 전체 세포 또는 세포막에 결합된 라벨된 리간드의 양을 측정하고;

(c) (a) 단계의 전체 세포 또는 세포막 및 라벨된 리간드의 혼합물에 후보 화합물을 첨가하여 혼합물이 평형에 이르도록 하고;

(d) (c) 단계이후 전체 세포 또는 세포막에 결합된 라벨된 리간드의 양을 측정하고;

(e) (b)와 (d) 단계에서 결합된 라벨된 리간드의 차이를 비교하고, (d) 단계에서 결합을 감소시키는 화합물은 항진제 또는 길항체로 간주된다.

상기한 특정 구체예에서, 간단한 결합 분석법을 이용할 수 있는데, 여기서 폴리펩티드를 보유하는 표면에 시험 화합물의 부착은 시험 화합물과 직-간접으로 결합된 라벨로 감지하거나 라벨된 경쟁물질과의 경쟁을 수반하는 분석법으로 감지한다. 또 다른 구체예에서, 경쟁적인 약물 스크리닝 분석법을 이용할 수 있는데, 여기서 폴리펩티드에 결합할 수 있는 중화 항체는 시험 화합물과 결합에 대하여 특이적으로 경쟁한다. 이런 방식으로, 항체는 폴리펩티드에 특이적인 결합 친화성을 가지는 임의 시험 화합물의 존재를 감지하는데 이용될 수 있다.

당업자는 본 발명의 폴리펩티드의 조절물질을 확인하는 분석법을 고안할 수 있다. 이와 관련하여 길항제(본원의 경우에 중화 항체)를 확인하는 분석법의 예를 보고하는 Lokker NA et al., J Biol Chem 1997 Dec 26; 272(52): 33037-44를 참조한다.

세포에서 폴리펩티드를 인코드하는 mRNA 생산에 대한 첨가된 시험 화합물의 효과를 감지하는 검사 방법을 설계할 수도 있다. 가령, 당분야에 공지된 표준 방법에 의한 단클론 또는 다클론 항체를 이용하여 폴리펩티드의 분비되거나 세포-결합된 수준을 측정하는 ELISA를 실시할 수 있는데, 상기 방법은 적절하게 조작된 세포 또는 조직으로부터 폴리펩티드 생산을 저해하거나 강화하는 화합물을 검색하는 데에도 이용될 수 있다. 이후, 폴리펩티드와 시험 화합물간

복합체 형성을 측정한다.

본 발명에 포함되는 임의의 검정 방법은 과다발현 또는 제거 분석에서 본 발명의 유전자 및 폴리펩티드를 이용한다. 이런 검정은 세포에서 이들 유전자/폴리펩티드의 수준 조절 및 조작된 세포의 생리에 대한 이런 조작 과정의 충격 평가에 관계한다. 가령, 이런 실험은 특정 유전자/폴리펩티드가 연관되고, 연구된 폴리펩티드와 상호작용하는 폴리펩티드 실체에 대한 정보를 발생시키며, 관련된 유전자와 단백질이 조절되는 방법에 관한 단서를 제공하는 신호 생성과 대사 경로의 세부적인 내용을 밝힌다.

이용될 수 있는 다른 약물 스크리닝 기술은 관심있는 폴리펩티드에 적절한 결합 친화성을 가지는 화합물의 고 처리량 스크리닝을 제공한다(International patent application WO84/03564). 상기 방법에서, 다수의 상이한 작은 시험 화합물을 고형 기질상에 합성되는데, 이들은 이후 본 발명의 폴리펩티드와 반응시키고 세척한다. 폴리펩티드를 고정시키는 한가지 방법은 비-중화 항체를 이용하는 것이다. 이후, 당분야에 공지된 방법을 이용하여, 결합된 폴리펩티드를 감지한다. 정제된 폴리펩티드는 상기한 약물 스크리닝 기술에 이용되는 플레이트에 직접 피복될 수 있다.

본 발명의 폴리펩티드는 당분야에 공지된 표준 수용체 결합 기술, 예를 들면 폴리펩티드가 방사성 동위원소로 라벨되거나, 화학적으로 변형되거나, 또는 감지 또는 정제를 조장하고 추정 수용체 자원(예, 세포, 세포 막, 세포 현탁액, 조직 추출물, 체액의 조성물)과 함께 배양되는 리간드 결합과 교차 결합 기술을 통하여 막-결합되거나 가용성 수용체를 확인하는데 이용될 수 있다. 표면 플라스몬 공명(surface plasmon resonance) 및 스펙트로스코피(spectroscopy)와 같은 생체물리학적 기술을 이용하여 결합 효과를 측정할 수 있다. 결합 분석은 수용체의 정제 및 클로닝에 이용할 수 있을 뿐만 아니라 수용체에 대한 폴리펩티드의 결합과 경쟁하는 폴리펩티드의 항진제 또는 길항제를 확인할 수도 있다. 스크리닝 검사를 실행하는 표준 방법은 당분야에 공지되어 있다.

본 발명에는 상기한 항진제, 길항제, 리간드, 수용체, 기질, 효소를 확인하는 방법에 유용한 스크리닝 키트 역시 포함된다.

본 발명에는 앞서 기술된 방법에 의해 발견되는 본 발명 폴리펩티드의 활성 또는 항원성을 조절할 수 있는 항진제, 길항제, 리간드, 수용체, 기질, 효소가 포함된다.

또한, 본 발명은 적절한 약학적 운반체와 복합된 폴리펩티드, 핵산, 리간드 또는 화합물을 함유하는 약학 조성물을 제공한다. 이들 조성물은 하기에 상술된 바와 같이 치료 또는 진단 시약, 백신 또는 다른 면역 조성물로써 적합하다.

본 명세서에서 폴리펩티드, 핵산, 리간드 또는 화합물[X]를 함유하는 조성물에서 화합물[x]과 불순물[Y]의 총 중량당 적어도 80%가 X인 경우에, '실질적으로 불순물[Y]이 없는' 조성물이라고 할 수 있다. 적절하게는, X는 조성물에서 X+Y의 총 중량당 적어도 90%로 구성되고, 좀더 적절하게는 총 중량의 약 95%, 98%, 99%로 구성된다.

적절하게는, 약학 조성물은 본 발명의 폴리펩티드, 핵산 분자, 리간드 또는 화합물의 치료요법적 효과량을 함유한다. '치료요법적 효과량'이란 표적하는 질병 또는 질환을 치료, 경감 또는 예방하거나 감지가능한 치료 또는 예방 효과를 나타내는데 필요한 치료제의 함량을 의미한다. 임의의 화합물에서, 치료요법적 효과량은 예로써 신생물 세포, 또는 동물 모델, 일반적으로 생쥐, 토끼, 개 또는 돼지의 세포 배양 분석으로 최초 측정할 수 있다. 상기 동물 모델은 적절한 농도 범위와 투여 경로를 결정하는 데에도 이용될 수 있다. 이후, 이런 정보를 활용하여 사람에 서 유용한 약량 및 투여 경로를 결정할 수 있다.

사람 개체의 정확한 효과량은 질병의 심각성, 개체의 전반적인 건강 상태, 나이, 체중, 성별, 음식, 투여 빈도와 기간, 약물 복합, 반응 감응도, 치료에 대한 내성/감응성에 좌우된다. 통상적인 실험을 통하여 효과량을 결정할 수 있는데, 이는 임상 의사의 판단 범주에 속한다. 일반적으로, 효과량은 0.01 ㎎/kg 내지 50 ㎎/kg, 적절하게는 0.05 ㎎/kg 내지 10 ㎎/kg 범위이다. 조성물을 환자에 개별적으로 투여하거나, 다른 약물, 물질 또는 호르몬과 공동으로 투여할 수도 있다.

약학 조성물은 치료제 투여를 위한 약학적으로 수용가능한 담체를 함유할 수 있다. 이런 담체에는 항체, 다른 폴리펩티드, 유전자, 리포좀과 같은 다른 치료제가 포함되는데, 담체는 조성물을 제공받는 개체에 유해한 항체 생산을 유도하지 않아야 하고, 과도한 독성없이 투여되어야 한다.

적절한 담체는 크고 대사가 느린 거대분자, 예를 들면 단백질, 폴리사카라이드, 폴리락트산, 폴리글리콜산, 고분자 아미노산, 아미노산 고분자 및 비활성 바이러스 입자 등이다.

약학적으로 수용가능한 염에는 염화수소, 브롬화수소, 인산염, 황산염 등과 같은 무기산염; 아세테이트, 프로피오네이트, 말로네이트, 벤조에이트 등과 유기산 염이 포함된다. 약학적으로 이용가능한 담체는 Remington's Pharmaceutical Sciences(Mack Pub. Co., N.J. 1991)를 참고한다.

치료 조성물에서 약학적으로 수용가능한 담체에는 물, 염, 글리세롤, 에탄올이 추가로 포함될 수 있다. 이에 더하여, 가습제, 유화제, pH 완충제 등과 같은 보 조 물질이 조성물에 포함될 수 있다. 이런 담체를 이용하면, 약학 조성물은 정제, 알약, 당의정, 캡슐, 액체, 젤, 시럽, 슬러리, 현탁액 등 환자에게 복용가능 형태로 제제화될 수 있다.

일단 제제화되면, 본 발명의 조성물은 환자에게 직간접적으로 투여할 수 있다. 치료 대상은 동물 특히, 사람 개체이다.

본 발명에 이용가능한 약학 조성물은 경구, 정맥내, 근육내, 동맥내, 골수내, 척수강내, 심실내, 경피내 또는 경피하(예, W098/20734), 피하, 복막내, 비강, 장, 국소, 혀밑, 질내 또는 직장을 비롯한 다양한 루트로 투여될 수 있다. 유전자 총 또는 하이포스프레이(hyposprays)를 이용하여 본 발명의 조성물을 투여할 수도 있다. 전형적으로, 치료 조성물을 액체 용액 또는 현탁액 형태의 주사액으로 만들 진다; 주사하기 직전에 용액, 현탁액 또는 액체 운반체에 용해하기에 적합한 고형으로 만들어 질 수도 있다.

조성물을 직접 운반하는 방법은 주사, 피하, 복막, 정맥, 근육으로 주사하거나 조직내 간질(間質) 공간으로 운반하는 것이다. 조성물은 병소에 투여할 수도 있다. 약량은 1회 투약 과정 또는 다회 투약 과정으로 제공될 수 있다.

본 발명의 폴리펩티드 활성이 특정 질병 상태에서 과도하게 나타나는 경우에 여러 가지 방법을 이용할 수 있다. 한 가지 방법은 상기한 바와 같이 저해 화합물(길항제) 및 리간드, 기질, 효소, 수용체의 결합을 차단하거나 제 2 신호를 저해함으로써 폴리펩티드 기능을 저해하는데 효과량의 약학적으로 수용가능한 담체를 개체에 투여하여 비정상 상태를 경감시키는 것이다. 적절하게는, 길항제가 항체이다. 가장 적절하게는, 항체는 앞서 기술한 바와 같이 면역원성을 최소화하는 인화 또는 키메라 항체이다.

다른 방법으로, 리간드, 기질, 효소, 수용체에 대한 결합 친화력을 보유하는 가용성 형태의 폴리펩티드를 투여할 수도 있다. 전형적으로, 폴리펩티드는 유관한 부분을 보유하는 단편 형태로 투여된다.

또 다른 방법으로, 내부적으로 생성되거나 별도로 투여된 안티센스 핵산 분자의 이용과 같은 발현 저해 기술을 이용하여, 폴리펩티드를 인코드하는 유전자 발현을 저해할 수 있다. 폴리펩티드를 인코드하는 유전자의 통제, 5' 또는 조절 부분(신호 서열, 프로모터, 인헨서, 인트론)에 상보적인 서열 또는 안티센스 분자(DNA, RNA 또는 PNA)를 설계하여 유전자 발현을 변화시킬 수 있다. 유사하게, '삼중 나선' 염기쌍 방법을 이용하여 저해를 달성할 수 있다. 삼중 나선 염기쌍 형성 기술은 삼중 나선이 중합효소, 전사 인자 또는 조절 분자의 결합을 위하여 이중 나선이 충분히 개방되는 능력을 저해하기 때문에 유용하다. 삼중 나선 DNA를 이용한 최근 치료법의 발전이 기존 문헌에서 보고되고 있다(Gee, J.E. et al. (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). 상보적인 서열 또는 안티센스 분자를 설계하여 전사체가 리보솜에 결합하는 것을 차단함으로써 mRNA의 전사를 저해할 수 있다. 이런 올리고뉴클레오티드는 투여되거나 *in vivo* 발현으로 *in situ* 생성될 수 있다.

이에 더하여, 인코딩 mRNA 서열에 특이적인 리보자임을 이용하여 본 발명의 폴리펩티드 발현을 차단할 수도 있다. 리보자임은 천연 또는 합성 형태로 촉매 활 성을 가진 RNA이다(Usman, N, et al., Curr. Opin. Struct. Biol (1996) 6 (4), 527-33). 선택된 위치에서 mRNA를 특이적으로 절단할 수 있는 합성 리보자임을 설계하여, mRNA가 기능적 폴리펩티드로 번역되는 것을 방지할 수 있다. RNA 분자에서 정상적으로 발견되는 천연 리보즈 포스페이트 골격과 천연 염기로 리보자임을 합성할 수 있다. 대안으로, 리보자임은 비-천연 골격, 예를 들면 2'-O-메틸 RNA로 합성하여 리보뉴클레아제 분해되지 않고, 변형된 염기를 보유하도록 할 수도 있다.

RNA 분자를 변형시켜, 세포내 안정성과 반감기를 증가시킬 수 있다. 가능한 변형 방법에는 분자의 5' 및/또는 3' 말단에 인접 서열을 추가, 분자의 골격내에 포스포디에스테라제 연쇄보다는 포스포로티오에이트 또는 2'O-메틸의 이용 등이 포함되지만 이들에 국한되지 않는다. 이런 개념은 PNA 생산에 내재하고, 내인성 엔도뉴클레아제에 의해 용이하게 인지되지 않는 비-전통적 염기, 예를 들면 이노신, 큐에오신, 부토신을 비롯하여 아세틸-, 메틸-, 티오- 및 아데닌, 시티딘, 구아닌, 티민, 우리딘의 유사하게 변형된 형태의 포함으로 이들 분자 모두에서 확장될 수 있다.

본 발명의 폴리펩티드의 과소 발현 및 이의 활성과 관련된 비정상 상태의 치료에는 여러 방법이 가능하다. 한 가지 방법은 비정상적 상태를 완화시키기 위하여 폴리펩티드를 활성화시키는 화합물, 다시 말하면 상기한 바와 같은 항진제의 효과량을 개체에 투여하는 것이다. 대안으로, 적절한 약학 담체와 공동으로 폴리펩티드의 치료요법적 효과량을 투여하여 폴리펩티드의 적절한 생리학적 균형을 복원한다.

유전자 요법을 이용하여 개체에서 관련 세포에 의한 폴리펩티드의 내적 생산 에 달성할 수 있다. 유전자 요법은 결함 유전자를 교정된 치료 유전자로 대체하여 폴리펩티드의 부적절한 생산을 영구적으로 치유하는데 이용된다.

본 발명의 유전자 요법은 *in vivo* 또는 *ex vivo* 상태에서 실시될 수 있다. *Ex vivo* 유전자 요법은 환자 세포의 분리와 정제, 치유 유전자의 도입, 유전적으로 변경된 세포의 환자로의 역도입이 요구된다. 대조적으로, *in vivo* 유전자 요법은 환자 세포 분리와 정제 과정이 필요하지 않다.

일반적으로, 치료 유전자는 환자에 투여하기 위하여 '포장'된다. 유전자 운반체는 리포좀과 같은 비-바이러스성 물질, 또는 복제-결함 바이러스, 예를 들면 Berkner, K.L., in Curr. Top. Microbiol. Immunol., 158, 39-66(1992)에서 설명하는 아데노바이러스 또는 Muzyczka, N., in Curr. Top. Microbiol. Immunol., 158, 97-129(1992)와 U.S. Patent No. 5.252,479에서 설명하는 아데노-연합된 바이러스(AAV) 벡터이다. 가령, 본 발명의 폴리펩티드를 인코드하는 핵산 분자는 복제-결함성 레트로바이러스 벡터에서 발현을 위하여 조작될 수 있다. 이후, 이런 발현 구조체는 분리하고 폴리펩티드를 인코드하는 RNA를 포함하는 레트로바이러스성 플라스미드 벡터로 형질도입된 포장 세포에 도입하여, 포장 세포가 관심있는 유전자를 포함하는 감염성 바이러스성 입자를 생산할 수 있도록 한다. 이들 생산 세포를 개체에 투여하여, *in vivo* 상태에서 세포를 조작하고, 폴리펩티드를 발현시킨다(참조: Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, in Human Molecular Genetics(1996), T Strachan and A P Read, BIOS Scientific Publishers Ltd).

또 다른 방법은 '나신 DNA'의 투여인데, 여기서 치료 유전자는 혈류 또는 근육 조직에 직접 주사된다.

본 발명의 폴리펩티드 또는 핵산 분자가 질병-원인 물질인 경우에, 본 발명에서는 이들 폴리펩티드 또는 핵산 분자가 질병 원인 물질에 대한 항체를 만드는 백신으로 사용될 수 있도록 한다.

본 발명의 백신은 예방(감염 방지) 또는 치료(감염후 질병 치유)를 목적으로 한다. 이런 백신은 통상적으로 상기한 약학적으로 수용가능한 담체(조성물을 제공받은 개체에서 유해한 항체 생산을 유도하지 않은 담체 포함)와 함께 면역용 항원, 면역원, 폴리펩티드, 단백질 또는 핵산 분자를 함유한다. 게다가, 이들 담체는 면역자극 물질('어쥬번트')로 기능할 수 있다. 더 나아가, 항원 또는 면역원은 디프테리아, 파상풍(테타누스), 콜레라, *H. pylori* 및 다른 병원균과 같은 세균성 독소에 공액될 수 있다.

