



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년08월27일
(11) 등록번호 10-2014890
(24) 등록일자 2019년08월21일

(51) 국제특허분류(Int. Cl.)
G01N 33/574 (2006.01)
(21) 출원번호 10-2013-7026248
(22) 출원일자(국제) 2012년03월05일
 심사청구일자 2017년03월03일
(85) 번역문제출일자 2013년10월04일
(65) 공개번호 10-2014-0040118
(43) 공개일자 2014년04월02일
(86) 국제출원번호 PCT/GB2012/050483
(87) 국제공개번호 WO 2012/120288
 국제공개일자 2012년09월13일
(30) 우선권주장
 1103726.4 2011년03월04일 영국(GB)

(56) 선행기술조사문헌

JP2010522875 A*

(뒷면에 계속)

전체 청구항 수 : 총 6 항

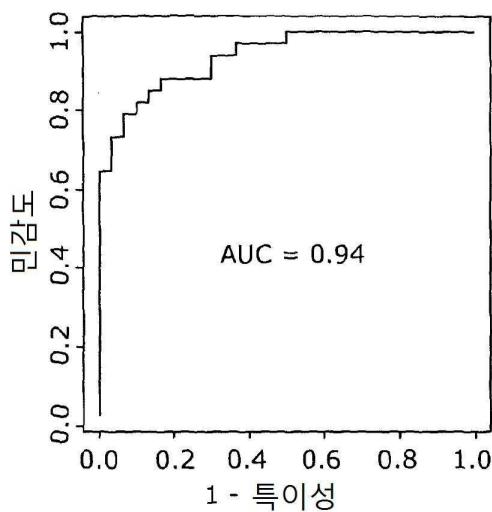
심사관 : 이수진

(54) 발명의 명칭 **췌장암의 존재를 결정하기 위한 방법, 어레이 및 그의 용도**

(57) 요 약

본 발명은 (a) 개체로부터 시험될 샘플을 제공하는 단계, 및 (b) 표 III에 정의된 군으로부터 선택된 하나 이상의 바이오마커의 시험 샘플에서의 발현을 측정함에 의해 시험 샘플의 바이오마커 시그니처를 결정하는 단계를 포함하거나 이로 이루어지는, 개체에서 췌장암의 존재를 결정하기 위한 방법으로서, 표 III에 정의된 군으로부터 선택된 하나 이상의 바이오마커의 시험 샘플에서의 발현은 췌장암을 갖는 개체를 나타내는 것인 방법에 관한 것이다. 본 발명은 또한 본 발명의 방법에서 사용하기 위한 어레이 및 키트를 포함한다.

대 표 도 - 도1c



(56) 선행기술조사문현

KR1020060112258 A

WO2009068857 A1

Chang Stephanie T et al, Journal of
Translational Medicine (2009), vol. 7, no.1,
pp. 105.

Grzesiak John J et al, Pancreas (2007), vol.
35, no.4, pp. 293-301.

WO2007107774 A2

WO2008117067 A2

WO2006113210 A2

*는 심사관에 의하여 인용된 문현

명세서

청구범위

청구항 1

개체에서의 췌장암의 존재를 결정하는 방법으로서:

- a) 상기 개체로부터의 시험 샘플을 제공하는 단계;
- b) 보체-1 에스테라제 억제제 (complement-1 esterase inhibitor), 보체 단백질 C3 (complement protein C3), 보체 단백질 C5, 인터류킨-4, 인터페론-감마, 및 혈관내피 성장 인자 (vascular endothelial growth factor) 바이오마커들의 시험 샘플에서의 발현을 측정하는 것에 의해 상기 시험 샘플의 바이오마커 시그니처 (biomarker signature)를 결정하는 단계를 포함하거나 이로 이루어지고,

단계 (b)에서 측정된 상기 바이오마커들의 시험 샘플에서의 발현은 상기 개체가 췌장암을 갖는다는 것을 나타내는 것인 방법.

청구항 2

청구항 1에 있어서, 단계 (b)는 프로페르딘의 상기 시험 샘플에서의 상기 발현을 측정하는 단계를 포함하는 방법.

청구항 3

청구항 1에 있어서, 단계 (b)는 프로페르딘, 보체-1 에스테라제 억제제, 보체 단백질 C3, 보체 단백질 C4, 보체 단백질 C5, 인터루킨-13, 인터루킨-4, 인터루킨-6, 인터페론-감마, 루이스 X, 시알릴 루이스 X 및 혈관내피 성장 인자의 상기 시험 샘플에서의 상기 발현을 측정하는 단계를 포함하는 방법.

청구항 4

청구항 1에 있어서, 상기 췌장암은 선암 (adenocarcinoma)인 것인 방법.

청구항 5

청구항 1에 있어서, 단계 (b)는 상기 하나 이상의 바이오마커에 결합할 수 있는 항체 또는 그 항원-결합 단편을 사용하여 수행되는 것인 방법.

청구항 6

청구항 5에 있어서, 단계 (b)는 어레이를 이용하여 수행되는 것인 방법.

청구항 7

삭제

청구항 8

삭제

청구항 9

삭제

청구항 10

삭제

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

발명의 설명

기술 분야

[0001] 본 발명은 췌장암의 진단 방법 및 그 방법에서의 사용을 위한 바이오마커 및 어레이에 관한 것이다.

배경 기술

[0002] 큰 효과에도 불구하고, 췌장암(PaC)은 여전히 불량한 예후(prognosis)를 보인다[1]. PaC는 10번째 흔한 암이긴 하지만, 미국에서는 암으로 인한 사망의 4번째 주요 원인이다[2-4]. 실제로, 5년 생존률이 <5%으로서, 모든 악성 종양(malignancy) 중 가장 낮다[2-3]. 그러나, 최근 데이터는 조기 PaC 절제(resection) 후, 30-60% (≤ 20 mm 크기의 암) 및 심지어 >75% (≤ 10 mm 크기의 암)의 5년 생존률로 보여지듯, 암이 대부분 1기에 있을 때, 조기 검출에 의해 그 결과가 놀라울 정도로 향상될 수 있다는 것을 보여주었다[2-4].

[0003] PaC는 빠른 암 진행(tumour progression), 조기 전이(early metastasization) 및 대부분의 기존 치료법에 대한 불응성(unresponsiveness)으로 특징지워진다[1, 5]. 불량한 예후는 주로 질병 진행의 후기에 질병-특이적 임상 증상이 발생하는 것과 효과적인 조기 진단의 부족 때문이다. 진단 시, 종양은 종종 30-40 mm의 크기에 이르고, 모든 환자의 대다수(52%)는 이미 전이(metastasis)를 갖고, 26%는 국부적으로 진행된 암(advanced cancer)을 갖고, 7%만이 췌장에 한정된 종양을 갖는다[2, 4]. 이 때, 환자의 약 15%는 아직 수술이 가능하지만 생존률 중앙값(median survival)은 20개월에 불과하다.

[0004] (내시경) 초음파, 컴퓨터 단층 촬영(computed tomography) 및/또는 내시경 역행성 췌담관 조영(endoscopic retrograde cholangio-pancreatography)을 포함한, 다양한 비침습성 방법(non-invasive methodology)이 PaC 진단을 위해 이용된다[1-2, 6]. 강력하기는 하나, 이러한 방법들은 PaC에 특이한 것이 아니고, 종양이 아직 작고 잠재적으로 치료가 가능한 때 조기 검출을 위해 설계된 것이 아니다. 이러한 상황은 PaC가 현재 이용 가능한 진단 장비를 사용하여, 만성 췌장염(chronic pancreatitis)과 같은 양성 상태(benign condition)로부터 구별하는 것이 어렵다는 사실로 인해 더 복잡해진다[2]. 따라서, PaC의 특이적인 조기 검출을 위한 바이오마커의 이용은 매우 귀중한 임상적 이점이 될 것이다.

[0005] 주요한 노력에도 불구하고, 특히, 미정제(crude), 미분획(non-fractionated) 혈청 및 혈장으로부터의, PaC와 관련된 분자 지문(molecular fingerprint)은 해독될 것이 남아있다[2, 7-9]. 예를 들어 CRP, CA 242, GDF-15, 합토글로빈(haptoglobin), M2-피루베이트 키나제(M2-pyruvate kinase), 혈청 아밀로이드(serum amyloid) A, IGFBP-1을 포함한, 지금까지 나타난 많은 수의 주요한 단일 바이오마커 중에서, CA 19-9 보다 임상적으로 더 뛰어나다고 입증된 것은 없다[2, 8-10]. 또한, CA 19-9의 이용은 그것이 i) 비-악성(non-malignant) 상태(예를 들어, 췌장염(pancreatitis) 및 급성 담관염(acute cholangitis)) 및 기타 위장관 암(gastro-intestinal cancers)(예를 들어, 위암(gastric cancer) 및 결장암(colorectal cancer)) 모두에서 상승되고, ii) 조기 PaC에 대해 민감도가 부족하며, iii) 인구의 약 10%에서 존재하지 않는 것으로 밝혀진 사실로 인해 상당히 제한된다[2, 8-10]. PaC에 대해 스크리닝할 때, CA 19-9는 단지 중간 정도의 민감도(69% 내지 98%의 범위) 및 특이성(46% 내지 98%)을 보였다[2, 9-11].

[0006] 이러한 배경에서, 본 발명자는 WO 2008/117067에서 암의 예후 진단(prognostic diagnosis)에 대한 프로테옴(proteomic) 접근법을 개발하였고, 이에 의해 췌장암의 검출 및 생존율 예측을 위한 혈청 바이오마커의 첫 번째 세트가 확인되었다.

발명의 내용

해결하려는 과제

[0007] 본 발명자가 친화성 프로테오믹스(affinity proteomics)[12-13]가 후보 PaC 혈청 바이오마커 시그니처(biomarker signature)[14]를 정확히 찾을 수 있게 되었다는 것을 보여준 것인 최근 연구로 자극받아, 본 발명자는 추가로 PaC의 혈청 프로테옴(proteome)을 해독하였다.

[0008] 이러한 연구에서, 본 발명자는 처음으로 PaC 진단용 다중화(multiplexed) 혈청 바이오마커 시그니처를 예비 검증하여(pre-validated), 높은 특이성 및 민감도를 보이는, 진단 정보(diagnostic information)가 미정제 혈액 샘플로부터 추출될 수 있다는 것을 보여주었다. 이는 강화된 PaC 진단을 제공하고, 이에 의해 향상된 예후를 제공하여, 유의하게 부가된 임상적 가치를 가져올 뿐만 아니라, 근원적인, 복잡한(intricate) 질병 생물학을 추가적으로 규명한다.

과제의 해결 수단

[0009] 따라서, 본 발명의 제1 양태는 다음 단계를 포함하거나 다음 단계들로 이루어지고:

[0010] a) 개체로부터 시험될 샘플을 제공하는 단계;

[0011] b) 표 III에 정의된 군으로부터 선택된 하나 이상의 바이오마커의 시험 샘플에서의 발현을 측정하여 시험 샘플의 바이오마커 시그니처를 결정하는 단계;

[0012] 표 III에 정의된 군으로부터 선택된 하나 이상의 바이오마커의 시험 샘플에서의 발현은 상기 개체가 체장암을 갖는 것임을 나타내는 것인 개체에서 체장암의 존재를 결정하는 방법을 제공한다

[0013] "시험될 샘플(sample to be tested)", "시험 샘플(test sample)" 또는 "대조군 샘플(control sample)"은, 적절하게, 시험될 개체 또는 대조군 개체로부터의 조직(tissue), 유체(fluid), 프로테옴(proteome) 및/또는 익스프레솜(expressome) 샘플을 포함한다.

[0014] "발현(expression)"은 mRNA 또는 단백질과 같은 유전자 생성물(gene product)의 레벨이나 양을 의미한다.

[0015] 단백질 및/또는 핵산의 농도를 검출 및/또는 측정하는 방법은 통상의 기술자에게 잘 알려져 있고, 예를 들어 Sambrook and Russell, 2001, Cold Spring Harbor Laboratory Press를 참조한다.

[0016] "바이오마커(biomarker)"는 자연-발생 생물학적 분자 또는 그 구성 요소(component) 또는 단편(fragment)을 의미하고, 그의 측정은 체장암의 예후에 유용한 정보를 제공할 수 있다. 예를 들어, 바이오마커는 자연-발생 단백질 또는 탄수화물 모이어티, 또는 그의 항원성 구성요소(antigenic component) 또는 단편(fragment)일 수 있다.

[0017] 일 구현예에서, 상기 방법은 단계 (a) 및 (b)와 추가로 다음 단계를 포함하거나 이들로 이루어지고:

[0018] c) 체장암을 앓지 않는 개체(즉, 음성 대조군)로부터의 대조군 샘플을 제공하는 단계;

[0019] d) 단계 (b)에서 측정된 하나 이상의 바이오마커의 대조군 샘플에서의 발현을 측정하여 대조군 샘플의 바이오마커 시그니처를 결정하는 단계;

[0020] 여기에서 단계 (b)에서 측정된 하나 이상의 바이오마커의 시험 샘플에서의 발현이 단계 (d)에서 측정된 하나 이상의 바이오마커의 대조군 샘플에서의 발현과 상이한 경우 체장암의 존재가 확인된다.

[0021] 다른 구현예에서, 상기 방법은 단계 (a), (b), (c) 및 (d)와 다음의 추가 단계를 포함하거나 이들로 이루어지고:

[0022] e) 체장암을 앓는 개체(즉, 양성 대조군)로부터의 대조군 샘플을 제공하는 단계;

[0023] f) 단계 (b)에서 측정된 하나 이상의 바이오마커의 대조군 샘플에서의 발현을 측정하여 대조군 샘플의 바이오마커 시그니처를 결정하는 단계;

[0024] 단계 (b)에서 측정된 하나 이상의 바이오마커의 시험 샘플에서의 발현이 단계 (f)에서 측정된 하나 이상의 바이오마커의 대조군 샘플에서의 발현과 일치하는 경우 체장암의 존재가 확인된다.

[0025] "대조군 샘플에서의 발현과 일치(corresponds to the expression in the control sample)"는 시험될 샘플에서의 하나 이상의 바이오마커의 발현이 양성 대조군 샘플의 하나 이상의 바이오마커의 발현과 동일하거나 유사한

것을 포함한다. 바람직하게는 시험될 샘플에서의 하나 이상의 바이오마커의 발현이 양성 대조군 샘플의 하나 이상의 바이오마커의 발현과 동일하다.

[0026] 바이오마커의 차등적(differential) 발현(상향 조절 또는 하향 조절) 또는 그의 부족은 통상의 기술자에게 알려진 적절한 수단으로 결정(determine)될 수 있다. 상이한 발현은 적어도 0.05 이하의 p 값($p = < 0.05$), 예를 들어 적어도 <0.04 , <0.03 , <0.02 , <0.01 , <0.009 , <0.005 , <0.001 , <0.0001 , <0.00001 또는 적어도 <0.000001 로 결정된다. 바람직하게는, 상이한 발현은 서포트 벡터 머신(support vector machine, SVM)을 사용하여 결정된다. 바람직하게는, SVM은 하기 기재된 것과 같은 SVM이다. 가장 바람직하게는, SVM은 하기 표 V(A)에 기재된 SVM이다.

[0027] 당업자는 차등적 발현이 단일 바이오마커 또는 조합되어 고려되는 다중 바이오마커(즉, 바이오마커 시그니처로서)와 관련될 수 있다는 것을 이해할 것이다. 따라서, p 값은 단일 바이오마커 또는 일 군의 바이오마커와 관련될 수 있다. 실제로, 개별적으로 고려될 경우 0.05 초과의 p 값의 차등적 발현을 갖는 단백질은 그럼에도 불구하고 그 발현 수준이 하나 이상의 다른 바이오마커와 조합하여 고려될 경우 본 발명에 따라 바이오마커로서 여전히 유용할 수 있다.

[0028] 일 구현예에서, 단계 (b)는 표 IV(A)에 열거된 하나 이상의 바이오 마커, 예를 들어, 표 IV(A)에 열거된 바이오마커 중 2개 이상의 바이오마커의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0029] 수반된 실시예에 예시된 바와 같이, 혈액, 혈청 또는 혈장(plasma) 시험 샘플에서의 특정 단백질의 발현은 개체에서의 췌장암을 나타내는 것일 수 있다. 예를 들어, 단일 시험 샘플에서의 특정 혈청 단백질의 상대적인 발현은 개체에서의 췌장암의 존재를 나타내는 것일 수 있다.

[0030] 바람직하게는, 상기 개체는 인간이다. 그러나, 시험되는 개체는 사육되는 포유류와 같은 포유류(바람직하게는 말, 돼지, 소, 양, 개 및 고양이를 포함한 농업 또는 상업적 중요성이 있는 것)일 수 있다.

[0031] 바람직하게는, 단계 (b)는 인터루킨-7(IL-7) 및/또는 인테그린 알파-10의 발현을 측정하는 단계, 예를 들어 인터루킨-7의 발현을 측정하는 단계, 인테그린 알파-10의 발현을 측정하는 단계, 또는 인터루킨-7 및 인테그린 알파-10의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 가장 바람직하게는, 단계 (b)는 표 IV(A)에 열거된 각 바이오마커의 발현을 측정하는 단계를 포함하거나 이로 이루어질 수 있다.

[0032] 일 구현예에서, 단계 (b)는 표 IV(B)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어, 표 IV(B)에 열거된 바이오마커 중 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 또는 12개 이상의 바이오마커의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 따라서, 단계 (b)는 바람직하게는 표 IV(B)에 열거된 바이오마커 모두의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0033] 다른 구현예에서, 단계 (b)는 표 IV(C)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 IV(C)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 또는 41개의 바이오마커의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는 표 IV(C)에 열거된 바이오마커 모두의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0034] 바람직하게는, 단계 (b)는 표 IV에 정의된 바이오마커 모두의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0035] 일 구현예에서, 상기 방법은 췌장암(PaC)과, 다른 질병 상태를 구별하기 위한 것이다.

[0036] 바람직하게는, 단계 (b)는 표 V(A)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어, 표 V(A)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는 또한 표 V(B)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(B)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 또는 24개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또한 단계 (b)가 표 V(C)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(C)에 열거된 바이오마커 중 적어도 2, 3, 4 또는 5개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어지는 것이 바람직하다. 바람직하게는, 단계 (b)는 표 V(D)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(D)에 열거된 바이오마커 중 적어도 2 또는 3개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는

표 V(F)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(F)에 열거된 바이오마커 중 적어도 2, 3, 4, 5 또는 6개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는 단계 (b)는 표 V(A), 표 V(B), 표 V(C), 표 V(D) 및/또는 표 V(F)에 열거된 모든 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0037] "췌장암(PaC)과 다른 질병 상태를 구별(differentiating between pancreatic cancer(Pac) and another disease state)"은 PaC와 건강한 상태를 포함한, 다른 상태를 구별하는 것을 포함한다.

[0038] 일 구현예에서, 다른 질병 상태 또는 상태들은 만성 췌장염(ChP), 급성 염증성 췌장염(acute inflammatory pancreatitis, AIP) 및/또는 정상, 예를 들어 다른 질병 상태 또는 질병 상태들은 만성 췌장염 단독; 급성 염증성 췌장염 단독; 만성 췌장염 및 급성 염증성 췌장염; 만성 췌장염 및 정상; 급성 염증성 췌장염 및 정상; 또는 만성 췌장염, 급성 염증성 췌장염 및 정상일 수 있다.

[0039] "정상(normal)" 질병 상태를 지칭할 경우, 만성 췌장염(ChP) 또는 급성 염증성 췌장염(AIP)를 앓지 않는 개체를 포함한다. 바람직하게는 상기 개체는 어떠한 췌장 질병 또는 장애도 앓지 않는다. 가장 바람직하게는, 상기 개체는 건강한 개체이고, 즉 어떠한 질병 또는 장애도 앓지 않는다.

[0040] 다른 구현예에서, 상기 방법은 췌장암과 만성 췌장염(ChP)을 구별하기 위한 것이다. 바람직하게는, 단계 (b)는 표 V(A)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(A)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 단계 (b)는 표 V(C)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(C)에 열거된 바이오마커 중 적어도 2, 3, 4 또는 5개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어질 수 있다. 단계 (b)는 표 V(A) 및/또는 표 V(C)에 열거된 모든 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어질 수 있다.

[0041] 추가적인/대안적인 구현예에서, 상기 방법은 췌장암과 급성 염증성 췌장염(AIP)을 구별하기 위한 것이고, 단계 (b)는 표 V(A)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(A)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는 표 V(B)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(B)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 또는 24개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함한다. 단계 (b)는 표 V(C)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(C)에 열거된 바이오마커 중 적어도 2, 3, 4 또는 5개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어질 수 있다. 바람직하게는, 단계 (b)는 표 V(E)에 열거된 바이오마커로부터 하나 이상의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는 표 V(F)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어, 표 V(F)에 열거된 바이오마커 중 적어도 2, 3, 4, 5 또는 6개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는 표 V(H)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어, 표 V(H)에 열거된 바이오마커 중 적어도 2 또는 3개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 따라서, 단계 (b)는 바람직하게는 표 V(A), 표 V(B), 표 V(C), 표 V(E), 표 V(F) 및/또는 표 IV(H)에 열거된 모든 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0042] 일 구현예에서, 상기 방법은 췌장암과 정상(N)을 구별하기 위한 것이다. "정상(normal)" 질병 상태의 정의에 대해서는, 상기를 참조하라. 바람직하게는 단계 (b)는 표 V(A)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(A)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9 또는 10개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는 표 V(B)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(B)에 열거된 바이오마커 중 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 또는 24개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 단계 (b)는 표 V(D)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(D)에 열거된 바이오마커 중 적어도 2 또는 3개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 바람직하게는, 여기에서 단계 (b)는 표 V(E)에 열거된 바이오마커로부터 하나 이상의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또한 단계 (b)가 표 V(G)에 열거된 바이오마커로부터 하나 이상의 바이오마커, 예를 들어 표 V(G)에 열거된 바이오마커 중 적어도 2 또는 3개의 바이오마커의 시험 샘플에서의 발현을 측정하는 단계

를 포함하거나 이로 이루어지는 것이 바람직하다. 따라서, 단계 (b)는 표 V(A), 표 V(B), 표 V(D), 표 V(E) 및/또는 표 IV(G)에 열거된 모든 바이오마커의 시험 샘플에서의 발현을 측정하는 단계를 포함하거나 이로 이루어질 수 있다.

[0043]

일 구현예에서, 단계 (b)는 IL-3의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 인테그린 α -10의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 뮤신-1(Mucin-1)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 C1s의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 MCP-3의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서, 단계 (b)는 안지오모틴(Angiomotin)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 BTK의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 C1q의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 CD40 리간드의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 GM-CSF의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서, 단계 (b)는 IgM의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 IL-11의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 IL-16의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 IL-1- α 의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 IL-1 β 의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 IL-2의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 IL-7의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 IL-9의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 JAK3의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 렘틴(Leptin)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 루이스 y(Lewis y)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 MCP-4의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서, 단계 (b)는 프로카테psin W(Procathepsin W)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 프로페리딘(Properdin)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 PSA의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 RANTES의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 시알릴(Sialyl) 루이스 x의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서, 단계 (b)는 TM 펩티드의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 TNF- α 의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 C4의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 β -갈락토시다제 (β -galactosidase)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0044]

추가 구현예에서, 단계 (b)는 IL-12의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서, 단계 (b)는 TGF- β 1의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 VEGF의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 IL-8의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 C3의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 IFN- γ 의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서, 단계 (b)는 IL-10의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 IL-13의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 추가 구현예에서, 단계 (b)는 IL-18의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 IL-6의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 루이스 x의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서, 단계 (b)는 에오팩신(Eotaxin)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 C1 에스테라제(esterase) 억제제의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서, 단계 (b)는 MCP-1의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 TNF- β 의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 GLP-1의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 일 구현예에서 단계 (b)는 IL-5의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서 단계 (b)는 IL-4의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 또 다른 구현예에서 단계 (b)는 인자 B

(Factor B)의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 다른 구현예에서, 단계 (b)는 C5의 발현을 측정하는 단계를 포함하거나 이로 이루어진다. 추가 구현예에서, 단계 (b)는 CD40의 발현을 측정하는 단계를 포함하거나 이로 이루어진다.

[0045] 일 구현예에서 단계 (b)는 IL-3의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 인테그린 α -10의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 뮤신-1의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 C1s의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 MCP-3의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 안지오모틴의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 BTK의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 C1q의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 CD40 리간드의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 GM-CSF의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 IgM의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 IL-11의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 IL-1-*ra*의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 IL-1 α 의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 IL-1 β 의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 IL-2의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 IL-9의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 INF- γ 의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 인테그린 α -11의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 JAK3의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 렙틴의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 MCP-4의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 프로카텝신 W의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 프로퍼딘의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 PSA의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 RANTES의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 시알릴 루이스 x의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 TM 웨პ티드의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 TNF- α 의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 C4의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 β -갈اكت오시다제의 발현을 측정하는 단계를 포함하지 않는다.

