(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2024 년 5 월 16 일 (16.05.2024)

WIPO | PCT

(10) 국제공개번호
WO 2024/101856 A1

(51) 국제특허분류:
G16B 5/00 (2019.01) *G16B 40/00* (2019.01)
G16B 20/20 (2019.01)

(21) 국제출원번호: PCT/KR2023/017786

(22) 국제출원일: 2023 년 11 월 7 일 (07.11.2023)

(25) 출원언어: 한국어

(26) 공개언어: 한국어

(30) 우선권정보:
10-2022-0148030 2022 년 11 월 8 일 (08.11.2022) KR

(71) 출원인: 주식회사 제로믹스 (**GEROMICS INC**) [KR/KR]; 16229 경기도 수원시 영통구 창룡대로 260, 602, Gyeonggi-do (KR). 인바이츠지노믹스 주식회사 (**INVITES GENOMICS CO., LTD.**) [KR/KR]; 63243 제주특별자치도 제주시 첨단로8길 40, 204, Jeju-do (KR).

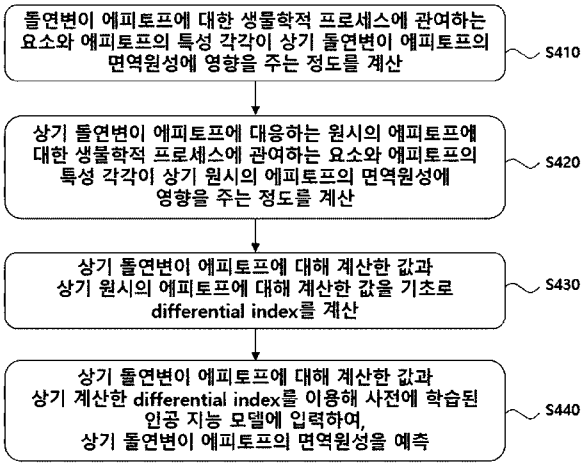
(72) 발명자: 조윤성 (**CHO, Yun Sung**); 13539 경기도 성남시 분당구 판교백현로 38, 111-1502, Gyeonggi-do (KR). 전재훈 (**JUN, Je Hoon**); 16239 경기도 수원시 팔달구 경수대로642번길 42-48, 201, Gyeonggi-do (KR). 이황열 (**LEE, Hwang Yeol**); 16953 경기도 용인시 기흥구 흥덕1로79번길 37, 504-1201, Gyeonggi-do (KR). 정옥성 (**CHUNG, Ok Sung**); 18499 경기도 화성시 동탄대로1길 66, 2847-203, Gyeonggi-do (KR).

(74) 대리인: 특허법인 정진 (**JUNGJIN IP LAW FIRM**); 08381 서울특별시 구로구 디지털로 29길 38, 505, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH,

(54) **Title:** METHOD FOR PREDICTING IMMUNOGENICITY OF NEOANTIGEN EPITOPE, AND DEVICE USING SAME

(54) 발명의 명칭: 신생항원 에피토프의 면역원성을 예측하는 방법 및 이를 이용한 장치



S410 ... Calculate degree to which each of factors, involved in biological process for mutant epitope, and characteristics of epitope affects immunogenicity of mutant epitope

S420 ... Calculate degree to which each of factors, involved in biological process for non-mutated epitope corresponding to mutant epitope, and characteristics of epitope affects immunogenicity of non-mutated epitope

S430 ... Calculate differential index on basis of value calculated for mutant epitope and value calculated for non-mutated epitope

S440 ... Use calculated value for mutant epitope and calculated differential index so as to input same into pre-trained artificial intelligence model, thereby predicting immunogenicity of mutant epitope

(57) **Abstract:** The present invention relates to a method for predicting the immunogenicity of a neoantigen epitope, and a device using same. The method for predicting the immunogenicity of a mutant epitope, according to the present invention, may comprise the steps of: calculating the degree to which each of factors, involved in a biological process for a mutant epitope, and characteristics of the epitope affects the immunogenicity of the mutant epitope; calculating the degree to which each of factors, involved in a biological process for a non-mutated epitope corresponding to the mutant epitope, and characteristics of the epitope affects the immunogenicity of the non-mutated epitope; calculating a differential index on the basis of a value calculated for the mutant epitope and a value calculated for the non-mutated epitope; and using the calculated value for the mutant epitope and the calculated differential index so as to input same into a pre-trained artificial intelligence model, thereby predicting the immunogenicity of the mutant epitope.

(57) **요약서:** 본 발명은 신생항원 에피토프의 면역원성을 예측하는 방법 및 이를 이용한 장치에 관한 것으로, 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법은 돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계, 상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index 를 계산하는 단계, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index 를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측하는 단계를 포함할 수 있다.

WO 2024/101856 A1

[다음 쪽 계속]

CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **지정국** (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 신생항원 에피토프의 면역원성을 예측하는 방법 및 이를 이용한 장치

기술분야

[1] 본 발명은 신생항원 에피토프의 면역원성을 예측하는 방법 및 이를 이용한 장치에 관한 것으로, 보다 구체적으로는 돌연변이가 발생한 신생항원 에피토프와 원시의 에피토프를 비교하여 신생항원 에피토프의 면역원성을 예측하는 방법 및 이를 이용한 장치에 관한 것이다.

배경기술

[2] 신생항원(neoantigen)은 돌연변이(mutation)가 일어난 항원(antigen)으로, 종양(tumor) 세포/조직(tissue)에서 발현되고, 정상 세포에서는 발현되지 않는 특징을 갖는다. 돌연변이에 의해 정상 서열이 변형되면, 면역 세포는 변형된 서열을 새로운 항원/에피토프로 인식해 면역반응을 일으킨다. 이러한 신생항원은 치료용 백신으로 활용되고 있으며, 면역치료 및 반응성 예측 등에 사용되고 있다. MHC 결합 친화성(MHC binding affinity), 프로테아좀 분절(proteasomal cleavage), TAP 수송체 효율(TAP transporter efficiency), pMHC 안정도(pMHC stability), T-세포 수용체 상호작용(T-cell receptor interaction) 등 다수의 생물학적 프로세스에 관여하는 여러가지 요소 및 소수성(hydrophobicity), 알려진 에피토프와의 유사성(similarity to known epitopes), 자기-단백체와의 비유사성(dissimilarity to self-proteome) 등 신생항원 에피토프가 가지는 특성들은 인실리코(in silico) 방식으로 CD8+ T-세포 면역원성(immunogenicity)을 보이는 신생항원 에피토프를 예측하는데 영향을 미칠 수 있다. 이에 돌연변이에 의해 발생되는 생물학적 프로세스에 관여하는 여러가지 요소 및/또는 신생항원 에피토프의 특성을 이용해 신생항원 에피토프의 면역원성을 예측했다.