폴리펩티드는 위에서 분해되기 때문에, 폴리펩티드를 함유하는 백신은 장관외(가령, 피하, 근육, 정맥, 경피 주사) 투여하는 것이 바람직하다. 장관외 투여에 적합한 조성물에는 항산화제, 완충제, 정균제, 개체의 혈액에서 조성물에 등장성을 부여하는 용매를 함유하는 수용성과 비-수용성 무균 주사 용맥 및 현탁제 또는 농후제를 함유하는 수용성과 비-수용성 무균 현탁액이 포함된다.

본 발명의 백신 조성물은 단위 약형 또는 다중 약형 용기로 제공될 수 있다. 가령, 밀봉된 앰플과 바이알은 냉동 건조 상태에 보관하는데, 이는 사용 직전에 무균 액상 담체만 추가하면 된다. 약량은 백신의 특정 활성에 따라 좌우되는데, 통상 적인 실험으로 결정할 수 있다.

본 발명에 따른 폴리펩티드에 결합하는 항체의 유전자 전달은 국제 특허 출원 WO 98/55607에 기술된 바와 같이 달성할 수도 있다. 제트 주사로 지칭되는 이 기술(참조: www.powderjet.com) 역시 백신 조성물의 제제화에 유용하다.

면역화에 적합한 다양한 방법 및 백신 전달 시스템은 국제 특허 출원 WO 00/29428에서 기술한다.

또한, 본 발명은 본 발명에 따른 핵산 분자의 진단 시약으로의 용도에 관한한다. 기능이상과 연관된 본 발명의 핵산 분자에 의해 특성화되는 유전자의 돌연변이 형태의 감지는 유전자의 과소 발현, 과다 발현, 변형된 공간적 또는 시간적 발현 등으로 인한 질병을 정의하거나 진단하거나 또는 질병에 대한 민감성을 진단하는 진단 도구를 제공한다. 다양한 기술을 이용하여, DNA 수준에서 유전자에 돌연변이를 보유하는 개체를 확인할 수 있다.

진단용 핵산 분자는 혈액, 뇨, 타액, 조직 생검 또는 부검 물질과 같은 개체 세포에서 얻을 수 있다. 게놈 DNA는 감지에 직접적으로 이용하거나 분석에 앞서 PCR, 리가아제 연쇄 반응(LCR), 가닥 치환 증폭(SDA) 또는 다른 증폭 기술을 이용하여 효소적으로 증폭할 수 있다(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)).

한 구체예에서, 본 발명은 환자에서 질병을 진단하는 방법을 제공하는데, 상 기 방법은 본 발명의 폴리펩티드를 인코드하는 천연 유전자의 발현 수준을 평가하고, 발현 수준을 대조 수준과 비교하는 과정으로 구성되는데, 이때 대조 수준과 상이한 수준은 질병을 암시한다. 상기 방법은 아래의 단계로 구성된다;

a) 본 발명의 핵산 분자와 프로브간 하이브리드 복합체 형성이 가능한 엄밀도 조건하에서 환자로부터 얻은 조직 시료와 핵산 프로브를 접촉시키고;

b) (a) 단계에서 이용된 조건하에서 프로브와 대조 시료를 접촉시키고;

c) 시료에서 하이브리드 복합체의 존재를 감지하고, 이때 대조 시료에서 하이브리드 복합체의 수준과 비교하여 환자 시료에서 하이브리드 복합체의 상이한 수준의 감지는 질병을 암시한다.

본 발명은 아래의 단계로 구성된 진단 방법을 제공한다;

a) 질병을 검사받는 환자로부터 조직 시료를 얻고;

b) 조직 시료로부터 본 발명의 핵산 분자를 분리하고;

c) 핵산 분자에서 질병과 연관된 돌연변이의 존재를 감지하여 환자에서 질병을 진단한다.

상기한 방법에서 핵산 분자의 감지를 보조하기 위하여, 예로써 PCR을 이용한 증폭 단계가 포함될 수 있다.

정상 유전자형과 비교하여 증폭된 생성물의 크기에서 변화를 이용하여 결손 및 삽입을 확인할 수 있다. 점 돌연변이는 증폭된 DNA를 본 발명의 라벨된 RNA 또는 본 발명의 라벨된 안티센스 DNA 서열에 혼성화시켜 확인할 수 있다. RNase 절단에 의해 또는 용융 온도에서 차이를 확인하여, 미스매치된 이중나선과 완전하게 매 치된 서열을 구별할 수 있다. 환자에서 돌연변이 유무는 하이브리드-이중 가닥 분자를 형성하는 엄밀도 조건하에서 DNA에 혼성화되는 핵산 프로브와 DNA를 접촉시키고, 여기서 하이브리드 이중-가닥 분자는 질병과 연관된 돌연변이에 상응하는 임의의 부분에서 핵산 프로브 가닥의 혼성화되지 않은 부분을 보유하고; DNA 가닥의 상응하는 부분에 질병-연관된 돌연변이 유무의 지표로써 프로브 가닥의 혼성화되지 않은 부분의 유무를 감지함으로써 확인할 수 있다.

이런 진단은 태아 및 신생아에서 특히 유용하다.

참고 유전자와 '돌연변이 유전자'사이에 점 돌연변이 및 다른 서열 차이는 다른 널리 공지된 기술, 예를 들면 직접 DNA 서열화 또는 단일-가닥 형태 다형성으로 확인할 수 있다(참조: Orita et al., Genomics, 5, 874-879 (1989)). 가령, 서열화 프라이머는 이중 가닥 PCR 생성물 또는 변형된 PCR에 의해 생성된 단일-가닥 주형 분자와 함께 사용할 수 있다. 서열 결정은 방사능라벨된 뉴클레오티드를 이용한 통상의 과정 또는 형광-태그를 이용한 자동화 서열화 과정으로 실행한다. 클론된 DNA 절편을 프로브로 이용하여 특정 DNA 단편을 감지할 수도 있다. PCR과 병용하는 경우에 상기 방법의 감응성은 현저하게 향상된다. 게다가, 예로써 단일 뉴클레오티드에 의해 구별되는 서열의 PCR 증폭을 위한 대립형질-특이적 올리고뉴클레오티드를 이용하여 점 돌연변이 및 다형성과 같은 다른 서열 변이, 예를 들면 다형성을 감지할 수 있다.

또한, 변성제 유무에 관계없이 겔에서 DNA 단편의 전기영동 이동성 변화 또는 직접DNA 서열화로 DNA 서열 차이를 확인할 수 있다(Myers et al., Science (1985) 230:1242). 특정 위치에서 서열 변화는 RNase 및 S1 보호와 같은 뉴클레아제 보호 검정 또는 화학적 절단 방법으로 알 수 있다(Cotton et al., Proc. Natl. Acad. Sci. USA (1985) 85: 4397-4401).

통상의 겔 전기영동 및 DNA 서열화 방법 이외에, 마이크로딜리션(microdeletions), 에뉴플로이디즈(aneuploidies), 전치(translocations), 역위(inversion)와 같은 돌연변이는 *in situ* 분석으로 감지할 수 있다(Keller et al., DNA Probes, 2nd Ed., Stockton Press, New York, N.Y., USA (1993)), 다시 말하면 세포에서 DNA 또는 RNA 서열은 분리 또는 막에 고정할 필요없이 돌연변이를 분석할 수 있다. 형광 *in situ* 혼성화 반응(FISH)이 현재 가장 일반적으로 이용된다(Trachuck et al., Science, 250, 559-562 (1990); Trask et al., Trends, Genet., 7, 149-154 (1991)).

본 발명의 또 다른 구체예에서, 본 발명의 핵산 분자를 포함하는 올리고뉴클레오티드 프로브 어레이를 작제하여 유전자 변이, 돌연변이, 다형성을 효과적으로 스크리닝할 수 있다. 어레이 기술은 널리 공지되어 있고 보편적으로 이용할 수 있는데, 유전자 발현, 유전적 연쇄, 유전자 변이성을 비롯한 분자 유전학에서 다양한 문제를 해결하고 있다(M. Chee et al., Science(1996), Vol 274, pp 610-613).

한 구체예에서, 어레이는 PCT 출원 W095/11995(Chee et al); Lockhart, D. J. et al. (1996) Nat. Biotech. 14. 1675-1680); Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93: 10614-10619)에 기술된 방법으로 제조하고 이용한다. 올리고뉴클레오티드 쌍은 2 내지 1만개 범위이다. 올리고머는 광-직접 화학 공정을 이용 하여 기질상의 지정된 지역에서 합성한다. 기질은 종이, 나일론 또는 다른 종류의 막, 필터, 칩, 유리 슬라이드 또는 다른 적절한 고형 서포트이다. 다른 측면에서, PCT 출원 W095/251116(Baldeschweiler et al)에 기술된 바와 같이 화학 결합 과정 및 잉크젯 적용 장치를 이용하여 기질 표면상에 올리고뉴클레오티드를 합성할 수 있다. 또 다른 측면에서, 도트(또는 슬롯) 블랏과 유사한 '격자' 어레이를 이용하고, 진공 시스템, 열, UV, 기계적 또는 화학적 결합 공정에 따라 지질 표면상에 cDNA 단편 또는 올리고뉴클레오티드를 배열하고 연결할 수 있다. 본 명세서에서 어레이는 수동으로 또는 이용가능한 장치(슬롯 블랏 또는 도트 블랏 장치), 재료(임의의 적절한 고형 서포트), 기계(로봇 장치)를 이용하여 만들 수 있고 8개, 24개, 96개, 384개, 1536개 또는 6144개 올리고뉴클레오티드를 포함하거나 시판되는 장치를 효과적으로 이용할 수 있는 2 내지 1만개 이상 범위의 올리고뉴클레오티드를 포함한다.

상기한 방법에 더하여, 개체에서 유래된 샘플로부터 폴리펩티드 또는 mRNA의 비정상적인 증가 또는 감소를 결정하

여 개체의 질병을 진단할 수 있다. 폴리펩티드를 정량하는 당분야에 공지된 임의의 방법, 예를 들면 PCR, RT-PCR, RNase 보호, 노던 블랏, 다른 하이브리드 방법과 같은 핵산 증폭을 이용하여 RNA 수준에서 발현의 증가 또는 감소를 결정할 수 있다.

숙주에서 유래된 시료에서 본 발명의 폴리펩티드 수준을 결정하는데 이용되는 검정 기술은 당분야에 공지되어 있고, 이들중 일부(방사능면역검사, 경쟁-결합 검사, 웨스턴 블랏 분석, ELISA 검사)는 앞서 상술한다. 본 발명의 이런 특징은 아 래의 단계로 구성되는 진단 방법을 제공한다;

(a) 리간드-폴리펩티드 복합체의 형성에 적합한 조건하에서 생물학적 시료와 상기한 리간드를 접촉시키고;

(b) 상기 복합체를 감지한다.

폴리펩티드 수준을 측정하는 ELISA, RIA, FACS와 같은 프로토콜은 폴리펩티드 발현의 변화되거나 비정상적인 수준에 대한 기초를 추가적으로 제공한다. 폴리펩티드 발현에 대한 정상 또는 표준 값은 복합체 형성에 적합한 조건하에서 폴리펩티드에 대한 항체와 정상적인 포유류 개체, 바람직하게는 사람에서 취한 세포 추출물 또는 체액을 복합시켜 확립한다. 표준 복합체 형성 양은 광도계 수단과 같은 다양한 방법으로 정량할 수 있다.

본 발명의 폴리펩티드에 특이적으로 결합하는 항체는 폴리펩티드 발현으로 특성화되는 질병 또는 상태의 진단, 또는 본 발명의 폴리펩티드, 핵산 분자, 리간드, 다른 화합물로 치료되는 환자를 모니터하는 분석에 이용될 수 있다. 진단 목적에 유용한 항체는 치료요법에서와 동일한 방식으로 만들 수 있다. 폴리펩티드의 진단 분석에는 사람 체액, 또는 세포 또는 조직의 추출물에서 폴리펩티드를 감지하는 라벨과 항체를 이용하는 방법이 포함된다. 항체는 변형하거나 변형하지 않고 이용할 수 있고, 리포터 분자와 공유적으로 또는 비-공유적으로 결합시켜 라벨링할 수도 있다. 당분야에 공지된 다양한 리포터 분자를 이용할 수 있으며, 일부는 앞서 기술한다.

개체, 대조 및 생검 조직에서 얻은 질병 시료에서 발현된 폴리펩티드의 양은 표준 수치와 비교한다. 표준과 개체 수치간 편차는 질병 진단을 위한 파라미터를 확립한다. 진단 검사를 이용하여 폴리펩티드 발현의 부재, 존재, 과다 발현을 구별하고, 치료 개입 동안 폴리펩티드 수준의 조절을 모니터할 수 있다. 이런 검사를 이용하여 동물 연구, 임상 시험, 개별 환자의 치료 모니터링에서 특정 치료 섭생의 효과를 평가할 수도 있다.

본 발명의 진단 키트는 아래와 같은 구성된다;

(a) 본 발명의 핵산 분자;

(b) 본 발명의 폴리펩티드;

(c) 본 발명의 리간드.

본 발명의 한 특징에서, 진단 키트는 엄밀도 조건하에서 본 발명의 핵산과 혼성화될 수 있는 핵산을 보유하는 제 1 용기; 핵산 분자를 증폭하는데 유용한 프라이머를 보유하는 제 2 용기; 질병 진단을 조장하는 프로브와 프라이머의 사용설명서를 포함한다. 키트에는 혼성화되지 않은 RNA를 절단하는 약품을 보유하는 제 3 용기가 추가로 포함될 수 있다.

본 발명의 다른 특징에서, 진단 키트에는 핵산 분자 어레이가 포함되는데, 이중 적어도 하나는 본 발명의 핵산 분자이다.

본 발명의 폴리펩티드를 감지하기 위하여, 진단 키트는 본 발명의 폴리펩티드에 결합하는 한 가지 이상의 항체; 항체와 폴리펩티드 사이에 결합 반응을 감지하는데 유용한 시약을 포함한다.

이런 키트는 세포 증식성 질환, 자가면역/염증 질환, 심혈관 질환, 신경과 정신 질환, 발달 장애, 유전 질환, 대사 질환, 감염 및 다른 병인성 질환을 포함하지만 이들에 국한되지 않는 질병 또는 질병에 대한 감수성을 진단하는데 유용하다. 적절하게는, 이들 질환은 신생물, 암, 뇌종양, 신경교종, 골암, 폐암, 유암암, 전립선암, 결장암, 혈관종, 골수증식성 질환, 백혈병, 혈액학적 질환, 호중구 감소증, 혈소판 감소증, 혈관신생 질환, 피부 질환, 노화, 상처, 화상, 섬유증, 심혈관 질환, 재협착증, 심장 질환, 말초 혈관 질환, 관상 동맥 질환, 부종, 혈전색전증, 월경통, 자궁내막증, 전자간증, 폐 질환, COPD, 천식 골 질환, 신장 질환, 사구체신염, 간 질환, 크론병, 위염, 궤양성 대장염, 궤양, 면역 질환, 자가면역 질환, 관절염, 류머티스 관절염, 건선, 수포성 표피박리증, 전신 홍반성 루퍼스, 강직성 척추염, 라임병, 다발성 경화증, 신경퇴행, 뇌졸중, 뇌/척수 손상, 알츠하이머병, 파킨슨병, 운동 신경 질환, 신경근육 질환, HIV, AIDS, 시토메갈로바이러스 감염, 진균 감염, 안구 질환, 망막황반 이상, 녹내장, 당뇨성 망막병증, 안구 고혈압 및 면역글로불린 도메인-포함 세포 표면 인식 분자가 관여하는 다른 이상 등이다.

본 발명의 다양한 특징 및 구체예는 특히 INSP052와 INSP055 폴리펩티드와 관련된 실시예를 통하여 상술한다.

당업자가 인지하는 바와 같이, 상세한 설명의 개변은 본 발명에 범위에 포섭된다.

도면의 간단한 설명

도 1: 전장 INSP052 폴리펩티드 서열을 이용한 NCBI 비-중복 데이터베이스에서 BLAST의 결과.

도 2: 전장 INSP052 폴리펩티드 서열 및 가장 밀접하게 관련된 서열, 담즙 당단백질 H(생쥐)간 BLAST에 의한 정렬.

도 3: 전장 INSP055 폴리펩티드 서열을 이용한 NCBI 비-중복 데이터베이스에서 BLAST의 결과.

도 4: 전장 INSP055 폴리펩티드 서열 및 가장 밀접하게 관련된 서열, 담즙 당단백질 H(생쥐)간 BLAST에 의한 정렬.

도 5: 번역되는 INSP052의 예상 뉴클레오티드 서열. 밑줄된 서열은 예상 신호 펩티드를 나타낸다. 박스로 표시된 서열은 예상 막통과 도메인을 나타낸다.

도 6: 게놈 DNA에서 INSP052 코딩 엑손 조직. 아래쪽 = INSP052.cDNA, 1251 bp. 위쪽 = chr11.게놈-DNA. 추정 세포외 도메인을 인코드하는 서열은 밑줄로 표시한다. 출발 코돈과 종결 코돈은 굵은 글씨로 표시한다.

도 7: 클론된 INSP052 세포외 도메인의 뉴클레오티드 서열과 번역

도 8: pENTR-INSP052-EC-6HIS의 유전자 지도

도 9: pEAK12d-INSP052-EC-6HIS의 유전자 지도

실시예

실시예 1: INSP052와 INSP055

INSP052로부터 구성적 엑손의 번역을 대표하는 SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16을 결합시켜 파생된 폴리펩티드 서열은 사람 게놈 서열로부터 유래된다. INSP055를 나타내는 폴리뉴클레오티드와 폴리펩티드 서열 SEQ ID NO 17과 SEQ ID 18은 각각 INSP052의 생쥐 오르토로그(orthologue)의 폴리뉴클레오티드와 폴리펩티드이다. 생쥐 오르토로그의 존재는 사람 서열 INSP052의 유전자 모델을 뒷받침한다.