[0046] 추가 구현예에서, 단계 (b)는 IL-12의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 TGF- β 1의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 VEGF의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 IL-8의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 C3의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 IFN- γ 의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 IL-10의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 IL-13의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 IL-18의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 IL-6의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 루이스 x의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 에오팍신의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 C1 에스테라제 억제제의 발현을 측정하는 단계를 포함하지 않는다. 또 다른 구현예에서, 단계 (b)는 MCP-1의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 TNF- β 의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 GLP-1의 발현을 측정하는 단계를 포함하지 않는다. 일 구현예에서, 단계 (b)는 IL-5의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 인자 B의 발현을 측정하는 단계를 포함하지 않는다. 다른 구현예에서, 단계 (b)는 C5의 발현을 측정하는 단계를 포함하지 않는다. 추가 구현예에서, 단계 (b)는 CD40의 발현을 측정하는 단계를 포함하지 않는다.

[0047] "TM 웨პ티드(TM peptide)"는 10TM 단백질로부터 유래한 웨პ티드를 의미하고, 여기에 하기 서열번호(SEQ ID NO) 1의 scFv 항체 구조물(antibody construct)이 특이성을 갖는다(여기에서 CDR 서열은 볼드, 이탈릭체의 문자로 표시된다)

- [0048] MAEVQLLESGGGLVQPGGLRLSCAASGFT**FSSYGFHWRQAPGK**GLEWV**SLISWDGGSTYYADSVKGRFTI**SRDNSKNTLYLQMNSLRAEDTAVYYCARGT**WFDPWGQGTLVTVSSGGGSGGGGSGGGGSQSLTQPPSASGTPGQRTI**SCS**GSSSNIGNNAVWYQQLPGTAPKLLIYRNNQRPSGVPDRFSGSKSGTSA**SLAISGLRSEDEADYY**CAAWDDSLSWVFGGTKLTVLG**
- [0049] [서열번호: 1]
- [0050] 따라서, 이러한 scFv, 또는 10TM 단백질과 결합하기 위해 이러한 scFv와 경쟁하는, 항체 또는 그의 항원 결합 단편(antigenic binding fragment)이 사용될 수 있다. 예를 들어, 항체 또는 그의 항원 결합 단편은 서열번호 1에 존재하는 것과 동일한 CDR을 포함할 수 있다.
- [0051] 그러한 항체가 정제 목적으로 친화성 태그(affinity tag)(예를 들어, C-말단에서)와 함께 생성될 수 있다는 것이 통상의 기술자에게 이해될 것이다. 예를 들어, 하기 서열번호 2의 친화성 태그가 이용될 수 있다.
- [0052] DYKDHDGDYKDHDIDYKDDDKAAAHHHHHH
- [0053] [서열번호: 2]
- [0054] 일 구현예에서, 췌장암의 존재는 IL-3, 인테그린 α -10, 뮤신-1, C1s, GLP-1R, MCP-3, 안지오모틴, BTK, CD40 리간드, GM-CSF, IgM, IL-11, IL-16, IL-1-*ra*, IL-1 α IL-1 β IL-2, IL-7, IL-9, INF- γ , 인테그린 α -11, JAK3, 텝틴, 루이스 γ , MCP-4, 프로카텝신 W, PSA, RANTES, 시알릴 루이스 x, TM 웹티드, TNF- α , C4, β -갈락토시다제, IL-12, TGF- β 1, VEGF, IL-8, C3, IFN- γ , IL-10, IL-13, IL-18, IL-6, 루이스 x, 에오틱신, C1 에스테라제 억제제, MCP-1, TNF- β , GLP-1, IL-5, IL-4, 인자 B, C5 및/또는 CD40의 시험 샘플에서의 발현이 음성 대조군(들)과 비교하여 상향 조절(up-regulated)되거나 및/또는 양성 대조군(들)의 발현에 일치하는 경우에 확인된다.
- [0055] 다른 구현예에서, 췌장암의 존재는 C1q 및/또는 프로퍼딘의 시험 샘플에서의 발현이 음성 대조군(들)과 비교하여 하향 조절(down-regulated)되거나 및/또는 양성 대조군(들)의 발현과 일치하는 경우에 확인된다. 일반적으로, 진단은 적어도 0.55의 ROC AUC, 예를 들어 적어도 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 0.96, 0.97, 0.98, 0.99의 ROC AUC 또는 1.00의 ROC AUC에 의해 이루어진다. 바람직하게는, 진단은 적어도 0.85의 ROC AUC 및 가장 바람직하게는 1의 ROC AUC로 이루어진다.
- [0056] 일반적으로, 진단은 <http://cran.r-project.org/web/packages/e1071/index.html> (예를 들어, e1071 1.5-24)로부터 이용가능한 것들과 같은, 서포트 벡터 머신(SVM)을 사용하여 수행된다. 그러나, 다른 적절한 수단도 사용될 수 있다.
- [0057] 서포트 벡터 머신(SVM)은 분류(classification) 및 회귀(regression)를 위해 사용되는 관련 감독 학습법(supervised learning method)의 한 세트이다. 각각 둘 중 하나의 카테고리에 속하는 것으로 표시되는, 한 세트의 훈련 예(training example)가 주어지면, SVM 훈련 알고리즘은 새로운 예가 하나의 카테고리 또는 다른 카테고리에 속하는지 여부를 예측하는 모델을 정립한다. 직관적으로, SVM 모델은 맵핑된, 공간의 포인트로서 예를 표시하고, 따라서 개별적인 카테고리의 예는 가능한 넓은 명확한 캡으로 나뉜다. 그 다음 새로운 예가 동일한 공간으로 맵핑되고 그들이 속하는 캡의 쪽에 기반한 카테고리에 속하는 것으로 예측된다.
- [0058] 더욱 형식적으로, 서포트 벡터 머신은 높은 또는 무한 차원 공간(dimensional space)에서 초평면(hyperplane) 또는 초평면 세트를 구성하는데, 이는 분류, 회귀 또는 다른 작업을 위해 사용될 수 있다. 직관적으로, 우수한 분리(separation)는 클래스의 가장 가까운 훈련 데이터 포인트(training datapoint)에서 가장 먼 거리에 있는 초평면(소위 기능적 여유도(functional margin))에 의해 달성되는데, 이는 일반적으로 여유도가 더 클수록 분류자(classifier)의 일반화 에러(generalization error)는 더 낮기 때문이다. SVM에 관한 더 많은 정보에 대해서는 예를 들어 Burges, 1998, *Data Mining and Knowledge Discovery*, 2:121-167을 참조한다.
- [0059] 본 발명의 일 구현예에서, SVM은 알려진 질병 상태를 갖는 개체(예를 들어, 췌장암을 갖는 것으로 알려진 개체, 급성 염증성 췌장염을 갖는 것으로 알려진 개체, 만성 췌장염을 갖는 것으로 알려진 개체 또는 건강한 것으로

알려진 개체)로부터의 바이오마커 프로파일(profile)을 사용하여 본 발명의 방법을 수행하기에 앞서 훈련(train)된다. 그러한 훈련 샘플을 운영하여, SVM은 어떤 바이오마커 프로파일이 체장암과 연관되었는지를 알 수 있다. 일단 훈련 과정이 완료되면, SVM은 시험된 바이오마커 샘플이 체장암을 갖는 개체로부터 유래된 것인지 아닌지를 알 수 있다.

[0060] 그러나, 이러한 훈련 방법은 SVM을 필요한 훈련 파라미터로 예비-프로그래밍(pre-programming)함에 의해 우회될 수 있다. 예를 들어, 진단은 표 III 또는 표 IV에 열거된 바이오마커 중 하나 또는 모두의 측정에 기초하여, 표 V에 상세히 기재된 SVM 알고리즘을 사용하여 알려진 SVM 파라미터에 따라 수행될 수 있다.

[0061] 적절한 SVM 파라미터는 데이터의 적절한 선택으로 SVM 머신을 훈련시킴에 의해 표 III 또는 표 IV에 열거된 바이오마커의 어떠한 조합에 대해서도 결정될 수 있다는 것을 통상의 기술자는 이해할 것이다(즉, 알려진 체장암 상태를 갖는 개체로부터 바이오마커 측정).

[0062] 바람직하게, 본 발명의 방법은 적어도 60%의 정확도(accuracy), 예를 들어 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%의 정확도를 갖는다.

[0063] 바람직하게, 본 발명의 방법은 적어도 60%의 민감도(sensitivity), 예를 들어 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%의 민감도를 갖는다.

[0064] 바람직하게, 본 발명의 방법은 적어도 60%의 특이성(specificity), 예를 들어 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% 또는 100%의 특이성을 갖는다.

[0065] "정확도"는 방법의 올바른 결과의 비율을 의미하고, "민감도"는 양성으로 올바르게 분류된 모든 PaC 양성 샘플의 비율을 의미하며, "특이성"은 음성으로 올바르게 분류된 모든 PaC 음성 샘플의 비율을 의미한다.

[0066] 일 구현예에서, 체장암을 앓지 않는 개체는 체장암(PaC), 만성 체장염(ChP) 또는 급성 염증성 체장염(AIP)을 앓지 않는다. 더욱 바람직하게, 건강한 개체는 어떠한 체장 질병 또는 질환도 앓지 않는다. 훨씬 더 바람직하게는, 체장암을 앓지 않는 개체는 어떠한 질병 또는 질환도 앓지 않는다. 가장 바람직하게는, 체장암을 앓지 않는 개체는 건강한 개체이다. "건강한 개체(healthy individual)"는 신체적으로 활발하고 신체적 질병이 없는 것으로 통상의 기술자에 의해 간주되는 개체를 포함한다.

[0067] 그러나, 다른 구현예에서, 체장암을 앓지 않는 개체는 만성 체장염을 앓는다. 또 다른 구현예에서, 체장암을 앓지 않는 개체는 급성 염증성 체장염을 앓는다.

[0068] 이미 언급한 바와 같이, 본원의 방법은 개체에서 체장암의 존재를 결정하기 위한 것이다. 일 구현예에서, 체장암은 선암(adenocarcinoma), 선편평세포 암종(adenosquamous carcinoma), 인환세포 암종(signet ring cell carcinoma), 간세포양 암종(hepatoid carcinoma), 콜로이드 암종(colloid carcinoma), 미분화 암종(undifferentiated carcinoma) 및 용골세포형 거대세포(osteoclast-like giant cells)를 갖는 미분화 암종으로 이루어진 군으로부터 선택된다. 바람직하게는, 체장암은 체장 선암(pancreatic adenocarcinoma)이다. 더욱 바람직하게는, 체장암은 외분비 체장암(exocrine 체장암)으로도 알려진 체장 관세포암(pancreatic ductal adenocarcinoma)이다.

[0069] 다른 구현예에서, 단계 (b), (d) 및/또는 단계 (f)는 하나 이상의 바이오마커와 결합할 수 있는 제1 결합제(binding agent)를 사용하여 수행된다. 제1 결합제는 단백질 바이오마커 한 개에 대한 특이성을 갖는 단일 종 또는 각각이 상이한 단백질 바이오마커에 대한 특이성을 갖는, 복수의 상이한 종을 포함하거나 이로 이루어질 수 있다는 것이 통상의 기술자에 의해 이해될 것이다.

[0070] 적절한 결합제(결합 분자로도 지칭됨)는 하기에서 논의되는 바와 같이, 주어진 모티브(motif)와 결합하는 그의 능력에 기초하여, 라이브러리로부터 선택될 수 있다.

[0071] 결합제의 적어도 하나의 타입 및 더욱 일반적으로 모든 타입이 항체 또는 그의 항원-결합 단편 또는 그의 변이체(variant)를 포함하거나 이로 이루어질 수 있다.

[0072] 항체의 생산 및 사용 방법은 해당 기술분야에 잘 알려져 있고, 예를 들어 *Antibodies: A Laboratory Manual*, 1988, Harlow & Lane, Cold Spring Harbor Press, ISBN-13: 978-0879693145, *Using Antibodies: A Laboratory*

Manual, 1998, Harlow & Lane, Cold Spring Harbor Press, ISBN-13: 978-0879695446 및 *Making and Using Antibodies: A Practical Handbook*, 2006, Howard & Kaser, CRC Press, ISBN-13: 978-0849335280를 참조한다 (이들 개시는 참조로서 본 명세서에 포함된다).

[0073] 따라서, 단편은 하나 이상의 가변 중쇄(V_H) 또는 가변 경쇄(V_L) 도메인을 포함할 수 있다. 예를 들어, 용어 항체 단편(antibody fragment)은 Fab-유사 분자(Better *et al* (1988) *Science* **240**, 1041); Fv 분자(Skerra *et al* (1988) *Science* **240**, 1038); V_H 및 V_L 파트너 도메인이 유연한 올리고펩티드(flexible oligopeptide)를 통해 연결된 단쇄 Fv(ScFv) 분자(Bird *et al* (1988) *Science* **242**, 423; Huston *et al* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5879) 및 단리된 V 도메인을 포함하는 단일 도메인 항체(dAbs) (Ward *et al* (1989) *Nature* **341**, 544)를 포함한다.

[0074] 용어 "항체 변이체(antibody variant)"는 합성 항체, 재조합 항체 또는 항체 하이브리드(antibody hybrid), 예를 들어 면역글로불린 경쇄 및/또는 중쇄 가변 및/또는 불변 영역(constant region)의 파지-디스플레이(phage-display)에 의해 생성된 단쇄 항체 분자, 또는 통상의 기술자에게 알려진 면역 분석(immunoassay format) 포맷에서 항원과 결합할 수 있는 기타 면역상호작용(immunointeractive) 분자를 포함하나 이에 제한되지 않는다.

[0075] 그의 특이적 결합 부위(binding site)를 보유하는 항체 단편의 합성에 관련된 기술의 일반적 검토는 Winter & Milstein (1991) *Nature* **349**, 293-299에서 발견된다.

[0076] 항체 라이브러리(Clackson *et al*, 1991, *Nature* **352**, 624-628; Marks *et al*, 1991, *J Mol Biol* **222**(3): 581-97), 펩티드 라이브러리(Smith, 1985, *Science* **228**(4705): 1315-7), 발현된 cDNA 라이브러리(Santi *et al* (2000) *J Mol Biol* **296**(2): 497-508), 어피바디(affibodies)와 같은 항체 프레임워크(framework)가 아닌 기타 지지체(scaffold)에 대한 라이브러리(Gunneriusson *et al*, 1999, *Appl Environ Microbiol* **65**(9): 4134-40) 또는 앱타머(aptamer)에 기초한 라이브러리(Kenan *et al*, 1999, *Methods Mol Biol* **118**, 217-31)가 이로부터 주어진 모티브에 특이적인 결합 분자가 본 발명의 방법에서의 사용을 위해 선택되는 소스(source)로 사용될 수 있다.

[0077] 분자 라이브러리는 원핵세포(prokaryotic cell) (Clackson *et al*, 1991, *op. cit.*; Marks *et al*, 1991, *op. cit.*) 또는 진핵세포(eukaryotic cell) (Kieke *et al*, 1999, *Proc Natl Acad Sci USA*, **96**(10):5651-6)에서 인비보(*in vivo*)에서 발현될 수 있거나, 세포의 관여 없이 인비트로(*in vitro*)에서 발현될 수 있다(Hanes & Pluckthun, 1997, *Proc Natl Acad Sci USA* **94**(10):4937-42; He & Taussig, 1997, *Nucleic Acids Res* **25**(24):5132-4; Nemoto *et al*, 1997, *FEBS Lett*, **414**(2):405-8).

[0078] 단백질 기반 라이브러리가 사용될 경우, 종종 잠재적 결합 분자의 라이브러리를 코딩하는 유전자가 바이러스에 팩키징(packaged)되고 잠재적 결합 분자는 바이러스의 표면에 표시(display)된다(Clackson *et al*, 1991, *op. cit.*; Marks *et al*, 1991, *op. cit.*; Smith, 1985, *op. cit.*).

[0079] 오늘날 가장 흔하게 사용되는 그러한 시스템은 그의 표면에 항체 단편을 표시하는 필라멘트성 박테리오파지(filamentous bacteriophage)이고, 항체 단편은 박테리오파지의 소수 외피 단백질(minor coat protein)로의 융합체(fusion)로서 발현된다(Clackson *et al*, 1991, *op. cit.*; Marks *et al*, 1991, *op. cit.*). 그러나, 또한 다른 바이러스(EP 39578), 박테리아(Gunneriusson *et al*, 1999, *op. cit.*; Daugherty *et al*, 1998, *Protein Eng* **11**(9):825-32; Daugherty *et al*, 1999, *Protein Eng* **12**(7):613-21), 및 효모(Shusta *et al*, 1999, *J Mol Biol* **292**(5):949-56)를 사용하여 표시하기 위한 다른 시스템이 사용되고 있다.

[0080] 또한, 표시 시스템(display system)은 소위 리보솜(ribosome) 표시 시스템에서 폴리펩티드 생성물과 그것을 코딩 mRNA와의 연관(linkage) (Hanes & Pluckthun, 1997, *op. cit.*; He & Taussig, 1997, *op. cit.*; Nemoto *et al*, 1997, *op. cit.*) 또는 대안적으로 폴리펩티드 생성물과 그를 코딩 DNA와의 연관(US Patent No. 5,856,090 및 WO 98/37186 참조)을 이용하여 개발되었다.

[0081] 잠재적 결합 분자가 라이브러리로부터 선택될 경우, 정의된 모티브를 갖는 하나 또는 수 개의 셀렉터 펩티드(selector peptide)가 일반적으로 채용된다. 구조, 펩티드에서 감소하는 유연성(flexibility) 또는 결합 분자와의 상호작용을 허용하는 하전된, 극성 또는 소수성(hydrophobic) 사슬을 제공하는 아미노산 잔기(residue)가 셀렉터 펩티드를 위한 모티브의 설계에 사용될 수 있다.

[0082] 예를 들어:

- [0083] (i) 프롤린은 그의 측쇄(side chain)가 알파 탄소 및 질소 모두와 결합할 때 웹티드 구조를 안정화할 수 있고;
- [0084] (ii) 페닐알라닌, 티로신 및 트립토판은 방향족 측쇄를 갖고 높은 소수성인 반면, 류신 및 이소류신은 지방족(aliphatic) 측쇄를 갖고 또한 소수성이고;
- [0085] (iii) 리신, 아르기닌 및 히스티딘은 염기성 측쇄를 갖고 중성 pH에서 양으로 하전될 것인 반면, 아스파르테이트 및 글루타메이트는 산성 측쇄를 갖고 중성 pH에서 음으로 하전될 것이고;
- [0086] (iv) 아스파라긴 및 글루타민은 중성 pH에서 중성이지만, 수소 결합에 참여할 수 있는 아미노기를 포함하고;
- [0087] (v) 세린, 트레오닌 및 티로신 측쇄는 히드록실기를 포함하고, 이는 수소 결합에 참여할 수 있다.
- [0088] 일반적으로, 결합제의 선택은 결합 분자의 타입에 해당하는 스팟(spot)에의 결합을 분석하기 위한 어레이 기술 및 시스템의 용도를 포함할 수 있다.
- [0089] 일 구현예에서, 제1 결합제(들)은 표면(예를 들어, 멀티웰 플레이트 또는 어레이)에 고정화된다.
- [0090] 항체의 가변 중쇄(V_H) 도메인 및 가변 경쇄(V_L) 도메인은 항원 인식(antigen recognition)에 관련되며, 이는 조기 프로테아제 소화 실험(early protease digestion experiments)에 의해 처음으로 인식된다는 사실이다. 설치류 항체(rodent antibody)의 "인간화(humanisation)"에 의해 추가 확인이 이루어졌다. 설치류 기원의 가변 도메인은 결과물 항체는 설치류 부모인(rodent parented) 항체의 항원 특이성(antigenic specificity)을 보유하도록 인간 기원의 불변 도메인과 융합시킬 수 있다(Morrison *et al* (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6851-6855).
- [0091] 항원 특이성은 가변 도메인에 의해 주어지고, 불변 도메인과 독립적이라는 점은 모두가 하나 이상의 가변 도메인을 포함하는, 항체 단편의 박테리아 발현을 수반하는 실험으로부터 알려졌다. 이러한 분자는 Fab-유사 분자 (Better *et al* (1988) *Science* **240**, 1041); Fv 분자 (Skerra *et al* (1988) *Science* **240**, 1038); V_H 및 V_L 파트너 도메인이 유연한 올리고웹티드를 통해 연결되는 단쇄 Fv(ScFv) 분자(Bird *et al* (1988) *Science* **242**, 423; Huston *et al* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5879) 및 단리된 V 도메인을 포함하는 단일 도메인 항체 (dAbs) (Ward *et al* (1989) *Nature* **341**, 544)를 포함한다. 특이적 결합 부위를 보유한 항체 단편의 합성에 관련된 기술의 일반적인 검토는 Winter & Milstein (1991) *Nature* **349**, 293-299에서 발견된다.
- [0092] "ScFv 분자(ScFv molecules)"은 V_H 파트너 영역과 V_L 파트너 영역이 유연한 올리고웹티드를 통해 연결된 것인 분자를 의미한다.
- [0093] 전체 항체가 아닌 항체 단편을 사용하는 장점은 여러 가지이다. 단편의 더 작은 크기는 고형 조직(solid tissue)의 더 나은 침투(penetration)와 같은, 향상된 약리적 특성을 가져올 수 있다. 보체 결합(complement binding)과 같은, 전체 항체의 효과기 기능(Effecter function)은 제거된다. Fab, Fv, ScFv 및 dAb 항체 단편은 모두 대장균(*E. coli*)에서 발현되고 대장균으로부터 분비될 수 있고, 따라서 많은 양의 상기 단편의 용이한 생산(facile production)을 가능하게 한다.
- [0094] 전체 항체 및 $F(ab')_2$ 단편은 "이가(bivalent)"이다. "이가"는 상기 항체 및 $F(ab')_2$ 단편이 두 개의 항원 결합 부위를 가짐을 의미한다. 반대로, Fab, Fv, ScFv 및 dAb 단편은 일가(monovalent)이고, 오직 하나의 항원 결합 부위를 갖는다.
- [0095] 항체는 모노클론(monoclonal) 또는 폴리클론(polyclonal)일 수 있다. 적절한 모노클론 항체는 공지된 기법, 예를 들어 모두 참조로서 본 명세서에 포함된, "Monoclonal Antibodies: A manual of techniques" H Zola (CRC Press, 1988) 및 "Monoclonal Hybridoma Antibodies: Techniques and applications", J G R Hurrell (CRC Press, 1982)에 개시된 것에 의해 제조될 수 있다.
- [0096] 일 구현예에서, 제1 결합제는 표면(예를 들어, 멀티웰 플레이트 또는 어레이)에 고정된다.
- [0097] 항체의 가변 중쇄(V_H) 도메인 및 가변 경쇄(V_L) 도메인은 항원 인식에 관련되며, 이는 조기 프로테아제 소화 실험에 의해 처음으로 인식된다는 사실이다. 설치류 항체(rodent antibody)의 "인간화"에 의해 추가 확인이 되었다. 설치류 기원의 가변 도메인은 결과물 항체가 설치류가 부모인(rodent parented) 항체의 항원 특이성(antigenic specificity)을 보유하도록 인간 오리진의 불변 도메인과 융합시킬 수 있다(Morrison *et al* (1984) *Proc. Natl. Acad. Sci. USA* **81**, 6851-6855).