발명의 상세한 설명

기술적 과제

[3] 현재까지 다수의 생물학적 프로세스에 관여하는 여러가지 요소 및/또는 에피토프 특성 각각을 개별적으로 이용하거나, 소수개의 요소 및/또는 특성을 결합하여 면역 반응을 유발하는 에피토프를 예측하는 방법이 개발되어 왔다. DAI(differential agretopicity index, 이하 DAI)도 에피토프의 면역원성을 예측하기 위해 이용된 요소 중 하나이다.

[4] DAI는 돌연변이 에피토프(또는, '펩티드'로 칭해질 수 있다.)의 MHC 친화성과 돌연변이 되지 않은 원시의(non-mutated or wild-type) 에피토프의 MHC 친화성의 비율(ratio)로 정의된다. 돌연변이 에피토프(neoepitope or mutant epitope)의 MHC 친화성만을 이용해 면역원성을 예측하기보다는 DAI를 이용해 면역원성을 예측하는 것이 더 효과적일 수 있다. 그 이유는, 돌연변이 에피토프가 MHC 결합 친화

도 등 특정 요소가 선호된다고 해서 모두 면역원성을 보이는 것은 아니다. 원시의 에피토프 자체가 원래 MHC 결합이 잘 되는 성향을 가졌다고 하더라도, 면역 시스템의 음성 선택(negative selection)/중심 관용(central tolerance) 과정 등을 통해 자기 에피토프에 대해 면역반응을 일으키지 않도록, 즉 정상 세포들이 면역 세포에 의해 공격당하지 않도록 작용할 수 있다. 따라서, 정상인 원시의 에피토프 대비 돌연변이에 의해 특성이 어떻게 변화된 것인지 그 변화의 차이를 측정하는 것이 에피토프의 면역원성을 예측하는데 매우 중요할 수 있다.

[5] 돌연변이 에피토프와 원시의 에피토프의 차이는 결합 친화도에만 영향을 미치는 것이 아니라, 에피토프가 면역반응을 일으키는데까지 관여하는 다양한 생물학적 프로세스에 영향을 미칠 수 있다. 예를 들어, 원시의 에피토프와 돌연변이 에피토프의 MHC 결합 친화도는 유사한 반면, 돌연변이에 의해 프로테아좀 분절, TAP 수송체 효율, pMHC 안정도, T-세포 수용체 상호작용 등의 효율/반응 등이 달라질 수 있다.

[6] 이에, 본 발명에서는 다수의 생물학적 프로세스 및 에피토프/네오에피토프 특성에 대해 differential index(즉, 돌연변이 에피토프와 원시의 에피토프의 측정값의 비율 혹은 차이)를 이용하여 돌연변이 에피토프의 면역원성을 예측하는 방법 및 이를 이용하는 장치에 대해 제안한다.

과제 해결 수단

[7] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법은 돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계, 상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index를 계산하는 단계, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측하는 단계를 포함할 수 있다.

[8] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법은 상기 학습된 인공 지능 모델에 계산한 값을 입력하기 전, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index에 대해 표준화 및 정규화 중 적어도 하나를 수행하는 단계를 더 포함할 수 있다.

[9] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법에서 상기 differential index는, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 비율(RATIO), 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 차이(DIFFERENCE) 중 적어도 하나를 포함할 수 있다.

[10] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법에서 상기 differential index는, 인산화, 소수성, 알려진 에피토프와의 유사성, 자기-단백체와의 비유사성, 에피토프 안정도, 항원 처리, MHC I 결합 친화도, pMHC 안정도, 면역, 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 중 적어도 하나에 대한 differential index일 수 있다.

[11] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법에서 상기 돌연변이 에피토프와 상기 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 면역원성에 영향을 주는 정도를 계산하는 단계는, 알고리즘, 프로그램, 또는 툴을 이용해 계산하는 단계일 수 있다.

[12] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법에서 상기 생물학적 프로세스는, 항원 처리 단계, 항원 제시 단계, 면역 단계, 및 종양 미세 환경으로 구성되고, 상기 종양 미세 환경에 관여하는 요소는 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 중 적어도 하나일 수 있다.

[13] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 컴퓨터 판독 가능한 기록 매체는 앞서 설명한 방법 중 적어도 하나를 실행하기 위한 컴퓨터 프로그램을 기록한 것일 수 있다.

[14] 본 발명에 따른 돌연변이 에피토프의 면역원성을 예측하는 장치는, 돌연변이 에피토프를 입력받거나, 상기 돌연변이 에피토프의 면역원성을 예측한 결과를 출력하는 입출력 장치, 돌연변이 에피토프와 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성을 기초로 계산된 differential index와 돌연변이 에피토프에 대해 생물학적 프로세스에 관여하는 요소와 에피토프의 특성을 기초로 계산된 값을 이용하여 돌연변이 에피토프의 면역원성 예측하도록 학습된 인공 지능 모델을 저장하는 저장 장치, 및 돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산하고, 상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산하고, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index를 계산하고, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측하는 연산 장치를 포함할 수 있다.

발명의 효과

[15] 본 발명에 따르면, 돌연변이에 의해 발생하는 다수의 생물학적 프로세스와 에피토프/네오에피토프 특성을 다양한 측면을 고려할 수 있어 돌연변이 에피토프만을 단독으로 사용하여 면역원성을 예측했을 때보다 더 높은 정확도로 돌연변이 에피토프의 면역원성을 예측할 수 있다. 높은 정확도로 예측된 면역원성 에피

토프는 신생항원 기반의 백신 개발시 더 높은 효용성과 낮은 부작용 등을 기대할 수 있다.