각각 SEQ ID NO 16과 SEQ ID NO 18로 대표되는 INSP052와 INSP055 폴리펩티드 서열은 신호 펩티드 서열 및 막통과 범위 도메인을 보유할 것으로 예상된다.

INSP052로부터 구성적 엑손의 번역을 대표하는 SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, SEQ ID NO:16을 결합시켜 파생된 폴리펩티드 서열은 NCBI 비-중복 서열 데이터베이스에서 BLAST 쿼리로 이용하였다. 상위 10개의 매치는 도 1에 도시하는데, 이들 모두 면역글로불린 도메인-포함 단백질이다.

도 2에서는 INSP052 쿼리 서열 및 가장 근접하는 공지된 단백질, 담즙 당단백질 H(생쥐)의 서열의 정렬을 도시한다.

폴리펩티드 서열 INSP055는 NCBI 비-중복 서열 데이터베이스에서 BLAST 쿼리로 이용하였다. 상위 10개의 매치는 도 3에 도시한다. 도 4에서는 INSP055 쿼리 서열 및 가장 근접하는 공지된 단백질, 담즙 당단백질 H(생쥐)의 서열의 정렬을 도시한다.

사람과 생쥐에서 INSP052와 INSP055 전사체를 나타내는 발현된 서열 태그(EST)는 아래의 cDNA 라이브러리: 소뇌, 외피, 해마상, 시상하부, 연수를 비롯한 뇌, 내이(inner ear), 유방으로부터 기원한다. 또한, 전사체는 핍지교종, 교모세포종, 다발성 경화증 병소로부터 EST로 대표된다. 이는 INSP052가 상기 조직으로부터 클론될 수 있고 상기 조직의 질병과 연관한다는 것을 암시한다. 따라서, 본원에 개시된 폴리펩티드, 항체 및 다른 성분은 상기 조직에서 질병을 치료하는데 유용하다.

실시예 2: 엑손 어셈블리에 의한 INSP052 세포외 도메인의 클로닝

예상 INSP052 전장은 면역글로불린 대과에 속하는 VEGF/PDGF 수용체와 관련된 416개 아미노산의 I형 막 단백질을 인코드한다. 개시 ATG 코돈에서부터 폴리 A 꼬리까지 상기 예상 뉴클레오티드 서열은 2025개 뉴클레오티드를 보유한다(도 5). 코딩 서열(cd)은 7개 엑손에 걸쳐있다(도 6). 추정 신호 서열(아미노산 1-33을 인코딩)은 엑손 1에 위치한다. 예상 막통과(TM) 도메인(아미노산 241 내지 263)을 인코드하는 서열은 엑손 3-4 경계에 위치한다.

아미노산 1-240을 인코드하는 세포외(EC) 도메인은 게놈 DNA로부터 엑손 어셈블리로 클론하였다. 엑손 어셈블리 방법의 개관은 아래에 요약한다:

- 개별 엑손 1, 2, 3은 PCR로 게놈 DNA로부터 증폭하였다. 엑손 3에 대한 역행 프라이머는 엑손 4의 5' 서열과 11개 염기가 중복되었다.

- 겔-정제된 엑손을 혼합하고 2차 PCR 반응을 실시하여 재-조합된 DNA를 증폭하였다.

- INSP052 EC 도메인에 상응하는 전장 PCR 산물은 겔-정제하고, 이후 Gateway ™ 방법을 이용하여 pDONR 201 (Gateway 진입 벡터) 및 pEAK12d(발현 벡터)에 서브클론하였다.

1. 게놈 DNA로부터 INSP052의 세포외 도메인을 인코드하는 엑손의 PCR 증폭.

엑손 1, 2, 3을 개별적으로 증폭하는 PCR 프라이머를 설계하였다(표 1). 엑손 1에 대한 선행 프라이머(INSP052-B1P-exon1F)는 Gateway attB1 부위(5' GCAGGCTTC)와 Kozak 서열(5' GCCACC)의 일부 서열을 또한 포함한다. 엑손 1에 대한 역행 프라이머(INSP052-exon1R)는 5' 말단에서 엑손 2와 18개 염기가 중복된다. 엑손 2에 대한 선행 프라이머(INSP052-exon2F) 5' 말단에서 엑손 1과 18개 bp가 중복된다. 엑손 2에 대한 역행 프라이머(INSP052-exon2R)는 5' 말단에서 엑손 3과 18개 염기가 중복된다. 엑손 3에 대한 선행 프라이머(INSP052-exon3F)는 5' 말단에서 엑손 2와 17 bp가 중복된다. 엑손 3에 대한 역행 프라이머(INSP052-exon3R)는 5' 말단에서 엑손 4와 11개 염기가 중복된다.

INSP052 엑손 1의 증폭을 위하여, PCR 반응은 50 μℓ의 최종 용적으로 실시하고, 1.5 μℓ의 사람 게놈 DNA(0.1 μg/μℓ, Novagen cat. no. 69237), 2 μℓ의 5 mM dNTP(Amersham Pharmacia Biotech), 6 μℓ의 INSP052-B1P-exon1F(10 ㅅM), 6 μℓ의 INSP052-exon1R, 5 μℓ의 10X Pwo 완충액, 0.5 μℓ Pwo 중합효소(5 U/μℓ)(Roche, cat. no. 1 644 955)가 포함되었다. PCR 조건은 94℃에서 2분; 94℃에서 30초, 60℃에서 30초, 72℃에서 1분의 35회 사이클; 72℃에서 5분의 추가 신장 사이클; 4℃에서 유지 사이클이었다. 반응 산물은 1.5% 아가로즈 겔(1X TAE)에 적하하고, 정확한 크기(118 bp)의 PCR 산물은 Qiaquick Gel Extraction Kit(Qiagen cat. no. 28704)로 겔-정제하고 50 μℓ 용리 완충액(Qiagen)에서 용리하였다.

엑손 2는 동일한 반응 조건하에 프라이머 INSP052-exon2F와 INSP052-exon2R 로 증폭하였다. 378 bp의 PCR 산물은 상기한 바와 같이 겔-정제하였다.

엑손 3은 동일한 반응 조건하에 프라이머 INSP052-exon3F와 INSP052-exon3R로 증폭하였다. 321 bp의 PCR 산물은 상기한 바와 같이 정제하였다.

2. INSP052의 세포외 도메인-인코딩 엑손의 어셈블리

엑손 1, 2 및 3-4는 5 μℓ의 각 겔-정제된 엑손, 2 μℓ의 5 mM dNTP, 6 μℓ의 INSP052-B1P-exon1F(10 μM), 6 μℓ의 INSP052-5HIS-R(10 μM), 5 μℓ의 10X Pfu 완충액, 14.5 μℓ H_2O, 0.5 μℓ Pfu 중합효소(3 U/μℓ; Promega cat. no. M774B)를 함유하는 PCR 반응물에서 재-조합하였다. 반응 조건은 아래와 같았다: 94℃에서 4분; 94℃에서 30초, 48℃에서 30초, 70℃에서 2분의 25회 사이클; 70℃에서 10분의 추가 신장 단계; 4℃에서 유지 사이클. 반응 산물은 1.5 % 아가로즈 겔(1X TAE)에서 분석하였다. 정확한 크기(750 bp)의 PCR 산물은 Qiaquick Gel Extraction Kit(Qiagen cat. no. 28704)로 겔-정제하고 50 μℓ의 용리 완충액(Qiagen)에서 용리하였다. 생성 산물(INSP052 EC ORF)은 5' 말단에서 attB1 부위와 Kozak 서열, 3' 말단에서 5HIS 태그 인코딩 서열이 측면에서 접하는 INSP052 EC 도메인의 ORF를 보유한다.

3. INSP052 EC 도메인 ORF의 pDONR201로의 서브클로닝

AttB1과 attB2 재조합 부위는 50 μℓ의 최종 용적으로 2 μℓ의 각 겔-정제된 INSP052 EC ORF, 2 μℓ의 5 mM dNTP(Amersham Pharmacia Biotech), 6 μℓ의 GCP-Forward(10 μM), 6 μℓ의 GCP-Reverse(10 μM), 5 μℓ의 10X Vent 완충액, 0.5 μℓ의 Vent DNA 중합효소(2 U/μℓ)(New England Biolabs, cat. no. M0254S)를 함유하는 PCR 반응물

에서 전장 INSP052 EC 도메인 서열의 5'와 3' 말단에 추가하였다. PCR 조건은 94℃에서 2분; 94℃에서 30초, 55℃에서 30초, 72℃에서 1분의 30회 사이클; 72℃에서 3분의 추가 신장 단계; 4℃에서 유지 사이클이었다. 반응 산물은 1.5% 아가로즈 겔(1X TAE)에서 분석하고, 정확한 크기(808 bp)의 PCR 산물은 Qiaquick Gel Extraction Kit(Qiagen cat. no. 28704)로 겔-정제하고 50 ㎕ 용리 완충액(Qiagen)에서 용리하였다. 정제된 PCR 산물(Gateway-변형된 INSP052 EC 도메인)은 아래와 같이 BP 클로나아제를 이용하여 pDONR201로 이전하였다: 5 ㎕의 Gateway-변형된 INSP052 EC 도메인은 실온에서 1시간동안 1.5 ㎕ pDONR201(0.1 ㎍/㎕), 2 ㎕ BP 완충액, 1.5 ㎕ BP 클로나아제 효소 혼합물(Invitrogen)과 함께 배양하였다. 반응물은 단백분해효소 K(2 ㎍)를 첨가하여 반응을 중단시키고, 37℃에서 추가로 10분동안 배양하였다. 상기 반응물의 분량(1 ㎕)은 Biorad Gene Pulser를 이용한 전기천공으로 20 ㎕의 대장균(*E. coli*) DH10B 세포(H_20에서 1/5 희석)에 형질전환시켰다. 전기천공된 세포는 1 ㎖ SOC 배지를 첨가하여 희석하고 37℃에서 1시간동안 배양하였다. 형질전환체는 LB-카나마이신 평판에 도말하고 37℃에서 하룻밤동안 배양하였다. 플라스미드 mini prep DNA는 Qiaprep Turbo 9600 자동화 시스템(Qiagen)을 이용하여 1-10개 생성 콜로니로부터 분리하고, 제조업자의 지시에 따라 BigDyeTerminator 시스템(Applied Biosystems cat. no. 4390246)을 이용하여 pENTR-F1과 pENTR-R1 서열화 프라이머로 DNA 서열분석하였다. 서열화 반응물은 Dye-Ex 칼럼(Qiagen) 또는 Montage SEQ 96 정화 평판(Millipore cat. no. LSKS09624)으로 정제하고, 이후 Applied Biosystems 3700 서열분석기에서 분석하였다.

4. INSP052 EC 도메인 ORF의 발현 벡터 pEAK12d로의 서브클로닝

INSP052 EC 도메인의 정확한 서열을 보유하는 pDONR201 클론(플라스미드 ID # 13497)의 플라스미드 용출액(1.5 ㎕)은 10 ㎕의 최종 용적으로 1.5 ㎕ pEAK12d(0.1 ㎍/㎕), 2 ㎕ LR 완충액, 1.5 ㎕ LR 클로나아제(Invitrogen)를 함유하는 재조합 반응물에 사용하였다. 혼합물은 실온에서 1시간동안 배양하고, 단백분해효소 K(2 ㎍)를 첨가하여 반응을 중단시키며, 37℃에서 추가로 10분동안 배양하였다. 상기 반응물의 분량(1 ㎕)은 전술한 바와 같이 전기천공으로 대장균(*E. coli*) DH10B 세포를 형질전환시켰다. 전기천공된 세포는 1 ㎖ SOC 배지를 첨가하여 희석하고 37℃에서 1시간동안 배양하였다. 형질전환체는 LB-암피실린 평판에 도말하고 37℃에서 하룻밤동안 배양하였다. mini prep DNA는 Qiaprep Turbo 9600 자동화 시스템(Qiagen)을 이용하여 4개 콜로니로부터 분리하고 50 ㎕ 용리 완충액에서 용리하였다. 이후, 2 ㎕의 각 miniprep은 2 ㎕의 5 mM dNTP, 6 ㎕의 10 μM pEAK12-F, 6 ㎕의 10 μM pEAK12-R, 5 ㎕의 10X AmpliTaq ™ 완충액, 0.5 ㎕ AmpliTaq ™ (Applied Biosystems cat. no. N808-0155)를 함유하는 50 ㎕의 전체 반응 용적에서 PCR 처리하였다. 사이클링 조건은 아래와 같았다: 94℃에서 2분; 94℃에서 30초, 55℃에서 30초, 72℃에서 1분의 30회 사이클; 72℃에서 3분의 1회 사이클. 이후, 시료는 추 가 분석에 앞서 4℃에 유지하였다(유지 사이클).

플라스미드 mini prep DNA는 예상 PCR 산물 크기(1074 bp)를 제공하는 클론으로부터 분리하고, 이후 pEAK12-F와 pEAK12-R 서열화 프라이머로 DNA 서열분석하였다.

플라스미드 pEAK12d-INSP052EC-6HIS(플라스미드 ID # 13495)의 CsCl 구배 정제된(gradient purified) maxi-prep DNA는 서열 확인된 클론의 500 ㎖ 배양물로부터 준비하고(Sambrook J. et al., in Molecular Cloning, a Laboratory Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press) 무균수에서 1㎍/㎕ 농도로 재부유시키며 -20℃에 보관하였다.

표 2

INSP052 EC 도메인 클로닝과 서열화를 위한 프라이머

프라이머	서열 (5'-3')
GCP Forward	G GGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GCC ACC
GCP Reverse	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTT **TCA** *ATG GTG ATG GTG ATG GTG*
INSP052-B1P-exon1F	GCA GGC TTC GCC ACC ATG AAG AGA GAA AGG GGA GCC CTG TC
INSP052- exon1R	TCA CCC CCT CCA GGG GGT **CTG TCT GGA TCA GAA GAA**
INSP052- exon2F	**TTC TTC TGA TCC AGA CAG** ACC CCC TGG AGG GGG TGA
INSP052- exon2R	GTG GCC TCG AAA TGG GCA **CAT CTA CAG TAA GGT TGA**
INSP052- exon3F	**CAA CCT TAC TGT AGA TG** T GCC CAT TTC GAG GCC ACA
INSP052- exon3R	**GGA GCT TCT TC** T GTA TAC GGT GAT CTT GAC AG
INSP052-5HIS-R	*GTG ATG GTG ATG GTG* GGA GCT TCT TCT GTA TAC GG

pEAK12-F	GCC AGC TTG GCA CTT GAT GT
pEAK12-R	GAT GGA GGT GGA CGT GTC AG
pENTR-F1	TCG CGT TAA CGC TAG CAT GGA TCT C
pENTR-R1	GTA ACA TCA GAG ATT TTG AGA CAC

밑줄된 서열 = Kozak 서열

굵은 글씨 = 종결 코돈

이탤릭체 서열 = His tag

밑줄된 굵은 글씨 = 인접 엑손과 중복

실시예 3: 클론된 His-태그된 INSP052-6His-V1(플라스미드 no. 13495)의 포유류 세포에서 발현

Epstein-Barr 바이러스 핵 항원(HEK293-EBNA, Invitrogen)을 발현하는 사람 배아 신장 293 세포는 Ex-cell VPRO 무-혈청 배지(원종, 유지용 배지, JRH) 현탁액에 유지시켰다. 트랜스펙션(D-1) 16 내지 20시간 전에, 세포는 2x T225 플라스크(2X10 ^{5} 세포/㎖ 밀도로 2% FBS 접종 배지(JRH)를 함유하는 DMEM/F12(1:1)에서 50㎖/플라스크)에 접종하였다. 다음날(트랜스펙션 당일), JetPEI ™ 시약(2㎕/㎍ 플라스미드 DNA, PolyPlus-트랜스펙션)을 이용하여 트랜스펙션을 실시하였다. 각 플라스크에서, 113 ㎍ cDNA(플라스미드 No. 13495)를 2.3 ㎍ GFP(형광 리포터 유전자)와 함께 공동 트랜스펙션시켰다. 이후, 트랜스펙션 혼합물은 2xT225 플라스크에 첨가하고, 6일간 37℃(5% CO_2)에서 배양하였다. 좀더 많은 재료를 얻기 위하여, 여분의 플라스크에서 이런 과정을 반복하여 총 200 ㎖를 수득하였다. 1일과 6일에 정성적 형광 검사로 양성 트랜스펙션을 확증하였다(Axiovert 1O Zeiss).

6일에(수확일), 4개 플라스크로부터 상층액(200 ㎖)을 모으고 원심분리하며(4℃, 400g) 독특한 확인물질을 포함하는 용기에 위치시켰다.

6His-태그 단백질(internal bioprocessing QC)의 QC를 위하여 1 분량(500㎕)을 보관하였다.

정제 과정

C-말단 6His 태그를 보유하는 재조합 단백질을 포함하는 200 ㎖ 배양 배지 시료는 냉각 완충액 A(50 mM NaH_2PO_4; 600 mM NaCl; 8.7%(w/v) 글리세롤, pH 7.5)를 이용하여 200㎖ 최종 용적으로 희석하였다. 시료는 0.22㎛ 멸균 필터(Millipore, 500㎖ 필터 단위)를 통과시켜 여과하고 250 ㎖ 무균 사각 배지 병(Nalgene)에서 4℃로 유지시켰다.