- [0098] 항원 특이성은 가변 도메인에 의해 주어지고, 불변 도메인과 독립적이라는 점은 모두가 하나 이상의 가변 영역을 포함하는, 항체 단편의 박테리아 발현을 수반하는 실현으로부터 알려졌다. 이러한 분자는 Fab-유사 분자 (Better *et al* (1988) *Science* **240**, 1041); Fv 분자 (Skerra *et al* (1988) *Science* **240**, 1038); V_H 및 V_L 파트너 도메인이 유연한 올리고펩티드를 통해 연결된 것인 단쇄 Fv (ScFv) 분자 (Bird *et al* (1988) *Science* **242**, 423; Huston *et al* (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5879) 및 단리된 V 영역을 포함하는 단일 도메인 항체(dAbs) (Ward *et al* (1989) *Nature* **341**, 544)를 포함한다. 특이적 결합 부위를 보유한 항체 단편의 합성에 관련된 기법의 일반적인 검토는 Winter & Milstein (1991) *Nature* **349**, 293-299에서 발견된다.
- [0099] "ScFv 분자"는 V_H 파트너 도메인과 V_L 파트너 도메인이 유연한 올리고펩티드를 통해 연결된 것인 분자를 의미한다.
- [0100] 전체 항체가 아닌 항체 단편을 사용하는 장점은 여러 가지이다. 단편의 더 작은 크기는 고형 조직의 더 나은 침투와 같은, 향상된 약리적 특성을 가져올 수 있다. 보체 결합과 같은, 전체 항체의 효과기 기능은 제거된다. Fab, Fv, ScFv 및 dAb 항체 단편은 모두 대장균(*E. coli*)에서 발현되고 대장균으로부터 분비될 수 있고, 따라서 많은 양의 상기 단편의 용이한 생산(facile production)을 가능하게 한다.
- [0101] 전체 항체 및 F(ab')₂ 단편은 "이가"이다. "이가"는 상기 항체 및 F(ab')₂ 단편이 두 개의 항원 결합 부위를 가짐을 의미한다. 대조적으로, Fab, Fv, ScFv 및 dAb 단편은 일가(monovalent)이고, 오직 하나의 항원 결합 부위를 갖는다.
- [0102] 항체는 모노클론 또는 폴리클론일 수 있다. 적절한 모노클론 항체는 알려진 기법, 예를 들어 모두 참조로서 본 명세서에 포함된, "Monoclonal Antibodies: A manual of techniques" H Zola (CRC Press, 1988) 및 "Monoclonal Hybridoma Antibodies: Techniques and applications", J G R Hurrell (CRC Press, 1982)에 개시된 것에 의해 제조될 수 있다.
- [0103] 따라서, 제1 결합체는 항체 또는 그의 항원-결합 단편을 포함하거나 이로 이루어질 수 있다. 바람직하게는, 항체 또는 그의 항원-결합 단편은 재조합 항체 또는 그의 항원-결합 단편이다. 항체 또는 그의 항원-결합 단편은 다음으로 이루어진 군으로부터 선택될 수 있다: scFv, Fab 및 면역글로불린 분자의 결합 도메인.
- [0104] 바람직하게는, 제1 결합체는 표면에 고정화된다.
- [0105] 시험 샘플 중 하나 이상의 바이오마커는 검출가능한 모이어티로 표지될 수 있다.
- [0106] "검출가능한 모이어티(detectable moiety)"는 모이어티가 검출될 수 있는 것이고, 모이어티의 상대적 양 및/또는 위치(예를 들어, 어레이 상의 위치)가 결정될 수 있다는 의미를 포함한다.
- [0107] 적절한 검출가능한 모이어티는 해당 기술분야에 잘 알려져 있다.
- [0108] 따라서, 검출가능한 모이어티는 특정 조건에 노출될 경우, 검출될 수 있는, 형광(fluorescent) 및/또는 발광(luminescent) 및/또는 화학발광(chemiluminescent) 모이어티일 수 있다. 예를 들어, 형광 모이어티는 형광 모이어티의 여기(excitation)를 야기하는 특정 파장 및 강도에서 방사선(radiation)(즉, 빛)에 노출될 필요가 있을 수 있고, 이에 의해 검출될 수 있는 특정 파장(wavelength)에서 검출가능한 형광을 발산하게 할 수 있다.
- [0109] 대안적으로, 검출가능한 모이어티는 (바람직하게는 검출될 수 없는) 기질(substrate)을 시각화 및/또는 검출될 수 있는 검출가능한 생성물로 전환할 수 있는 효소일 수 있다. 적절한 효소의 예는 예를 들어, ELISA 분석과 비교하여 하기에 상술된다.
- [0110] 대안적으로, 검출가능한 모이어티는 이미징(imaging)에 유용한 방사성 원자(radioactive atom)일 수 있다. 적절한 방사성 원자는 신티그래픽 (scintigraphic) 연구를 위해 ^{99m}Tc 및 ¹²³I 를 포함한다. 기타 용이하게 검출가능한 모이어티는 ¹²³I, ¹³¹I, ¹¹¹In, ¹⁹F, ¹³C, ¹⁵N, ¹⁷O, 가돌리늄(gadolinium), 망간(manganese) 또는 철(iron)과 같은 자기 공명 이미징(magnetic resonance imaging, MRI)을 위한 스핀 라벨 спин label)을 포함한다. 명확하게, 검출될 물질(agent)(예를 들어, 본 명세서에 기재된 시험 샘플 및/또는 대조군 샘플에서의 하나 이상의 바이오마커 및/또는 선택된 단백질을 검출하는데 사용하기 위한 항체 분자와 같은 물질)은 검출가능한 모이어티가 용이하게 검출가능하도록 적절한 동위 원자(atomic isotope)를 충분히 가져야만 한다.
- [0111] 방사성 표지 또는 다른 표지는 알려진 방법으로 본 발명의 물질(즉, 본 발명의 방법의 샘플에 존재하는 단백질 및/또는 본 발명의 결합체)로 혼입(incorporate)될 수 있다. 예를 들어, 만일 결합 모이어티가 폴리펩티드라면,

이는 생합성되거나, 또는 예를 들어, 수소를 대신하여 불소-19(fluorine-19)를 수반하는 적절한 아미노산 전구체(precursor)를 사용하여 화학적 아미노산 합성에 의해 합성될 수 있다. 99m Tc, 123 I, 186 Rh, 188 Rh 및 111 In과 같은 표지는 예를 들어, 결합 모이어티에서 시스테인 잔기를 통해 부착될 수 있다. 이트륨-90(Yttrium-90)은 리신 잔기를 통해 부착될 수 있다. IODOGEN 방법(Fraker *et al* (1978) *Biochem. Biophys. Res. Comm.* **80**, 49-57)은 123 I를 혼입시키기 위해 사용될 수 있다. 참조 문헌("Monoclonal Antibodies in Immunoscintigraphy", J-F Chatal, CRC Press, 1989)은 상세하게 다른 방법을 기재한다. 다른 검출가능한 모이어티(예를 들면, 효소, 형광, 발광, 화학발광 또는 방사성 모이어티)를 단백질에 접합(conjugate)하는 방법은 해당 기술분야에 잘 알려져 있다.

[0112] 바람직하게는, 대조군 샘플(들)에서 하나 이상의 바이오마커는 검출가능한 모이어티로 표지된다. 검출가능한 모이어티는 다음으로 이루어진 군으로부터 선택될 수 있다: 형광 모이어티; 발광 모이어티; 화학발광 모이어티; 방사성 모이어티; 효소 모이어티. 그러나, 검출가능한 모이어티는 비오틴(biotin)인 것이 바람직하다.

[0113] 추가 구현예에서, 단계 (b), (d) 및/또는 단계 (f)는 하나 이상의 바이오마커와 결합할 수 있는 제2 결합제를 포함하는 분석을 이용하여 수행되고, 제2 결합제는 검출가능한 모이어티를 포함한다. 바람직하게는, 제2 결합제는 항체 또는 그의 항원-결합 단편을 포함하거나 이로 이루어진다. 바람직하게는, 항체 또는 그의 항원-결합 단편은 재조합 항체 또는 그의 항원-결합 단편이다. 가장 바람직하게는, 항체 또는 그의 항원-결합 단편은 scFv, Fab 및 면역글로불린 분자의 결합 도메인으로 이루어진 군으로부터 선택된다. 일 구현예에서, 검출가능한 모이어티는 다음으로 이루어진 군으로부터 선택된다: 형광 모이어티; 발광 모이어티; 화학발광 모이어티; 방사성 모이어티 및 효소 모이어티. 바람직하게는, 검출가능한 모이어티는 형광 모이어티(예를 들어, Alexa Fluor 염료, 예를 들어 Alexa647).

[0114] 일 구현예에서, 본 발명의 제1 형태의 방법은 ELISA (Enzyme Linked Immunosorbent Assay)를 포함하거나 이로 이루어진다.

[0115] 혈청 또는 혈장 단백질을 검출하기 위한 바람직한 분석은 모노클론 및/또는 폴리클론 항체를 이용한 샌드위치 분석법(sandwich assay)을 포함한, 면역효소 분석법(immunoenzymatic assay, IEMA), 효소 결합 면역흡착 분석법(ELISA), 방사면역측정법(radioimmunoassay, RIA) 및 면역방사측정 분석법(immunoradiometric assay, IRMA)을 포함한다. 예시적 샌드위치 분석법은 David 등의 미국특허 제4,376,110호 및 제4,486,530호에 기재되어 있고, 이들은 이에 의해 참조로서 포함된다. 슬라이드 상의 세포의 항체 염색(staining)은 통상의 기술자에게 잘 알려진 것처럼, 세포학 실험(cytology laboratory) 진단 시험에서 잘 알려진 방법에 사용될 수 있다.

[0116] 일반적으로, 상기 분석법은 고체상(solid phase) 분석에서, 유색 반응 생성물(coloured reaction product)을 제공하는 효소의 사용을 일반적으로 수반하는 ELISA(Enzyme Linked Immunosorbent Assay)이다. 호스라디시 퍼옥시다제(horseradish peroxidase) 및 포스파타제(phosphatase)와 같은 효소가 널리 채용되고 있다. 포스파타제 반응을 증폭하는 방법은 제2 효소 시스템을 위한 조효소로서 현재 역할을 하는 NAD를 생성하는 기질로서 NADP를 사용하는 것이다. 대장균의 피로포스파타제(pyrophosphatase)는 그 효소가 조직에 존재하지 않기 때문에 우수한 접합체를 제공하고, 안정하며, 우수한 반응색(reaction colour)을 제공한다. 루시퍼라제(luciferase)와 같은 효소에 기초한 화학발광 시스템이 또한 사용될 수 있다.

[0117] ELISA 방법은 해당 기술분야에서 잘 알려져 있고, 예를 들어 The ELISA Guidebook (Methods in Molecular Biology), 2000, Crowther, Humana Press, ISBN-13: 978-0896037281(그의 개시는 참조로서 포함된다)를 참조한다.

[0118] 비타민 비오틴과의 접합은 비오틴이 높은 특이성 및 친화성을 갖고 결합하는 효소-결합 아비딘(avidin) 또는 스트렙타비딘(streptavidin)과의 반응에 의해 용이하게 검출될 수 있기 때문에, 자주 사용된다.

[0119] 그러나, 단계 (b), (d) 및/또는 단계 (f)는 대안적으로 어레이를 사용하여 수행된다. 어레이 자체는 해당 기술 분야에 잘 알려져 있다. 일반적으로 각각은 고체 지지체의 표면상에 형성된 유한 영역(finite area)을 갖는 간격을 두고 떨어진(즉, 별개의) 영역("스팟(spot)")을 갖는 선형 또는 2차원 구조로 형성된다. 어레이는 각 비드(bead)가 분자 코드 또는 컬러 코드에 의해 확인되거나 연속 흐름(continuous flow)에서 확인될 수 있는 것인 비드 구조(bead structure)일 수 있다. 각각이 용액으로부터 분자의 클래스(class of molecule)를 흡착하는 일련의 스팟 위로 샘플을 통과시키는 것인 분석이 순차적으로 수행될 수도 있다. 고체 지지체는 일반적으로 유리(glass) 또는 폴리미이고, 가장 흔하게 사용되는 폴리미는 셀룰로스(cellulose), 폴리아크릴아미드(polyacrylamide), 나일론(nylon), 폴리스티렌(polystyrene), 염화 폴리비닐(polyvinyl chloride) 또는 폴리프

로필렌(polypropylene)이다. 고체 지지체는 튜브, 비드, 디스크, 실리콘 칩, 마이크로플레이트, 폴리비닐리텐 디플루오라이드(polyvinylidene difluoride, PVDF) 멤브레인, 니트로셀룰로스 멤브레인, 나일론 멤브레인, 기타 다공성 멤브레인, 비다공성 멤브레인(예를 들어, 특히, 플라스틱, 폴리머, 퍼스펙스(perspex), 실리콘), 복수의 폴리머 핀(polymeric pin), 또는 복수의 마이크로타이터 웰(microtitre well), 또는 단백질, 폴리뉴클레오티드 (polynucleotide) 및 다른 적합한 분자를 고정화시키고 및/또는 면역분석법을 실행하기에 적합한 다른 표면의 형태일 수 있다. 결합 과정은 당해 기술분야에 잘 알려져 있고 일반적으로 고체 지지체에 단백질 분자, 폴리뉴클레오티드 등을 공유적으로 결합 또는 물리적으로 흡착(adsorbing)시키는 가교(cross-linking)로 이루어진다. 접촉 또는 비접촉 프린팅, 매스킹(masking) 또는 포토리소그래피와 같은, 잘 알려진 기술을 이용하여, 각 스팟의 위치가 정의될 수 있다. 검토를 위하여, Jenkins, R.E., Pennington, S.R. (2001, *Proteomics*, 2, 13-29) 및 Lal 등 (2002, *Drug Discov Today* 15;7(18 Suppl):S143-9)를 참조한다.

[0120] 일반적으로, 어레이는 마이크로어레이이다. "마이크로어레이"는 적어도 약 $100/\text{cm}^2$ 및 바람직하게는 적어도 약 $1000/\text{cm}^2$ 의 별개의 영역(discrete region)의 밀도를 갖는 영역의 어레이의 의미를 포함한다. 마이크로어레이의 영역은 전형적인 크기(typical dimension), 예를 들어, 약 $10\text{--}250 \mu\text{m}$ 사이의 범위의 직경을 갖고, 대략 동일 거리에 의해 어레이의 다른 영역과 분리된다. 어레이는 또한 매크로어레이(macroarray) 또는 나노어레이(nanoarray)일 수 있다.

[0121] 일단 적절한 결합 분자(상기에서 논의됨)가 확인되고 단리되면 통상의 기술자는 분자생물학 분야에서 잘 알려진 방법을 이용하여 어레이를 제조할 수 있다.

[0122] 따라서, 상기 어레이는 비드-기반 어레이 또는 표면-기반 어레이일 수 있다. 바람직하게는 상기 어레이는 다음으로부터 이루어진 군으로부터 선택된다: 매크로어레이, 마이크로어레이 및 나노어레이.

[0123] 일 구현예에서, 본 발명의 일 양태에 따른 방법은:

[0124] (i) 샘플에 존재하는 바이오마커를 비오틴으로 표지하는 단계;

[0125] (ii) 비오틴-표지된 단백질을 그의 표면에 별개의 위치(discrete location)에 고정화된 복수의 scFv를 포함하는 어레이와 접촉시키는 단계로서, scFv는 표 III의 하나 이상의 단백질에 특이성을 갖는 것인 단계;

[0126] (iii) 고정화된 scFv를 형광 염료(dye)를 포함하는 스트렙타비딘 컨쥬게이트와 접촉시키는 단계; 및

[0127] (iv) 어레이 표면 상의 별개의 위치에서 염료의 존재를 검출하는 단계를 포함하고,

[0128] 어레이 표면 상의 염료의 발현은 샘플 중 표 III의 바이오마커의 발현을 나타낸다.

[0129] 대안적인 구현예에서, 단계 (b), (d) 및/또는 (f)는 하나 이상의 바이오마커를 코딩하는 핵산 분자의 발현을 측정하는 단계를 포함한다. 바람직하게는 핵산 분자는 cDNA 분자 또는 mRNA 분자이다. 가장 바람직하게는 핵산 분자는 mRNA 분자이다.

[0130] 따라서, 단계 (b), (d) 및/또는 (f)에서 하나 이상의 바이오마커의 발현은 서던 혼성화(Southern hybridisation), 노던 혼성화(Northern hybridisation), 종합 효소 연쇄 반응(polymerase chain reaction, PCR), 역전사 효소(reverse transcriptase) PCR (RT-PCR), 정량적 실시간(quantitative real-time) PCR (qRT-PCR), 나노어레이(nanoarray), 마이크로어레이(microarray), 매크로어레이(macroarray), 자기방사선술(autoradiography) 및 인 시투 혼성화(*in situ* hybridisation)로 이루어진 군으로부터 선택되는 방법을 이용하여 수행될 수 있다. 바람직하게는, 단계 (b)에서 하나 이상의 바이오마커(들)의 발현은 DNA 마이크로어레이를 사용하여 결정된다.

[0131] 일 구현예에서, 단계 (b), (d) 및/또는 (f)에서 하나 이상의 바이오마커(들)의 발현의 측정은 각각이 개별적으로 표 III에서 확인된 바이오마커 중 하나를 코딩하는 핵산 분자와 선택적으로 결합할 수 있는, 하나 이상의 결합 모이어티를 사용하여 수행된다.

[0132] 추가 구현예에서, 하나 이상의 결합 모이어티 각각은 핵산 분자를 포함하거나 이로 이루어진다. 따라서, 하나 이상의 결합 모이어티는 각각 DNA, RNA, PNA, LNA, GNA, TNA 또는 PMO를 포함하거나 이로 이루어질 수 있다. 그러나, 하나 이상의 결합 모이어티 각각은 DNA를 포함하거나 이로 이루어지는 것이 바람직하다.

[0133] 바람직하게는, 하나 이상의 결합 모이어티는 길이가 5 내지 100 개의 뉴클레오티드이다. 더욱 바람직하게는, 하나 이상의 핵산 분자는 길이가 15 내지 35 개의 뉴클레이오티드이다. 더욱 더 바람직하게는, 결합 모이어티는

검출가능한 모이어티를 포함한다.

[0134] 추가 구현예에서, 검출가능한 모이어티는 다음으로 이루어진 군으로부터 선택된다: 형광 모이어티; 발광 모이어티; 화학발광 모이어티; 방사성 모이어티(예를 들어, 방사성 원자); 및 효소 모이어티. 바람직하게는, 검출가능한 모이어티는 방사성 원자를 포함하거나 이로 이루어진다. 방사성 원자는 테크네튬(technetium)-99m, 요오드(iodine)-123, 요오드-125, 요오드-131, 인듐(indium)-111, 불소-19, 탄소-13, 질소-15, 산소-17, 인(phosphorus)-32, 황-35, 중수소(deuterium), 삼중수소(tritium), 레늄(rhenium)-186, 레늄-188 및 이트륨(yttrium)-90으로 이루어진 군으로부터 선택될 수 있다.

[0135] 그러나, 결합 모이어티의 검출가능한 모이어티는 형광 모이어티(예를 들어 Alexa Fluor 염료, 예를 들어 Alexa647)일 수 있다.

[0136] 일 구현예에서, 단계 (b), (d) 및/또는 (f)에서 제공되는 샘플은 미분획된 혈액(unfractionated blood), 혈장(plasma), 혈청(serum), 조직액(tissue fluid), 췌장 조직(pancreatic tissue), 췌액(pancreatic juice), 담즙(bile) 및 소변(urine)으로 이루어진 군으로부터 선택된다. 바람직하게는, 단계 (b), (d) 및/또는 (f)에서 제공되는 샘플은 미분획된 혈액, 혈장 및 혈청으로 이루어진 군으로부터 선택된다. 더욱 바람직하게는 단계 (b), (d) 및/또는 (f)에서 제공되는 샘플은 혈장이다. 다른 바람직한 구현예에서, 단계 (b), (d) 및/또는 (f)에서 제공되는 샘플은 혈청이다.

[0137] 본 발명의 제2 양태는 본 발명의 제1 양태에 정의된 바와 같이, 하나 이상의 결합제를 포함하는, 개체에서의 췌장암의 존재를 결정하기 위한 어레이를 제공한다.

[0138] 본 발명의 방법에 사용되기에 적합한 어레이는 상기에 논의된다.

[0139] 바람직하게는 하나 이상의 결합제는 표 III에 정의된 모든 단백질과 결합할 수 있다.

[0140] 본 발명의 제3 양태는 개체에서의 췌장암의 존재를 결정하기 위한 진단 마커(diagnostic marker)로서 본 발명의 제1 양태에 정의된 군으로부터 선택된 하나 이상의 바이오마커의 용도를 제공한다. 바람직하게는, 표 III에 정의된 모든 단백질이 개체에서의 췌장암의 존재를 결정하기 위한 진단 마커로서 사용된다.

[0141] 본 발명의 제4 양태는 다음을 포함하는, 췌장암의 존재를 결정하기 위한 키트를 제공한다:

[0142] A) 본 발명의 제1 양태에 따른 하나 이상의 제1 결합제 또는 본 발명의 제2 양태에 따른 어레이; 및

[0143] B) 본 발명의 제1 양태에 따른 방법을 수행하기 위한 설명서(instructions).

[0144] 바람직하게는, 본 발명의 특정 양태를 채용하는 비제한적 예시가 하기 표 및 도면을 참조로 기재될 것이다:

도면의 간단한 설명

도면 1: PaC vs. N의 분류

A) PaC 혈청과 혼성화(hybridized)된 항체 마이크로어레이의 스캔된 이미지. 전체로서, 위치 마커(position marker) 및 대조군을 포함한, 160개의 프로브가 슬라이드 당 8개의 20x8 서브어레이로 프린트되었다. B) PaC vs. N에 대해 상이하게 발현된($p<0.05$) 혈청 분석물(analyte). C) 모든 항체에 기초한, 즉 필터되지 않은 데이터를 사용한, PaC vs. N에 대한 ROC 커브. D) 모든 항체에 기초한 SVM 예측값을 이용한, PaC vs. N의 분류(적색 점-PaC, 청색 점-N). 상위 20개의 상이하게 발현된($p<0.02$) 비잉여적(non redundant) 분석물의 상대적 발현 레벨이 열지도(heatmap)에 나타난다. 적색-상향 조절, 녹색-하향 조절, 검은색-동일 레벨. (E) 278개의 인간 단백질 어레이를 사용한, 매우 차등적으로 발현된($p=0.005$) 분석물에 대해 예시된, scFv 항체 특이성의 검증(Validation). F) 278개의 인간 단백질 어레이를 사용한, 적당하게 차등적으로(modestly differentially) 발현된($p=0.04$) 분석물에 대해 예시된, scFv 항체 특이성의 확인.

도 2: PaC vs. N 분류에 대한 바이오마커 시그니처의 예비 검증(Pre-validation)

(A) 후진 제거 전략(backward elimination strategy)과 조합된 LOO 방법을 이용한 제1 환자 집단(cohort)에서 PaC vs. N 분류를 위한 바이오마커 시그니처의 압축(Condensation). 관찰된 ROC AUC 값은 항체의 남은 수에 대하여 플롯팅(plot)되었다. (C) 제1 환자 집단으로부터 얻어진 압축된 18-분석물 비잉여적(non-redundant) 바이오마커 시그니처. (D) 제1 환자 집단은 훈련 세트로 이용되었고, 그 다음에 산출 분류자(output classifier)가 새로운, 독립적인 환자군인, 제2 환자 집단에 대하여 시험되었다. (E) 시험 세트에 대한 분류자에 대해 얻어진

ROC 커브에 의해 나타난 PaC vs. N 분류에 대한 바이오마커 시그니처의 예비 검증.

도 3: PaC와 체장염을 구별하는 후보 혈청 바이오마커 시그니처

(A) PaC vs. ChP, AIP 또는 ChP+AIP+N 각각에 대한 차등적으로 발현된($p<0.05$) 혈청 분석물. (B) 모든 항체에 기초한, 즉 필터되지 않은 데이터의 사용한 PaC vs ChP, AIP, 또는 ChP+AIP+N에 대한 ROC 커브. (D) 10-plex 사이토카인(cytokine) 샌드위치 항체 마이크로어레이(MSD)를 사용하여 선택된 분석물의 항체 마이크로어레이 데이터의 검증. 데이터는 단지 분석물, IL-8에 대해서만 표시되고, 이에 대해 관찰된 신호의 대부분이 MSD 분석에 대한 검출의 하한 한계(lower limit)보다 위에 있었다.

도 4: PaC 진단을 위한 후보 혈청 바이오마커 시그니처의 예비 검증

(A) PaC, N, ChP 및 AIP로 구성된 제1 환자 집단을 훈련 세트(3분의 2) 및 시험 세트(3분의 1)로 나누었다. (B) 후진 제거 전략을 사용한 훈련 세트에 대해 얻어진 압축된 25개의 비잉여적 혈청 바이오마커 시그니처. (C) 시험 세트에 대한 분류자에 대해 얻어진 ROC 커브에 의해 예시된, PaC 진단을 위한 압축된 25-분석물 바이오마커 시그니처의 예비 검증. (D) i) 1000개의 랜덤 25-마커 시그니처(백색 원(open circle)), ii) 가장 낮은 p 값(파선(dashed lines)), 또는 iii) 가장 높은 배수변화(fold-change)(점선(dotted lines)) 중 하나에 의해 얻어진 25-분석물 시그니처의 성과 대비, 후진 제거 전략에 의해 얻어진 압축된 바이오마커 시그니처의 ROC AUC 값으로 표현된, 성과(performance). (E) 샘플 주석(annotation)이 옳거나(correct)(실선(solid line)) 또는 1000 번 반복된 경우(백색 원), 시험 세트에 대한 압축된 25-분석물 바이오마커 시그니처에 대해 얻어진 ROC AUC의 비교.

도 5: 항체 마이크로어레이 전략의 방식적 유팽

도 6: (A) 체장암과 (B) 정상, 만성 체장염, 및/또는 급성 염증성 체장염 간의 구별을 위한 ROC-AUC 값

ROC-AUC 값은 표 IV(A)(즉, 핵심) 및 표(B)(즉, 선호되는) 마커 모두와 증가되는 갯수의 표(C)(즉, 선택적) 마커를 갖는 마커 시그니처에 대해 도시된다. 가장 우수한 ROC AUC 값(0.90)은 29 분석물 시그니처, 즉 핵심 마커 + 선호되는 마커 + 15개의 선택적 마커에 대해 얻어진다. 그러나 모든 마커 조합은 상당한 예측력을 갖는다.

발명을 실시하기 위한 구체적인 내용

[0146]

재료 및 방법

[0147]

혈청 샘플

[0148]

혈청 샘플을 진단시, 즉 치료를 시작하기 전에, 두 개의 독립 환자 집단으로부터 수집하고 -80°C 에서 저장하였다. 제1 집단에서, 체장암 (PaC) (n=34), 만성 체장염 (ChP) (n=16), 자가면역성 체장염(autoimmune pancreatitis)(AIP)(n=23), 또는 건강한 개체 (N) (n=30) (임상적 증상 없음)로 진단받은 103 명의 환자로부터 얻은 혈청 샘플을 스크리닝하였다. 환자 인구통계학(demographics)을 표 1에 기술하였다. 이 집단을 무작위적으로 나누고 훈련 세트 및 시험 세트로 이용하였다. PaC (n=25) 또는 N (n=20) (환자 인구통계학을 위해, [14] 참조)로 진단받은 45 명의 환자로 구성된, 제2 집단을 최근 수집된 항체 마이크로어레이 데이터를 이용하여, 독립 시험 세트로만 이용하였다[14]. 샘플 집단의 크기는 진단시에 수집된 잘 규명된(well-characterized) 혈청 샘플의 이용가능성에 의해 제한되었다.