도면의 간단한 설명

[16] 도 1은 종양 세포의 생물학적 프로세스의 일 예를 나타낸 것이다.

[17] 도 2는 신생항원의 면역원성에 영향을 미칠 수 있는 다수의 생물학적 프로세스의 요소 및 신생항원의 특성을 나타낸 것이다.

[18] 도 3은 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 학습하는 과정의 일 예를 나타낸 것이다.

[19] 도 4는 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 예측하는 과정의 일 예를 나타낸 것이다.

[20] 도 5는 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 예측하는 장치의 구성도의 일 예를 나타낸 것이다.

발명의 실시를 위한 형태

[21] 이하, 실시예는 오로지 본 발명을 보다 구체적으로 설명하기 위한 것으로서, 본 발명의 요지에 따라 본 발명의 범위가 이들 실시예에 의해 제한되지 않는다는 것은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 있어서 자명할 것이다. 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[22] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 설명된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[23] 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정/단계들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정/단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.

[24]

[25] 도 1은 종양 세포의 생물학적 프로세스의 일 예를 나타낸 것이다.

[26] 도 1을 참조하면, 종양 세포의 생물학적 프로세스는 항원 처리(antigen processing) 단계, 항원 제시(antigen presentation) 단계, 및 면역원성 (Immunogenicity) 단계로 구성될 수 있다. 종양 세포를 둘러싼 종양 미세 환경 (tumor microenvironment, TME)도 생물학적 프로세스에 영향을 미칠 수 있어 종양 세포의 생물학적 프로세스의 일부로 볼 수 있다. 종양 세포의 생물학적 프로

세스는 일반 세포의 생물학적 프로세스와 동일 또는 유사하나, 각 단계에서 다양한 요소들이 에피토프 생성 및/또는 특성에 영향을 미칠 수 있다.

[27] 항원 처리 단계는 종양 세포의 핵 내의 DNA가 mRNA가 전사되어 세포질로 나오면, 전사된 mRNA가 펩티드로 변형되어 MHC 클래스 I 과 결합하는 단계이다. 종양 세포의 DNA는 돌연변이 등의 원인으로 변형된 상태일 수 있다. mRNA는 번역되어 단백질로 합성된 후, 프로테아좀에 의해 펩티드로 분절될 수 있다. 분절된 펩티드는 TAP 수송체에 의해 조면소포체 내부로 수송되어 MHC 클래스 I 과 결합한다.

[28] 항원 제시 단계는 항원 처리 단계에서 생성된 펩티드가 MHC 클래스 I 과 결합하여 펩티드-MHC 복합체로서 세포 표면에 제시되는 단계이다.

[29] 면역원성 단계는 세포 표면에 제시된 펩티드-MHC 복합체가 T 세포에 의해 인지되는 단계이다.

[30] 종양 미세 환경(TME)은 세포를 둘러싼 환경으로, 세포는 혈관, 면역 세포, 섬유아세포, 골수 유래 염증 세포, 림프구, 신호 전달 분자, 및 세포외 기질 등에 의해 둘러싸일 수 있다. 종양 세포는 세포외 신호 방출, 종양 세포 신생혈관 촉진, 및 말초 면역 관용 등을 유도하여 종양 미세 환경에 영향을 줄 수 있고, 종양 미세 환경도 종양 세포의 성장 또는/및 전이에 영향을 줄 수 있다. 예를 들어, 면역 반응을 유도하는 질환이 암인 경우, 시료의 클론성이 에피토프의 면역원성에 영향을 줄 수 있으며, 면역 세포의 분포/비율, 종양 침윤 림프구(tumor infiltrating lymphocyte), RNA 발현(expression) 등도 생성되는 에피토프의 면역원성에 영향을 줄 수 있다.

[31]

[32] 도 2는 에피토프의 면역원성에 영향을 미칠 수 있는 다수의 생물학적 프로세스의 요소 및 에피토프 특성을 나타낸 것이다.

[33] 에피토프의 면역원성에는 도 1에서 설명한 생물학적 프로세스 외에 에피토프의 특성(properties)도 영향을 줄 수 있다. 예를 들어, 에피토프의 독성(toxicity) 및 알러지원성(allegenicity) 등 인체에 미치는 부정적인 요소들은 에피토프의 면역원성에 영향을 줄 수 있다. 에피토프의 면역원성에 영향을 미치는 요소들은 한 번에 하나씩 영향을 미칠 수도 있지만 여러가지가 동시에 또는/및 순차적으로 영향을 미칠 수도 있다.

[34] 이하에서는 에피토프 특성 및 생물학적 프로세스에서 에피토프의 면역원성에 영향을 줄 수 있는 여러가지 요소들에 대해 자세히 설명한다.

[35] 에피토프 특성 중 면역원성에 영향을 줄 수 있는 요소로는 돌연변이에 의한 기능적 영향(functional impact by mutations), 인산화(phosphorylation, PTM), 소수성(hydrophobicity or physicochemical hydrophobicity), 알려진 에피토프와의 유사성(Similarity to known epitopes), 사람을 참조로 한 자기-단백체와의 비유사성(Dissimilarity to self (human reference)-proteome), 펩티드 안정성(Peptide stability) 등이 있다.

[36] 이하에서는 각 요소가 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 특정 프로그램/알고리즘/툴 등을 예로 설명하나, 이에 한정되는 것은 아니다.

[37] 돌연변이에 의한 기능적 영향은 종양의 원인이 되는 체세포 돌연변이에 의해 단백질의 기능에 미치는 영향을 나타내는 요소이다. 기능적 영향이 큰 돌연변이를 포함한 에피토프는 기존 면역 세포에 의해 자연스럽게 제거되나, 기능적 영향이 상대적으로 적은 돌연변이를 포함한 에피토프는 더 높은 빈도로 면역원성을 보일 수 있다. 돌연변이에 의한 기능적 영향이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, PROVEAN 또는/및 PolyPhen-2라는 프로그램을 이용할 수 있다.

[38] 인산화는 면역원성을 예측하려는 에피토프의 MHC 클래스 I 리간드가 인산화된 정도를 나타내는 요소이다. 인산화가 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, NetMHCphosPan 프로그램을 이용할 수 있다.