정제는 자동 시료 적하장치(Labomatic)에 연결된 VISION workstation (Applied Biosystems)에서 4℃에서 실시하였다. 정제 과정은 2단계, Ni 이온이 충전된 Poros 20 MC(Applied Biosystems) 칼럼(4.6 x 50 mm, 0.83 ㎖)에서 금속 친화성 크로마토그래피, 이후 Sephadex G-25 배지(Amersharn Pharmacia) 칼럼(1.0 x 10 cm)에서 겔 여과로 구성되었다.

제 1 크로마토그래피 단계에서, 금속 친화력 칼럼은 30 칼럼 용적의 EDTA 용액(100 mM EDTA; 1M NaCl; pH 8.0)으로 재생시키고, 15 칼럼 용적의 100 mM $NiSO_4$ 용액으로 세척하여 Ni 이온으로 재충전하며, 10 칼럼 용적의 완충액 A와 7 칼럼 용적의 완충액 B(50 mM NaH_2PO_4; 600 mM NaCl; 8.7%(w/v) 글리세롤, 400 mM; 이미다졸, pH 7.5)로 순차적으로 세척하고, 최종적으로 15 mM 이미다졸을 함유하는 15 칼럼 용적의 완충액 A로 평형시켰다. 시료는 Labomatic 시료 적하장치로 200 ㎖ 시료 루프로 옮기고, 후속으로 분당 10 ㎖ 유속으로 Ni 금속 친화성 칼럼에 충전시켰다. 칼럼은 12 칼럼 용적의 완충액 A, 이후 20 mM 이미다졸을 함유하는 28 칼럼 용적의 완충액 A로 세척하였다. 20 mM 이미다졸 세척 동안, 느슨하게 결합된 오염 단백질이 칼럼으로부터 용리된다. 최종적으로, 재조합 His-태그된 단백질은 분당 2㎖ 유속으로 10 칼럼 용적의 완충액 B로 용리시키고, 용리된 단백질은 1.6 ㎖ 분획물로 수집하였다.

제 2 크로마토그래피 단계에서, Sephadex G-25 겔-여과 칼럼은 2㎖ 완충액 D(1.137M NaCl; 2.7 mM KCl; 1.5 mM KH_2PO_4; 8mM Na_2HP0_4; pH 7.2)로 재생시키고, 후속으로 4 칼럼 용적의 완충액 C(137 mM NaCl; 2.7 mM KCl, 1.5 mM KH_2PO_4; 8 mM Na_2HP0_4; 20%(w/v) 글리세롤; pH 7.4)로 평형시켰다. Ni-칼럼으로부터

용리되는 피크 분획은 VISION에 통합된 시료 적하장치를 통하여 Sephadex G-25 칼럼으로 자동적으로 적하하고, 단백질은 분당 2㎖ 유속으로 완충액 C로 용리하였다. 탈염된 시료는 2.2㎖ 분획물에 수집하였다. 상기 분획물은 0.22 ㎛ 무균 원심분리 필터(Millipore)를 통하여 여과하고 동결시키며 -80℃에서 보관하였다. 시료의 1 분량은 SDS-PAGE(4-12 % NuPAGE gel; Novex)에서 코마시 블루 착색 및 항-His 항체를 이용한 웨스턴 블랏으로 분석하였다.

코마시 착색: NuPAGE 겔은 실온에서 1시간동안 0.1% 코마시 블루 R250 착색 용액(30% 메탄올, 10% 아세트산)으로 착색하고, 후속으로 배경 부분이 투명해지고 단백질 밴드가 선명하게 보일 때까지 20% 메탄올, 7.5% 아세트산에서 탈색시켰다.

웨스턴 블랏: 전기영동이후, 단백질은 4℃에서 1시간동안 겔로부터 니트로셀룰로오즈 막으로 290mA에서 전기적으로 이동시켰다. 막은 완충액 E(137 mM NaCl; 2.7 mM KCl; 1.5 mM KH_2PO_4; 8 mM Na_2HP0_4; 0.1% Tween 20, pH 7.4)에서 5% 우유 분말로 1시간동안 실온에서 차단하고, 후속으로 2가지 토끼 다클론 항-His 항체(G-18 및 H-15, 각각 0.2 ㎍/㎖; Santa Cruz)/2.5% 우유 분말/완충액 E의 혼합물과 함께 4℃에서 하룻밤동안 배양하였다. 실온에서 1시간 배양이후, 막은 완충액 E(3x10 분)로 세척하고, 이후 2.5% 우유 분말을 함유하는 완충액 E에서 1/3000으로 희석된 2차 HRP-공액된 항-토끼 항체(DAKO, HRP 0399)와 함께 실온에서 2시간동안 배양하였다. 완충액 E(3 x 10 분)로 세척한 이후, 막은 ECL 키트(Amersham Pharmacia)로 1분간 현상시켰다. 막은 후속으로 Hyperfilm (Amersham Pharmacia)에 노출시키고, 필름은 현상하고, 웨스턴 블랏 이미지는 시각적으로 분석하였다.

단백질 분석: 단백질 농도는 소 혈청 알부민을 표준으로 하는 BCA 단백질 분석 키트(Pierce)를 이용하여 결정하였는데, 890 ㎍ 정제된 단백질이 200 ㎖ 배양 배지로부터 회수되었다.

(57) 청구의 범위

청구항 1.

아래에서 선택되는 폴리펩티드:

(i) SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 또는 SEQ ID NO:16에 상술된 아미노산 서열, 또는 도 7에 제시된 INSP052의 세포외 도메인의 아미노산 서열을 포함하거나 이들 서열로 구성되는 폴리펩티드;

(ii) (i)에 따른 폴리펩티드의 활성을 보유하거나 (i)에 따른 폴리펩티드와 공통의 항원 결정부위를 보유하는 단편;

(iii) (i) 또는 (ii)의 기능적 등가체.

청구항 2.

제 1 항에 있어서, SEQ ID NO:16에 상술된 아미노산 서열을 포함하거나 이런 서열로 구성되는 것을 특징으로 하는 폴리펩티드.

청구항 3.

제 2 항에 있어서, SEQ ID NO:16에 상술된 아미노산 서열로 구성되는 것을 특징으로 하는 폴리펩티드.

청구항 4.

SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 또는 SEQ ID NO:16에 상술된 아미노산 서열에 상동하고 면역글로불린 도메인-포함 세포 표면 인식 분자 활성을 보유하는 제 1 항의 (iii)에 따른 폴리펩티드의 기능적 등가체.

청구항 5.

SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 또는 SEQ ID NO:16에 상술된 아미노산 서열, 또는 이의 활성 단편과 84% 이상, 바람직하게는 85%, 90%, 95%, 98% 또는 99% 이상의 서열 상동성을 보유하는 전술한 항중 어느 한 항에 따른 폴리펩티드의 기능적 등가체 또는 활성 단편.

청구항 6.

SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14 또는 SEQ ID NO:16에 상술된 아미노산 서열을 보유하는 폴리펩티드와 상당한 구조적 유사성을 나타내는 전술한 항중 어느 한 항에 따른 폴리펩티드의 기능적 등가체.

청구항 7.
제 1 항의 (i)에 따른 폴리펩티드와 공통의 항원 결정부위를 보유하고, SEQ ID NO:16에 상술된 아미노산 서열로부터 7개 또는 그 이상(예, 8개, 10개, 12개, 14개, 16개, 18개, 20개 또는 그 이상)의 아미노산 잔기로 구성되는 전술한 항중 어느 한 항에 따른 폴리펩티드의 단편.

청구항 8.
전술한 한중 어느 한 항에 따른 폴리펩티드를 인코드하는 정제된 핵산 분자.

청구항 9.
제 8 항에 있어서, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13 또는 SEQ ID NO:15에 상술된 핵산 서열, 이의 등가체 또는 단편을 포함하는 것을 특징으로 하는 핵산 분자.

청구항 10.
제 9 항에 있어서, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13 또는 SEQ ID NO:15에 상술된 핵산 서열, 이의 등가체 또는 단편으로 구성되는 것을 특징으로 하는 핵산 분자.

청구항 11.
높은 엄밀도 조건하에서 제 8 항 내지 10 항중 어느 한 항에 따른 핵산 분자와 혼성화되는 정제된 핵산 분자.

청구항 12.
제 8 내지 11 항중 어느 한 항에 따른 핵산 분자를 포함하는 벡터.

청구항 13.
제 12 항에 따른 벡터로 형질전환된 숙주 세포.

청구항 14.
제 1 내지 7 항중 어느 한 항에 따른 폴리펩티드에 특이적으로 결합하여 이의 활성을 저해하는 리간드.

청구항 15.
제 14 항에 있어서, 항체인 것을 특징으로 하는 리간드.

청구항 16.
제 1 내지 7 항중 어느 한 항에 따른 폴리펩티드의 발현이나 활성 수준을 증가 또는 감소시키는 화합물.

청구항 17.
제 16 항에 있어서, 제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드에 결합하지만 폴리펩티드의 생물학적 효과를 유도하지 않는 것을 특징으로 하는 화합물.

청구항 18.
제 17 항에 있어서, 천연이나 변형된 기질, 리간드, 효소, 수용체, 구조적 또는 기능적 모방체인 것을 특징으로 하는 화합물.

청구항 19.
질병의 치료 또는 진단에 사용되는 제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드, 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자, 제 12 항에 따른 벡터, 제 13 항에 따른 숙주 세포, 제 14 항 또는 15 항에 따른 리간드, 또는 제 16 항 내지 18 항중 어느 한 항에 따른 화합물.

청구항 20.
환자에서 질병을 진단하는 방법에 있어서, 아래의 단계의 구성되는 것을 특징으로 하는 방법:

환자의 조직에서 제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드를 인코드하는 천연 유전자의 발현 수준을 평가하거나 제 1 항 내지 제 7 항중 어느 한 항에 따른 폴리펩티드의 활성 수준을 평가하고;

발현이나 활성 수준을 대조 수준과 비교하고;

대조 수준과 상이한 수준은 질병을 암시한다.

청구항 21.
제 20 항에 있어서, *in vivo* 로 실시되는 것을 특징으로 하는 방법.

청구항 22.
제 20 항 또는 21 항에 있어서, 아래의 단계로 구성되는 것을 특징으로 하는 방법:

(a) 리간드-폴리펩티드 복합체의 형성에 적합한 조건하에서 제 14 항 또는 15 항에 따른 리간드를 생물학적 시료와 접촉시키고;

(b) 상기 복합체를 감지한다.

청구항 23.
제 20 항 또는 21 항에 있어서, 아래의 단계로 구성되는 것을 특징으로 하는 방법:

(a) 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자와 프로브간 하이브리드 복합체의 형성에 적합한 엄밀도 조건하에서 핵산 프로브와 환자로부터 취한 조직 시료를 접촉시키고;

b) a) 단계에서와 동일한 조건하에서 대조 시료와 프로브를 접촉시키고;

c) 시료에서 하이브리드 복합체의 존재를 감지하고;

대조 시료에서 하이브리드 복합체의 수준과 비교하여 환자 시료에서 하이브리드 복합체의 상이한 수준의 감지는 질병을 암시한다.

청구항 24.
제 21 항 또는 22 항에 있어서, 아래의 단계로 구성되는 것을 특징으로 하는 방법:

(a) 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자와 프라이머간 하이브리드 복합체의 형성에 적합한 엄밀도 조건하에서 핵산 프라이머와 환자 조직의 핵산 시료를 접촉시키고;

b) a) 단계에서와 동일한 조건하에서 대조 시료와 프라이머를 접촉시키고;

c) 견본으로 조사된 핵산을 증폭하고;

d) 환자 시료와 대조 시료로부터 증폭된 핵산의 수준을 감지하고;

대조 시료에서 증폭된 핵산 수준과 비교하여 환자 시료에서 증폭된 핵산의 현저하게 상이한 수준의 감지는 질병을 암시한다.

청구항 25.
제 21 항 또는 22 항에 있어서, 아래의 단계로 구성되는 것을 특징으로 하는 방법:

a) 질병을 검사받는 환자로부터 조직 시료를 얻고;

b) 조직 시료로부터 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자를 분리하고;

c) 질병의 지표로써 핵산 분자에서 질병과 연관된 돌연변이의 존재를 감지하여 환자에서 질병을 진단한다.

청구항 26.
제 25 항에 있어서, 핵산 분자를 증폭시켜 증폭 산물을 만들고, 증폭 산물에서 돌연변이 유무를 감지하는 단계가 추가로 포함되는 것을 특징으로 하는 방법.

청구항 27.
제 25 항 또는 26 항에 있어서, 하이브리드-이중 가닥 분자를 형성하는 엄밀도 조건하에서 핵산 분자에 혼성화되는 핵산 프로브와 핵산 분자를 접촉시키고, 여기서 하이브리드 이중-가닥 분자는 질병과 연관된 돌연변이에 상응하는

임의의 부분에서 핵산 프로브 가닥의 혼성화되지 않은 부분을 보유하고; 질병-연관된 돌연변이 유무의 지표로써 프로브 가닥의 혼성화되지 않은 부분의 유무를 감지함으로써 환자에서 돌연변이 유무를 확인하는 것을 특징으로 하는 방법.

청구항 28.

제 20 항 내지 27 항중 어느 한 항에 있어서, 질병은 세포 증식성 질환, 자가면역/염증 질환, 심혈관 질환, 신경 질환, 정신 질환, 발달 장애, 유전 질환, 대사 질환, 감염 또는 다른 병인성 질환인 것을 특징으로 하는 방법.

청구항 29.

면역글로불린 도메인-포함 세포 표면 인식 분자로써 제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드의 용도.

청구항 30.

제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드, 제 8 항 내지 11 항 중 어느 한 항에 따른 핵산 분자, 제 12 항에 따른 벡터, 제 13 항에 따른 숙주 세포, 제 14 항 또는 15 항에 따른 리간드, 또는 제 16 내지 18 항중 어느 한 항에 따른 화합물을 함유하는 약학 조성물.

청구항 31.

제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드 또는 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자를 함유하는 백신 조성물.

청구항 32.

세포 증식성 질환, 자가면역/염증 질환, 심혈관 질환, 신경 질환, 정신 질환, 발달 장애, 유전 질환, 대사 질환, 감염, 다른 병인성 질환에서 선택되는 질병 치료용 약물의 제조에 이용되는 제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드, 제 8 항 내지 11 항 중 어느 한 항에 따른 핵산 분자, 제 12 항에 따른 벡터, 제 13 항에 따른 숙주 세포, 제 14 항 또는 15 항에 따른 리간드, 또는 제 16 내지 18 항중 어느 한 항에 따른 화합물.

청구항 33.

환자에서 질병을 치료하는 방법에 있어서, 제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드, 제 8 항 내지 11 항 중 어느 한 항에 따른 핵산 분자, 제 12 항에 따른 벡터, 제 13 항에 따른 숙주 세포, 제 14 항 또는 15 항에 따른 리간드, 제 16 내지 18 항중 어느 한 항에 따른 화합물, 또는 제 30 항에 따른 약학 조성물을 환자에 투여하는 것을 특징으로 하는 방법.

청구항 34.

제 33 항에 있어서, 건강한 사람에서 발현이나 활성 수준과 비교하여 천연 유전자의 발현 또는 폴리펩티드의 활성이 병든 환자에서 감소하는 질병의 경우에, 환자에 투여되는 폴리펩티드, 핵산 분자, 벡터, 리간드, 화합물 또는 조성물은 항진제인 것을 특징으로 하는 방법.

청구항 35.

제 33 항에 있어서, 건강한 사람에서 발현이나 활성 수준과 비교하여 천연 유전자의 발현 또는 폴리펩티드의 활성이 병든 환자에서 증가하는 질병의 경우에, 환자에 투여되는 폴리펩티드, 핵산 분자, 벡터, 리간드, 화합물 또는 조성물은 길항제인 것을 특징으로 하는 방법.

청구항 36.

환자에서 질병 치료를 모니터하는 방법에 있어서, 환자의 조직에서 일정 기간동안 제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드의 발현이나 활성 수준, 또는 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자의 발현 수준을 모니터하고, 시간이 경과함에 따라 활성이나 발현 수준이 대조 수치에 근접하는 것은 질병으로부터 회복을 암시하는 것을 특징으로 하는 방법.

청구항 37.

질병의 진단 또는 치료에 효과적인 화합물을 확인하는 방법에 있어서, 아래의 단계로 구성되는 것을 특징으로 하는 방법:

제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드, 또는 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자를 상기 핵산 분자 또는 폴리펩티드에 결합 친화성을 갖는 것으로 생각되는 한가지이상의 화합물과 접촉시키고;

이들 핵산 분자 또는 폴리펩티드에 특이적으로 결합하는 화합물을 선별한다.

청구항 38.
질병을 진단하는데 유용한 키트에 있어서,

엄밀도 조건하에서 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자와 혼성화되는 핵산 프로브를 포함하는 제 1 용기;

핵산 분자를 증폭하는데 유용한 프라이머를 포함하는 제 2 용기;

질병 진단을 조장하는 프로브와 프라이머의 사용 설명서로 구성되는 것을 특징으로 하는 키트.

청구항 39.
제 38 항에 있어서, 혼성화되지 않은 RNA를 절단하는 약품을 보유하는 제 3 용기가 추가로 포함되는 것을 특징으로 하는 키트.

청구항 40.
핵산 분자 어레이로 구성되고, 어레이중 적어도 하나는 제 8 항 내지 11 항중 어느 한 항에 따른 핵산 분자인 키트.

청구항 41.
제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드에 결합하는 한가지이상의 항체 및 항체와 폴리펩티드 사이의 결합 반응을 감지하는데 유용한 시약으로 구성되는 진단 키트.

청구항 42.
제 1 항 내지 7 항중 어느 한 항에 따른 폴리펩티드를 좀더 높거나 낮은 수준으로 발현하거나 이런 폴리펩티드를 발현하지 않도록 형질전환된 유전자도입 또 는 유전자적중(knockout)된 사람-아닌 동물.