[0149]

항체 마이크로어레이 분석

[0150]

재조합 항체 마이크로어레이 분석을 자체 최적화 프로토콜(in-house optimized protocol)을 이용하여 수행하였다[12-15](하기 참조). 간단히, 57개의 주요 면역조절성 분석물을 표적화하는, 121개의 인간 재조합 단쇄 Fv (scFv) 항체를 프로브로 사용하였다. 파지-디스플레이 유래 scFv의 특이성, 친화성(nM 범위), 및 온칩(on-chip) 기능성[16]은 i) 엄격한 선택 프로토콜(stringent selection protocol)[16], ii) 타겟 분석물당 다중 클론(≤ 4) 및 iii) 분자 설계에 의해 개조된 scFv 라이브러리 마이크로어레이(REF)를 이용하여 보장되었다. 평면 항체 마이크로어레이(planar antibody microarray)(어레이 크기; 160×8 , $<0.5 \text{ cm}^2$)를 비접촉 디스펜서(dispenser)를 사용하여 항체 및 대조군을 하나씩(330 pL/drop) 분배(dispensing)하여 제조하였다. 비오티닐화(biotinylated) 혈청 샘플을 별개로 스크리닝하고 특이적으로 결합된(bound) 분석물을 콘포칼 형광 스캐너(confocal fluorescence scanner)를 사용하여, 형광 표지된 스트렙타비딘을 첨가하여 시각화하였다. 각각의 개별적인 어레

이 데이터 포인트는 4개의 재현률(replicate)의 평균 값을 나타낸다. 칩 대 칩 정규화(Chip-to-chip normalization)를 반전역(semi-global) 정규화 접근을 이용하여 수행하였다. 이전 연구에 따라[12, 15, 17], 스팟 대 스팟 재현성(spot to spot reproducibility) 및 어레이 대 어레이 재현성에 대한 상관 계수(correlation coefficient)는 각각 0.99 및 0.94였다. 선택된 항체 특이성 및 마이크로어레이 데이터를 234개 인간 단백질 어레이(human protein array) 및 10-plex 사이토카인 샌드위치 항체 어레이를 각각 사용하여 검증(validate)하였다(표II). 또한, 여러 개의 항체 특이성은 ELISA, 단백질 어레이, 블로킹/스파이킹(blocking/spiking) 시험 및/또는 질량분석기(mass spectrometry)를 사용하여 이전에 검증하였다(표 II).

[0151] 마이크로어레이 데이터 분석

데이터 분석을 R에서 수행하였다(하기 참조). 간단히, 제한조건의 값(cost of constraint)을 1로 설정한 선형 커널(linear kernel)을 이용하여, 두 개의 정의된 그룹의 하나(예를 들어, 암 vs. 건강한 개체)에 속하는 샘플을 분류하는데 서포트 벡터 머신(SVM)을 채용하였다. 오버피팅(over-fitting)의 위험을 피하기 위해 그것을 조정(tune)하려는 시도를 하지 않았다. L00(leave-one-out) 교차 검증 방법(cross validation procedure)을 이용하여 SVM을 훈련하고 시험하였다. 비교 중 두 개에서, 이 훈련 부분은 훈련 세트에서 가장 높은 판별력(discriminatory power)을 보이는 항체를 선택하는 것에 의한 항체 서브 패널(sub-panel)의 생성을 포함했다. 이러한 항체의 선택은 직접적 또는 교차 검증된 후진 제거 전략 중 하나를 이용하여 이루어졌다. 이러한 접근법을 사용하여, 압축된 후보 바이오마커 시그니처를 확인하고, 뒤이어 독립적인 시험 세트에 대해 평가하였다.

[0153] 민감도 및 특이성 값을 0의 역치 레벨(threshold level)을 이용하여, SVM 결정 값(decision value)으로부터 계산하였다. SVM 결정 값을 이용하여 ROC(receiver operating characteristic) 커브를 작성하였다. 곡선 하 면적(AUC)을 계산하고 예측 성능(prediction performance)의 척도(measure)로 이용하였다. 또한, Wilcoxon p-값 및 배수변화를 각 항체에 대해 계산하였다. 후보 바이오마커 시그니처를 종양 마커 예후 연구를 위한 권고(recommendation)에 따라 보고하였다[18].

[0154] 혈청 샘플

고지된 동의(informed consent) 후에, 혈청 샘플을 진단 시, 즉 치료의 시작 전에 두 개의 독립적인 환자 집단으로부터 수집하고 -80°C에 보관하였다. PC를 조직학(histology)으로 확인하였다. 환자 집단 1은 췌장 관세포암(PaC) (n=34), 만성 췌장염(hCP) (n=16), 자가 면역성 췌장염(AIP) (n=23), 또는 건강한 개체(대조군; N) (n=30) (임상적 증상 없음)로 진단 받은, 103명의 환자(Mannheim 대학 병원, 독일)로부터 얻어진 혈청 샘플로 구성하였다. 환자 인구통계학이 표 1에 기술되어 있다. 이 집단을 또한 무작위적으로 나누고 훈련 세트(샘플의 3분의 2) 및 시험 세트(3분의 1)로서 이용하였다. 환자 집단 2는 PaC (n=25), 또는 N (n=20) (Stockholm South 종합 병원 및 Lund 대학 병원, 스웨덴) (환자 인구통계학에 대해 [39] 참조)으로 진단된, 45 명의 환자로 구성하고 최근 기재된 바와 같이 항체 마이크로어레이 데이터를 사용하여 채택하였다[39]. 샘플 집단의 크기가 80% 보다 높은 통계적 검증력(statistical power)을 제공하기에 충분하다는 것을 확인하기 위해 검증력 분석(power analysis)(하기 참조)을 수행하였다. 주요 실험을 환자 집단 1에 대해 수행하고, 집단 2는 한 실험에서 검증을 위한 독립 데이터 세트로서 이용하였다(도 5 참조).

[0156] 혈청 샘플의 표지

혈청 샘플을 혈청 프로테옴에 대한 미리 최적화된 표지화 프로토콜(labelling protocol)을 이용하여 표지하였다[39-43]. 간단히, 미정제 혈청 샘플을 얼음 위에서 해동(thaw)하고, 30 μ L 분취액(aliquot)을 원심분리하였다(4°C에서 20분간 16 000 x g). 상청액(supernatant)의 5 μ L를 PBS로 45배 희석하여, 대략 2 mg/mL의 단백질 농도가 되게 하였다. 샘플을 매 20분마다 약하게 불택싱(vortexing)하면서 얼음 위에서 2 시간 동안 0.6 mM의 최종 농도에서 EZ-link® Sulfo-NHS-LC-Biotin (Pierce, Rockford, IL, USA)으로 표지하였다. 유리(free) 비오틴을 3.5 kDa MW 컷-오프(cut-off) 투석(dialysis) 유닛(Thermo Scientific, Rockford, IL, USA)을 사용하여 4°C에서 72 시간 동안 PBS에 대한 투석(dialysis)에 의해 제거하였다. 샘플을 분취하고 -20°C에서 저장하였다.

[0158] *scFv*의 생성 및 정제

전체적으로, 57개의 주요한 면역조절성 생물분자를 표적화하는, 121개의 인간 재조합 단체 항체 단편을 n-CoDeR 라이브러리 [43]로부터 선택하고 스웨덴 Lund의 BioInvent International AB에 의해 친절하게 제공되거나 Mats Ohlin 교수(Lund 대학, 스웨덴) (유신-1에 대한 5개의 클론)에 의해 제공되었다. 파지-디스플레이 유래 *scFv*의 특이성, 친화성(nM 범위) 및 온침 기능성을 i) 염색한 선택 프로토콜[43], ii) 표적 분자 당 다중 클론(≤ 4) 및 iii) 분자 설계에 의해 개조된 *scFv* 라이브러리 마이크로어레이를 사용함에 의해 보장(ensure)하였다. 항체 단편을 100 mL 대장균 배양에서 생산하고 Ni-NTA 아가로스(Qiagen, Hilden, 독일)에서 친화성 크로마토그래피(affinity chromatography)를 사용하여, 발현 상정액 또는 세포 주변세포질(cell periplasm)로부터 정제하였다. 결합된 분자(bound molecule)를 250 mM 이미다졸로 용출(elute)하고 PBS에 대해 광범위하게 (extensively) 투석하고, 마이크로어레이 제작(fabrication)에 사용될 때까지 4°C에서 저장하였다. 항체 농도를 280 nm에서 흡광도(absorbance)를 측정하는 것에 의해 결정하였다(평균 500 μ g/mL, 범위 50-1840 μ g/mL).

[0160] 항체 마이크로어레이의 제작 및 가공

평면 항체 마이크로어레이의 생산을 위하여, 미리 최적화되고 검증된 설정(set-up)을 사용하였다[39-43, 46]. 간단히, 피에조(piezo) 기술을 이용하여, 대략 330 pL 드롭(drop)을 침착(depositing)시키는 것에 의해, 비-접촉 프린터(BioChip Arrayer, PerkinElmer Life & Analytical Sciences, Wellesley, MA, USA)를 사용하여 *scFv*를 흑색(black) 폴리머 Maxisorb 마이크로어레이 슬라이드(NUNC, Roskilde, Denmark)상에 배열(array)하였다. 두 개의 드롭을 각 위치에서 스팟팅(spot)하여, 첫 번째 드롭이 두 번째 드롭이 분배되기 전에 건조되게 하였다. 평균, 5 fmol 항체(범위 1.5-25)를 위치마다 침적시켰다. 적절한 통계값을 보장하고 부분 결함(local defect)을 설명하기 위해, 각 프로브를 8개의 재현물에 프린팅하였다. 전체적으로, 위치 마커(position marker) 및 대조군 *scFv*를 포함한, 160개의 프로브를 슬라이드마다 프린트하고, 8개의 20 x 8 서브어레이로 배향시켰다. 정량(quantification) 동안 격자 정렬(grid alignment)을 보조하기 위해, 한 줄(row)의 Alexa647 접합(conjugated) 스트렙타비딘(Invitrogen, Carlsbad, CA, USA) (10 μ g/mL)을 선택된 위치에 스팟팅하였다. 어레이를 밤새 PBS 중 5%(w/v) 탈지 분유(fat-free milk powder)로 블로킹하였다.

마이크로어레이 슬라이드를 전술한 프로토콜에 따라 ProteinArray Workstation (PerkinElmer Life & Analytical Sciences)에서 가공하였다[42]. 간단히, 어레이를 PBS 중 0.5% (v/v) Tween-20(PBS-T)으로 60 μ l/min으로 4분간 세척하고 나서 PBS 중 1% (w/v) 탈지 분유 및 1% (v/v) Tween-20 중, 75 μ l 비오티닐화 혈청 샘플(1:2 희석하여, 1:90의 최종 혈청 희석이 됨)과, 매 15 초마다 교반(agitation)을 하면서 1시간 동안, 인큐베이션하였다. 그 후, 어레이를 다시 PBS-T로 세척하고, PBS_MT 중 1 μ g/mL Alexa-647 접합 스트렙타비딘과 1시간 동안 인큐베이션시켰다. 마지막으로, 어레이를 PBS-T로 세척하고, 질소 가스의 흐름 하에서 건조시키고, PMT 게인(gain) 및 레이저 파워의 4개의 상이한 스캐너 설정을 이용하여, 공초점(confocal) 마이크로어레이 스캐너(PerkinElmer Life & Analytical Sciences)로 10 μ m 해상도에서 스캔하였다. 각 스팟의 강도를 고정된 원(fixed circle) 방법을 이용하여, ScanArray Express 소프트웨어 v.4.0 (PerkinElmer Life & Analytical Sciences)에서 정량하였다. 국소 백그라운드(local background)는 차감(subtract)하였다. 가능한 국소 결함(local defect)을 보상하기 위하여, 두 개의 가장 높은 재현물 및 두 개의 가장 낮은 재현물을 자동적으로 제외시키고 남은 4개의 재현물의 평균 값을 이용하였다. 포화 신호(saturated signal)를 보이는 항체에 대해, 더 낮은 스캐너 설정으로부터의 값을 조정(scale)하고 대신 사용하였다. 칩 대 칩 정규화를 전술한 반전역 정규화 접근법을 이용하여 수행하였다[39, 40, 42]. 첫째, 모든 샘플에 대해 각 프로브별 CV를 계산하고 순위를 정했다(rank). 둘째, 모든 샘플에 걸쳐 가장 낮은 CV 값을 보이는 프로브의 15%를 확인하고 각 어레이에 대한 칩 대 칩 정규화를 계산하는데 사용하였다. 정규화 인자(normalization factor) N_i 를 식 $N_i = S_i / \mu$ 에 의해 계산하였고, 식 중에서 S_i 는 모든 샘플에 걸쳐 평균된, 사용된 항체에 대한 신호 강도의 합이고, μ 는 S_i 의 샘플 평균이다. 통계 분석 전에, 강도를 log2 값으로 재계산하였다.

[0163] 항체 특이성의 검증

두 개의 선택된 *scFv*(항-IL-6 (2) 및 항-IL-10 (1))의 특이성을 RayBio® 278 Human Protein Array G series (Norcross, GA, USA)를 제조사가 제공하는 프로토콜에 따라, 사용하여 시험하였다. *scFv*를 비오틴의 3.5 배 몰 초과(molar excess)로 얼음 위에서 2 시간 동안 EZ-link® Sulfo-NHS-LC-Biotin (Pierce)으로 표지하였다. 비

결합 비오틴을 PBS에 대해 72시간 투석으로 제거하였다. 전체적으로, 5 μg 의 항체를 각 어레이에 첨가하였다. 결합을 1 $\mu\text{g}/\text{mL}$ Alexa647 접합(conjugated) 스트렙타비딘 (Invitrogen)을 사용하여 검출하였다. 스트렙타비딘의 비특이성 결합을 체크하기 위해 음성 대조군으로서 하나의 어레이에 PBS를 첨가하였다. 어레이를 스캔하고 음성 대조군 어레이로부터의 신호를 차감하였다. 또한, 여러 개의 항체 특이성이 잘 규명되고, 표준화된 혈청 샘플 및 독립적인 방법, 예를 들어 질량 분석, ELISA, MSD, 및 CBA를 사용하고, 스파이킹(spiking) 및 블로킹(blocking) 실험을 사용하여 미리 검증하였다(표 II).

[0165] 어레이 데이터의 검증

[0166] 인간 Th1/Th2 10-plex MSD (Meso Scale Discovery, Gaithersburg, MD, USA) 분석을 항체 마이크로어레이 결과를 검증하는 시도로 수행하였다. MSD 96-플레이트의 각 웰(well)을 공간적으로 구별된 전극 스팟(electrode spot)에서, IFN- γ , IL-1 β , IL-2, IL-4, IL-5, IL-8, IL-10, IL-12p70, IL-13 및 TNF- α 에 대한 항체로 예비-기능화(pre-functionalize)하였다. 11개의 PaC, 11개의 건강한 개체 샘플, 9개의 ChP 및 3개의 AIP 샘플 (AIP의 낮은 수는 그 서브그룹에서의 제한된 샘플 부족때문이었다)을 포함한, 총 34개의 혈청 샘플(희석되지 않음)을 분석하였다. 제조사가 제공하는 프로토콜에 따라 분석을 수행하고, 전기화학발광-기반 (electrochemiluminescence-based) 판독(readout)을 MSD SECTOR® 기구에서 수행하였다.

[0167] 마이크로어레이 데이터 분석

[0168] 모든 통계값 및 데이터 분석을 R에서 수행하였다(<http://www.r-project.org>). 간단히, 제한 조건의 값을 1로 설정한 선형 커넬을 이용하여, 두 개의 정의된 그룹 중 하나(예를 들어, 암 또는 건강)에 속하는 것으로 샘플을 분류하기 위하여, 서포트 벡터 머신(SVM)을 채용하였다. 오버피팅(overfitting)의 위험을 피하기 위하여, 그것을 조정하려는 시도는 없었다. SVM을 훈련하고 LOO(leave-one-out) 교차 검증 방법을 이용하여 시험하였다 [42]. 두 개의 비교에서, 이 훈련 부분은 훈련 세트에서 가장 높은 판별력(discriminatory power)을 나타내는 항체를 선택하는 것에 의한 항체 서브 패널(sub-panel)의 생성을 포함한다. 이러한 항체의 선택은 직접적 또는 교차 검증된 후진 제거 전략을 이용하여 이루어졌다. 이러한 접근법을 사용하여, 압축된 후보 바이오마커 시그니처를 확인하고 뒤이어 독립 시험 세트에 대해 평가하였다.

[0169] 민감도 및 특이성 값을 0의 역치 레벨을 이용하여, SVM 결정 값(decision value)으로부터 계산하였다. ROC(receiver operating characteristic) 커브를 SVM 결정 값을 사용하여 작성하였다. 곡선 하 면적(AUC)을 계산하고 예측 성능의 척도로 이용하였다. 또한, Wilcoxon p-값 및 배수 변화를 각 항체에 대해 계산하였다. 후보 바이오마커 시그니처를 종양 마커 예후 연구를 위한 추천에 따라 보고하였다[47].

[0170] 바이오마커 시그니처 확인

[0171] 건강한 개체로부터 PaC을 구별하기 위한 바이오마커 시그니처를 확인하기 위하여 후진 제거 과정을 사용하였다. 이 접근법에서, 그 시간에 하나의 샘플을 데이터세트에서 제거하였다. 남은 샘플을 그 시간에 하나의 항체를 제거하고, 남은 항체를 사용하여 분류를 수행하는 것에 의해 SVM을 훈련하기 위해 사용하였다. 모든 항체가 한번에 배제되는 경우, 분류를 위해 가장 작은 Kullback-Leibler (KL) 에러를 얻을 때, 배제된 항체를 최소 정보적 항체(informative antibody)로 정의하고, 데이터 세트에서 제거하였다. LOO 과정을 오직 하나의 항체가 남고 항체를 제거하는 순서가 기록될 때까지 반복하였다. 새로운 샘플을 제외하는 것에 의해 과정을 반복하고, 모든 샘플이 일단 제외될 때까지 반복을 계속하였다. 항체를 제거하는 순서의 리스트를, 샘플을 제외하는 각 시점에 대해 생성하였다. 마지막에, 모든 샘플을 일단 제거하고, 컨센서스 리스트(consensus list)를 생성하였고, 여기서 각 항체가 수행된 모든 반복(iteration)에 걸쳐 평균된, 제거 프로세스를 얼마나 오래 견디었는지에 기초한 스코어를 각 항체에 부과하였다. 그 과정을 통해, 각각의 새로운 항체 서브패널 길이에 대해 수립된 SVM 모델을 시험하는데 각각의 제거된(left out) 샘플을 사용하여, 성능(performance)에 대응하는 결정값을 제공(return)하였다. 그 결과, 주어진 서브패널 길이에 대한 모든 샘플의 결정값을 수집하였다. 상응하는 ROC 면적을 데이터 세트의 강도 및 제거 전략을 평가하는 수단으로서 항체 숫자에 대하여 플롯팅하였다. 18개의 분석물의 압축된 시그니처를 컨센서스 리스트로부터 선택하고, 25개의 PaC 및 20개의 N 혈청 샘플 [39]의 항체 마이크로어레이 분석으로부터의 독립적인 데이터 세트를 후보 시그니처의 예비검증을 위한 시험 세트로 이용하였다. 시그니처 분석물을 시험 세트에서 SVM LOO 교차 검증 방법에서 사용하고, N으로부터 PaC를 구별하는 시그니처의 능력을

ROC 커브에 나타내었다.

[0172] 제2 후보 바이오마커 시그니처를 N, ChP 및 AIP 모두 종에서 PaC의 분류를 위해 생성하였다. 먼저, 데이터를 무작위적으로 훈련 세트(샘플의 3분의 2) 및 시험 세트(3분의 1)로 나누었다. 변형된 (훨씬 더 엄격한) 후진 제거 전략을 이용하였다. 그 시간에 하나의 샘플을 제거하는 대신, 훈련 세트에서 모든 샘플을 사용하여, SVM을 단지 한번 훈련하였다. 결과적으로, 하나의 제거 리스트를 생성하였고, 이로부터 25개의 분석물의 압축된 패널을 선택하고 훈련 세트에서 SVM을 수립하기 위해 이용하였다. 그 모델을 독립 시험 세트에 적용하고, ROC 커브를 생성하였다. 또한, 통계적 검증력 분석(stastical power analysis)을 수행하여, R에서 함수 "파워.t.시험(power.t.test)"을 이용하여, 시험 세트에서 요구되는 환자의 수를 추정하였다(Shapiro-Wilk 시험에 의해 제안되는 바와 같은 결정 값은 정규 분포를 따르는 것으로 가정됨). 훈련 세트에서 SVM 분석으로부터의 관찰된 결정 값은 2.87의 표준 편차(standard deviation) 및 3.47의 그룹 간의 델타 값(delta value)(평균 값 사이의 차이)을 나타내었다. 알파 레벨(유의 수준(level of significance))을 0.05로 설정하였다. 또한, 이 후진 제거 과정의 유효성을 각각 선택된 시그니처의 성능을 동일한 길이의 1000개의 무작위적으로 생성된 시그니처 및 가장 낮은 p-값 및 가장 높은 배수변화의 항체를 선택함에 의해 생성된 시그니처와 비교하는 것에 의해 시험하였다. 마지막으로, 분류자 및 데이터 세트의 강도를 시험 데이터 세트에서의 샘플 주석(annotation)의 1000개의 순열(permuation)을 생성함에 의해, 시험 데이터세트에서의 시그니처 성능을 랜덤 데이터와 비교하는 것에 의해 시험하였다.

결과

PaC vs. 건강한 대조군의 분류

[0175] PaC와 관련된 혈청 바이오마커 시그니처를 확인하기 위하여, 첫번째 환자 집단을 이용하여, PaC (n=34) vs. N (n=30)의 차등적인 혈청 단백질 발현 프로파일링(profiling)을 수행하였다. 역동적 신호 강도, 적절한 스팟 형태(morphology), 및 낮은 비특이적 백그라운드 결합을 얻음을 보여주는, 항체 마이크로어레이의 대표적인 이미지가 도 1a에 도시된다. 그 결과는 예를 들면, Th1 및 Th2 사이토카인 모두를 포함하는, 33개의 비잉여적(non-redundant) 단백질 분석물이 차등적으로 발현되고($p<0.05$), 그 중 보체 단백질 C1q 및 프로페딘을 제외한, 모두가 PaC에서 상향 조절되는 것으로 확인되었다는 것을 보여주었다(도 1b).

[0176] PaC와 N을 구별할 수 있는지 여부를 조사하기 위하여, 모든 항체에 기초하여, 즉 필터되지 않은 데이터를 이용하여, SVM L00 교차 검증을 수행하였다. 데이터는 환자 집단을 0.94의 ROC AUC 값으로 분류할 수 있음을 보여주었다(도 1c). 도 1d에서, 샘플을 SVM 결정 값을 감소시킴에 의해 플롯팅하고, 상위 20개의 차등적으로 발현되는 분석물($p<0.02$)의 상대적인 발현 패턴을 열 지도(heat map)에 보여준다. 0의 역치(디폴트값(default value))를 이용하여, 분석은 PaC vs. N이 각각 82% 및 87%의 민감도 및 특이성으로 분류될 수 있음을 보여주었다.

[0177] 다음으로, 278개의 인간 단백질 어레이를 선택된 항체 특이성의 검증을 위해 사용하였다(도 1e 및 1f). 이를 위해, 하나의 고도로 차등적으로 발현되는 분석물, IL-6 ($p=0.005$) 및 하나의 적당히(modest) 차등적으로 발현되는 분석물, IL-10 ($p=0.04$)에 대한 scFv 항체를 선택하였다. 두 경우 모두, 단백질 어레이 분석은 scFv 항체 단편이 그의 표적 단백질과 특이적으로 결합함을 보여주었다.

PaC vs. N 분류를 위한 압축된 바이오마커 시그니처의 예비 검증

[0179] 제1 환자 집단으로부터 유래한 분류의 강도(strength)를 시험하기 위하여, 먼저 L00 과정을 반복적 후진 제거 전략과 조합하는 것에 의해, 분석물의 총 수를 분류에 가장 많이 기여하는, 18개의 비잉여적 바이오마커로 압축하였다. 이 과정에서, Kullback-Leibler 발산(divergence) 에러를 최소화하고 하나씩 항체의 단계적 제거를 위한 가이드로서 이용하였다. 각 라운드(round) 후에, SVM 결정 값을 수집하고 대응하는 ROC 커브 및 AUC 값을 계산하였다. 도 2a에서, AUC 값을 낮은 항체의 갯수에 대하여 플롯팅하고, 이는 단지 수 개의 항체를 포함할 경우에도, 높고 안정한 분류를 나타낸다. 다양한 분석물, 예를 들어 사이토카인, 보체 단백질 및 효소로 구성된, 18-분석물 압축된 후보 혈청 바이오마커 시그니처를 도 2b에 나타낸다. 다음으로, 이 18-분석물 분류자를 새로운 독립 시험 그룹인, 제2 환자 집단(n=45)에 적용하였다(도 2c). 그 결과는 분류자가 88%의 민감도 및 85%의 특이성에 대응되는, 0.95의 ROC AUC 값으로, 환자의 PaC vs. N으로의 계층화(stratification)를 가능하게 한다는 것을 보여주었다(도 2d). 따라서, 데이터는 PaC 진단을 위한, 첫번째 예비 검증된 혈청 바이오마커 시그니처

를 묘사하였다(out line).