[39] 소수성은 면역원성을 예측하려는 에피토프가 소수성인지를 나타내는 요소이다. 소수성이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 ProtFP라는 프로그램을 이용할 수 있다.

[40] 알려진 에피토프와의 유사성은 면역원성을 예측하려는 에피토프와 면역원성(또는 비면역원성)으로 알려진 에피토프와 유사한지를 나타내는 요소이다. 알려진 에피토프와의 유사성이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, blastP 프로그램을 이용할 수 있다.

[41] 사람을 참조로 한 자기-단백체와의 비유사성은 면역원성을 예측하려는 에피토프가 자기-단백체와의 비유사한지를 나타내는 요소이다. 여기서 자기-단백체는 사람을 참조로 한 자기-단백체일 수 있다. 사람을 참조로 한 자기-단백체와의 비유사성이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, 파이썬 바이오 패키지(Python Bio package)의 pairwise2 모듈을 이용할 수 있다.

[42] 에피토프 안정성(또는 펩티드 안정성)은 에피토프 자체의 안정성을 나타내는 요소이다. 기존에는 펩티드-MHC의 결합 안정성이 사용되고 있었으나, 에피토프 자체의 안정성이 사용되지는 않았다. 에피토프의 안정성이 낮을 경우, MHC 분자에 결합/상호 작용할 기회가 줄어들어 면역원성을 가질 가능성이 낮을 수 있다. 에피토프 안정성이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, ProtParam을 이용할 수 있다.

[43] 생물학적 프로세스 중 항원 처리 단계에서는 프로테아좀 분절(Proteasomal cleavage) 및/또는 TAP 수송체 효율(TAP transporter efficiency)을 에피토프의 면역원성을 예측하는데 이용할 수 있다. 프로테아좀 분절과 TAP 수송체 효율이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, NetCTLpan 프로그램을 이용할 수 있다.

[44] 생물학적 프로세스 중 항원 제시 단계에서는 MHC I 결합 친화도 및/또는 pMHC 안정도 중 적어도 일부가 면역 반응을 일으키는 에피토프를 예측하는데 이용될 수 있다.

[45] MHC 클래스 I 결합 친화도는 면역원성을 예측하려는 에피토프가 MHC 클래스 I와 결합하려는 정도를 나타내는 요소이다. MHC 클래스 I 결합 친화도가 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, 프로그램 NetMHCpan 또는/및 MHCflurry을 이용할 수 있다.

[46] pMHC 안정도는 에피토프의 펩티드-MHC 안정도를 나타내는 요소이다. pMHC 안정도가 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, NetMHCstabpan 프로그램을 이용할 수 있다.

[47] 생물학적 프로세스 중 면역원성 단계에서는 면역원성이 면역 반응을 일으키는 에피토프를 예측하는데 이용될 수 있다. 면역원성이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, IEDB, PRIME, DeepHLApan, DeepImmuno 중 적어도 하나를 이용할 수 있다.

[48] 종양 미세 환경에서는 염증 반응(inflammatory response), B-세포 선형(linear) 에피토프, B-세포 형태 이상(conformational) 에피토프 중 적어도 하나가 면역 반응을 유발하는 에피토프를 예측하는 요소로 고려될 수 있다.

[49] 면역 세포는 사이토카인(cytokine) 등 염증 물질을 분비해 면역 반응을 조절할 수 있다. 에피토프가 이러한 사이토카인을 유도할 수 있는 성질을 계산해, 에피토프가 사이토카인을 과다하게/과소하게 유도할 경우, 면역 반응이 억제될 수 있다. 염증 반응이 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, PIP-EL, AIPpred, IFNepitope 중 적어도 하나를 이용할 수 있다.

[50] 종양의 원인이 되는 체세포 돌연변이(somatic mutation)는 B-세포에 의해 인식될 수 있다. B-세포는 펩티드 형태의 선형 에피토프와 단백질 구조적인 형태 이상 에피토프를 모두 인식할 수 있다. B-세포 선형 에피토프는 선형 에피토프 가능성을 나타낸 요소이고, B-세포 형태 이상 에피토프는 형태 이상 에피토프 가능성을 나타낸 요소이다. B-세포 선형 에피토프가 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, Bepipred 프로그램을 이용할 수 있고, B-세포 형태 이상 에피토프가 에피토프의 면역원성에 영향을 주는 정도를 확인하기 위해 예를 들어, Discotope를 이용할 수 있다.

[51]

[52] 도 3은 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 학습하는 과정의 일 예를 나타낸 것이다.

[53] 신생항원의 CD8+ T-세포 면역원성을 학습하고 평가하기 위해, 면역원성 돌연변이 에피토프와 비면역원성 돌연변이 에피토프가 이용될 수 있다. 면역원성 돌연변이 에피토프와 비면역원성 돌연변이 에피토프는 기 공개된 데이터 및/또는 자체 보유/생산한 데이터일 수 있다. 기 공개된 데이터는 대표적으로, dbPepNeo, NEPdb 등을 이용해 획득할 수 있으며, 돌연변이 에피토프와 원시의 에피토프의 차이(differential index)를 계산해야 하기 때문에, 데이터는 종양 신생항원 에피토프로 한정하여 수집할 수 있다. 즉, 돌연변이 에피토프와 원시의 에피토프의 정

보를 직간접적으로 추출할 수 없는 병원체(pathogen) 유래 에피토프는 differential index를 계산할 수 없기 때문에 제외할 수 있다.