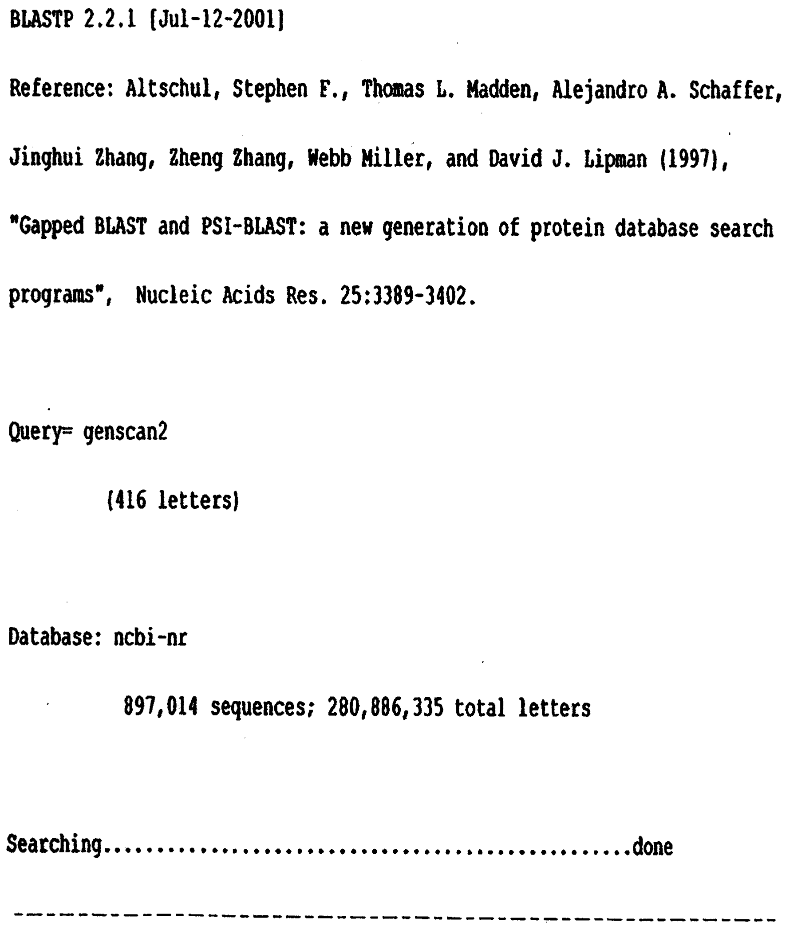
청구항 43.
질병을 치료하는데 효과적인 화합물을 스크리닝하는 방법에 있어서, 아래의 단계로 구성되는 것을 특징으로 하는 방법:

제 42 항에 따른 사람-아닌 유전자도입 동물과 후보 화합물을 접촉시키고;

동물에서 질병에 대한 화합물의 효과를 결정한다.

도면

도면1a



```
BLASTP 2.2.1 [Jul-12-2001]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer,

Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997),

"Gapped BLAST and PSI-BLAST: a new generation of protein database search

programs",  Nucleic Acids Res. 25:3389-3402.


Query= genscan2

         (416 letters)


Database: ncbi-nr

          897,014 sequences; 280,886,335 total letters


Searching................................................done
---------------------------------------------------------------
```

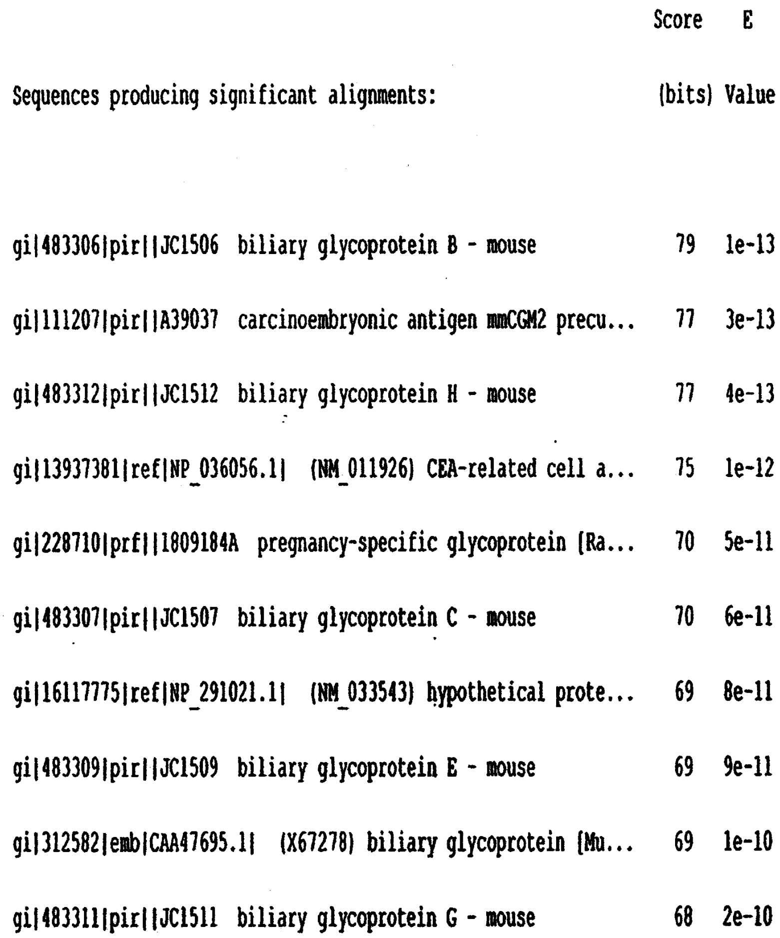
도면1b

```
---------------------------------------------------------------
                                                     Score    E
Sequences producing significant alignments:          (bits)  Value


pir||JC1512 biliary glycoprotein H - mouse              81  2e-14

pir||JC1506 biliary glycoprotein B - mouse              81  2e-14

pir||A39037 carcinoembryonic antigen mmCGM2 precursor - mouse >g...   79  9e-14

ref|NP_036056.1| (NM_011926) CEA-related cell adhesion molecule ...   79  9e-14

pir||JC1509 biliary glycoprotein E - mouse              73  5e-12

ref|NP_001758.1| (NM_001767) CD2 antigen (p50), sheep red blood ...   73  6e-12

ref|NP_113943.1| (NM_031755) carcinoembryonic antigen-related ce...   72  8e-12

pir||RWHUC2 T-cell surface glycoprotein CD2 precursor - human >g...   72  8e-12

gb|AAA51946.1| (M16336) CD2 surface antigen [Homo sapiens]   72  8e-12

ref|NP_291021.1| (NM_033543) hypothetical protein R29124_1 [Homo...   72  1e-11

pir||JC1507 biliary glycoprotein C - mouse              71  2e-11

emb|CAA47697.1| (X67280) biliary glycoprotein [Mus musculus]   71  2e-11

pir||S34338 biliary glycoprotein F - mouse >gi|312586|emb|CAA476...   71  2e-11

pir||JC1511 biliary glycoprotein G - mouse              71  2e-11
```

도면2

```
>pir||JC1512 biliary glycoprotein H - mouse

          Length = 341


 Score = 80.9 bits (198), Expect = 2e-14

 Identities = 54/168 (32%), Positives = 86/168 (51%), Gaps = 9/168 (5%)


Query: 73  RDKPVTVVQSIGTEVIGTLR----PDYRDRIRLFENGSLLLSDLQLADEGTYEVEISITD 128

            +  PV+    I  +V GT +    P +  R  ++ NGSLL+  + + D G Y +E+  TD

Sbjct: 69  KGNPVSTNAEIVHQVTGTNKTTTGPAHSGRETVYSNGSLLIQRVTVKDTGVYTIEM--TD 126


Query: 129 DTFTG-EKTINLTVDVPISRPQVLVASTTVLELSEAFTLNCSHENGTKPSYTWLKDGKPL 187

            + F   E T+   V  P+++P + V +TTV EL ++ TL C   N    +  WL + + L

Sbjct: 127 ENFRRTEATVQFHVHQPVTQPSLQVTNTTVKEL-DSVTLTCL-SNDIGANIQWLFNSQSL 184


Query: 188 LNDSRMLLSPDQKVLTITRVLMEDDDLYSCMVENPISQGRSLPVKITV 235

                RM LS +  +L I  +  ED   Y C + NP+S  RS  +K+ +

Sbjct: 185 QLTERMTLSQNNSILRIDPIKREDAGEYQCEISNPVSVKRSNSIKLDI 232
```

도면3



```
                                                              Score    E
Sequences producing significant alignments:                   (bits) Value

gi|483306|pir||JC1506  biliary glycoprotein B - mouse            79   1e-13
gi|111207|pir||A39037  carcinoembryonic antigen mmCGM2 precu...  77   3e-13
gi|483312|pir||JC1512  biliary glycoprotein H - mouse            77   4e-13
gi|13937381|ref|NP_036056.1|  (NM_011926) CEA-related cell a...  75   1e-12
gi|228710|prf||1809184A  pregnancy-specific glycoprotein (Ra...  70   5e-11
gi|483307|pir||JC1507  biliary glycoprotein C - mouse            70   6e-11
gi|16117775|ref|NP_291021.1|  (NM_033543) hypothetical prote...  69   8e-11
gi|483309|pir||JC1509  biliary glycoprotein E - mouse            69   9e-11
gi|312582|emb|CAA47695.1|  (X67278) biliary glycoprotein (Mu...  69   1e-10
gi|483311|pir||JC1511  biliary glycoprotein G - mouse            68   2e-10
```

도면4

```
>gi|4833061pir||JC1506  biliary glycoprotein B - mouse

          Length = 278


 Score = 78.6 bits (192), Expect = 1e-13

 Identities = 54/168 (32%), Positives = 86/168 (51%), Gaps = 9/168 (5%)


Query: 73  RDKPVTVVQSIGTEVIGTLR----PDYRDRIRLFENGSLLLSDLQLADEGTYEVEISITD 128

           +  PV+    I  +V GT +    P +  R  ++ NGSLL+  + + D G Y +E  +TD

Sbjct: 69  KGNPVSTNAEIVHQVTGTNKTTTGPAHSGRETVYSNGSLLIQRVTVKDTGVYTIE--MTD 126


Query: 129 DTF-TGEKTINLTVDVPISRPQVLVASTTVLELSEAFTLNCSHENGTKPSYTWLKDGKPL 187

           + F   E T+   V  P+++P + V +TTV EL ++ TL C   N    + WL + + L

Sbjct: 127 ENFRRTEATVQFHVHQPVTQPSLQVTNTTVKEL-DSVTLTCL-SNDIGANIQWLFNSQSL 184


Query: 188 LNDSRMLLSPDQKVLTITRVLMEDDDLYSCVVENPISQVRSLPVKITV 235

               RM LS +  +L I  +  ED   Y C + NP+S  RS  +K+ +

Sbjct: 185 QLTERMTLSQNNSILRIDPIKREDAGEYQCEISNPVSVKRSNSIKLDI 232
```

도면5a

```
  1  ATGAAGAGAG AAAGGGGAGC CCTGTCCAGA GCCTCCAGGG CCCTGCGCCT TGCTCCTTTT
       m  k  r   e  r  g   a  l  s   r   a  s  r   a  l  r   l  a  p  f

 61  GTCTACCTTC TTCTGATCCA GACAGACCCC CTGGAGGGGG TGAACATCAC CAGCCCCGTG
       v  y  l   l  l  i   q  t  d   p   l  e  g   v  n  i   t  s  p  v

121  CGCCTGATCC ATGGCACCGT GGGGAAGTCG GCTCTGCTTT CTGTGCAGTA CAGCAGTACC
       r  l  i   h  g  t   v  g  k   s   a  l  l   s  v  q   y  s  s  t

181  AGCAGCGACA GGCCTGTAGT GAAGTGGCAG CTGAAGCGGG ACAAGCCAGT GACCGTGGTG
       s  s  d   r  p  v   v  k  w   q   l  k  r   d  k  p   v  t  v  v

241  CAGTCCATTG GCACAGAGGT CATCGGCACC CTGCGGCCTG ACTATCGAGA CCGTATCCGA
       q  s  i   g  t  e   v  i  g   t   l  r  p   d  y  r   d  r  i  r

301  CTCTTTGAAA ATGGCTCCCT GCTTCTCAGC GACCTGCAGC TGGCCGATGA GGGCACCTAT
       l  f  e   n  g  s   l  l  l   s   d  l  q   l  a  d   e  g  t  y

361  GAGGTCGAGA TCTCCATCAC CGACGACACC TTCACTGGGG AGAAGACCAT CAACCTTACT
       e  v  e   i  s  i   t  d  d   t   f  t  g   e  k  t   i  n  l  t

421  GTAGATGTGC CCATTTCGAG GCCACAGGTG TTGGTGGCTT CAACCACTGT GCTGGAGCTC
       v  d  v   p  i  s   r  p  q   v   l  v  a   s  t  t   v  l  e  l

481  AGCGAGGCCT TCACCTTGAA CTGCTCACAT GAGAATGGCA CCAAGCCCAG CTACACCTGG
       s  e  a   f  t  l   n  c  s   h   e  n  g   t  k  p   s  y  t  w

541  CTGAAGGATG GCAAGCCCCT CCTCAATGAC TCGAGAATGC TCCTGTCCCC CGACCAAAAG
       l  k  d   g  k  p   l  l  n   d   s  r  m   l  l  s   p  d  q  k

601  GTGCTCACCA TCACCCGCGT GCTCATGGAG GATGACGACC TGTACAGCTG CATGGTGGAG
       v  l  t   i  t  r   v  l  m   e   d  d  d   l  y  s   c  m  v  e

661  AACCCCATCA GCCAGGGCCG CAGCCTGCCT GTCAAGATCA CCGTATACAG AAGAAGCTCC
       n  p  i   s  q  g   r  s  l   p   v  k  i   t  v  y   r  r  s  s

721  CTTTACATCA TCTTGTCTAC AGGAGGCATC TTCCTCCTTG TGACCTTGGT GACAGTCTGT
       l  y  i   i  l  s   t  g  g   i   f  l  l   v  t  l   v  t  v  c

781  GCCTGCTGGA AACCCTCCAA AAGGAAACAG AAGAAGCTAG AAAAGCAAAA CTCCCTGGAA
       a  c  w   k  p  s   k  r  k   q   k  k  l   e  k  q   n  s  l  e
---------------------------------------------------------------------
```

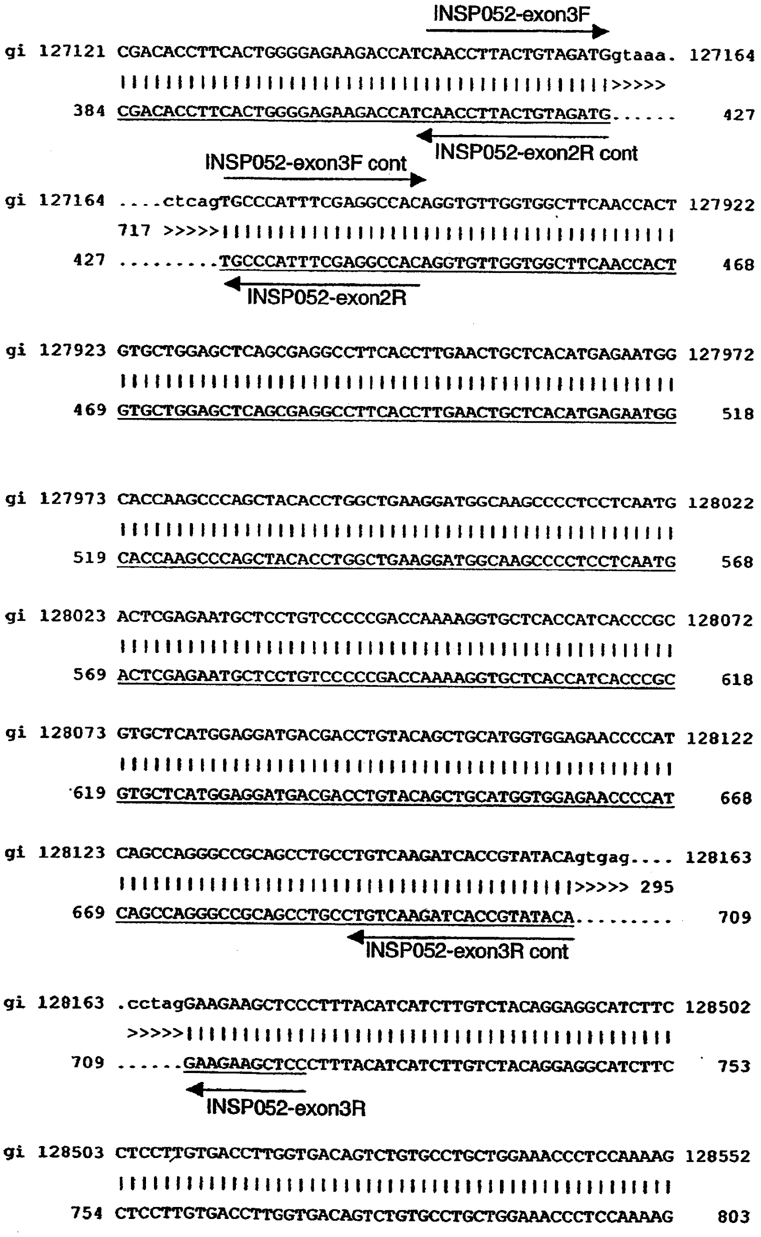
도면5b

```
---------------------------------------------------------------------

 841  TACATGGATC AGAATGATGA CCGCCTGAAA CCAGAAGCAG ACACCCTCCC TCGAAGTGGT
       y  m  d   q  n  d   d  r  l  k   p  e  a   d  t  l   p  r  s  g

 901  GAGCAGGAAC GGAAGAACCC CATGGCACTC TATATCCTGA AGGACAAGGA CTCCCCGGAG
       e  q  e   r  k  n   p  m  a  l   y  i  l   k  d  k   d  s  p  e

 961  ACCGAGGAGA ACCCGGCCCC GGAGCCTCGA AGCGCGACGG AGCCCGGCCC GCCCGGCTAC
       t  e  e   n  p  a   p  e  p  r   s  a  t   e  p  g   p  p  g  y

1021  TCCGTGTCTC CCGCCGTGCC CGGCCGCTCG CCGGGGCTGC CCATCCGCTC TGCCCGCCGC
       s  v  s   p  a  v   p  g  r  s   p  g  l   p  i  r   s  a  r  r

1081  TACCCGCGCT CCCCAGCGCG CTCCCCAGCC ACCGGCCGGA CACACTCGTC GCCGCCCAGG
       y  p  r   s  p  a   r  s  p  a   t  g  r   t  h  s   s  p  p  r

1141  GCCCCGAGCT CGCCCGGCCG CTCGCGCAGC GCCTCGCGCA CACTGCGGAC TGCGGGCGTG
       a  p  s   s  p  g   r  s  r  s   a  s  r   t  l  r   t  a  g  v

1201  CACATAATCC GCGAGCAAGA CGAGGCCGGC CCGGTGGAGA TCAGCGCCTG AGCCGCCTCG
       h  i  i   r  e  q   d  e  a  g   p  v  e   i  s  a

1261  GGATCCCCTG AGAGGCGCCC GCGGTCTGCG GCCAGTGGCC CGGGGGAAAG CTGGGGCTGG
1321  GAAGCCCGGG CGCGGCGCGC TGGGGACGAG GGGAGGTCCC GGGGGGGCGC TGGTGTCTCG
1381  GGTGTGAACG TGTATGAGCA TGCGCAGACG GAGGCGGGTG CGCGGAGGCG GCAGTGTTGA
1441  TATGGTGAAA CCGGGTCGCA TTTGCTTCCG GTTTACTGGC TGTGTCCTCA CTTGGTATAG
1501  GTTGTGCCCT CTTAGGACCA CATAGATTAT TACATTTCTG GCCCAATACC CAAAAGGGTT
1561  TTATGGAAAC TAACATCAGT AACCTAACCC CCGTGACTAT CCTGTGCTCT TCCTAGGGAG
1621  CTGTGTTGTT TCCCACCCAC CACCCTTCCC TCTGAACAAA TGCCTGAGTG CTGGGGCACT
1681  TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT GCAAGTTCAG ATTAGAGAGG CCACTTTCCC
1741  AGAATCCACA GCTGCACTAA GCTAAGGAGA AGCCAGATGC CGGTTACTGG GTGTGCAGGG
1801  GCTGTTCTGA GCTGGGGGGA TCATTGTGAA GGCCTTCTTC CCTGGGCACC TGGTACCTGG
1861  GGACCTACAA GGTGGTGAGG GAAGGGTACG AGTACATTCC TTTTCCCTCT GACCTGGGCG
1921  CTAGCAAGGG CAAAGAACCC GAGCCTGCCA GCTTGGCCTC CTCCCACAGC CTCCCTCGGA
1981  GGCATGCCAT GCCAAGCACT CTTTCTGTCT CTGTTCATGA ATAAA
```