[0180] *PaC vs. 혜장염을 구별하는 바이오마커 시그니처*

암이 혜장에서 양성 상태(benign condition)로부터 구별될 수 있는지 여부를 시험하기 위하여, 제1 환자 집단을 이용하여, PaC의 혈청 단백질 발현 프로파일(n=34)을 ChP (n=16) 또는 AIP (n=23)의 것과 비교하였다. PaC vs. ChP의 경우에서, 15개의 비잉여적이고 차등적으로 발현된 ($p<0.05$) 혈청 분석물을 찾아냈고, 그 중 두 개(IL-4 및 IL-12)를 제외한 모두는 PaC에서 상향 조절되었다(도 3a). 필터되지 않은 데이터에 기초하여, 그 결과는 PaC 및 ChP가 97% 민감도 및 69% 특이성에 대응하는, 0.86의 ROC AUC 값으로 구별될 수 있음을 보여주었다(도 3b). 총 49개의 비잉여적 혈청 분석물이 PaC vs. AIP에서 상이하게 발현되는 것으로 나타났고, C1q 및 프로페딘을 제외한 모두가 PaC에서 상향조절된다(도 3a). 또한, 필터되지 않은 데이터에 기초하여, 그 결과는 PaC vs. AIP가 각각 97% 및 91%의 민감도 및 특이성에 기초하여, 0.99의 ROC AUC(도 3b)으로 분류될 수 있음을 보여주었다.

[0182] 임상적 현실(clinical reality)을 더욱 잘 반영하기 위하여, 제1 환자 집단(n=103)을 이용하여, PaC과 ChP+AIP+N의 조합된, 이질적(heterogeneous) 환자 그룹 간에 차이가 판독(decipher)될 수 있는지 여부를 조사하였다. 그 결과는 47개의 비잉여적 혈청 단백질이 상이하게 발현($p<0.05$)된다는 것을 보여주었다(도 3a). 대부분의 분석물(47개 중 45개)은 PaC에서 상향 조절되는 것으로 나타났고, 이들은 광범위한 단백질을 포함한다. 필터되지 않은 데이터에 기초하여, 그 결과는 PaC가 0.85의 ROC AUC 값으로, 이질적인 환자 그룹으로부터 구별될 수 있음을 보여주었다(도 3b).

[0183] 어레이 데이터를 검증하려는 시도에서, 독립적인 10-plex 사이토카인 샌드위치 항체 마이크로어레이(MSD)를 적용하였다(도 3c). 그러나, 10개의 표적화된 혈청 분석물 중 1개인, IL-8만이 대다수의 샘플에서 MSD 분석의 검출 하한보다 위에 있었다. 또한, PaC vs. N, ChP, AIP 및 그의 조합된 집단에서 관찰된 IL-8상향 조절은 PaC vs. AIP ($p=0.29$)를 제외한, 모든 케이스에서 MSD 분석에 의해, 통계적으로 확인되었다($p<0.05$).

[0184] *PaC 진단을 위한 정제된(refined) 바이오마커 시그니처*

[0185] PaC, N, ChP 및 AIP를 포함하는, 전체의 제1 환자 집단(n=103)의 분류의 강도를 시험하기 위하여, 집단을 훈련 세트(3분의 2) 및 시험 세트(3분의 1)로 나누었다(도 4a). 다음으로, 훈련 세트에서 분류에 가장 많이 기여하는 25개의 비잉여적 분석물로 구성된 압축된 혈청 바이오마커 시그니처를 직접적, 반복적 후진 제거 전략을 이용하여 판독하였다. 예를 들어 사이토카인 및 보체 단백질로 구성된, 25-분석물 압축 바이오마커 시그니처가 도 4b에 도시된다. 다음으로 독립적인 시험 세트 상에 이 25-분석물 분류자(classifier)를 적용하였다(도 4c). 그 데이터는 PaC가 민감도 및 특이성이 각각 73% 및 75%로 묘사되는, 0.88의 ROC AUC 값으로 찾아질 수 있음을 보여주었다.

[0186] 분류자를 더 검사(challenge)하기 위하여, 그의 판별력을 통계적으로 평가하였다. 첫째, 동일한 길이의 1000개의 무작위 시그니처(25개의 항체)를 훈련 세트에서 생성하고 시험 세트에 적용하였다. 그 결과는 무작위 시그니처에 대한 AUC 값이 케이스의 95%에서 분류자 바이오마커 시그니처(AUC=0.88)의 것보다 더 낮다는 것을 보여주었다(도 4d). 또한, 가장 낮은 p -값(AUC=0.77) 또는 가장 높은 배수 변화(AUC=0.78)에 기초하여 선택된 대응하는 25-분석물 시그니처에 대한 AUC 값은 분류자 시그니처의 것보다 유의성 있게 더 낮았다. 따라서, 그 데이터는 압축된, 고성능(high-performing) 시그니처를 정의하기 위한, 분류자의 판별력 및 후진 제거 전략의 적용성(applicability)을 나타내었다. 둘째, 특이적 분류를 샘플의 동일한 수의 무작위 분류와 비교하기 위해, 시험 세트의 샘플 주석을 1000회 교환(permute)하였다. 그 결과는 무작위 주석을 적용할 때보다 정확한 샘플 주석을 사용할 때 유의하게 더 높은 AUC 값(0.88 vs 0.19-0.86, 0.5의 중앙값)이 얻어짐을 보여주었고, 이는 또한 분류의 강도를 나타낸다.

[0187] 결론

[0188] 이 연구에서, PaC를 표적화하기 위해 면역계의 진단력을 활용(harness)하기 위해, 친화성 프로테오믹스(affinity proteomics)를 적용하였다. 이러한 접근은 면역계가 질병에서 기인하는, 개체의 건강 상태에서의 변화(alteration)에 정교하게(exquisitely) 민감하고, 이러한 변화를 특히, 면역조절 분석물의 레벨에서의 변동(fluctuation)을 통해 반영(register)한다는 개념에 근거했다. 이를 위해, 본 발명자들은 이러한 종류의 주요한

조절성 혈청 분석물을 우세하게(predominantly) 표적화하는 항체 마이크로어레이를 설계하였다. 그 데이터는 높은 진단력을 나타내는 PaC-연관 후보 바이오마커 시그니처가 해석(de-convolute)될 수 있음을 보여주었다. 유사한 방식으로, 이 친화성 프로테오믹스 접근법은 최근 다른 암 징후(indication)와 건강한 대조군을 구별하는 여러 개의 혈청형(serological) 바이오마커 시그니처의 확인을 가능하게 하였고[14-15, 17, 19], 이는 또한 플랫폼의 강도를 나타내었다.

[0189] 본 발명자들은 처음으로 높은 신뢰도로, 혈청이 PaC의 잘 규정된(well-defined) 환자 집단 vs. 대조군 및 PaC vs. 체장염뿐만 아니라 PaC vs. 대조군 조합 집단과 체장염 환자 간을 구별하게 하는 정보를 저장함을 보여주었다. 후보 바이오마커 시그니처가 이질적인 환자 그룹이 스크리닝될 임상적 환경(clinical setting)에서도 잘 기능해야 하기 때문에, 후자의 발견은 특히 중요하였다.

[0190] 검증된 혈청형 판별자(discriminator)가 아직 없기 때문에 고성능 PaC 분류자의 임상적 영향(clinical impact)은 끝 것이다[2, 7-9, 20-21]. 최상의(golden) 분류자가 확립(establish)되는 것을 기다리는 동안, CA-19-9이 PaC 진단을 위한 가장 유용한 분자 마커로 남는다[2, 8-10]. 특히, 본 데이터는 CA-19-9에 대해 일관되게 판찰된 것보다 PaC 진단에 대해 유의하게 더 높은 중위(median) 민감도(88%) 및 특이성(85%)을 보여주었고, 이는 유의성 있는 임상적 부가가치를 나타낸다. 또한, 본 발명자들은 최근 위험 환자(risk patient)에 대한 비용, 생존 및 삶의 질에 대한 새로운 진단적 가능성의 영향(impact)을 모델링하였고, 친화성 프로테오믹스가 PaC 스크리닝에서 비용-효율적(cost-effective) 도구가 될 것이라는 유망한 전망(prospect)을 가짐을 보여주었다 (Bolin *et al.*, ms in prep.).

[0191] 분류자는 종양이 아직 작고 수술가능한 경우, 조기 진단이 수행될 수 있다면, 최고로 작용할 것이다[2, 9].

[0192] 암 바이오마커에 대한 탐색(quest)에서, 전신성 염증(systemic inflammation)이 잠재적 교란(confounding) 인자로서 자주 강조되는데[23], 이는 암 발생(cancer development)과 염증이 연관되었기 때문이다. 친화성 프로테오믹스에 기초한 초기 작업에서, 그 결과는 또한 종종 암-특이적 지문(fingerprint)보다는 일반적 질병(염증성) 시그니처가 기술(delineate)됨을 보여주었다[24-26]. 특히, 본 발명자들은 여기서 PaC 및 체장염이 높은 신뢰도로 판별될 수 있음을 보여주었다. 또한, 판찰된 시그니처(들)는 유의한 차이, 즉 다른 다양한 염증성 질환 참조[19, 27] (Carlsson *et al.*, ms in prep.) 및 다른 암 [14-15, 17, 19]에 대해 판찰되는 것과, 단지 작은 오버랩을 보여주었고, 이는 PaC-특이적 시그니처가 판독되었다는 개념을 더 뒷받침하였다.

[0193] 혈청 면역 시그니처는 시험 시 환자에서 면역계의 활성의 스냅샷(snapshot)으로서 고려될 수 있다. 이러한 지문은 암에 대한 반응에서 직접 및 간접(전신적) 효과의 조합을 반영할 것이다. 사이토카인 발현 프로파일에 초점을 맞추면, 이전의 보고는 체장암 세포주가 예를 들어, IL-6, IL-8, IL-10, IL-12, IL-13, IL-18, 및 TGF- β 1을 포함한, 이 연구에서 또한 과발현되는 것으로 밝혀진 사이토카인의 세트를 발현함을 보여주었다[28]. 이들 및 다른 사이토카인(예를 들어, VEGF 및 IL-7)중 수개는 또한 PaC 종양 조직 및/또는 혈청/혈장에서 과발현되는 것으로 밝혀졌고[29-33], 이는 또한 본 판찰을 뒷받침한다. 사이토카인이 면역계에서 중추적 역할(pivotal role)을 하지만, 이러한 분석물이 종 다수는 다면발현적(pleiotropic) 기능을 나타내고 PaC는 특이한(peculiar) 사이토카인 발현 패턴을 특징으로 하기 때문에, 생물학적 맥락상(biological context) 이러한 복잡한 발현 패턴을 해석하는 것은 부담이 된다[29]. 예를 들어, IFN- γ 의 발현이, 시도된 항-종양 면역 반응을 시사(signal)할 수 있으나[29], 항염증 사이토카인(예를 들어, TGF- β 및 IL-10) 및 잠재적으로 불활성인 전염증성(proinflammatory) 사이토카인(예를 들어, IL-12 및 IL-18)의 수반되는(concomitant) 발현에서 보여지듯, PaC의 면역 환경은 종종 면역억제(immunosuppressive) 부위에 있는 것으로 밝혀졌다[29]. 세포 면역억제는 많은 환자에서 판찰되는 PaC의 두드러진 생물학적 특성이다[34]. Th2 편향(skewed) 반응이 보고된 반면, 여기에 나타난 Th1/Th2 밸런스(balance)가 또한 판찰되었다[29, 31, 35]. 사이토카인 발현 패턴은 또한 생존률과 같은, 다른 파라미터를 반영하는 것으로 밝혀졌다[14, 29]. 비-사이토카인 마커의 일부를 볼 때, 종양에 대한 면역 감시(surveillance)에서 기능하는 것으로 제안된, C3과 같은 여러 개의 보체 단백질[36-37] 및 탄수화물 항원 루이스 x가 또한 이전에 PaC와 관련이 있는 것으로 밝혀졌다[38].

[0194] 종합하면, 본 발명자들은 임상적 요구를 해소하고 면역시그니처링(immunosignaturing)이 PaC 진단을 위한 제1의 예비 검증된 혈청형의 바이오마커 시그니처를 해독하기 위한 강력한 접근법임을 보여주었다. 이는 고성능 플랫폼, 잘 조절된 샘플 및 엄격한 바이오인포메틱스(bioinformatic) 및 검증 접근법을 통해 달성되었다. 예측자 시그니처(predictor signature)의 잠재력은 후속 연구에서 더 검증될 것이고, 여기서 독립 샘플 집단이 프로파일링될 것이다. 결국, 이러한 발견은 향상된 PaC 진단 및 이에 의해 PaC의 강화된 예후 및 임상 관리(management)를 위한 신규한 기회를 제공할 것이다.

[0195] 참고문헌

- [0196] 1. Hidalgo, M., *Pancreatic cancer*. N Engl J Med, 2010. **362**(17): p. 1605-17.
- [0197] 2. Chu, D., W. Kohlmann, and D.G. Adler, *Identification and screening of individuals at increased risk for pancreatic cancer with emphasis on known environmental and genetic factors and hereditary syndromes*. JOP, 2010. **11**(3): p. 203-12.
- [0198] 3. Jemal A, S.R., Ward E, Hao Y, Xu J, Thun MJ., *Cancer statistics, 2009*. CA Cancer J Clin, 2009. **59**(4).
- [0199] 4. Pannala, R., et al., *New-onset diabetes: a potential clue to the early diagnosis of pancreatic cancer*. Lancet Oncol, 2009. **10**(1): p. 88-95.
- [0200] 5. Warshaw, A.L. and C. Fernandez-del Castillo, *Pancreatic carcinoma*. N Engl J Med, 1992. **326**(7): p. 455-65.
- [0201] 6. Galasso, D., A. Carnuccio, and A. Larghi, *Pancreatic cancer: diagnosis and endoscopic staging*. Eur Rev Med Pharmacol Sci, 2010. **14**(4): p. 375-85.
- [0202] 7. Chen, R., et al., *Proteomics studies of pancreatic cancer*. Proteomics Clin Appl, 2007. **1**(12): p. 1582-1591.
- [0203] 8. Duffy, M.J., et al., *Tumor markers in pancreatic cancer: a European Group on Tumor Markers (EGTM) status report*. Ann Oncol, 2010. **21**(3): p. 441-7.
- [0204] 9. Fry LC, M.K., Malfertheiner P., *Molecular markers of pancreatic cancer: development and clinical relevance*. Langenbecks Arch Surg., 2008. **393**(6).
- [0205] 10. Koopmann J, R.C., Zhang Z, Canto MI, Brown DA, Hunter M, Yeo C, Chan DW, Breit SN, Goggins M., *Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19-9*. Clin Cancer Res., 2006. **15**(12).
- [0206] 11. Boeck S, S.P., Holdenrieder S, Wilkowsky R, Heinemann V., *Prognostic and therapeutic significance of carbohydrate antigen 19-9 as tumor marker in patients with pancreatic cancer*. Oncology, 2006. **70**(4).
- [0207] 12. Ingvarsson J, L.A., Sjoholm AG, Truedsson L, Jansson B, Borrebaeck CA, Wingren C., *Design of recombinant antibody microarrays for serum protein profiling: targeting of complement proteins*. J Proteome Res, 2007. **6**(9).
- [0208] 13. Wingren C, I.J., Dexlin L, Szul D, Borrebaeck CA., *Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support*. Proteomics, 2007. **7**(17).
- [0209] 14. Ingvarsson J, W.C., Carlsson A, Ellmark P, Wahren B, Engstrom G, Harmenberg U, Krogh M, Peterson C, Borrebaeck CA., *Detection of pancreatic cancer using antibody microarray-based serum protein profiling*. Proteomics, 2008. **8**(11).
- [0210] 15. Carlsson, A., et al., *Plasma proteome profiling reveals biomarker patterns associated with prognosis and therapy selection in glioblastoma multiforme patients*. Proteomics Clinical Applications, 2010. **4**(6-7): p. 591-602.
- [0211] 16. Soderlind E, S.L., Jirholt P, Kobayashi N, Alexeiva V, Aberg AM, Nilsson A, Jansson B, Ohlin M, Wingren C, Danielsson L, Carlsson R, Borrebaeck CA., *Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries*. Nat Biotechnol., 2000. **18**(8).
- [0212] 17. Carlsson A, W.C., Ingvarsson J, Ellmark P, Baldertorp B, Ferno M, Olsson H, Borrebaeck CA., *Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays*. Eur J

Cancer, 2008. **44**(3).

- [0213] 18. McShane, L.M., et al., *REporting recommendations for tumor MARKer prognostic studies (REMARK)*. Nat Clin Pract Oncol, 2005. **2**(8): p. 416-22.
- [0214] 19. Ellmark P, I.J., Carlsson A, Lundin BS, Wingren C, Borrebaeck CA., *Identification of protein expression signatures associated with Helicobacter pylori infection and gastric adenocarcinoma using recombinant antibody microarrays*. Mol Cell Proteomics., 2006. **5**(9).
- [0215] 20. Garcea, G., et al., *Molecular prognostic markers in pancreatic cancer: a systematic review*. Eur J Cancer, 2005. **41**(15): p. 2213-36.
- [0216] 21. Rustgi, A.K., *Pancreatic cancer: novel approaches to diagnosis and therapy*. Gastroenterology, 2005. **129**(4): p. 1344-7.
- [0217] 22. Biankin, A.V., et al., *Molecular pathogenesis of precursor lesions of pancreatic ductal adenocarcinoma*. Pathology, 2003. **35**(1): p. 14-24.
- [0218] 23. Chechlinska, M., M. Kowalewska, and R. Nowak, *Systemic inflammation as a confounding factor in cancer biomarker discovery and validation*. Nat Rev Cancer, 2010. **10**(1): p. 2-3.
- [0219] 24. Orzechowski, R., et al., *Antibody microarray profiling reveals individual and combined serum proteins associated with pancreatic cancer*. Cancer Res, 2005. **65**(23): p. 11193-202.
- [0220] 25. Gao, W.M., et al., *Distinctive serum protein profiles involving abundant proteins in lung cancer patients based upon antibody microarray analysis*. BMC Cancer, 2005. **5**: p. 110.
- [0221] 26. Miller, J.C., et al., *Antibody microarray profiling of human prostate cancer sera: antibody screening and identification of potential biomarkers*. Proteomics, 2003. **3**(1): p. 56-63.
- [0222] 27. Dexlin-Mellby, L., et al., *Tissue proteomic profiling of preeclamptic placenta tissue using recombinant antibody microarrays*. Proteomics - Clinical Applications, 2010. **4**(10-11): p. 794-807.
- [0223] 28. Bellone, G., et al., *Cytokine expression profile in human pancreatic carcinoma cells and in surgical specimens: implications for survival*. Cancer Immunol Immunother, 2006. **55**(6): p. 684-98.
- [0224] 29. Bellone G, S.C., Mauri FA, Tonel E, Carbone A, Buffolino A, Dughera L, Robecchi A, Pirisi M, Emanuelli G., *Cytokine expression profile in human pancreatic carcinoma cells and in surgical specimens: implications for survival*. Cancer Immunol Immunother., 2006. **55**(6).
- [0225] 30. Chang ST, Z.J., Horecka J, Kunz PL, Ford JM, Fisher GA, Le QT, Chang DT, Ji H, Koong AC., *Identification of a biomarker panel using a multiplexed proximity ligation assay improves accuracy of pancreatic cancer diagnosis*. J Transl Med., 2009. **7**(105).
- [0226] 31. Poch B, L.E., Ramadani M, Gansauge S, Beger HG, Gansauge F., *Systemic immune dysfunction in pancreatic cancer patients*. Langenbecks Arch Surg., 2007. **392**(3).
- [0227] 32. Wigmore SJ, F.K., Sangster K, Maingay JP, Garden OJ, Ross JA., *Cytokine regulation of constitutive production of interleukin-8 and -6 by human pancreatic cancer cell lines and serum cytokine concentrations in patients with pancreatic cancer*. Int J Oncol., 2002. **21**(4).
- [0228] 33. Frick VO, R.C., Wagner M, Graeber S, Grimm H, Kopp B, Rau BM, Schilling MK., *Enhanced ENA-78 and IL-8 expression in patients with malignant pancreatic diseases*. Pancreatology., 2008. **8**(4-5).
- [0229] 34. Ungefroren, H., et al., *Immunological escape mechanisms in pancreatic carcinoma*. Ann N Y Acad Sci, 1999. **880**: p. 243-51.
- [0230] 35. Schmitz-Winnenthal FH, V.C., Z'graggen K, Galindo L, Nummer D, Ziouta Y, Bucur M, Weitz J, Schirrmacher V, Buchler MW, Beckhove P., *High frequencies of functional tumor-reactive T cells in bone marrow and blood of pancreatic cancer patients*. Cancer Res., 2005. **65**(21).
- [0231] 36. Chen R, P.S., Cooke K, Moyes KW, Bronner MP, Goodlett DR, Aebersold R, Brentnall TA.,

Comparison of pancreas juice proteins from cancer versus pancreatitis using quantitative proteomic analysis. *Pancreas*, 2007. **34**(1).

- [0232] 37. Yu KH, R.A., Blair IA., *Characterization of proteins in human pancreatic cancer serum using differential gel electrophoresis and tandem mass spectrometry*. *J Proteome Res.*, 2005. **4**(5).
- [0233] 38. Nakano M, N.T., Ito T, Kitada T, Hijioka T, Kasahara A, Tajiri M, Wada Y, Taniguchi N, Miyoshi E., *Site-specific analysis of N-glycans on haptoglobin in sera of patients with pancreatic cancer: a novel approach for the development of tumor markers*. *Int J Cancer*, 2008. **122**(10).
- [0234] 39. Ingvarsson J, Wingren C, Carlsson A, et al. *Detection of pancreatic cancer using antibody microarray-based serum protein profiling*. *Proteomics* 2008;8.
- [0235] 40. Carlsson A, Wingren C, Ingvarsson J, et al. *Serum proteome profiling of metastatic breast cancer using recombinant antibody microarrays*. *Eur J Cancer* 2008;44: 472-80.
- [0236] 41. Wingren C, Ingvarsson J, Dexlin L, Szul D, Borrebaeck CAK. *Design of recombinant antibody microarrays for complex proteome analysis: choice of sample labeling-tag and solid support*. *Proteomics* 2007;7:3055-65.
- [0237] 42. Carlsson A, Persson O, Ingvarsson J, et al. *Plasma proteome profiling reveals biomarker patterns associated with prognosis and therapy selection in glioblastoma multiforme patients*. *Proteomics Clin Appl* 2010;4:1-12.
- [0238] 43. Soderlind E, Strandberg L, Jirholt P, et al. *Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries*. *Nat Biotechnol* 2000;18:852-6.
- [0239] 44. Borrebaeck CAK, Wingren C. *Design of high-density antibody microarrays for disease proteomics: key technological issues*. *J Proteomics* 2009;72:928-35.
- [0240] 45. Borrebaeck CAK, Wingren C. *High-throughput proteomics using antibody microarrays: an update*. *Expert Rev Mol Diagn* 2007;7: 673-86.
- [0241] 46. Ellmark P; Ingvarsson J, Carlsson A, Lundin BS, Wingren C, Borrebaeck CAK. *Identification of protein expression signatures associated with Helicobacter pylori infection and gastric adenocarcinoma using recombinant antibody microarrays*. *Mol Cell Proteomics* 2006;5:1638-46.
- [0242] 47. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. *REporting recommendations for tumor MARKer prognostic studies (REMARK)*. *Nat Clin Pract Oncol* 2005;2:416-22.