[54] 돌연변이 에피토프의 면역원성에 영향을 미치는 다수의 생물학적 프로세스에 관여하는 요소 및 돌연변이 에피토프의 특성 등은 크게 돌연변이 에피토프의 특성, 항원 처리, 항원 제시, 면역(또는 T-세포 인식), 종양 미세 환경(TME) 등이 있으며, 자세하게는 돌연변이에 의한 기능적 영향, PTM(인산화 반응, 유비퀴틴화, 글리코실화반응 등), 알려진 에피토프와 유사한 서열, 자기-단백체와의 비유사성, 에피토프(펩티드) 안정도, 프로테아좀 분절, TAP 수송체 효율, MHC 클래스 I/II 결합 친화도, pMHC 안정도, 면역(T-세포 수용체 인식), 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 등이 있다. 만약, 다루려는 질환이 암인 경우, 시료의 클론성(clonality)이 더 포함될 수 있으며, 염증 외의 TME 요소(면역세포의 분포/비율, 종양 침윤성 림프구, RNA 발현 등)도 더 포함될 수 있다. 또한, 에피토프의 독성(toxicity) 및/또는 알러지원성(allegenicity) 등 인체에 미치는 부정적인 요소들도 더 포함될 수 있다. 일 실시예에 따르면, 다수의 생물학적 프로세스 및 에피토프 특성들 중에 부분적으로 선택되어 활용되어 질 수 있다.

[55] 돌연변이 에피토프의 면역원성에 관여하는 상세 요소들 중, differential index를 계산할 수 있는 요소는 인산화, 소수성, 알려진 에피토프와의 유사성, 자기-단백체와의 비유사성, 에피토프 안정도, 항원 처리, MHC I 결합 친화도, pMHC 안정도, 면역, 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 중 적어도 하나일 수 있다. 돌연변이에 의한 기능적 영향은 differential index를 계산할 수 있는 요소에서 제외되는데 그 이유는 돌연변이에 의한 기능적 영향을 측정한 값 자체가 원시의 에피토프 대비 돌연변이 에피토프가 단백질에 미치는 영향을 측정하고 있기 때문이다. 일 실시예에 따르면, 항원 처리는 하나의 요소가 될 수도 있지만, 프로테아좀 분절과 TAP 수송체 효율 두 개의 요소로 나뉠 수도 있다.

[56] 학습은 모든 요소를 동시에 통합하여 수행될 수도 있지만 각 요소별로 따로 학습될 수도 있다. 예를 들어, 모든 요소가 매칭된 데이터가 부족한 경우, 데이터가 존재하는 요소에 대해서만 학습이 수행될 수도 있다. 학습에 입력될 데이터는 각 요소를 프로그램/알고리즘/툴을 이용해 측정한 값이 될 수 있다. 프로그램/알고리즘/툴은 도 2에서 설명한 것 중 적어도 하나일 수 있고, 또는 별도로 개발된 것일 수도 있다. 만약, 복수의 프로그램/알고리즘/툴로 하나의 요소를 측정하는 경우에는 각각을 입력으로 할 수도 있다. 또 다른 예에서는 하나의 프로그램/알고리즘/툴로 복수의 요소를 측정하여 입력할 수도 있다.

[57] 학습에 이용할 면역원성/비면역원성 에피토프는 돌연변이 에피토프와 원시의 에피토프 쌍으로 구성되어 있으며, 수집 또는 개발한 알고리즘/툴을 활용하여 각 요소별 돌연변이 에피토프와 원시의 에피토프에 대해 도출한 값을 이용하여 돌연변이 에피토프의 면역원성을 예측할 수 있다.

[58] 본 발명에서는, 각 요소에 대한 돌연변이 에피토프와 원시의 에피토프의 differential index를 RATIO(비율), DIFFERENCE(차이)로 계산하여 학습에 입력으

로 할 수 있으며, 하나를 이용하거나 두가지를 모두 이용할 수 있다. 아래의 [수학식 1]과 [수학식 2]는 RATIO와 DIFFERENCE를 수식으로 나타낸 것이다.

[59] [수학식 1]

[60] $$RATIO = \frac{\text{돌연변이 에피토프에 대해 프로그램/알고리즘/툴로 측정한 값}}{\text{원시의 에피토프에 대해 프로그램/알고리즘/툴로 측정한 값}}$$

[61] [수학식 2]

[62] DIFFERENCE = 돌연변이 에피토프에 대해 프로그램/알고리즘/툴로 측정한 값 - 원시의 에피토프에 대해 프로그램/알고리즘/툴로 측정한 값

[63] 다른 실시예에서는, [수학식 1]과 [수학식 2] 외에 기본 수식의 파생형인 LOG를 적용한 값을 이용할 수도 있다.

[64] 돌연변이 에피토프의 면역원성에 영향을 미치는 다수의 생물학적 프로세스에 관여하는 여러가지 요소 및 돌연변이 에피토프의 특성 등에 대한 differential index는 아래와 같이 명명할 수 있으나, 이에 제한되는 것은 아니다.

[65] - Phosphorylation (one of PTM)(인산화) : Differential phosphorylation index (DPI)

[66] - Hydrophobicity(소수성) : Differential hydrophobicity index (DHI)

[67] - Similarity to known epitopes(알려진 에피토프와의 유사성) : Differential sequence similarity index (DSSI)

[68] - Dissimilarity to self-proteome(자기-단백체와의 비유사성) : Differential dissimilarity index (DDI)

[69] - Peptide (epitope) stability(에피토프 안정도) : Differential peptide stability index (DPSI)

[70] - Proteasomal cleavage(프로테아좀 분절) : Differential cleavage index (DCI)

[71] - TAP transport efficiency(TAP 수송체 효율) : Differential TAP index (DTI)

[72] - Antigen processing(항원 처리) : Differential antigen processing index (DAPI)

[73] - MHC I binding affinity(MHC I 결합 친화도) : Differential agretopicity index (DAI)

[74] - pMHC stability(pMHC 안정도) : Differential stability index (DSI)

[75] - Immunogenicity(T-cell receptor interaction)(면역) : Differential immunogenicity index (DII)

[76] - Inflammatory response(염증 반응) : Differential inflammatory response index (DIRI)

[77] - B-cell linear epitope(B-세포 선형 에피토프) : Differential B-cell linear epitope index (DBLI)

[78] - B-cell conformationl epitope(B-세포 형태 이상 에피토프) : Differential B-cell conformationl epitope index (DBCI)

[79] 돌연변이 에피토프에 대해 프로그램/알고리즘/툴을 이용하여 다양한 생물학적 프로세스에 관여하는 요소 및 에피토프의 특성을 측정한 값, [수학식 1] 및 [수학식 2]로 계산된 값 중 통계적으로 유의한 요소들을 선별하고 인코딩하여 기계 학

습 또는 통계 방식으로 학습시킨다. 인코딩 전에 각 요소들, [수학식 1] 및 [수학식 2]로 계산된 값에 대해 표준화(standardization) 및/또는 정규화(normalization)를 수행할 수 있다. 프로그램/알고리즘/툴을 이용해 도출한 값은 각 요소들 또는 적용한 프로그램/알고리즘/툴에 따라 값의 범위나 분해능(resolution)이 다를 수 있어 표준화 또는/및 정규화하여 이용할 수 있다.