도면6a

```
gi 115886 ATGAAGAGAGAAAGGGGAGCCCTGTCCAGAGCCTCCAGGGCCCTGCGCCT 115935
          ||||||||||||||||||||||||||||||||||||||||||||||||||
        1 ATGAAGAGAGAAAGGGGAGCCCTGTCCAGAGCCTCCAGGGCCCTGCGCCT      50
          ─────────────────────────►
               INSP052-BIP-exon2F
                          INSP052-exon2F
                          ───────────────►
gi 115936 TGCTCCTTTTGTCTACCTTCTTCTGATCCAGACAGgtagg.......cac 115970
          |||||||||||||||||||||||||||||||||||>>>>> 10852 >>>
       51 TGCTCCTTTTGTCTACCTTCTTCTGATCCAGACAG...............      85
                          ◄────────────
                           INSP052-exon1R cont
          INSP052-exon2F cont
            ─────────────────►
gi 115970 agACCCCCTGGAGGGGGTGAACATCACCAGCCCCGTGCGCCTGATCCATG 126870
          >>||||||||||||||||||||||||||||||||||||||||||||||||
       85 ..ACCCCCTGGAGGGGGTGAACATCACCAGCCCCGTGCGCCTGATCCATG     133
           ◄─────────────
             INSP052-exon1R

gi 126871 GCACCGTGGGGAAGTCGGCTCTGCTTTCTGTGCAGTACAGCAGTACCAGC 126920
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      134 GCACCGTGGGGAAGTCGGCTCTGCTTTCTGTGCAGTACAGCAGTACCAGC     183

gi 126921 AGCGACAGGCCTGTAGTGAAGTGGCAGCTGAAGCGGGACAAGCCAGTGAC 126970
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      184 AGCGACAGGCCTGTAGTGAAGTGGCAGCTGAAGCGGGACAAGCCAGTGAC     233

gi 126971 CGTGGTGCAGTCCATTGGCACAGAGGTCATCGGCACCCTGCGGCCTGACT 127020
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      234 CGTGGTGCAGTCCATTGGCACAGAGGTCATCGGCACCCTGCGGCCTGACT     283

gi 127021 ATCGAGACCGTATCCGACTCTTTGAAAATGGCTCCCTGCTTCTCAGCGAC 127070
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      284 ATCGAGACCGTATCCGACTCTTTGAAAATGGCTCCCTGCTTCTCAGCGAC     333

gi 127071 CTGCAGCTGGCCGATGAGGGCACCTATGAGGTCGAGATCTCCATCACCGA 127120
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      334 CTGCAGCTGGCCGATGAGGGCACCTATGAGGTCGAGATCTCCATCACCGA     383
```

도면6b



```
                                                   INSP052-exon3F
                                                   ------------->
gi 127121 CGACACCTTCACTGGGGAGAAGACCATCAACCTTACTGTAGATGgtaaa. 127164
          ||||||||||||||||||||||||||||||||||||||||||||>>>>>
      384 CGACACCTTCACTGGGGAGAAGACCATCAACCTTACTGTAGATG......    427
                                                  <------------
                 INSP052-exon3F cont              INSP052-exon2R cont
                 ------------->
gi 127164 ....ctcagTGCCCATTTCGAGGCCACAGGTGTTGGTGGCTTCAACCACT 127922
          717 >>>>>||||||||||||||||||||||||||||||||||||||||||
      427 .........TGCCCATTTCGAGGCCACAGGTGTTGGTGGCTTCAACCACT    468
                   <------------
                    INSP052-exon2R

gi 127923 GTGCTGGAGCTCAGCGAGGCCTTCACCTTGAACTGCTCACATGAGAATGG 127972
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      469 GTGCTGGAGCTCAGCGAGGCCTTCACCTTGAACTGCTCACATGAGAATGG    518


gi 127973 CACCAAGCCCAGCTACACCTGGCTGAAGGATGGCAAGCCCCTCCTCAATG 128022
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      519 CACCAAGCCCAGCTACACCTGGCTGAAGGATGGCAAGCCCCTCCTCAATG    568

gi 128023 ACTCGAGAATGCTCCTGTCCCCCGACCAAAAGGTGCTCACCATCACCCGC 128072
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      569 ACTCGAGAATGCTCCTGTCCCCCGACCAAAAGGTGCTCACCATCACCCGC    618

gi 128073 GTGCTCATGGAGGATGACGACCTGTACAGCTGCATGGTGGAGAACCCCAT 128122
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      619 GTGCTCATGGAGGATGACGACCTGTACAGCTGCATGGTGGAGAACCCCAT    668

gi 128123 CAGCCAGGGCCGCAGCCTGCCTGTCAAGATCACCGTATACAgtgag.... 128163
          |||||||||||||||||||||||||||||||||||||||||>>>>> 295
      669 CAGCCAGGGCCGCAGCCTGCCTGTCAAGATCACCGTATACA.........    709
                                      <------------
                                       INSP052-exon3R cont

gi 128163 .cctagGAAGAAGCTCCCTTTACATCATCTTGTCTACAGGAGGCATCTTC 128502
           >>>>>|||||||||||||||||||||||||||||||||||||||||||||
      709 ......GAAGAAGCTCCCTTTACATCATCTTGTCTACAGGAGGCATCTTC    753
               <----------
                INSP052-exon3R

gi 128503 CTCCTTGTGACCTTGGTGACAGTCTGTGCCTGCTGGAAACCCTCCAAAAG 128552
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      754 CTCCTTGTGACCTTGGTGACAGTCTGTGCCTGCTGGAAACCCTCCAAAAG    803
```

도면6c

```
gi 128552 gtctg.....cacagGAAACAGAAGAAGCTAGAAAAGCAAAACTCCCTGG 129108
          >>>>> 521 >>>>>||||||||||||||||||||||||||||||||||||
      803 ...............GAAACAGAAGAAGCTAGAAAAGCAAAACTCCCTGG    838

gi 129109 AATACATGGATCAGAATGATGACCGCCTGAAACCAGAAGgtgag.....t 129147
          |||||||||||||||||||||||||||||||||||||||>>>>> 286 >
      839 AATACATGGATCAGAATGATGACCGCCTGAAACCAGAAG...........    877

gi 129147 gcagCAGACACCCTCCCTCGAAGTGGTGAGCAGGAACGGAAGAACCCCAT 129479
          >>>>||||||||||||||||||||||||||||||||||||||||||||||
      877 ....CAGACACCCTCCCTCGAAGTGGTGAGCAGGAACGGAAGAACCCCAT    923

gi 129480 GGCACTCTATATCCTGAAGGACAAGgtgag.....tgcagGACTCCCCGG 130461
          |||||||||||||||||||||||||>>>>> 947 >>>>>||||||||||
      924 GGCACTCTATATCCTGAAGGACAAG...............GACTCCCCGG    958

gi 130462 AGACCGAGGAGAACCCGGCCCCGGAGCCTCGAAGCGCGACGGAGCCCGGC 130511
          ||||||||||||||||||||||||||||||||||||||||||||||||||
      959 AGACCGAGGAGAACCCGGCCCCGGAGCCTCGAAGCGCGACGGAGCCCGGC   1008

gi 130512 CCGCCCGGCTACTCCGTGTCTCCCGCCGTGCCCGGCCGCTCGCCGGGGCT 130561
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1009 CCGCCCGGCTACTCCGTGTCTCCCGCCGTGCCCGGCCGCTCGCCGGGGCT   1058

gi 130562 GCCCATCCGCTCTGCCCGCCGCTACCCGCGCTCCCCAGCGCGCTCCCCAG 130611
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1059 GCCCATCCGCTCTGCCCGCCGCTACCCGCGCTCCCCAGCGCGCTCCCCAG   1108

gi 130612 CCACCGGCCGGACACACTCGTCGCCGCCCAGGGCCCCGAGCTCGCCCGGC 130661
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1109 CCACCGGCCGGACACACTCGTCGCCGCCCAGGGCCCCGAGCTCGCCCGGC   1158

gi 130662 CGCTCGCGCAGCGCCTCGCGCACACTGCGGACTGCGGGCGTGCACATAAT 130711
          ||||||||||||||||||||||||||||||||||||||||||||||||||
     1159 CGCTCGCGCAGCGCCTCGCGCACACTGCGGACTGCGGGCGTGCACATAAT   1208

gi 130712 CCGCGAGCAAGACGAGGCCGGCCCGGTGGAGATCAGCGCCTGA 130754
          |||||||||||||||||||||||||||||||||||||||||||
     1209 CCGCGAGCAAGACGAGGCCGGCCCGGTGGAGATCAGCGCCTGA   1251
```

도면7

```
  1  ACAAGTTTGT ACAAAAAAGC AGGCTTCGCC ACCATGAAGA GAGAAAGGGG AGCCCTGTCC
                                         m  k   r  e  r   g  a  l  s

 61  AGAGCCTCCA GGGCCCTGCG CCTTGCTCCT TTTGTCTACC TTCTTCTGAT CCAGACAGAC
      r  a  s   r  a  l   r  l  a  p   f  v  y   l  l  l   i  q  t  d

121  CCCCTGGAGG GGGTGAACAT CACCAGCCCC GTGCGCCTGA TCCATGGCAC CGTGGGGAAG
      p  l  e   g  v  n   i  t  s  p   v  r  l   i  h  g   t  v  g  k

181  TCGGCTCTGC TTTCTGTGCA GTACAGCAGT ACCAGCAGCG ACAGGCCTGT AGTGAAGTGG
      s  a  l   l  s  v   q  y  s  s   t  s  s   d  r  p   v  v  k  w

241  CAGCTGAAGC GGGACAAGCC AGTGACCGTG GTGCAGTCCA TTGGCACAGA GGTCATCGGC
      q  l  k   r  d  k   p  v  t  v   v  q  s   i  g  t   e  v  i  g

301  ACCCTGCGGC CTGACTATCG AGACCGTATC CGACTCTTTG AAAATGGCTC CCTGCTTCTC
      t  l  r   p  d  y   r  d  r  i   r  l  f   e  n  g   s  l  l  l

361  AGCGACCTGC AGCTGGCCGA TGAGGGCACC TATGAGGTCG AGATCTCCAT CACCGACGAC
      s  d  l   q  l  a   d  e  g  t   y  e  v   e  i  s   i  t  d  d

421  ACCTTCACTG GGGAGAAGAC CATCAACCTT ACTGTAGATG TGCCCATTTC GAGGCCACAG
      t  f  t   g  e  k   t  i  n  l   t  v  d   v  p  i   s  r  p  q

481  GTGTTGGTGG CTTCAACCAC TGTGCTGGAG CTCAGCGAGG CCTTCACCTT GAACTGCTCA
      v  l  v   a  s  t   t  v  l  e   l  s  e   a  f  t   l  n  c  s

541  CATGAGAATG GCACCAAGCC CAGCTACACC TGGCTGAAGG ATGGCAAGCC CCTCCTCAAT
      h  e  n   g  t  k   p  s  y  t   w  l  k   d  g  k   p  l  l  n

601  GACTCGAGAA TGCTCCTGTC CCCCGACCAA AAGGTGCTCA CCATCACCCG CGTGCTCATG
      d  s  r   m  l  l   s  p  d  q   k  v  l   t  i  t   r  v  l  m

661  GAGGATGACG ACCTGTACAG CTGCATGGTG GAGAACCCCA TCAGCCAGGG CCGCAGCCTG
      e  d  d   d  l  y   s  c  m  v   e  n  p   i  s  q   g  r  s  l

721  CCTGTCAAGA TCACCGTATA CAGAAGAAGC TCCCACCATC ACCATCACCA TTGAAACCCA
      p  v  k   i  t  v   y  r  r  s   s  h  h   h  h  h   h

781  GCTTTCTTGT ACAAAGTGGT
```

도면8a

Molecule: pENTR-INSP052- EC-6HIS, 3005 bps DNA Circular

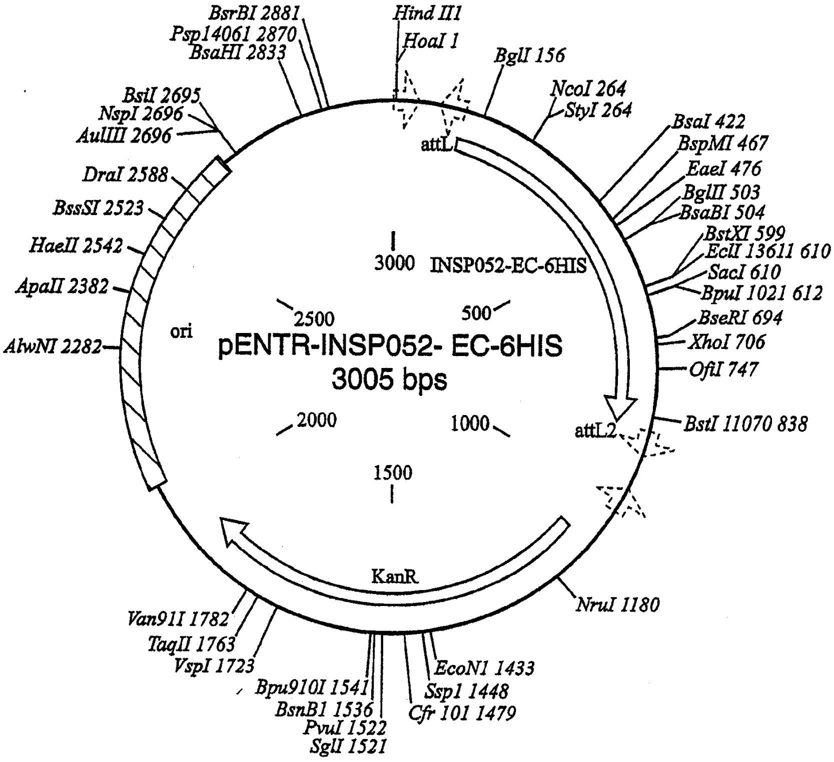
File Name: pENTR-INSP052-6HIS.cm5, dated 21 Feb 2003

Description: Ligation of Cons-6His.SEQ into pENTR-attL1-attL2

Molecule Features:

Type	Start	End	Name	Description
MARKER	21			pENTR-F1 primer
MARKER	110		C attL1	
GENE	136	873	INSP052-EC-6HIS	
MARKER	888		attL2	
MARKER	1001		C	pENTR-R1 primer
GENE	1100	1909	KanR	
REGION	2030	2669	ori	