[0243] [표 I] 제1 환자 집단의 환자 인구통계학

분류	환자의 수	성별 (M/F/미상)	나이	
			평균 (SD)	범위
PaC	34	18/12/4	65.0 (10.4)	42-93
N	30	15/15/0	33.2 (8.6)	24-53
ChP	16	12/4/0	48.8 (14.2)	32-73
AIP	23	11/11/1	42.4 (18.3)	14-74
총	103	56/42/5	48.2 (18.1)	14-93

[0244]

[0245] [표 II] 항체 마이크로어레이에 의해 분석된 혈청 바이오마커의 요약

항원 (클론의 수)	항원 (클론의 수)
안지오모틴 (2)	IL-2 (3)
	IL-3 (3)
β-갈락토시다제 (1)	IL-4 (4)**
브루톤(Bruton) 티로신 키나제	IL-5 (3)**
BTK (1)	IL-6 (4)**/**
C1 에스테라제 억제제 (1)	IL-7 (2)
C1q (1)**	IL-8 (3)**
C1s (1)	IL-9 (3)
C3 (2)**	인테그린 α10 (1)
C4 (1)**	인테그린 α11 (1)
C5 (2)**	렙틴 (1)
CD40 (4)	루이스 x (2)
CD40 리간드 (1)	루이스 y (1)
CT-17 (대조군) (1)	MCP-1 (3)**
디곡신(Digoxin) (대조군) (1)	MCP-3 (1)
에오틱신 (3)	MCP-4 (2)
인자 B (1)**	뮤신-1 (6)
GLP-1 (1)	프로카텝신 W (1)
GLP-1 R (1)	프로퍼딘 (1)**
GM-CSF (3)	PSA (1)
IFN-γ (2)	RANTES (2)
IgM (1)	시알릴 루이스 x (1)
IL-10 (3)*	TGF-β1 (3)
IL-11 (3)	TM 펩티드 (1)
IL-12 (4)**	TNF-α (2)
IL-13 (2)**	TNF-β (4)**
IL-16 (2)	티로신 단백질 키나제
IL-18 (3)	JAK3 (1)
IL-1α (3)**	VEGF (4)**
IL-1β (3)	
IL-1-ra (3)	

[0246]

* 단백질 어레이에 의해 결정된 항체 특이성

[0248]

** ELISA, 단백질 어레이, 블로킹/스파이킹 실험 및/또는 질량 분석에 의해 이전에 확인된 항체 특이성

[0249]

[표 III] 체장암 진단 바이오마커

바이오마커 명칭	예시적인 서열
인터루킨-7 (IL-7)	AK226000, AB102893, AB102885, P13232
인테그린 α-10	Hs158237; 075578
B-갈락토시다제	P16278
브루톤 의 티로신 키나제 (BTK)	Q06187
보체 단백질 C1q (C1q)	IPR001073, PR00007
보체 단백질 C1s (C1s)	P09871
B 세포 수용체 μ 사슬 (IgM)	예를 들어, P01871 (완전한 단백질 아님); Ramos B 세포 상의 IgM에 대해 이소타입-특이적 ¹⁾
인터루킨-9 (IL-9)	P15248
인테그린 α-11	Q9UKX5
야누스 키나제 3 단백질 티로신 키나제 (JAK3)	P52333
프로카텝신 W	P56202
프로퍼дин	P27918
TM 웹티드 (10TM 단백질)	NA – 상기 참조
종양 고사 인자-α (TNF-α)	P01375
안지오모틴	AAG01851; Q4VCS5
보체-1 애스테라제 억제제 (C1-INH)	P05155
보체 단백질 C3 (C3)	BC150179, BC150200; P01024
보체 단백질 C4 (C4)	BC151204, BC146673, AY379959, AL645922, AY379927, AY379926, AY379925
보체 단백질 C5 (C5)	BC113738, BC113740, DQ400449, AB209031, P01031
CD40	Q6P2H9
에오탁신	P51671
보체 인자 B (인자 B)	P00751
글루카곤 유사 웹티드-1 (GLP-1)	
글루카곤 유사 웹티드-1 수용체 (GLP-1 R)	P43220
과립구-매크로파지 콜로니-자극 인자 (GM-CSF)	P04141
인터루킨-10 (IL-10)	P22301

[0250]

바이오마커 명칭	예시적인 서열
인터루킨-11 (IL-11)	P20809
인터루킨-12 (IL-12)	O60595
인터루킨-13 (IL-13)	P35225
인터루킨-18 (IL-18)	Q14116
인터루킨-1 α (IL-1 α)	P01583
인터루킨-1 β (IL-1 β)	P01584
인터루킨-2 (IL-2)	P60568
인터루킨-3 (IL-3)	P08700
인터루킨-4 (IL-4)	P05112
인터루킨-5 (IL-5)	BC066282, CH471062, P05113
인터루킨-6 (IL-6)	P05231
인터루킨-8 (IL-8)	CR623827, CR623683, DQ893727, DQ890564, P10145
인터페론- γ (INF- γ)	P01579
렙틴	P41159
루이스 X /CD15	탄수화물 구조 (적용 없음)
루이스 y	탄수화물 구조 (적용 없음)
단핵구 화학주성 단백질-1 (Monocyte chemotactic protein-1, MCP-1)	P13500
뮤신-1	P15941
전립선 특이 항원(Prostate specific antigen, PSA)	P07288
RANTES	P13501
시알릴 루이스 x	탄수화물 구조 (적용 없음)
전환 성장 인자-1(Transforming growth factor-1, TGF-b1)	P01137
종양 괴사 인자- β (TNF- β)	P01374
혈관 내피 성장 인자 (Vascular endothelial growth factor, VEGF)	P15692, P49765, P49767, =O43915
CD40 리간드	P29965
인터루킨-16 (IL-16)	Q05BE6, Q8IUU6, B5TY35
인터루킨-1ra (IL-1ra)	P18510
단핵구 화학주성 단백질-3 (MCP-3)	BC112258, BC112260, BC092436, BC070240
단핵구 화학주성 단백질-4 (MCP-4)	Q99616

[0251]

[표 IV] 췌장암 진단 바이오마커

(A) 핵심 바이오마커

바이오마커 명칭
인터루킨-7 (IL-7)
인테그린 α -10

[0254]

[0255]

(B) 바람직한 바이오마커

바이오마커 명칭
B-갈락토시다제
브루톤의 티로신 키나제 (BTK)
보체 단백질 C1q (C1q)
보체 단백질 C1s (C1s)
B 세포 수용체 μ 사슬 (IgM)
인터루킨-9 (IL-9)
인테그린 α -11
아누스 키나제 3 단백질 티로신 키나제 (JAK3)
프로카텝신 W
프로퍼дин
TM 패티드
증양 고사 인자- α (TNF- α)

[0256]

(C) 선택적 추가 바이오마커

바이오마커 명칭
안지오모틴
보체-1 에스테라제 억제제 (C1-INH)
보체 단백질 C3 (C3)
보체 단백질 C4 (C4)
보체 단백질 C5 (C5)
CD40
에오틱신
보체 인자 B (인자 B)

[0258]

바이오마커 명칭
글루카곤 유사 펩티드-1 (GLP-1)
글루카곤 유사 펩티드-1 수용체 (GLP-1 R)
과립구-매크로파지 콜로니-자극 인자 (GM-CSF)
인터루킨-10 (IL-10)
인터루킨-11 (IL-11)
인터루킨-12 (IL-12)
인터루킨-13 (IL-13)
인터루킨-18 (IL-18)
인터루킨-1 α (IL-1 α)
인터루킨-1 β (IL-1 β)
인터루킨-2 (IL-2)
인터루킨-3 (IL-3)
인터루킨-4 (IL-4)
인터루킨-5 (IL-5)
인터루킨-6 (IL-6)
인터루킨-8 (IL-8)
인터페론- γ (IFN- γ)
렙틴
루이스 X /CD15
루이스 y
단핵구 화학주성 단백질-1 (MCP-1)
뮤신-1
전립선 특이 항원 (PSA)
Rantes
시알릴 루이스 x
전환 성장 인자-1 (TGF- β 1)
종양 괴사 인자- β (TNF- β)
혈관 내피 성장 인자 (VEGF)
CD40 리간드
인터루킨-16 (IL-16)
인터루킨-1ra (IL-1ra)
단핵구 화학주성 단백질-3 (MCP-3)
단핵구 화학주성 단백질-4 (MCP-4)

[0259]

[0260]

[표 V] 췌장암 진단 바이오마커 서브셋

바이오마커 명칭		PaC vs N		PaC vs N+Chp+AlP		PaC vs ChP		PaC vs AlP	
		PaC 전	후진	PaC 전	후진	PaC 전	후진	PaC 전	후진
A									
CD40	X			X	X	X	X	X	X
인터루간-12 (IL-12)		X	X	X	X	X	X	X	X
인터루간-3 (IL-3)		X	X	X	X	X	X	X	X
인터루간-4 (IL-4)		X	X	X	X	X	X	X	X
인터루간-8 (IL-8)	X			X	X	X	X	X	X
단핵구 회화주성 단백질-1 (MCP-1)		X	X	X	X	X	X	X	X
뮤신-1		X	X	X	X	X	X	X	X
전활 성장 인자, 베타-1 (TGF-β1)	X	X	X	X	X	X	X	X	X
중양 괴사 인자-β (TNF-β)	X	X	X	X	X	X	X	X	X
혈관 내피 성장 인자 (VEGF)	X	X	X	X	X	X	X	X	X
B									
B-갈락토시다제	X			X	X	X	X	X	X
브루톤의 디로신 칸나제 (BTK)		X	X	X	X	X	X	X	X
CD40 리간드		X	X	X	X	X	X	X	X
보체 단백질 C1q (C1q)		X	X	X	X	X	X	X	X
보체 단백질 C3 (C3)	X			X	X	X	X	X	X
글루카곤 유사 펩티드-1 (GLP-1)		X		X		X		X	
B 세포 수용체 H-사슬 (IgM)		X		X		X		X	
인터루간-10 (IL-10)		X		X		X		X	
인터루간-11 (IL-11)		X		X		X		X	
인터루간-13 (IL-13)		X		X		X		X	

[0261]

바이오마커 명칭	PaC vs N		PaC vs N+Chp+AlP		PaC vs ChP		PaC vs AlP	
	P \geq 1	후진	P \geq 1	후진	P \geq 1	후진	P \geq 1	후진
인터루간-16 (IL-16)	x	x	x	x	x	x	x	x
인터루간-18 (IL-18)	x	x	x	x	x	x	x	x
인터루간-1α (IL-1α)	x	x	x	x	x	x	x	x
인터루간-1ra (IL-1ra)	x	x	x	x	x	x	x	x
인터루간-5 (IL-5)	x	x	x	x	x	x	x	x
인터루간-6 (IL-6)	x	x	x	x	x	x	x	x
인터루간-7 (IL-7)	x	x	x	x	x	x	x	x
인터파로 γ (INF- γ)	x	x	x	x	x	x	x	x
인터그린 α -11	x	x	x	x	x	x	x	x
아누스 키나제 3 단백질 티로신 키나제 (JAK3)	x	x	x	x	x	x	x	x
루이스 x / CD15	x	x	x	x	x	x	x	x
프로카탈산 W	x	x	x	x	x	x	x	x
프로페린	x	x	x	x	x	x	x	x
시알릴루이스 x	x	x	x	x	x	x	x	x
C								
보체단백질 C1s (C1s)			x	x	x	x	x	x
에오틱신			x	x	x	x	x	x
글루카곤 유사 펩티드-1 수용체 (GLP-1 R)			x	x	x	x	x	x
인터그린 α -10			x	x	x	x	x	x
단백구체학주성 단백질-3 (MCP-3)			x	x	x	x	x	x
D								
보체-1 에스테라제 억제제 (C1-INH)	x			x				
보체 단백질 C5 (C5)	x	x	x	x	x	x	x	x
중성 고사 인자- α (TNF- α)	x	x	x	x	x	x	x	x
E								

비이오마커 명칭		PaC vs N		PaC vs N+Chp+AlP		PaC vs ChP		PaC vs AlP	
		P \geq 1	후진	P \geq 1	후진	P \geq 1	후진	P \geq 1	P \geq 1
인터루킨-9 (IL-9)			X						X
F									
과립구-메크로파지 콜로니-자극 인자 (GM-CSF)					X				X
인터루킨-2 (IL-2)				X	X			X	X
랩틴				X				X	X
루이스 y				X				X	X
진립신 특이 항원 (PSA)				X	X			X	X
Rantes				X				X	X
G									
안지오모틴			X						
보체 단백질 C4 (C4)				X					
보체 인자 B (인자 B)				X					
H									
인터루킨-1 β (IL-1 β)								X	
단핵구 회핵주성 단백질-4 (MCP-4)								X	
TM 펩티드									X

[0263]

[표 VI] 훈련된 SVM 프로그램

[0264]

다음 파라미터는 <http://cran.r-project.org/web/packages/e1071/index.html>로부터 이용가능한, e1071 1.5-24 SVM를 이용하여 얻어졌다.

[0266] (A) 후진 제거 전략을 이용한, PaC vs 모두 (N+Chp +AIP)에 대한 압축된 바이오마커 시그니처의 정의

```

filnamn <- "PaC_vs_all_training_set.txt"
group1 <- "other"
group2 <- "PaC"

# Include
source("NaiveBayesian")
library(e1071)

# Hämta data
rawfile <- read.delim(filnamn)

# Läs in grupper
groups <- rawfile[,2]

# Hämta provnamn i datafilen
samplenames <- as.character(rawfile[,1])

# Skapa dataset ur råfilen
data <- t(rawfile[,-c(1,2)])

# Log
# data <- log(data)/log(2)

# antal prover
nsamples <- ncol(data)

```

[0267]

```

# Skapa antikroppsnamnlista ur NYA datafilen
ProteinNames <- read.delim(filnamn,header=FALSE)
ProteinNames <- as.character(as.matrix(ProteinNames)[1,])
ProteinNames <- ProteinNames[-(1:2)]

# Kolla antal Ab i nya datasetet
antal <- length(ProteinNames)

# Ge rätt prov- och Ab-namn
rownames(data) <- ProteinNames
colnames(data) <- samplenames

# Skapa subsets
subset1 <- is.element(groups , strsplit(group1,",")[[1]])
subset2 <- is.element(groups , strsplit(group2,",")[[1]])

# Skapa factorlista
svmfac <-
factor(rep('rest',ncol(data )),levels=c(group1,group2,'rest'))
svmfac[subset1] <- group1
svmfac[subset2] <- group2
svmfac <- svmfac[subset1|subset2]

```

[0268]

```
# Skapa vektor för K-L felet där det minsta för varje signaturlängd
sparas
smallestErrorPerLength <- rep(NA,antal)
```

```
# Beräkna medelvärde för varje Ab över alla prov som är med
averages <- apply(data, 1, mean)
```

```
# Skapa vektor för Ab-ordningen efter K-L felet som erhållits när
# respektive antikropp var satt till medelvärde.
abOrder <- rep(NA,antal)
```

```
# Skapa ett dataset att eliminera i
elimData <- data[,subset1|subset2]
```

```
# Lista att förvara SVM-modellerna i
models <- numeric(nsamples)
```

```
# Skapa variabel för att hålla reda på hur många Ab som tagits bort
borttagna <- 0
```

[0269]

```
#####
# BEGIN BACKELIM      #####
#####
#####
```

```
print(Sys.time())
```

```
# Kör tills bara två analyter återstår
for(j in 1:(antal-1))
```

```
{
  # Check if groups are given in correct order
  control <- as.numeric(svmfac)
  if(sum(control[subset1]) > sum(control[subset2]))
  {
    print("ERROR: Change order of your group1 and group2!!!")
    break
  }
  # För varje signaturlängd, där alla är med från början, träna en
  modell för
```

```
  # varje N-1 kombination av prover med den data som finns i elimData
  for (i in 1:nsamples)
  {
    # Modellerna sparas i en array av listor kallad models
    models[i] <- list(svm(t(elimData[,-i]), svmfac[-i],
    kernel="linear"))
  }
```

```
[0270]
  # Nu är alla modeller som behövs för LOO tränade och ska testas på
  # elimData.
  # I elimData sätts först en analyt till medelvärde, sen testas var och
  # en av
  # modellerna med det prov som var borttaget när den tränades.
  # När alla modellerna är testade en gång beräknas KL-fel som sparas i
  # errors.
  # Nu sätts nästa analyt till medelvärde och testprocessen görs om,
  # tills alla
  # analyter varit medelvärdeseliminerade en gång. Resultatet blir en
  # KL-fel
```

[0270]

[0271]

```

# lista lika lång som antalet analyter som är kvar i datasetet.

# Skapa en lista med K-L fel en viss signaturlängd (antal + 1 - j
lång)
# där areorna för varje körning där en Ab i taget har satts till
medelvärde
errors <- testModels(models, elimData, averages)

# Lägg namnet på Ab med sämst inverkan på felet i abOrder
abOrder[j] <- getWorstAb(errors, row.names(elimData))

# Lägger till värdet på det minsta felet
smallestErrorPerLength[j] <- getSmallestError(errors)

# Tar bort sämsta Ab ur medelvärdeslistan
averages <- getNewAverages(errors, averages)

# Tar bort sämsta Ab ur elimData
elimData <- getNewElimData(errors, elimData)

# Noterar att en Ab tagits bort
borttagna <- borttagna + 1

```

[0272]

```

# Ange hur många analyter som elimineras, samt vad klockan är.
print(paste(j, "analytes eliminated @", Sys.time()), sep="")
}

```

```

# Lägg till namnet på sista analyten, som aldrig blen elimineras
abOrder[length(abOrder)] <- setdiff(ProteinNames, abOrder)

```

```

# Spara resultatet till fil
filename <- paste("Backward elimination
result(", rnorm(1)+1, ".txt", sep="")
write.table(cbind(smallestErrorPerLength, abOrder), file=filename,
sep="\t", quote = F, row.names = F)

```

[0273]

```

#####
# FUNCTIONS
#####
#####
```

```

# getWorstAb: Rapporterar namnet på antikroppen som kommer tas bort
# (den där ROC-arean var som störst)
getWorstAb <- function(errors, abNames)
{
  return(abNames[order(errors, decreasing = F)[1]])
}

```

```

# testModels: testar alla modeller som finns i 'models' med alla
# analyser satta till medelvärde en gång
testModels <- function(models, elimData, averages)
{
  nsamples <- ncol(elimData)
  d <- as.numeric(svmfac)-1
  y <- numeric(nsamples)
  E <- numeric(nsamples)
}

```

[0274]

```

analytes <- nrow(elimData)
errors <- numeric(nrow(elimData))

for(k in 1:analytes)
{
  # Sätt analyt k till medelvärde i elimData
  # Men spara först analytens orginalvärde
  backup <- elimData[k,]
  elimData[k,] <- averages[k]

  # Gör LOO loop för datasetet med de redan färdiga modellerna
  for (i in 1:nsamples)
  {
    pred <- predict(models[[i]] , t(elimData[,i]),
decision.values=TRUE)
    #spara decision values
    y[i] <- as.numeric(attributes(pred)$decision.values)
  }

  # Beräkna "sannolikheterna"
  y = 1-(1/(1 + exp(-y)))

  # Beräkna KL-fel när aktuell analyt är elimineras
  for (i in 1:nsamples)
  {
    E[i] <- -(d[i]*log(y[i])+(1-d[i])*log(1-y[i]))
  }

  # Spara felet
  errors[k] <- sum(E)

  # Lägg tillbaka analyten
  elimData[k,] <- backup
}
return( errors )
}

```

```

# getNewElimData: Väljer vilken antikropp som ska tas bort ur
# träningsdata och tar bort den
getNewElimData <- function(errors, elimData)
{
  # Positionen för det minsta felet
  tasBort <- order(errors,decreasing = F)[1]
  return(elimData[-tasBort,])
}

```

```

# getSmallestError: Rapporteras minsta K-L felet
getSmallestError <- function(errors)
{
  return(min(errors))
}

```

```

# getNewAverages: skapar en ny lista med medelvärden efter att en analyt
# elimineras.
getNewAverages <- function(errors, averages)
{

```

[0275]

[0276]

[0277]

```

# Positionen för det minsta felet
tasBort <- order(errors, decreasing = F)[1]
return(averages[-tasBort])
}

# getRemovedAb: tar fram ID på analyt som elimineras
getRemovedAb <- function(errors, abNames)
{
  return(abNames[order(errors, decreasing = T)[1]])
}

```

[0278]

(B) 변형된 후진 제거 전략을 이용한 PaC vs N에 대한 압축된 바이오마커 시그니처의 정의

```

# Datafil och grupper
filnamn <- "PaC_vs_N_dataset.txt"

group1 <- "N"
group2 <- "PaC"

# Läs in datafil
rawfileORG <- read.delim(filnamn)

# Läs in grupper
groupsORG <- rawfileORG[,2]

# Läs in data
dataORG <- log(t(rawfileORG[,-c(1,2)]))

# Läs in Ab-namn
ProteinNames <- read.delim(filnamn, header=FALSE)
ProteinNames <- as.character(as.matrix(ProteinNames)[1,])
ProteinNames <- ProteinNames[-(1:2)]

# Kalla Ab rätt namn
rownames(dataORG) <- ProteinNames

# Kalla prover rätt namn
samplenamesORG <- as.character(rawfileORG[,1])
colnames(dataORG) <- samplenamesORG

# Kontrollera antalet prover
NoSamples <- dim(rawfileORG)[1]

```

[0280]

```

# Kontrollera antalet Ab
NoAntibodies <- dim(rawfileORG) [2] - 2

# Skapa subsets utifrån grupper
subsetORG1 <- is.element(groupsORG , strsplit(group1,"") [[1]])
subsetORG2 <- is.element(groupsORG , strsplit(group2,"") [[1]])

# Skapa faktorer utifrån subsets
svmfacORG <-
factor(rep('rest',ncol(dataORG )),levels=c(group1,group2,'rest'))
svmfacORG[subsetORG1] <- group1
svmfacORG[subsetORG2] <- group2

# Skapa vektor och array får ROC-areor respektive Signaturer
# från varje körning utan A provet
BestROCsForEachRun <- rep(NA,NoSamples*NoAntibodies)
dim(BestROCsForEachRun) <- c(NoSamples,NoAntibodies)
AbRemovalOrderForEachRun <- rep(NA,NoSamples*NoAntibodies)
dim(AbRemovalOrderForEachRun) <- c(NoSamples,NoAntibodies)

```

```

# För varje prov i datasetet:
for(A in 1:NoSamples) # for(A in NoSamples:1)
{

```

```

# Hämta data från orginal-råfilen för alla prover utom A
rawfile<-rawfileORG[-A,]

```

```

# Hämta provnamn i NYA datafilen
samplenames <- as.character(rawfile[,1])

```

```

# Hämta grupper i NYA datafilen
groups <- rawfile[,2]

```

```

# Skapa dataset ur NYA datasete
runData <- t(rawfile[,-c(1,2)])

```

```

# Skapa antikroppsnamnlista ur NYA datafilen
ProteinNames <- read.delim(filnamn,header=FALSE)
ProteinNames <- as.character(as.matrix(ProteinNames) [1,])
ProteinNames <- ProteinNames[-(1:2)]

```

```

# Kolla antal Ab i nya datasetet
antal <- length(ProteinNames)

```

```

# Ge rätt prov- och Ab-namn
rownames(runData) <- ProteinNames
colnames(runData) <- samplenames

```

[0281]

[0282]

```

# Skapa nya subsets
subset1 <- is.element(groups, strsplit(group1, ",") [[1]])
subset2 <- is.element(groups, strsplit(group2, ",") [[1]])

# Skapa ny factorlista
svmfac <-
factor(rep('rest', ncol(runData)), levels=c(group1, group2, 'rest'))
svmfac[subset1] <- group1
svmfac[subset2] <- group2

# Skapa vektor för ROC-areor där den bästa för varje signaturlängd
sparas
bestRocPerLength <- rep(NA, antal)

# Beräkna medelvärde för varje Ab över alla prov som är med
averages <- apply(runData, 1, mean)

# Skapa vektor för Ab-ordningen efter ROC-areorna som erhållits när
# respektive antikropp var satt till medelvärde.
abOrder <- rep(NA, antal)

# Skapa tränings och testset att köra
trainData <- runData
testData <- runData

```

[0283]

```

# Skapa variabel för att hålla reda på hur många Ab som tagits bort
borttagna <- 0

# Kör lika många gånger som antalet Ab - 1
for(j in 1:(antal-1))
{
  # Skapa en lista med ROC-areor en viss signaturlängd (antal + 1 - j
  # längd)
  # där areorna för varje körning där en Ab i taget har satts till
  # medelvärde
  ROCList <- svmForAbList(antal-borttagna, trainData, svmfac,
  averages)

```

[0284]

```

  # Lägg Ab med sämst inverkan på ROC-area
  abOrder[j] <- getRemovedAb(ROCList, row.names(trainData))

  # Skapa ny träningsdata där sämsta Ab tas bort
  trainData <- getNewTrainData(ROCList, trainData)

  # Tar bort sämsta Ab ur medelvärdeslistan
  averages <- getNewAverages(ROCList, averages)

  # Noterar att en Ab tagits bort
  borttagna <- borttagna+1

  # Lägger till värdet på den bästa ROC-arean
  bestRocPerLength[j] <- getBestROC(ROCList)
}

# Lägg till den bästa arean för aktuell längd i en lista
BestROCsForEachRun[A,] <- bestRocPerLength

# Lägg till vilken Ab som togs bort för aktuell längd i en lista
AbRemovalOrderForEachRun[A,] <- abOrder

# Skriv vilken körning som genomförts till prompten
print(paste(j, "in", A, "of", NoSamples, "at", Sys.time()))
}
```

[0285]

```

# Include
source("NaiveBayesian")
library(e1071)

# Skapar en listamed ROC-areor för en vända med ett antal antikroppar
# där alla antikroppar satts till medelvärde en gång
svmForAbList <- function(abNumber, trainData, svmfac, averages)
{
  testData <- trainData
  ROClist <- rep(NA, abNumber)
  for(k in 1:abNumber) # Byter en variabel i träningsdata till
  medelvärden,
  {
    testData[k,] <- averages[k] # kör svmloo, byter tillbaka till orginalvärdena.
    #ROClis[k] <- svmLOOvaluesBE(trainData, testData, svmfac)
    ROClist[k] <- svmLOOvaluesProb(trainData, testData, svmfac)

    testData[k,] <- trainData[k,]
  }
  return(ROClis)
}

```

[0286]

```

# Rapporterar namnet på antikroppen som kommer tas bort
# (den där ROC-arean var som störst)
getRemovedAb <- function(ROClis, abNames)
{
  return(abNames[order(ROClis, decreasing = T)[1]])
}

# Rapporterar största ROC-arean
getBestROC <- function(ROClis)
{

```

[0287]

```

  return(max(ROClis))
}

# Väljer vilken antikropp som ska tas bort ur träningsdata och tar bort
# den
getNewTrainData <- function(ROClis, trainData)
{
  # Positionen för den största ROC-arean
  tasBort <- order(ROClis, decreasing=T)[1]

  return(trainData[-tasBort,])
}

# Väljer vilken antikropp som ska tas bort ur averages
getNewAverages <- function(ROClis, averages)
{
  # Positionen för den största ROC-arean
  tasBort <- order(ROClis, decreasing=T)[1]
  return(averages[-tasBort])
}