[80] 각 요소들에 대한 RATIO 및/또는 DIFFERENCE는 모두 또는 통계 평가를 통해 부분적으로 선택되어 활용될 수 있다. 예를 들어, 모든 요소들에 대한 RATIO와 DIFFERENCE를 동시에 사용하여 학습할 수도 있고, 또는/및 그 중 통계적 유의성을 검토/평가하여, 부분적으로 유의한 값들만 사용하여 학습할 수도 있다. 일 실시예에 따르면, 프로그램/알고리즘/툴 또는/및 개별 요소를 가감하는 것은 대상으로 하는 질환에 따라 달라질 수 있다. 예를 들어, 암 종에 따라 서로 다른 요소 및/또는 프로그램/알고리즘/툴의 조합을 형성할 수 있다.

[81] 학습에 이용될 수 있는 인공 지능 모델은 neural network나 deep learning, logistic regression, AdaBoost, LogitBoost, XgBoost, Support vector machine, random forest 등이 활용될 수 있으며, 특정 기계 학습 알고리즘이나 통계 방법으로 한정하지 않는다. 또한, 여러 기계 학습 알고리즘이나 통계 방법으로 통합하는 앙상블 기계 학습 방식(e.g., bagging, stacking, voting, boosting 등)도 활용될 수 있다. 이렇게 기계 학습 방식으로 통합 학습시킴으로써, 최종 목적인 면역원성 돌연변이 에피토프를 예측하는 요소들의 가중치(weight)가 결정되고, 모든 생물학적 프로세스에 관여하는 요소들, 에피토프의 특징과 더불어 돌연변이 에피토프와 원시의 에피토프의 차이를 이용하여 학습할 수 있게 된다.

[82] 이상의 학습을 통해 생성된 인공 지능 모델은, 기존 알려진 cross validation, leave-one-out validation 또는 독립 데이터를 활용한 검증(validation)을 통해 성능 평가를 할 수 있으며, 예측된 돌연변이 에피토프를 대상으로 in vitro, in vivo, 임상 등 실험적 방식으로 성능을 평가할 수 있다.

[83] 인공 지능 모델의 결과는 돌연변이 에피토프에 대한 면역원성의 예측치이다. 돌연변이 에피토프에 대한 면역원성의 예측치는 에피토프가 면역원성인지 또는 비면역원성인지를 확률로 나타낸 것으로 0에서 1 사이의 값일 수 있다. 예를 들어, 돌연변이 에피토프에 대한 면역원성의 예측치가 1에 가까우면 면역원성으로 판단할 수 있고, 0에 가까우면 비면역원성으로 판단할 수 있다.

[84] 다른 일 실시예에 따르면, 인공 지능 모델의 결과는 돌연변이 에피토프가 면역원성인지에 대해 판단된 결과일 수 있다. 인공 지능 모델은 돌연변이 에피토프에 대한 면역원성의 예측치를 기준값과 비교하여 면역원성 또는 비면역원성으로 판단할 수 있다. 기준값은 종양의 종류, 목적, 상황 중 어느 하나 또는 조합에 따라 달라질 수 있다.

[85]

[86] 도 4는 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 예측하는 과정의 일 예를 나타낸 것이다.

[87] 도 4를 참조하면, 본 발명의 일 실시예에 따른 돌연변이 에피토프의 면역원성을 예측하는 방법은 돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계/과정(S410)을 포함할 수 있다.

[88] 돌연변이 에피토프의 면역원성을 예측하는 방법은 상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계/과정(S420)을 포함할 수 있다.

[89] 돌연변이 에피토프의 면역원성을 예측하는 방법은 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index를 계산하는 단계/과정(S430)을 포함할 수 있다. differential index는, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 비율(RATIO), 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 차이(DIFFERENCE) 중 적어도 하나를 포함할 수 있고, 인 산화, 소수성, 알려진 에피토프와의 유사성, 자기-단백체와의 비유사성, 에피토프 안정도, 항원 처리, MHC I 결합 친화도, pMHC 안정도, 면역, 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 중 적어도 하나에 대한 differential index일 수 있다.

[90] 돌연변이 에피토프의 면역원성을 예측하는 방법은 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측하는 단계/과정(S440)을 포함할 수 있다. 돌연변이 에피토프의 면역원성을 예측하는 방법은 상기 학습된 인공 지능 모델에 계산한 값을 입력하기 전, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index에 대해 표준화 및 정규화 중 적어도 하나를 더 수행할 수 있다.

[91]

[92] 도 5는 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 예측하는 장치의 구성도의 일 예를 나타낸 것이다.

[93] 도 5를 참조하면, 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 예측하는 장치는 입력/출력 장치(510), 저장 장치(520) 및 연산 장치(530)로 구성될 수 있다.

[94] 입력/출력 장치(510)(또는, 입출력 장치)는 돌연변이 에피토프를 입력받거나 돌연변이 에피토프의 면역원성을 예측한 결과를 출력하기 위한 구성일 수 있다. 입력/출력 장치(510)는 사용자로부터 또는 다른 장치로부터 돌연변이 에피토프를 입력받기 위한 구성일 수 있다. 여기서는 입력/출력 장치(510)를 하나로 표시하나 별도의 장치로 구성될 수 있다. 예를 들어, 입력 장치는 마우스 또는/및 키보드가 될 수 있고, 출력 장치는 모니터와 같은 디스플레이 또는 스피커가 될 수 있다.