도면8b



도면9a

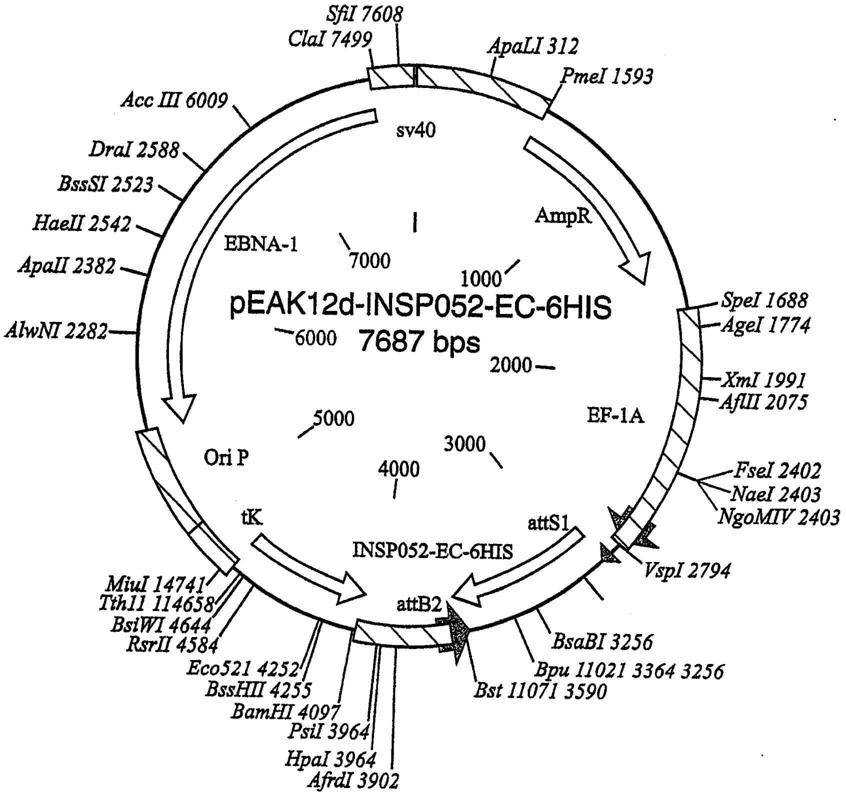
Molecule: pEAK12d-INSP052-EC-6HIS, 7687 bps DNA Circular
File Name: pEAK12d-INSP052-6HIS.cm5, dated 21 Feb 2003

Description: Ligation of Cons-6His.SEQ into pEAK12d-attB1-attB2

Molecule Features:

Type	Start	End	Name	Description
REGION	2	595		pmb-ori
GENE	596	1519	AmpR	
REGION	1690	2795	EF-1a	
MARKER	2703			pEAK12F primer
REGION	2855	2887	attB1	
GENE	2888	3625	INSP052-EC-6HIS (aa1-240)	
REGION	3629	3654	attB2	
MARKER	3656	C		pEAK12R primer
REGION	3661	4089		poly A/splice
GENE	4708	4090 C		PUROMYCIN resistance
REGION	4932	4709 C	tK	tK promoter
REGION	5427	4933 C	Ori P	
GENE	7479	5427 C	EBNA-1	
REGION	7480	7679	sv40	

도면9b



SEQUENCE LISTING

<110> Ares Trading S.A

<120> Immunoglobulin-Domain Containing Cell Surface Recognition Molecules

<130> P030459WO

<140> PCT/GB03/01851
<141> 2003-04-30

<150> GB0209884.6
<151> 2002-04-30

<160> 35

<170> SeqWin99, version 1.02

<210> 1

<211> 85

<212> DNA

<213> Homo sapiens

<400> 1

```
atgaagagag aaaggggagc cctgtccaga gcctccaggg ccctgcgcct tgctcctttt   60
gtctaccttc ttctgatcca gacag                                         85
```

<210> 2

<211> 29

<212> PRT

<213> Homo sapiens

<400> 2

```
Met Lys Arg Glu Arg Gly Ala Leu Ser Arg Ala Ser Arg Ala Leu Arg
1               5                   10                  15

Leu Ala Pro Phe Val Tyr Leu Leu Leu Ile Gln Thr Asp
            20                  25
```

<210> 3

<211> 342

<212> DNA

<213> Homo sapiens

<400> 3

```
accccctgga gggggtgaac atcaccagcc ccgtgcgcct gatccatggc accgtgggga   60
agtcggctct gctttctgtg cagtacagca gtaccagcag cgacaggcct gtagtgaagt  120
ggcagctgaa gcgggacaag ccagtgaccg tggtgcagtc cattggcaca gaggtcatcg  180
gcaccctgcg gcctgactat cgagaccgta tccgactctt tgaaaatggc tccctgcttc  240
tcagcgacct gcagctggcc gatgagggca cctatgaggt cgagatctcc atcaccgacg  300
acaccttcac tggggagaag accatcaacc ttactgtaga tg                      342
```

<210> 4

<211> 114

<212> PRT

<213> Homo sapiens

<400> 4

```
Pro Leu Glu Gly Val Asn Ile Thr Ser Pro Val Arg Leu Ile His Gly
1               5                   10                  15

Thr Val Gly Lys Ser Ala Leu Leu Ser Val Gln Tyr Ser Ser Thr Ser
            20                  25                  30

Ser Asp Arg Pro Val Val Lys Trp Gln Leu Lys Arg Asp Lys Pro Val
        35                  40                  45

Thr Val Val Gln Ser Ile Gly Thr Glu Val Ile Gly Thr Leu Arg Pro
    50                  55                  60

Asp Tyr Arg Asp Arg Ile Arg Leu Phe Glu Asn Gly Ser Leu Leu Leu
65                  70                  75                  80

Ser Asp Leu Gln Leu Ala Asp Glu Gly Thr Tyr Glu Val Glu Ile Ser
                85                  90                  95

Ile Thr Asp Asp Thr Phe Thr Gly Glu Lys Thr Ile Asn Leu Thr Val
            100                 105                 110

Asp Val
```

<210> 5

<211> 282

<212> DNA

<213> Homo sapiens

<400> 5

tgcccatttc gaggccacag gtgttggtgg cttcaaccac tgtgctggag ctcagcgagg 60

ccttcacctt gaactgctca catgagaatg gcaccaagcc cagctacacc tggctgaagg 120

atggcaagcc cctcctcaat gactcgagaa tgctcctgtc ccccgaccaa aaggtgctca 180

ccatcacccg cgtgctcatg gaggatgacg acctgtacag ctgcatggtg gagaacccca 240

tcagccaggg ccgcagcctg cctgtcaaga tcaccgtata ca 282

<210> 6

<211> 94

<212> PRT

<213> Homo sapiens

<400> 6

```
Pro Ile Ser Arg Pro Gln Val Leu Val Ala Ser Thr Thr Val Leu Glu
1               5                   10                  15

Leu Ser Glu Ala Phe Thr Leu Asn Cys Ser His Glu Asn Gly Thr Lys
            20                  25                  30

Pro Ser Tyr Thr Trp Leu Lys Asp Gly Lys Pro Leu Leu Asn Asp Ser
        35                  40                  45

Arg Met Leu Leu Ser Pro Asp Gln Lys Val Leu Thr Ile Thr Arg Val
    50                  55                  60

Leu Met Glu Asp Asp Asp Leu Tyr Ser Cys Met Val Glu Asn Pro Ile
65                  70                  75                  80

Ser Gln Gly Arg Ser Leu Pro Val Lys Ile Thr Val Tyr Arg
                85                  90
```

<210> 7

<211> 94

<212> DNA

<213> Homo sapiens

<400> 7

gaagaagctc cctttacatc atcttgtcta caggaggcat cttcctcctt gtgaccttgg 60

tgacagtctg tgcctgctgg aaaccctcca aaag 94

<210> 8

<211> 31

<212> PRT

<213> Homo sapiens

<400> 8

```
Arg Ser Ser Leu Tyr Ile Ile Leu Ser Thr Gly Gly Ile Phe Leu Leu
1               5                   10                  15

Val Thr Leu Val Thr Val Cys Ala Cys Trp Lys Pro Ser Lys Arg
            20                  25                  30
```

<210> 9

<211> 74

<212> DNA

<213> Homo sapiens

<400> 9

gaaacagaag aagctagaaa agcaaaactc cctggaatac atggatcaga atgatgaccg 60

cctgaaacca gaag 74

<210> 10

<211> 25

<212> PRT

<213> Homo sapiens

<400> 10

Lys Gln Lys Lys Leu Glu Lys Gln Asn Ser Leu Glu Tyr Met Asp Gln

1   5   10   15

Asn Asp Asp Arg Leu Lys Pro Glu Ala

            20   25

<210>   11

<211>   71

<212>   DNA

<213>   Homo sapiens

<400>   11

cagacaccct ccctcgaagt ggtgagcagg aacggaagaa ccccatggca ctctatatcc   60

tgaaggacaa g   71

<210>   12

<211>   23

<212>   PRT

<213>   Homo sapiens

<400>   12

Asp Thr Leu Pro Arg Ser Gly Glu Gln Glu Arg Lys Asn Pro Met Ala

1   5   10   15

Leu Tyr Ile Leu Lys Asp Lys

            20

<210>   13

<211>   303

<212>   DNA

<213>   Homo sapiens

<400>   13

gactccccgg agaccgagga gaacccggcc ccggagcctc gaagcgcgac ggagcccggc   60 ccgcccggct actccgtgtc tcccgccgtg cccggccgct cgccggggct gcccatccgc 120

tctgcccgcc gctacccgcg ctccccagcg cgctccccag ccaccggccg gacacactcg 180

tcgccgccca gggccccgag ctcgcccggc cgctcgcgca gcgcctcgcg cacactgcgg 240

actgcgggcg tgcacataat ccgcgagcaa gacgaggccg gcccggtgga gatcagcgcc 300

tga 303

<210> 14

<211> 100

<212> PRT

<213> Homo sapiens

<400> 14

Asp Ser Pro Glu Thr Glu Glu Asn Pro Ala Pro Glu Pro Arg Ser Ala

1 5 10 15

Thr Glu Pro Gly Pro Pro Gly Tyr Ser Val Ser Pro Ala Val Pro Gly

20 25 30

Arg Ser Pro Gly Leu Pro Ile Arg Ser Ala Arg Arg Tyr Pro Arg Ser

35 40 45

Pro Ala Arg Ser Pro Ala Thr Gly Arg Thr His Ser Ser Pro Pro Arg

50 55 60

Ala Pro Ser Ser Pro Gly Arg Ser Arg Ser Ala Ser Arg Thr Leu Arg

65 70 75 80

Thr Ala Gly Val His Ile Ile Arg Glu Gln Asp Glu Ala Gly Pro Val

85 90 95

Glu Ile Ser Ala

100

<210> 15

<211> 1251

<212> DNA

<213> Homo sapiens

<400> 15

atgaagagag aaaggggagc cctgtccaga gcctccaggg ccctgcgcct tgctcctttt 60

gtctaccttc ttctgatcca gacagacccc ctggaggggg tgaacatcac cagccccgtg 120

cgcctgatcc atggcaccgt ggggaagtcg gctctgcttt ctgtgcagta cagcagtacc 180

agcagcgaca ggcctgtagt gaagtggcag ctgaagcggg acaagccagt gaccgtggtg 240

cagtccattg gcacagaggt catcggcacc ctgcggcctg actatcgaga ccgtatccga 300

ctctttgaaa atggctccct gcttctcagc gacctgcagc tggccgatga gggcacctat 360

gaggtcgaga tctccatcac cgacgacacc ttcactgggg agaagaccat caaccttact 420

gtagatgtgc ccatttcgag gccacaggtg ttggtggctt caaccactgt gctggagctc 480

agcgaggcct tcaccttgaa ctgctcacat gagaatggca ccaagcccag ctacacctgg 540

ctgaaggatg gcaagcccct cctcaatgac tcgagaatgc tcctgtcccc cgaccaaaag 600

gtgctcacca tcacccgcgt gctcatggag gatgacgacc tgtacagctg catggtggag 660

aaccccatca gccagggccg cagcctgcct gtcaagatca ccgtatacag aagaagctcc 720

ctttacatca tcttgtctac aggaggcatc ttcctccttg tgaccttggt gacagtctgt 780

gcctgctgga aaccctccaa aaggaaacag aagaagctag aaaagcaaaa ctccctggaa 840

tacatggatc agaatgatga ccgcctgaaa ccagaagcag acaccctccc tcgaagtggt 900

gagcaggaac ggaagaaccc catggcactc tatatcctga aggacaagga ctccccggag 960

accgaggaga acccggcccc ggagcctcga agcgcgacgg agcccggccc gcccggctac 1020

tccgtgtctc ccgccgtgcc cggccgctcg ccggggctgc ccatccgctc tgcccgccgc 1080

tacccgcgct ccccagcgcg ctccccagcc accggccgga cacactcgtc gccgcccagg 1140

gccccgagct cgcccggccg ctcgcgcagc gcctcgcgca cactgcggac tgcgggcgtg 1200

cacataatcc gcgagcaaga cgaggccggc ccggtggaga tcagcgcctg a 1251

<210> 16

<211> 416

<212> PRT

<213> Homo sapiens

<400> 16

Met Lys Arg Glu Arg Gly Ala Leu Ser Arg Ala Ser Arg Ala Leu Arg

1 5 10 15

Leu Ala Pro Phe Val Tyr Leu Leu Leu Ile Gln Thr Asp Pro Leu Glu

20 25 30

Gly Val Asn Ile Thr Ser Pro Val Arg Leu Ile His Gly Thr Val Gly

35 40 45

Lys Ser Ala Leu Leu Ser Val Gln Tyr Ser Ser Thr Ser Ser Asp Arg

50 55 60

Pro Val Val Lys Trp Gln Leu Lys Arg Asp Lys Pro Val Thr Val Val

65 70 75 80

Gln Ser Ile Gly Thr Glu Val Ile Gly Thr Leu Arg Pro Asp Tyr Arg

85 90 95

Asp Arg Ile Arg Leu Phe Glu Asn Gly Ser Leu Leu Leu Ser Asp Leu

100 105 110

Gln Leu Ala Asp Glu Gly Thr Tyr Glu Val Glu Ile Ser Ile Thr Asp

115 120 125

Asp Thr Phe Thr Gly Glu Lys Thr Ile Asn Leu Thr Val Asp Val Pro

130 135 140

Ile Ser Arg Pro Gln Val Leu Val Ala Ser Thr Thr Val Leu Glu Leu

145 150 155 160

Ser Glu Ala Phe Thr Leu Asn Cys Ser His Glu Asn Gly Thr Lys Pro

165 170 175

Ser Tyr Thr Trp Leu Lys Asp Gly Lys Pro Leu Leu Asn Asp Ser Arg

180 185 190

Met Leu Leu Ser Pro Asp Gln Lys Val Leu Thr Ile Thr Arg Val Leu

195 200 205

Met Glu Asp Asp Asp Leu Tyr Ser Cys Met Val Glu Asn Pro Ile Ser

210 215 220

Gln Gly Arg Ser Leu Pro Val Lys Ile Thr Val Tyr Arg Arg Ser Ser

225 230 235 240

Leu Tyr Ile Ile Leu Ser Thr Gly Gly Ile Phe Leu Leu Val Thr Leu

245 250 255

Val Thr Val Cys Ala Cys Trp Lys Pro Ser Lys Arg Lys Gln Lys Lys

260 265 270

Leu Glu Lys Gln Asn Ser Leu Glu Tyr Met Asp Gln Asn Asp Asp Arg

275 280 285

Leu Lys Pro Glu Ala Asp Thr Leu Pro Arg Ser Gly Glu Gln Glu Arg

290 295 300

Lys Asn Pro Met Ala Leu Tyr Ile Leu Lys Asp Lys Asp Ser Pro Glu

305 310 315 320

Thr Glu Glu Asn Pro Ala Pro Glu Pro Arg Ser Ala Thr Glu Pro Gly

325 330 335

Pro Pro Gly Tyr Ser Val Ser Pro Ala Val Pro Gly Arg Ser Pro Gly

340 345 350

Leu Pro Ile Arg Ser Ala Arg Arg Tyr Pro Arg Ser Pro Ala Arg Ser

355 360 365

Pro Ala Thr Gly Arg Thr His Ser Ser Pro Pro Arg Ala Pro Ser Ser

370 375 380

Pro Gly Arg Ser Arg Ser Ala Ser Arg Thr Leu Arg Thr Ala Gly Val

385 390 395 400

His Ile Ile Arg Glu Gln Asp Glu Ala Gly Pro Val Glu Ile Ser Ala

405 410 415

<210> 17

<211> 1257

<212> DNA

<213> Homo sapiens

<400> 17

atgaagagag aaaggggagc cctgtcaaga gcctccaggg ctctgcgcct ctctcctttt 60

gtctacctgc ttctcatcca gccagtcccc ctggaggggg tgaacatcac cagcccagta 120

cgtctgatcc acggcacagt ggggaagtcg gccctgcttt ccgtgcagta cagtagcacc 180

agcagcgaca agcccgtggt gaagtggcag ctgaagcgtg acaagccagt gaccgtggtg 240

cagtctatag gcacagaggt cattggcact ctgcggcctg actatcgaga ccgtatccgg 300

ctctttgaaa atggctcctt gcttctcagc gacctgcagc tggcggatga gggaacctat 360

gaagtggaga tttccatcac tgacgacacc ttcaccgggg agaagaccat caacctcacc 420

gtggatgtgc ccatttcaag gccgcaggta ttagtggctt caaccactgt gctggagctc 480

agtgaggcct tcaccctcaa ctgctcccat gagaatggca ccaagcctag ctacacgtgg 540

ctgaaggatg gcaaacccct cctcaatgac tcccgaatgc tcctgtcccc tgaccaaaag 600

gtgctcacca tcacccgagt actcatggaa gatgacgacc tgtacagctg tgtggtggag 660

aaccccatca gccaggtccg cagcctgcct gtcaagatca ctgtgtatag aagaagctcc 720

ctctatatca tcttgtctac aggaggcatc ttcctccttg tgaccctggt gacagtttgt 780

gcctgctgga aaccctcaaa aaagtctagg aagaagagga agttggagaa gcaaaactcc 840

ttggaataca tggatcagaa tgatgaccgc ctaaaatcag aagcagatac cctaccccga 900

agtggagaac aggagcggaa gaacccaatg gcactctata tcctgaagga taaggattcc 960

tcagagccag atgaaaaccc tgctacagag ccacggagca ccacagaacc cggtcccccт 1020

ggctactccg tgtcgccgcc cgtgcccggc cgctctccgg ggcttcccat ccgctcagcc 1080

cgccgctacc cgcgctcccc agcacgttcc cctgccactg gccggacgca cacgtcgcca 1140
ccgcgggccc cgagctcgcc aggccgctcg cgcagctctt cgcgctcact gcggactgca 1200
ggcgtgcaga gaatccggga gcaggacgag tcagggcagg tggagatcag tgcctga 1257