```

[0288]

```

svmLOOvaluesBE <- function(trainData, testData, svmfac)
{
  nsamples <- ncol(trainData)
  res <- numeric(nsamples)
  sign <- numeric(nsamples)
  for (i in 1:nsamples)
  {
    svmtrain <- svm(t(trainData[,-i]), svmfac[-i], kernel="linear" )
    pred <- predict(svmtrain, t(testData[,i]), decision.values=TRUE)
    res[i] <- as.numeric(attributes(pred)$decision.values)
    fcn <- colnames(attributes(pred)$decision.values)[1]

```

[0289]

```

if(fcn==paste(levels(svmfac)[1],"/",levels(svmfac)[2],sep="")){sign[i]<-
1}

if(fcn==paste(levels(svmfac)[2],"/",levels(svmfac)[1],sep="")){sign[i]<-
-1}
}
res <- sign * res
ROCdata <- myROC(res,svmfac)
return(ROCdata[1])
}

```

[0290]

```

# Tränar svm med viss signatur på viss data och testar på ett prov
test1sample <- function(dataORG, svmfacORG, A, signatureAbs)
{
  svmtrain <- svm(t(dataORG[signatureAbs,-A]) , svmfacORG[-A] ,
kernel="linear")
  pred <- predict(svmtrain , t(dataORG[signatureAbs,A]) ,
decision.values=TRUE)
  res <- as.numeric(attributes(pred)$decision.values)
  return(res)
}

```

[0291]

```

svmLOOvaluesProb <- function(trainData, testData, svmfac)

{
  nsamples <- ncol(trainData)
  #res <- numeric(nsamples)
  #sign <- numeric(nsamples)
  d <- as.numeric(svmfac)-1
  y <- numeric(nsamples)
  E <- numeric(nsamples)
  for (i in 1:nsamples)
  {
    svmtrain <- svm(t(trainData[,-i]) , svmfac[-i] , kernel="linear")
    pred <- predict(svmtrain , t(testData[,i]), decision.values=TRUE)
    y[i] <- as.numeric(attributes(pred)$decision.values)
  }
  y = 1-(1/(1 + exp(-y)))

    for (i in 1:nsamples)
    {
      E[i] <- -(d[i]*log(y[i])+(1-d[i])*log(1-y[i]))
    }
    return(1/sum(E))
}

```

[0293]

[0294]

(C) 표 5(A)에 정의된 시그니처의 시험

```
apri<-c(
  "IL-3 (1)", "", "C3 (1)", "", "C5 (1)", "", "IL-7 (2)", "", "IL-4 (3)", "", "CD40 (2)", "", "TGF- $\beta$ 1 (1)", "", "IL-12 (1)", "", "GM-CSF (1)", "", "Properdin", "", "IgM (B)", "", "VEGF (3)", "", "IL-16 (1)", "", "MUC-1 (P3-15)", "", "IL-1 $\alpha$  (1)", "", "TNF- $\beta$  (1)", "", "Integrin  $\alpha$ -10", "", "C1 est. inh.", "", "MCP-1 (3)", "", "MCP-3 (2)", "", "IL-2 (3)", "", "Eotaxin (3)", "", "IL-11 (2)", "", "TNF- $\alpha$  (1)", "", "IL-1 $\alpha$  (3)", "")
```

[0295]

```
library(MASS)
library(gplots)
library(e1071)
source("C:/Program/R/R-2.8.1/library/NaiveBayesian")
```

[0296]

```
filnamn<-"PaC_all_data.txt"
rawfile <- read.delim(filnamn)
samplenames <- as.character(rawfile[,1])
groups <- rawfile[,2]
data <- t(rawfile[,-c(1,2)])
ProteinNames <- read.delim(filnamn,header=FALSE)
ProteinNames <- as.character(as.matrix(ProteinNames)[1,])
ProteinNames <- ProteinNames[-(1:2)]
rownames(data) <- ProteinNames
colnames(data) <- samplenames

group1 <- "other"
group2 <- "PaC"

nTrainingSamples <- 68
```

[0297]

```

nTestSamples <- 35

training <- data[,1:nTrainingSamples ]
test <- data[, (nTrainingSamples+1):(nTrainingSamples+nTestSamples) ]

aprioriBoolean <- is.element(rownames(data) , apri)
factTr <- factor(rep("rest",ncol(training)),levels=c(group1, group2,
"rest"))
subset1Tr <- is.element(groups[1:nTrainingSamples] , group1)
subset2Tr <- is.element(groups[1:nTrainingSamples] , group2)
factTr[subset1Tr] <- group1
factTr[subset2Tr] <- group2

factTe <- factor(rep("rest",ncol(test)),levels=c(group1, group2,
"rest"))
subset1Te <-
is.element(groups[(nTrainingSamples+1):(nTrainingSamples+nTestSamples)] ,
strsplit(group1,",") [[1]])
subset2Te <-
is.element(groups[(nTrainingSamples+1):(nTrainingSamples+nTestSamples)] ,
strsplit(group2,",") [[1]])
factTe[subset1Te] <- group1
factTe[subset2Te] <- group2

```

[0298]

```

svmtrain <- svm(t(training[aprioriBoolean,]) , factTr, kernel="linear")
pred <- predict(svmtrain, t(test[aprioriBoolean,]) , decision.values =
TRUE, probability = T)
res <- as.numeric(attributes(pred)$decision.values, probability = T)
facnames <- colnames(attributes(pred)$decision.values) [1]
ROCdata <- myROC(res,factTe)
ROCdata[1]

```

[0299]

```

SenSpe <- SensitivitySpecificity(res,factTe)
Sensitivity <- as.numeric(SenSpe[,1])
Specificity <- as.numeric(SenSpe[,2])
omSpecificity <- 1-Specificity
plot(omSpecificity, Sensitivity, ylab="Sensitivity", xlab="1-
Specificity",type="l")
mtext(side=1, line = -1.1, paste("ROC AUC = ",signif(ROCdata[1],
digits=2)))

```

[0300]


```

$y.scale
NULL

$nclasses
[1] 2

$levels
[1] "other" "PaC" "rest"

$tot.nSV
[1] 27

$nSV
[1] 17 10

$labels
[1] 1 2

[0306] $SV
      C1.est..inh.  C3..1.  C5..1.  CD40..2. Eotaxin..3.  GM.CSF..1.  IgM..B.  IL.11..2.
IL.12..1.  IL.16..1.  IL.1a..1.
Pa009 -0.75869601 -0.43683015 -0.20743017 -0.005289015 -0.38931564  0.095770897 -
0.93670642 -0.05169534 -0.30674506 -0.074417127  0.50792542
Pa038 -0.42410595 -0.73311709 -1.25244853 -0.692015029 -0.05078382 -0.003980786
5.67933331 -0.28192859 -0.28602139 -0.214764177  0.06911333
Pa006 -0.73654816  0.24109661 -0.46040093 -1.012551914 -1.15296122 -0.841258482 -
0.72678562 -0.78781462 -0.32368921 -0.848293554 -0.69234178
Pa013 -0.21700833  1.86892400  0.58834639 -0.264721795 -0.68552957 -0.153202159 -
0.04927657  0.05757379 -0.08137510  1.221751926  1.08098684
Pa024  1.58648919  1.44742521  0.75444772 -0.284927936 -0.69406449  2.493561060 -
0.44527793 -0.39055704  0.29383416 -0.455154737 -0.81120268
Pa056 -0.36431725 -0.46909174 -0.50130198 -0.554747166 -0.35106339 -0.494146089 -
0.17013264 -0.60888484 -0.20597864 -0.517666268  1.81952423
Pa125 -0.23467729  0.25849084  0.56420454  0.369299406  0.43426269  0.208651188
1.25308235 -0.20834865  0.08202813 -0.357037439 -0.55800518
Pa001 -0.47076992  0.56381381  0.06633327  0.435287817  0.08932416 -0.222587527 -
0.47610845 -0.27857151 -0.35763315 -0.269646986 -0.76364068
Pa010  0.05320001 -0.18419292 -0.59262780  0.130515458  0.29607807  0.356063766 -
0.15926715  0.20807177 -0.15588319  0.580881124 -1.35854055
Pa021  0.38072938 -0.51120893 -0.08552594  0.833802541  0.12119111  0.162821416
0.27919151  0.38595265 -0.02303639  0.687902307 -0.40739811
Pa029 -0.39903826  1.23107814  0.03028247 -0.646633317 -0.61805906 -0.141471627 -
0.55837741 -0.62561111 -0.08591402 -0.544224442 -0.39352990
Pa048 -0.26957423  0.75584809 -0.08955982 -0.255480195 -0.87836464 -0.745964552 -
0.45255194 -0.26172904  0.03931362 -0.449749768 -1.32821555
Pa058  6.10905531 -0.04325893  0.08698841  0.870476656  0.53283461  0.979088486 -
0.13691383 -0.05059637 -0.09088942  0.270927629 -0.42939318
Pa092 -0.24149853 -0.22137826 -0.86374589 -0.004986871 -0.21952427 -0.118794005
0.17005557  0.18556909 -0.28585022  0.326046954 -0.87098967
Pa106 -0.23439949 -0.29286642 -0.90705556  1.431821157  0.87865697  1.208446338 -
0.20106733  0.67417004 -0.14514808  1.186336278 -0.70633875
Pa128 -0.08417237  0.97358160  0.52585704 -0.062347659 -0.02788034 -0.411255681
0.46178541 -0.17013906 -0.12290537 -0.008370042 -0.89794811
Pa142 -0.50958252 -1.08085583 -1.23272636 -0.299928811  0.33070388 -0.396222532 -

```

[0307]

0.41999851 -0.39335783 -0.26785141 -0.089558380 -0.27466106
 Pa015 -0.71635090 0.14276634 -0.74693179 -0.473760238 -1.26913415 -0.964917983 -0.79188785 -0.58820570 -0.41838792 -0.616645822 0.34550134
 Pa025 -0.67904289 1.32071592 -0.29298820 -0.497584785 -0.89584008 -0.363827564 -0.60314141 -0.52579460 -0.42216397 -0.400049252 -0.05516486
 Pa027 -0.30026876 0.52136967 -0.53568667 -0.092663002 0.38297396 0.085123350 -0.14805742 -0.08863111 -0.32783954 -0.031648097 -0.41278449
 Pa045 2.32770652 1.11229339 2.68850620 -0.053703190 -0.83048219 -0.639849777 -0.14176235 -0.17565594 0.52874104 -0.418407376 0.50156594
 Pa047 -0.07301964 -0.60084555 -1.39164969 0.347757599 -0.01329602 -0.298201129 0.07255214 0.10285325 -0.08591223 0.019651353 -0.67791456
 Pa100 -0.43294548 -0.20545382 -0.28057938 0.237411951 -0.20538311 -0.003113888 0.42568691 -0.47951871 -0.21867907 -0.315650824 -0.68013358
 Pa121 -0.05039884 1.52624676 0.54898444 -0.250353390 0.26504091 -0.228812744 -0.10786287 -0.37683608 -0.30183383 -0.169528272 -0.60888710
 Pa129 -0.04672091 -1.20455598 -0.72519120 0.144364354 0.39842641 -0.176936684 1.07960964 0.09003457 -0.28256596 0.455159702 -1.07832336
 Pa137 -0.81063375 0.39213184 0.40088436 -1.227901001 -1.21269965 -0.956598086 -0.86084278 -0.79973661 7.63985656 -0.907598547 0.72267511
 Pa147 -0.13778847 0.76981229 -1.00254399 -0.050061794 0.15942688 -0.300030972 0.29380087 0.02270757 -0.15355286 -0.030274049 0.78409332

[0308]

IL.1.ra..3. IL.2..3. IL.3..1. IL.4..3. IL.7..2. Integrin.a.10 MCP.1..3. MCP.3..2. MUC.1..P3.15. Properdin TGF.b1..1.
 Pa009 -0.873764678 -0.411112998 -1.03076857 -0.78714053 -0.5233176752 -1.0569348707 2.27173362 -1.25407017 -1.0252649 0.1360317 -0.2661644
 Pa038 -0.206056847 0.009708982 -0.60314751 0.44373161 -0.3919296075 0.8585696202 -0.07305799 -0.03616494 0.8055718 1.6515214 -0.1293500
 Pa006 -0.961571105 -0.289056355 -0.34582654 -1.06718063 -0.6638369686 -0.7098872057 -1.23264385 -1.31169599 -0.6500758 -1.3448447 -0.9306547
 Pa013 0.234806863 2.175447589 0.01266928 0.54717541 -0.1957777100 -0.4801514210 -0.99759860 -0.91896270 0.6498548 -1.1706810 0.4960250
 Pa024 -0.305358326 -0.221631817 0.36267874 -0.18586281 -0.2572215174 -0.0456176001 0.08051256 0.56408243 -0.5609039 -1.1089805 -0.2477406
 Pa056 0.217707838 -0.154053300 -0.72632423 -0.17209153 -0.0467640708 -0.5423370601 -0.53963593 -0.61602551 0.9101779 -1.0850042 -0.7841200
 Pa125 0.275035954 -0.224926621 -0.11926723 1.08023147 -0.0206201972 -0.0008422566 1.28473810 0.48915923 0.6111125 -0.3934791 0.1846749
 Pa001 -0.091741993 -0.213105078 -0.47534929 -0.76802149 -0.0327458023 -0.4608177008 -0.98212847 -0.23474026 -0.7966689 -0.2220837 -0.4402696
 Pa010 -0.464155570 -0.068647798 0.64754337 -0.22005784 -0.0766204082 0.5695475081 -0.13335331 0.53436185 0.3457089 -0.1103537 -0.2059105
 Pa021 0.809674963 0.082111154 0.57592395 0.85625877 0.5975398206 0.3056060393 -0.07617649 1.04456867 2.3882107 0.2397382 0.4214540

[0309]

Pa029 -0.598511728 -0.226930137 -0.44135325 -0.88287360 -0.4211643576 -0.5463940807 -0.73863658 -0.13480309 -1.0338848 1.3801873 -0.9098994
 Pa048 -0.540799012 -0.132230617 -0.48802315 -0.40421411 -0.3511893489 -0.2933086328 -0.69288196 -0.32862131 -0.5117169 -0.5428015 -0.6331645
 Pa058 -0.216266096 -0.263744491 -0.02111594 0.02466775 0.7903127709 -0.4266999225 0.04047197 0.59968815 -0.6309139 0.2644843 0.4347455
 Pa092 0.188422574 -0.178937815 0.55239494 -0.03520631 0.0359698815 -0.1928812726 -0.60201879 -0.03639941 0.2006243 -0.6324198 -0.2174438
 Pa106 0.134678555 -0.023059591 0.13626501 0.32218390 -0.0635055923 -0.0806119478 0.09892621 0.72394826 -0.4466987 0.3305753 0.8539264
 Pa128 0.007572306 -0.139367554 0.21470674 -0.12000217 -0.1506709444 0.2875368605 -0.35548071 0.63819849 -0.4067314 1.0337810 0.1537436
 Pa142 -0.266469132 -0.281152595 0.82774995 -0.25295292 -0.3165063500 -0.3388735672 1.71680502 0.49242804 -0.2902914 1.5769876 -0.3204796
 Pa015 -0.414827646 -0.335150876 -0.59577082 -0.84721879 -0.6165233284 -0.5728379380 -1.06685693 -1.32707320 -0.8710250 -1.3049299 -0.7987132
 Pa025 -0.614322842 -0.351470829 -0.93514922 -1.21928380 -0.5784378396 -0.7578714150 -

[0310]

1.07065042 -0.36768309 -1.0996868 -1.1292653 -0.8399010
 Pa027 -0.289995023 -0.234718460 -0.34640819 -0.11023567 0.0001702044 0.0004392925 -
 0.35285696 -0.23601985 0.5504571 -0.8609139 0.5920421
 Pa045 -0.071721708 -0.122541540 1.74977696 0.33065121 -0.3344757232 -0.0134402551
 0.14558436 -0.40419305 0.5895674 -0.9058004 -0.2084024
 Pa047 0.579454777 -0.023082065 1.00137061 -0.02504929 0.3732057511 0.4029526915
 0.28311658 0.20414137 -0.3904702 1.3038092 -0.3220880
 Pa100 -0.409728805 -0.323942678 0.13563267 -0.32632146 0.2183483192 -0.5858327012 -
 0.32588781 0.36754365 -0.8761754 1.0424584 -0.3768059
 Pa121 -0.177281870 -0.189628137 0.04549011 -0.22387032 0.0960551709 -0.1245495609 -
 0.01148733 0.08500916 -0.2435519 -0.4560824 -0.2646761
 Pa129 0.211284366 -0.255781747 0.20501384 -0.08248423 0.7175913867 0.0496293721
 0.39857908 0.87914844 0.9371010 -0.3958375 0.6287985
 Pa137 -0.641370977 -0.393683315 -1.25425888 1.23623466 -0.7024639259 -0.6739023588 -
 0.83664778 -0.50259826 -1.1584413 -1.1886815 -1.1372398
 Pa147 0.091538371 -0.178739738 0.06208210 0.04644500 -0.0264655443 0.1878050701
 0.50826277 -0.60143916 1.5374629 1.3112835 0.0817073

TNF.a..1. TNF.b..1. VEGF..3.

Pa009 0.40813630 -0.29072629 0.21376386
 Pa038 -0.18826712 0.01047747 -0.29116552
 Pa006 -0.66615076 -1.03733849 -0.95344856
 Pa013 0.13709356 -0.43789203 -0.39786986
 Pa024 0.35985522 -0.69484995 -0.39811383
 Pa056 -0.32725609 -0.65537476 -0.81052618
 Pa125 0.13099818 0.09666528 0.25745881
 Pa001 -0.54730109 -0.26784028 -0.01851188
 Pa010 -0.07002966 0.14109466 0.47584857
 Pa021 0.35647693 0.84739885 1.41457100
 Pa029 -0.70424687 0.01067955 -0.12682955
 Pa048 -0.88347271 -0.37892814 0.03951195
 Pa058 1.71697123 0.18848291 0.01324853
 Pa092 -0.43475610 0.92880031 -0.27996250
 Pa106 1.01132754 1.45346413 0.67985051
 Pa128 -0.59595956 0.29156722 0.09509325
 Pa142 0.12253269 0.41393499 -0.31173738
 Pa015 -0.93543929 -0.62606091 -0.82759284
 Pa025 -0.16053458 -0.51839963 -0.21614866
 Pa027 0.20233072 0.41248908 -0.18909465
 Pa045 -0.17202350 -0.50984474 0.77281999
 Pa047 -0.29192335 0.22268877 0.33534050
 Pa100 0.48915718 -0.30606315 -0.21250898
 Pa121 -0.45976438 0.09615171 0.28373940
 Pa129 0.29426203 0.62249829 0.08120801
 Pa137 -0.91532270 -0.81985007 -1.12040867
 Pa147 0.11963137 0.13800477 0.29478773

\$index
 [1] 1 3 15 16 17 21 24 26 28 30 31 32 34 37 38 42 44 47 48 50 53 54 61 64 66 67 68

\$rho
 [1] -0.6668827

\$compprob
 [1] FALSE

\$probA
 NULL

\$probB
 NULL

\$sigma
 NULL

\$coefs

[,1]
 [1,] 0.30343126
 [2,] 0.09542226
 [3,] 1.00000000
 [4,] 0.20660644
 [5,] 0.05461905
 [6,] 0.82341233
 [7,] 0.50601183
 [8,] 1.00000000
 [9,] 0.24168457
 [10,] 0.33709008
 [11,] 0.15704187
 [12,] 0.50405791
 [13,] 0.24715788
 [14,] 0.85521666
 [15,] 0.20152940
 [16,] 1.00000000
 [17,] 0.50180868
 [18,] -1.00000000
 [19,] -1.00000000
 [20,] -1.00000000
 [21,] -0.49831606
 [22,] -1.00000000
 [23,] -1.00000000
 [24,] -1.00000000
 [25,] -1.00000000
 [26,] -0.10618995
 [27,] -0.43058420

[0315]

\$na.action
 NULL

 \$fitted
 Pa009 Pa012 Pa038 Pa042 Pa055 Pa063 Pa066 Pa069 Pa081 Pa097 Pa105 Pa108 Pa114
 Pa135 Pa006 Pa013 Pa024 Pa033 Pa039 Pa052 Pa056 Pa089 Pa096 Pa125 Pa138
 other
 other other other other other other other other other other other other other other other
 Pa001 Pa007 Pa010 Pa014 Pa021 Pa029 Pa048 Pa053 Pa058 Pa088 Pa091 Pa092 Pa106
 Pa110 Pa119 Pa126 Pa128 Pa133 Pa142 Pa145 Pa005 Pa015 Pa025 Pa026 Pa027
 other
 other other other PaC other PaC PaC PaC
 Pa034 Pa040 Pa045 Pa047 Pa070 Pa071 Pa079 Pa080 Pa085 Pa094 Pa100 Pa102 Pa113
 Pa121 Pa123 Pa129 Pa137 Pa147
 PaC PaC PaC PaC PaC PaC PaC PaC other PaC PaC PaC PaC
 other PaC PaC
 Levels: other PaC rest

[0316]

[0317]

[표 VII] (A) 췌장암과 (B) 정상, 만성 췌장염, 및/또는 급성 염증성 췌장염 간의 구별을 위한 ROC-AUC 값

ROC-AUC	바이오마커 시그니처
0.71	IL-7
0.69	인테그린 α -10
0.76	IL-7 + 인테그린 α -10 + 1 개의 표 IV B 마커
0.79	IL-7 + 인테그린 α -10 + 2 개의 표 IV B 마커
0.80	IL-7 + 인테그린 α -10 + 3 개의 표 IV B 마커
0.79	IL-7 + 인테그린 α -10 + 4 개의 표 IV B 마커
0.81	IL-7 + 인테그린 α -10 + 5 개의 표 IV B 마커
0.81	IL-7 + 인테그린 α -10 + 6 개의 표 IV B 마커
0.80	IL-7 + 인테그린 α -10 + 7 개의 표 IV B 마커
0.84	IL-7 + 인테그린 α-10 + 8 개의 표 IV B 마커
0.79	IL-7 + 인테그린 α -10 + 9 개의 표 IV B 마커
0.80	IL-7 + 인테그린 α -10 + 10 개의 표 IV B 마커
0.79	IL-7 + 인테그린 α -10 + 11 개의 표 IV B 마커
0.76	IL-7 + 인테그린 α -10 + 12 개의 표 IV B 마커

[0318]

[0319]

핵심(core) 마커 + 8 개의 선호되는 마커는 가장 좋은 ROC-AUC 값을 나타낸다.

[0320]

이 시그니처는 다음에 대응한다(핵심을 적색으로 표시):

[0321]

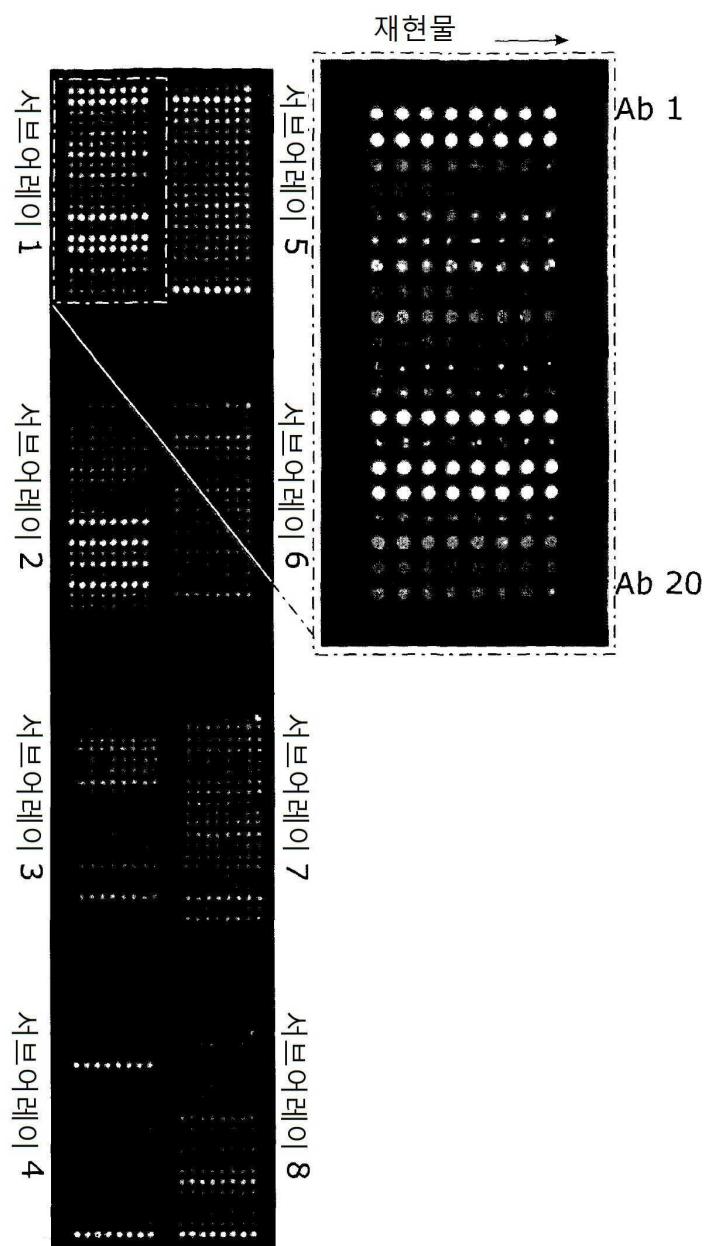
IL-7 + 인테그린 α -10 + BTK + C1q + IgM + IL-9 + 프로카세틴 W + 프로퍼딘 + TM 웨티드 + b-갈락토시다제.