[95] 저장 장치(520)는 돌연변이 에피토프와 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성을 기초로 계산된 differential index와 돌연변이 에피토프에 대해 생물학적 프로세스에 관여하는 요소와 에피토프의 특성을 기초로 계산된 값을 이용하여 돌연변이 에피토프의 면역원성 예측하도록 학습된 인공 지능 모델을 저장을 저장할 수 있다. 저장 장치(520)는 인공 지능 학습 모델 외에 데이터 처리에 필요한 프로그램/알고리즘/툴을 저장할 수 있다.

[96] 연산 장치(530)는 돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산하고, 상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산하고, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index를 계산하고, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측할 수 있다.

[97] 연산 장치(530)는 학습된 인공 지능 모델에 계산한 값을 입력하기 전, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index에 대해 표준화 및 정규화 중 적어도 하나를 더 수행할 수 있다.

[98] 여기서는 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 예측하는 장치가 입력/출력 장치(510), 저장 장치(520) 및 연산 장치(530)로 구성된 것으로 설명하나, 복수의 구성이 하나의 구성으로 될 수도 있고, 하나의 구성이 복수의 구성으로 될 수도 있다. 그 외에도, 인공 지능 모델을 이용해 돌연변이 에피토프의 면역원성을 예측하는 장치는 통신 장치 등을 더 포함할 수 있다.

[99] differential index는, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 비율(RATIO), 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 차이(DIFFERENCE) 중 적어도 하나를 포함할 수 있고, 인산화, 소수성, 알려진 에피토프와의 유사성, 자기-단백체와의 비유사성, 에피토프 안정도, 항원 처리, MHC I 결합 친화도, pMHC 안정도, 면역, 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 중 적어도 하나에 대한 differential index일 수 있다.

[100]

[101] 이하는 앞서 설명한 방법에 따라 돌연변이 에피토프의 면역원성을 예측하는 일 실시예를 나타낸 것이다.

[102] 에피토프의 면역원성을 학습하고 평가하기 위해, 공개된 데이터베이스 및/또는 논문 등으로부터 종양에서 발생한 네오에피토프에 대한 면역원성 에피토프와 비면역원성 에피토프에 관한 데이터를 수집하였다.

[103] 아래의 [표 1]은 학습 데이터 세트(학습 데이터 세트)와 평가 데이터 세트(평가용 데이터 세트)의 개수를 나타낸 것이다.

[104] [표1]

구분	CD8+ T-세포 면역원성 에피토프 (positive)	CD8+ T-세포 비면역원성 에피토프 (negative)
학습 데이터 세트	315	4067
평가 데이터 세트	79	3125

[105] 본 실시예에서는, [표 1]과 같이 dbPepNeo, NEPdb, PRIME 데이터, INeo-Epp 데이터로부터 수집한 총 315개의 면역원성 에피토프(positive)와, 4067개의 비면역원성 에피토프(negative)(즉, 에피토프 및 에피토프와 결합하는 MHC I 대립 형질)를 학습 데이터 세트(training set)로 사용하였고, McPAS-TCR, VDJdb, IEDB t-cell db, TESLA consortium 데이터로부터 수집한 79개의 면역원성 에피토프(positive)와, 3125개의 비면역원성 에피토프(negative)를 평가 데이터 세트(independent set)로 사용하였다. 에피토프의 면역원성을 예측하기 위해, 돌연변이 에피토프와 원시의 에피토프에 대해 [표 2]에서 제시한 프로그램 및 방법으로 각 생물학적 프로세스에 관여하는 요소 및 에피토프 특성을 계산하였다.

[106] [표2]

생물학적 프로세스/ 에피토프 특성	프로그램 / 알고리즘 / 툴	참고문헌
돌연변이에 의한 기능적 영향(Functional impact by the mutation)	PROVEAN	PMID: 23056405
인산화(Phosphorylation (phosphorylated MHC class I ligands))	NetMHCphosPan	10.1016/j.immuno.2021.100005
소수성(Hydrophobicity)	ProtFP	PMID: 24059694
알려진 에피토프와의 유사성 (Similarity to known epitopes)	blastP	-
Dissimilarity to self (human reference)-proteome	“pairwise2”module in Python Bio package	-
펩티드 안정성(Peptide stability)	ProtParam	PMID: 2075190
프로테아좀 분절(Proteasomal cleavage)	NetCTLpan	PMID: 20379710
TAP transporter efficiency	NetCTLpan	PMID: 20379710
Antigen processing(Proteasomal cleavage + TAP transporter efficiency)	MHCflurry (antigen processing score)	PMID: 32711842
MHC I binding affinity	NetMHCpan	PMID: 32406916

	MHCflurry (antigen presentation score)	PMID: 32711842
pMHC 안정(pMHC stability)	NetMHCstabpan	PMID: 27402703
Immunogenicity (T-cell interaction)	IEDB tools	PMID: 24204222
	PRIME	PMID: 33665637
Inflammatory response	PIP-EL	PMID: 30108593
B-cell linear epitope	Bepipred	PMID: 16635264
B-cell conformational epitope	Discotope	PMID: 17001032

[107] 이후, 돌연변이 에피토프와 원시의 에피토프에 대해 [표 2]에서 제시한 프로그램 및 방법으로 각각 계산한 값을 기초로 앞서 설명한 DIFFERENCE과 RATIO를 계산한다. 돌연변이 에피토프에 대해 [표 2]에서 제시한 프로그램 및 방법으로 계산된 값, DIFFERENCE, 및 RATIO 중 적어도 하나를 이용해 면역원성 에피토프와 비면역원성 에피토프를 구별할 수 있는 구분능을 AUC와 prAUC, P-value (Wilcoxon rank-sum test, $P<0.05$)를 기준으로 확인하였다. 이 중 통계적으로 의미있는 요소들만을 인공 지능 학습 모델에 입력하였다.