<210> 18
<211> 418
<212> PRT
<213> Homo sapiens

<400> 18
Met Lys Arg Glu Arg Gly Ala Leu Ser Arg Ala Ser Arg Ala Leu Arg
1 5 10 15

Leu Ser Pro Phe Val Tyr Leu Leu Leu Ile Gln Pro Val Pro Leu Glu
20 25 30

Gly Val Asn Ile Thr Ser Pro Val Arg Leu Ile His Gly Thr Val Gly
35 40 45

Lys Ser Ala Leu Leu Ser Val Gln Tyr Ser Ser Thr Ser Ser Asp Lys
50 55 60

Pro Val Val Lys Trp Gln Leu Lys Arg Asp Lys Pro Val Thr Val Val
65 70 75 80

Gln Ser Ile Gly Thr Glu Val Ile Gly Thr Leu Arg Pro Asp Tyr Arg
85 90 95

Asp Arg Ile Arg Leu Phe Glu Asn Gly Ser Leu Leu Leu Ser Asp Leu
100 105 110

Gln Leu Ala Asp Glu Gly Thr Tyr Glu Val Glu Ile Ser Ile Thr Asp
115 120 125

```
Asp Thr Phe Thr Gly Glu Lys Thr Ile Asn Leu Thr Val Asp Val Pro
    130                 135                 140

Ile Ser Arg Pro Gln Val Leu Val Ala Ser Thr Thr Val Leu Glu Leu
145                 150                 155                 160

Ser Glu Ala Phe Thr Leu Asn Cys Ser His Glu Asn Gly Thr Lys Pro
                165                 170                 175

Ser Tyr Thr Trp Leu Lys Asp Gly Lys Pro Leu Leu Asn Asp Ser Arg
            180                 185                 190

Met Leu Leu Ser Pro Asp Gln Lys Val Leu Thr Ile Thr Arg Val Leu
        195                 200                 205

Met Glu Asp Asp Asp Leu Tyr Ser Cys Val Val Glu Asn Pro Ile Ser
    210                 215                 220

Gln Val Arg Ser Leu Pro Val Lys Ile Thr Val Tyr Arg Arg Ser Ser
225                 230                 235                 240

Leu Tyr Ile Ile Leu Ser Thr Gly Gly Ile Phe Leu Leu Val Thr Leu
                245                 250                 255

Val Thr Val Cys Ala Cys Trp Lys Pro Ser Lys Lys Ser Arg Lys Lys
            260                 265                 270

Arg Lys Leu Glu Lys Gln Asn Ser Leu Glu Tyr Met Asp Gln Asn Asp
        275                 280                 285

Asp Arg Leu Lys Ser Glu Ala Asp Thr Leu Pro Arg Ser Gly Glu Gln
    290                 295                 300

Glu Arg Lys Asn Pro Met Ala Leu Tyr Ile Leu Lys Asp Lys Asp Ser
```

305 310 315 320

Ser Glu Pro Asp Glu Asn Pro Ala Thr Glu Pro Arg Ser Thr Thr Glu
325 330 335

Pro Gly Pro Pro Gly Tyr Ser Val Ser Pro Pro Val Pro Gly Arg Ser
340 345 350

Pro Gly Leu Pro Ile Arg Ser Ala Arg Arg Tyr Pro Arg Ser Pro Ala
355 360 365

Arg Ser Pro Ala Thr Gly Arg Thr His Thr Ser Pro Pro Arg Ala Pro
370 375 380

Ser Ser Pro Gly Arg Ser Arg Ser Ser Ser Arg Ser Leu Arg Thr Ala
385 390 395 400

Gly Val Gln Arg Ile Arg Glu Gln Asp Glu Ser Gly Gln Val Glu Ile
405 410 415

Ser Ala

<210> 19
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<220>
<221> GCP Forward
<223> GCP Forward Primer

<400> 19

gggggacaagt ttgtacaaaa aagcaggctt cgccacc 37

<210> 20

<211> 51

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> GCP Reverse

<223> GCP Reverse Primer

<400> 20

ggggaccact ttgtacaaga aagctgggtt tcaatggtga tggtgatggt g 51

<210> 21

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> INSP052-B1P-exon1F

<223> INSP052-B1P-exon1F Primer

<400> 21

gcaggcttcg ccaccatgaa gagagaaagg ggagccctgt c 41

<210> 22

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> INSP052-exon1R

<223> INSP052-exon1R Primer

<400> 22

tcaccccctc cagggggtct gtctggatca gaagaa 36

<210> 23

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> INSP052-exon2F

<223> INSP052-exon2F Primer

<400> 23

ttcttctgat ccagacagac cccctggagg gggtga 36

<210> 24

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> INSP052-exon2R

<223> INSP052-exon2R primer

<400> 24

gtggcctcga aatgggcaca tctacagtaa ggttga 36

<210> 25

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> INSP052-exon3F

<223> INSP052-exon3F primer

<400> 25

caaccttact gtagatgtgc ccatttcgag gccaca 36

<210> 26

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> INSP052-exon3R

<223> INSP052-exon3R primer

<400> 26

ggagcttctt ctgtatacgg tgatcttgac ag 32

<210> 27

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> INSP052-5HIS-R

<223> INSP052-5HIS-R primer

<400> 27

gtgatggtga tggtgggagc ttcttctgta tacgg 35

<210> 28

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> pEAK12-F

<223> pEAK12-F primer

<400> 28

gccagcttgg cacttgatgt 20

<210> 29

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> pEAK-12R

<223> pEAK-12R primer

<400> 29

gatggaggtg gacgtgtcag 20

<210> 30

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<220>

<221> pENTR-F1

<223> pENTR-F1 primer

<400> 30

tcgcgttaac gctagcatgg atctc 25

<210> 31

<211>   24
<212>   DNA
<213>   Primer

<220>
<221>   pENTR-R1
<223>   pENTR-R1 Primer

<400>   31
gtaacatcag agattttgag acac   24

<210>   32
<211>   2025
<212>   DNA
<213>   Homo sapiens

<400>   32
atgaagagag aaaggggagc cctgtccaga gcctccaggg ccctgcgcct tgctcctttt   60
gtctaccttc ttctgatcca gacagacccc ctggaggggg tgaacatcac cagccccgtg   120
cgcctgatcc atggcaccgt ggggaagtcg gctctgcttt ctgtgcagta cagcagtacc   180
agcagcgaca ggcctgtagt gaagtggcag ctgaagcggg acaagccagt gaccgtggtg   240
cagtccattg gcacagaggt catcggcacc ctgcggcctg actatcgaga ccgtatccga   300
ctctttgaaa atggctccct gcttctcagc gacctgcagc tggccgatga gggcacctat   360
gaggtcgaga tctccatcac cgacgacacc ttcactgggg agaagaccat caaccttact   420
gtagatgtgc ccatttcgag gccacaggtg ttggtggctt caaccactgt gctggagctc   480
agcgaggcct tcaccttgaa ctgctcacat gagaatggca ccaagcccag ctacacctgg   540
ctgaaggatg gcaagcccct cctcaatgac tcgagaatgc tcctgtcccc cgaccaaaag   600
gtgctcacca tcacccgcgt gctcatggag gatgacgacc tgtacagctg catggtggag   660
aaccccatca gccagggccg cagcctgcct gtcaagatca ccgtatacag aagaagctcc   720
ctttacatca tcttgtctac aggaggcatc ttcctccttg tgaccttggt gacagtctgt   780
gcctgctgga aaccctccaa aaggaaacag aagaagctag aaaagcaaaa ctccctggaa   840
tacatggatc agaatgatga ccgcctgaaa ccagaagcag acaccctccc tcgaagtggt   900
gagcaggaac ggaagaaccc catggcactc tatatcctga aggacaagga ctccccggag   960
accgaggaga acccggcccc ggagcctcga agcgcgacgg agcccggccc gcccggctac   1020

tccgtgtctc ccgccgtgcc cggccgctcg ccggggctgc ccatccgctc tgcccgccgc 1080

tacccgcgct ccccagcgcg ctccccagcc accggccgga cacactcgtc gccgcccagg 1140

gccccgagct cgcccggccg ctcgcgcagc gcctcgcgca cactgcggac tgcgggcgtg 1200

cacataatcc gcgagcaaga cgaggccggc ccggtggaga tcagcgcctg agccgcctcg 1260

ggatcccctg agaggcgccc gcggtctgcg gccagtggcc cgggggaaag ctggggctgg 1320

gaagcccggg cgcggcgcgc tggggacgag gggaggtccc gggggggcgc tggtgtctcg 1380

ggtgtgaacg tgtatgagca tgcgcagacg gaggcgggtg cgcggaggcg gcagtgttga 1440

tatggtgaaa ccgggtcgca tttgcttccg gtttactggc tgtgtcctca cttggtatag 1500

gttgtgccct cttaggacca catagattat tacatttctg gcccaatacc caaaagggtt 1560

ttatggaaac taacatcagt aacctaaccc ccgtgactat cctgtgctct tcctagggag 1620

ctgtgttgtt tcccacccac cacccttccc tctgaacaaa tgcctgagtg ctggggcact 1680

tttttttttt tttttttttt tttttttttt gcaagttcag attagagagg ccactttccc 1740

agaatccaca gctgcactaa gctaaggaga agccagatgc cggttactgg gtgtgcaggg 1800

gctgttctga gctgggggga tcattgtgaa ggccttcttc cctgggcacc tggtacctgg 1860

ggacctacaa ggtggtgagg gaagggtacg agtacattcc ttttccctct gacctgggcg 1920

ctagcaaggg caaagaaccc gagcctgcca gcttggcctc ctcccacagc ctccctcgga 1980

ggcatgccat gccaagcact ctttctgtct ctgttcatga ataaa 2025

<210> 33

<211> 416

<212> PRT

<213> Homo sapiens

<400> 33

Met Lys Arg Glu Arg Gly Ala Leu Ser Arg Ala Ser Arg Ala Leu Arg

1 5 10 15

Leu Ala Pro Phe Val Tyr Leu Leu Leu Ile Gln Thr Asp Pro Leu Glu

20 25 30

Gly Val Asn Ile Thr Ser Pro Val Arg Leu Ile His Gly Thr Val Gly

35 40 45

Lys Ser Ala Leu Leu Ser Val Gln Tyr Ser Ser Thr Ser Ser Asp Arg

50 55 60

Pro Val Val Lys Trp Gln Leu Lys Arg Asp Lys Pro Val Thr Val Val

65 70 75 80

Gln Ser Ile Gly Thr Glu Val Ile Gly Thr Leu Arg Pro Asp Tyr Arg

85 90 95

Asp Arg Ile Arg Leu Phe Glu Asn Gly Ser Leu Leu Leu Ser Asp Leu

100 105 110

Gln Leu Ala Asp Glu Gly Thr Tyr Glu Val Glu Ile Ser Ile Thr Asp

115 120 125

Asp Thr Phe Thr Gly Glu Lys Thr Ile Asn Leu Thr Val Asp Val Pro

130 135 140

Ile Ser Arg Pro Gln Val Leu Val Ala Ser Thr Thr Val Leu Glu Leu

145 150 155 160

Ser Glu Ala Phe Thr Leu Asn Cys Ser His Glu Asn Gly Thr Lys Pro

165 170 175

Ser Tyr Thr Trp Leu Lys Asp Gly Lys Pro Leu Leu Asn Asp Ser Arg

180 185 190

Met Leu Leu Ser Pro Asp Gln Lys Val Leu Thr Ile Thr Arg Val Leu

195 200 205

Met Glu Asp Asp Asp Leu Tyr Ser Cys Met Val Glu Asn Pro Ile Ser

210 215 220

Gln Gly Arg Ser Leu Pro Val Lys Ile Thr Val Tyr Arg Arg Ser Ser

225 230 235 240

```
Leu Tyr Ile Ile Leu Ser Thr Gly Gly Ile Phe Leu Leu Val Thr Leu
                245                 250                 255

Val Thr Val Cys Ala Cys Trp Lys Pro Ser Lys Arg Lys Gln Lys Lys
            260                 265                 270

Leu Glu Lys Gln Asn Ser Leu Glu Tyr Met Asp Gln Asn Asp Asp Arg
        275                 280                 285

Leu Lys Pro Glu Ala Asp Thr Leu Pro Arg Ser Gly Glu Gln Glu Arg
    290                 295                 300

Lys Asn Pro Met Ala Leu Tyr Ile Leu Lys Asp Lys Asp Ser Pro Glu
305                 310                 315                 320

Thr Glu Glu Asn Pro Ala Pro Glu Pro Arg Ser Ala Thr Glu Pro Gly
                325                 330                 335

Pro Pro Gly Tyr Ser Val Ser Pro Ala Val Pro Gly Arg Ser Pro Gly
            340                 345                 350

Leu Pro Ile Arg Ser Ala Arg Arg Tyr Pro Arg Ser Pro Ala Arg Ser
        355                 360                 365

Pro Ala Thr Gly Arg Thr His Ser Ser Pro Pro Arg Ala Pro Ser Ser
    370                 375                 380

Pro Gly Arg Ser Arg Ser Ala Ser Arg Thr Leu Arg Thr Ala Gly Val
385                 390                 395                 400

His Ile Ile Arg Glu Gln Asp Glu Ala Gly Pro Val Glu Ile Ser Ala
                405                 410                 415
```

<210> 34

<211> 860

<212> DNA

<213> Homo sapiens

<400> 34

acaagtttgt acaaaaaagc aggcttcgcc accatgaaga gagaaagggg agccctgtcc 60

agagcctcca gggccctgcg ccttgctcct tttgtctacc ttcttctgat ccagacagac 120

cccctggagg gggtgaacat caccagcccc gtgcgcctga tccatggcac cgtggggaag 180

tcggctctgc tttctgtgca gtacagcagt accagcagcg acaggcctgt agtgaagtgg 240

cagctgaagc gggacaagcc agtgaccgtg gtgcagtcca ttggcacaga ggtcatcggc 300

accctgcggc ctgactatcg agaccgtatc cgactctttg aaaatggctc cctgcttctc 360

agcgacctgc agctggccga tgagggcacc tatgaggtcg agatctccat caccgacgac 420

accttcactg gggagaagac catcaacctt actgtagatg tgcccatttc gaggccacag 480

accttcactg gggagaagac catcaacctt actgtagatg tgcccatttc gaggccacag 540

gtgttggtgg cttcaaccac tgtgctggag ctcagcgagg ccttcacctt gaactgctca 600

catgagaatg gcaccaagcc cagctacacc tggctgaagg atggcaagcc cctcctcaat 660

gactcgagaa tgctcctgtc cccgaccaa aaggtgctca ccatcacccg cgtgctcatg 720

gaggatgacg acctgtacag ctgcatggtg gagaacccca tcagccaggg ccgcagcctg 780

cctgtcaaga tcaccgtata cagaagaagc tcccaccatc accatcacca ttgaaaccca 840

gctttcttgt acaaagtggt 860

<210> 35

<211> 246

<212> PRT

<213> Homo sapiens

<400> 35

Met Lys Arg Glu Arg Gly Ala Leu Ser Arg Ala Ser Arg Ala Leu Arg

1 5 10 15

Leu Ala Pro Phe Val Tyr Leu Leu Leu Ile Gln Thr Asp Pro Leu Glu

20 25 30

Gly Val Asn Ile Thr Ser Pro Val Arg Leu Ile His Gly Thr Val Gly

35 40 45

Lys Ser Ala Leu Leu Ser Val Gln Tyr Ser Ser Thr Ser Ser Asp Arg

50 55 60

Pro Val Val Lys Trp Gln Leu Lys Arg Asp Lys Pro Val Thr Val Val

65 70 75 80

Gln Ser Ile Gly Thr Glu Val Ile Gly Thr Leu Arg Pro Asp Tyr Arg

85 90 95

Asp Arg Ile Arg Leu Phe Glu Asn Gly Ser Leu Leu Leu Ser Asp Leu

100 105 110

Gln Leu Ala Asp Glu Gly Thr Tyr Glu Val Glu Ile Ser Ile Thr Asp

115 120 125

Asp Thr Phe Thr Gly Glu Lys Thr Ile Asn Leu Thr Val Asp Val Pro

130 135 140

Ile Ser Arg Pro Gln Val Leu Val Ala Ser Thr Thr Val Leu Glu Leu

145 150 155 160

Ser Glu Ala Phe Thr Leu Asn Cys Ser His Glu Asn Gly Thr Lys Pro

165 170 175

Ser Tyr Thr Trp Leu Lys Asp Gly Lys Pro Leu Leu Asn Asp Ser Arg

180 185 190

Met Leu Leu Ser Pro Asp Gln Lys Val Leu Thr Ile Thr Arg Val Leu

195 200 205

Met Glu Asp Asp Asp Leu Tyr Ser Cys Met Val Glu Asn Pro Ile Ser

    210   215   220

Gln Gly Arg Ser Leu Pro Val Lys Ile Thr Val Tyr Arg Arg Ser Ser

225   230   235   240

His His His His His His

                245