[0322]

그러나, 모든 마커 조합이 실질적인 예측력을 갖는다.

도면

도면 1a

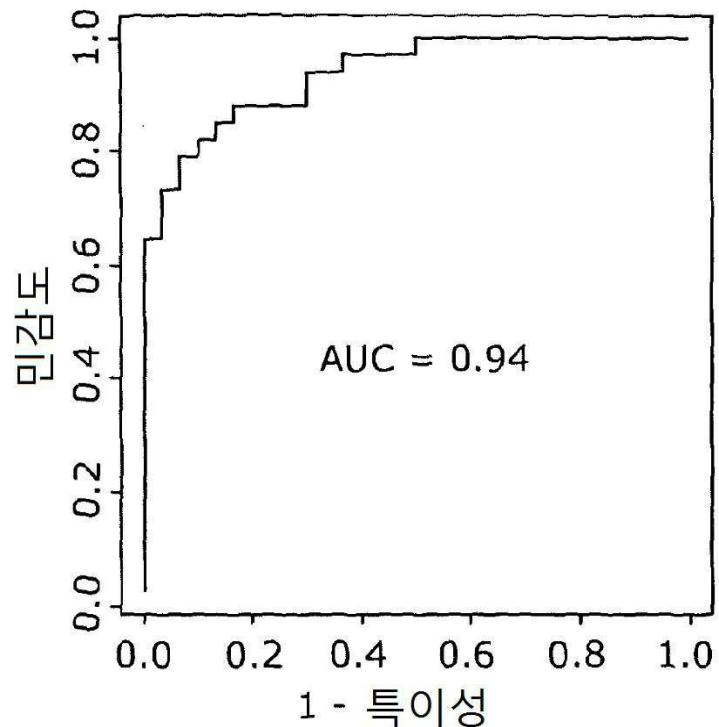


도면1b

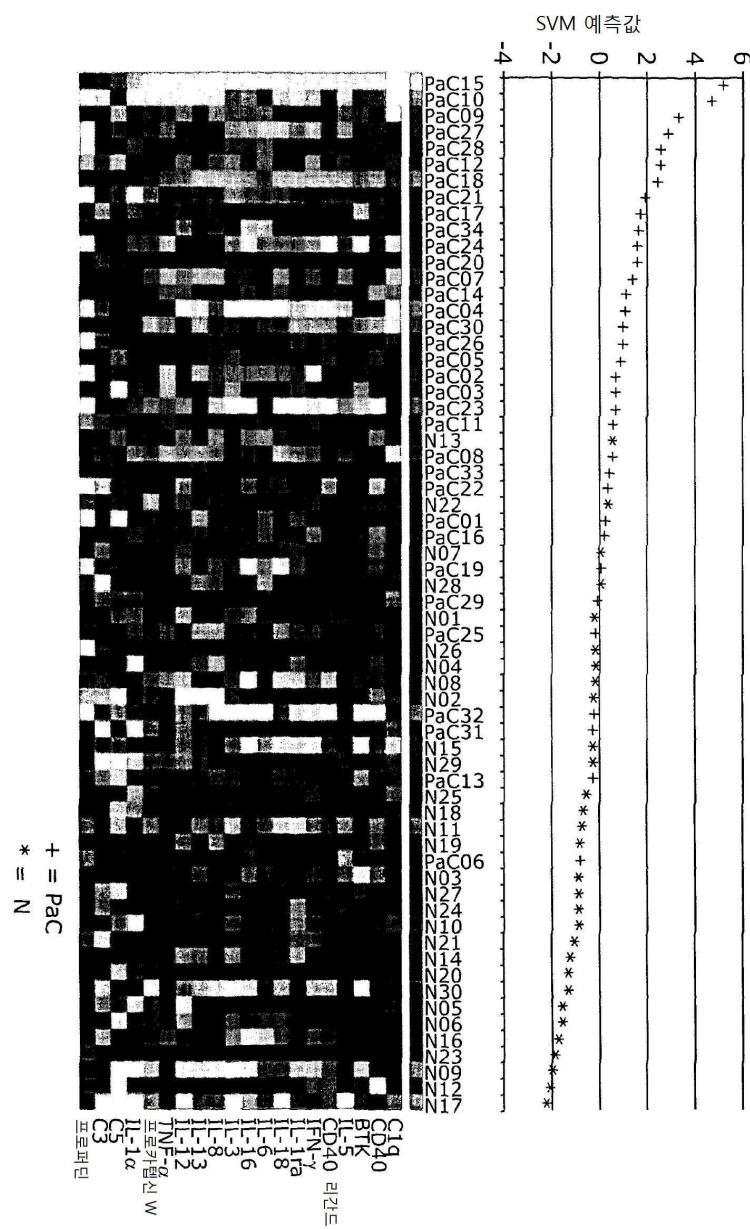
PaC vs. N

안지오모틴	
B-갈락토시다제	
BTK	
C1에스테라제 억제제	
C3	
C5	
CD40	
CD40	리간드
	인자 B
GLP-1	
IFN- γ	
IgM	
IL-10	
IL-11	
IL-12	
IL-13	
IL-16	
IL-18	
IL-1- <i>ra</i>	
IL-1 α	
IL-3	
IL-5	
IL-6	
IL-7	
IL-8	
인테그린 α -11	
프로카텝신 W	
시알릴 루이스 X	
TGF- β 1	
TNF- α	
VEGF	
C1q	
프로페딘	

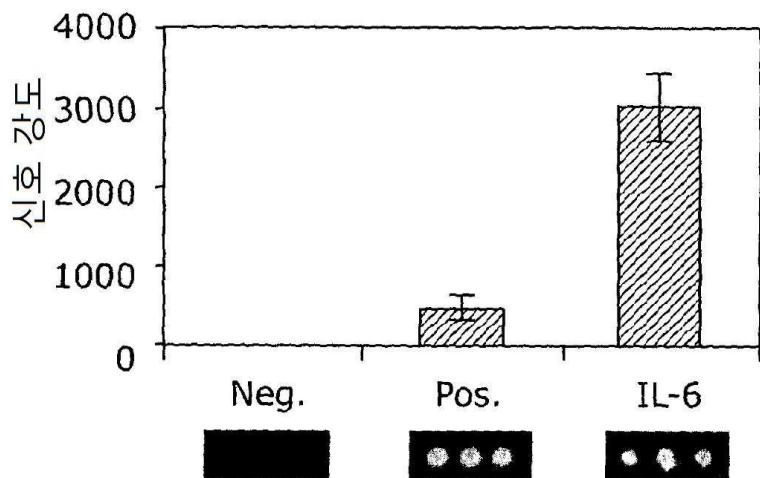
도면 1c



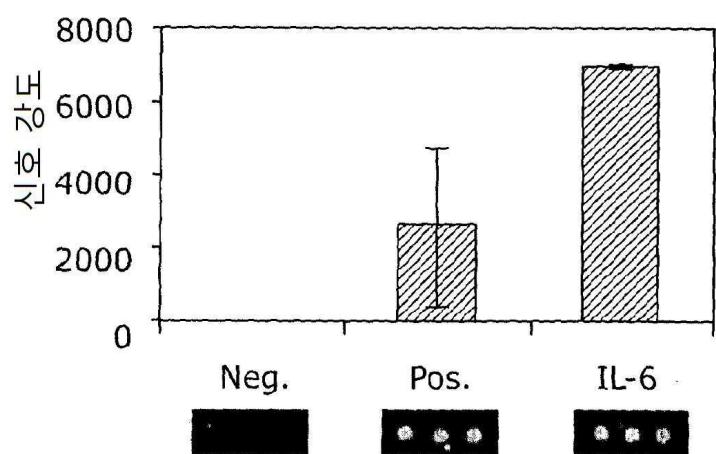
도면 1d



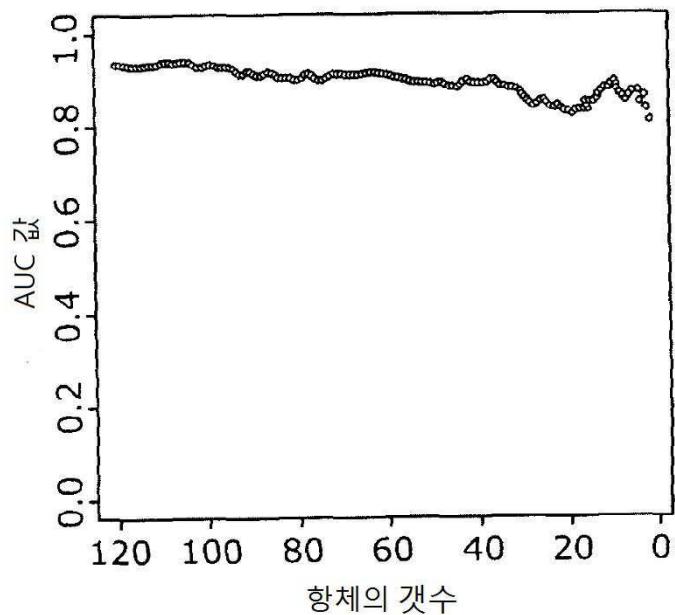
도면1e



도면1f



도면2a

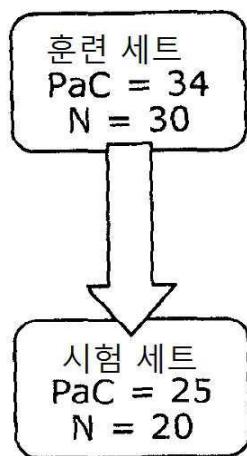


도면2b

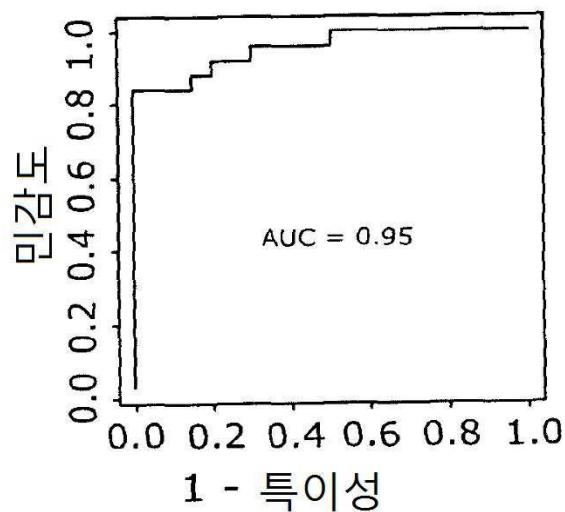
시험 시그니처

BTK
C4
C5
IgM
IL-12
IL-16
IL-18
IL-1 α
IL-3
IL-4
IL-5
IL-6
IL-7
루이스 X
MCP-1
TNF- α
TNF- β
VEGF

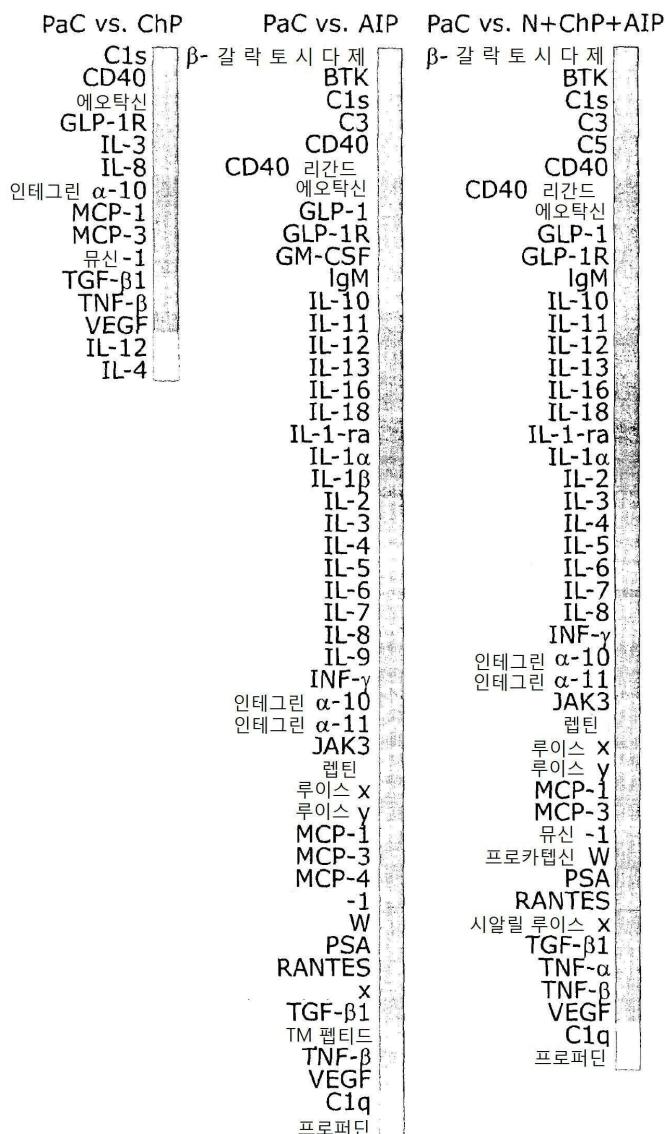
도면2c



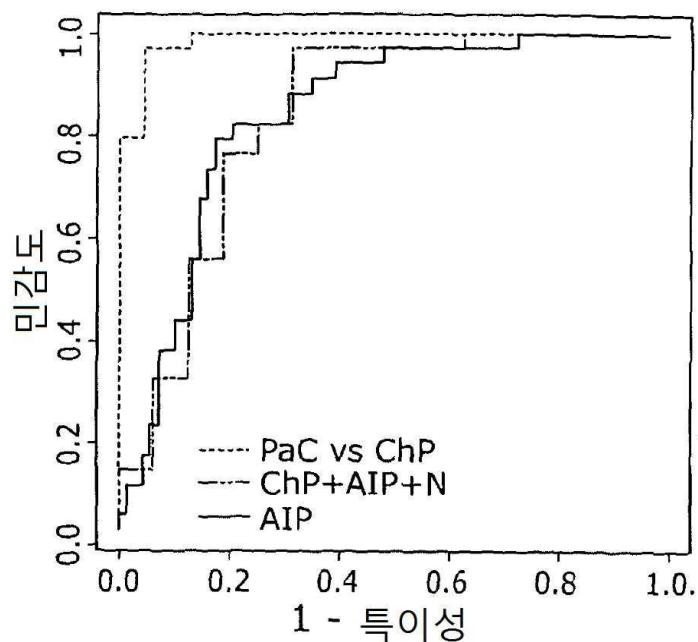
도면2d



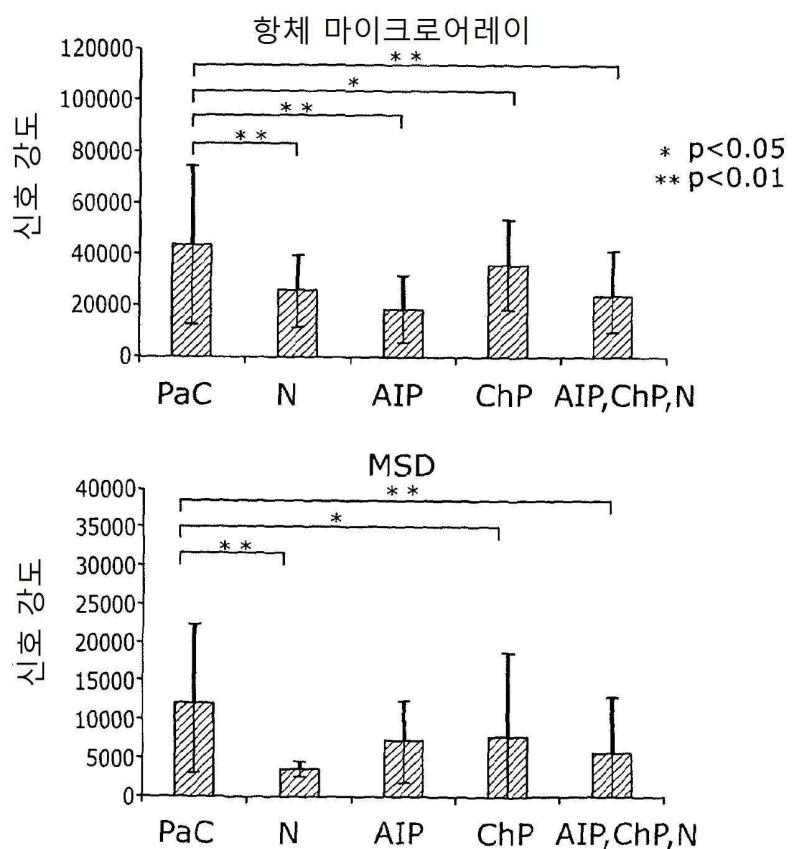
도면3a



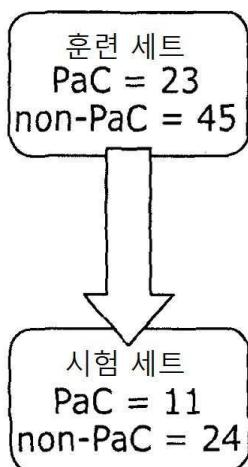
도면3b



도면3c



도면4a



도면4b

시험 시그니처

C1 에스테라제 억제제

C3

C5

CD40

에오탁신

GM-CSF

IgM

IL-11

IL-12

IL-16

IL-1 α

IL-1-ra

IL-2

IL-3

IL-4

IL-7

인테그린 α -10

MCP-1

MCP-3

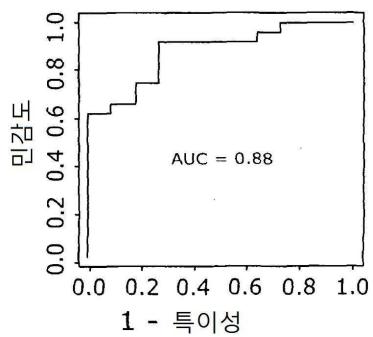
뮤신 -1

프로페딘

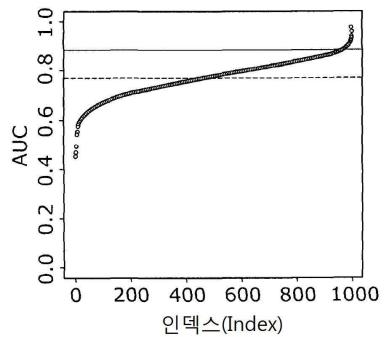
TGF- β 1TNF- α TNF- β

VEGF

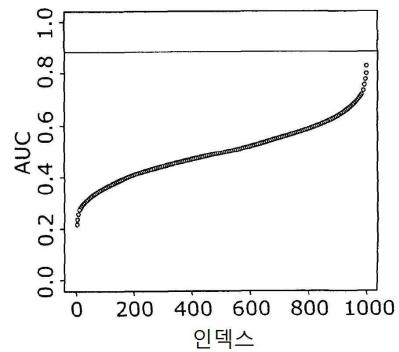
도면4c



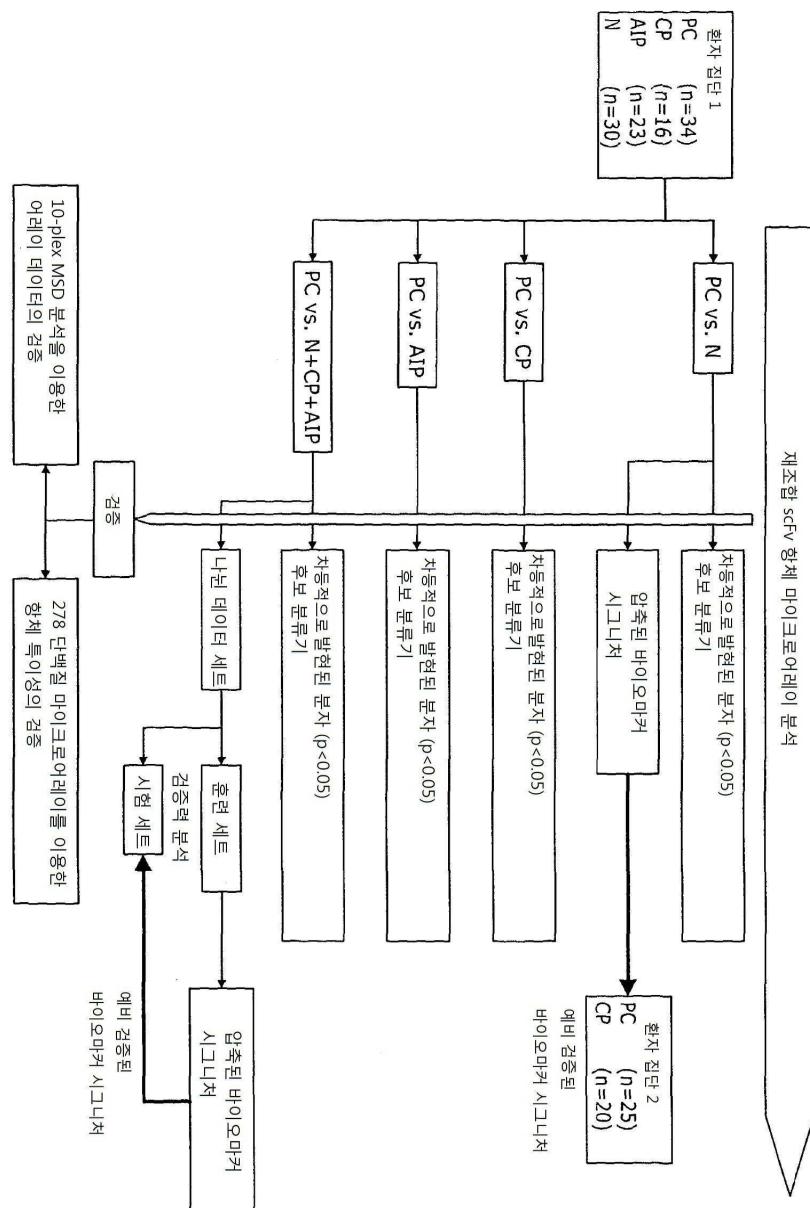
도면4d



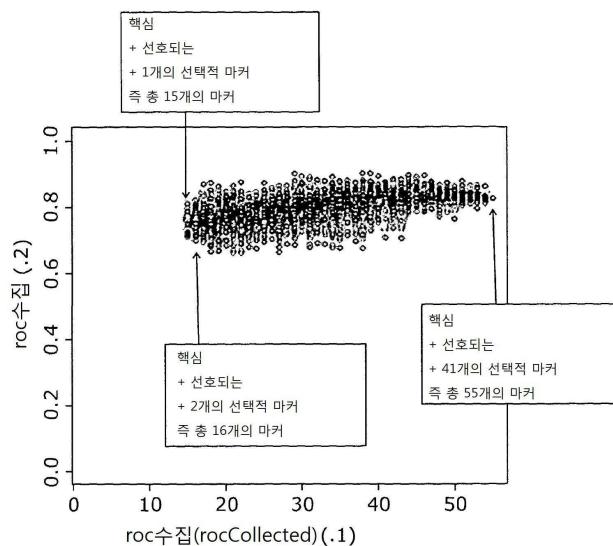
도면4e



도면5



도면6



서 열 목록

SEQUENCE LISTING

<110> Immunovia AB

<120> Method, array and use thereof

<140> PCT/GB2012/050483

<141> 2012-03-05

<150> GB1103726.4

<151> 2011-03-04

<160> 2

<170> SeqWin99

<210> 1

<211> 242

<212> PRT

<213> Homo sapiens

<400> 1

Met Ala Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro

1 5 10 15

Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser

20 25 30

Ser Tyr Gly Phe His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu

35 40 45

Trp Val Ser Leu Ile Ser Trp Asp Gly Gly Ser Thr Tyr Tyr Ala Asp
 50 55 60

Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr
 65 70 75 80

Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr
 85 90 95

Tyr Cys Ala Arg Gly Thr Trp Phe Asp Pro Trp Gly Gln Gly Thr Leu
 100 105 110

Val Thr Val Ser Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly
 115 120 125

Gly Gly Gly Ser Gln Ser Val Leu Thr Gln Pro Pro Ser Ala Ser Gly
 130 135 140

Thr Pro Gly Gln Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn
 145 150 155 160

Ile Gly Asn Asn Ala Val Asn Trp Tyr Gln Gln Leu Pro Gly Thr Ala
 165 170 175

Pro Lys Leu Leu Ile Tyr Arg Asn Asn Gln Arg Pro Ser Gly Val Pro
 180 185 190

Asp Arg Phe Ser Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile
 195 200 205

Ser Gly Leu Arg Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Ala Trp
 210 215 220

Asp Asp Ser Leu Ser Trp Val Phe Gly Gly Thr Lys Leu Thr Val
 225 230 235 240
 Leu Gly

<210> 2

<211> 31

<212> PRT

<213> Artificial Sequence

<220><223> 6-His affinity tag

<400> 2

Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr
1 5 10 15
Lys Asp Asp Asp Asp Lys Ala Ala Ala His His His His His
20 25 30