[108] [표3]

Methods/ Programs	Epitope			DIFFERENCE			RATIO		
	AUC	prAUC	P-value	AUC	prAUC	P-value	AUC	prAUC	P-value
Similarity to known epitopes	0.561052	0.14287	2.91E-04	0.518993	0.082119	0.261265	0.523862	0.079501	0.158083
Dissimilarity to self-proteome	0.459067	0.066936	0.015349	0.458521	0.065998	0.014032	0.541067	0.081895	0.015003
Hydrophobicity (ProtFP)	0.550065	0.078362	0.003028	0.538348	0.086689	0.022478	0.531443	0.078767	0.06256
NetCTLpan (Cleavage)	0.535254	0.083073	0.036788	0.489895	0.072992	0.552856	0.492999	0.071916	0.682194
NetCTLpan (TAP efficiency)	0.526048	0.079180	0.122752	0.501619	0.075759	0.806286	0.489829	0.071463	0.650069
NetCTLpan (Comb)	0.56176	0.126133	2.55E-04	0.517495	0.079112	0.300154	0.513233	0.072897	0.43322
NetMHCpan (%Rank_EL)	0.652255	0.167535	1.90E-19	0.506969	0.069157	0.679652	0.544356	0.105322	0.008624

NetMHCpan (%Rank_BA)	0.657736	0.148926	9.41E-21	0.49905	0.068368	0.955226	0.562486	0.099488	2.15E-04
MHCflurry (processing score)	0.650284	0.145857	5.56E-19	0.558957	0.103276	4.80E-04	0.53688	0.076464	0.028934
MHCflurry (presentation score)	0.673387	0.177378	9.75E-25	0.556889	0.097804	7.54E-04	0.521614	0.073229	0.200527
NetMHCStabPan (%Rank_Stab)	0.58387	0.122517	5.40E-07	0.518213	0.071701	0.280288	0.56374	0.09354	1.61E-04
PRIME (%Rank)	0.678614	0.162947	2.20E-26	0.558192	0.083278	4.00E-04	0.600326	0.09653	1.55E-09
DeepHLApan (immunogenic score)	0.552553	0.086783	0.001855	0.522287	0.076852	0.186621	0.510505	0.074257	0.533784
NetMHCphosPan (Rnk_EL)	0.625828	0.142283	9.12E-14	0.520097	0.075307	0.233684	0.561617	0.096367	2.62E-04
PIP-EL (inflammatory response)	0.460926	0.065636	0.020728	0.474419	0.066363	0.129793	0.473708	0.066579	0.119485
PROVEAN (functional impact by the mutation	0.431995	0.059677	5.64E-05	-	-	-	-	-	-
B-cell linear epitope (BcellPred)	**0.45141**	**0.062508**	**0.004006**	0.468083	0.069809	0.058817	0.493824	0.075081	0.714663
Peptide stability	**0.455594**	**0.065555**	**0.008785**	0.486047	0.070685	0.494792	0.479056	0.069242	0.272637
B-cell conformational epitope (Discotope)	**0.445959**	**0.06171**	**0.001374**	0.481587	0.072856	0.276434	0.514706	0.071238	0.38363

[109] [표 3]을 참조하면, 돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소 및 에피토프 특성 중 TAP 효율을 제외한 모든 요소들이 면역원성 에피토프인지를 판단하는데 있어 통계적으로 의미있는 요소이다. DIFFERENCE에 대해서는 생물학적 프로세스에 관여하는 요소 및 에피토프 특성 중 자기-단백체와의 비유사성, 소수성, MHCflurry (processing score), MHCflurry (presentation score), PRIME (%Rank)가 면역원성 에피토프인지를 판단하는데 있어 통계적으로 의미있는 요소이다. 또한, RATIO에 대해서는 생물학적 프로세스에 관여하는 요소 및 에피토프 특성 중 자기-단백체와의 비유사성, NetMHCpan (%Rank_EL), NetMHCpan (%Rank_BA), MHCflurry (processing score), NetMHCStabPan

(%Rank_Stab), PRIME (%Rank), NetMHCphosPan (Rnk_EL)이 면역원성 에피토프인지를 판단하는데 있어 통계적으로 의미있는 요소이다.

[110] 학습 데이터 세트/평가 데이터 세트의 에피토프에 대해 통계적으로 유의한 총 30개의 요소(면역원성 관련 생물학적 프로세스 및 에피토프 특성)을 2차원 매트릭스화하여, 기계 학습 분석(인공 지능 학습 모델을 이용한 분석)을 수행하였다. 본 실시예에서는 Weka 프로그램을 이용하여, 여러 기계 학습 알고리즘 분석을 수행하였고, 최종적으로 여러 기계 학습 알고리즘을 통합하여 선택하는 방식으로 최종 모델을 생성하였다. 본 실시예에서 최종 모델은 LogitBoost, BayesNet, CSForest, AdaBoost, Logistic, PART, NaiveBayes, RandomForest, SMO의 9개 기계 학습 알고리즘에 의한 모델 결과를 평균 확률을 기준으로 선택하였고, 본 발명의 성능 향상 정도를 비교 평가하기 위해, 동일한 데이터에 대해, 공개된 면역원성 관련 프로그램들을 수행하고 성능을 비교하였다.

[111] 아래의 [표 4]는 학습 데이터 세트에 대한 에피토프 면역원성 예측 성능 평가 결과를 나타낸 것이고, [표 5]는 평가 데이터 세트에 대한 에피토프 면역원성 예측 성능 평가 결과를 나타낸 것이다.

[112] [표4]

알고리즘	Threshold	Training DB (P:315, N:4067)					
		F-Measure	Precision	Recall	Accuracy	ROC Area (=AUC)	PRC Area (=prAUC)
NetMHCpan (%Rank_EL)	< 0.5	**0.193**	**0.116**	**0.565**	**0.660**	**0.652**	**0.168**
MhcFlurry-2.0 (presentation_score)	> 0.5	**0.179**	0.103	0.689	0.546	0.673	0.177
PRIME (%Rank)	< 0.5	**0.210**	0.131	0.524	0.716	0.679	0.163
DeepHLApan (Immunogenic_score)	> 0.5	**0.145**	0.083	0.590	0.499	0.553	0.087
IEDB tool	> 0.0	**0.137**	0.078	0.578	0.478	0.532	0.080
DeepImmuno	> 0.5	0.158	0.087	0.934	0.187	0.548	0.091
TESLA 컨소시움 기준	-	0.194	0.164	0.238	0.858	0.572	0.228
본 발명 (NeoPro-Onco)	> 0.5	0.302	0.259	0.362	0.880	0.724	0.234

[113] [표5]

알고리즘	Threshold	Independent test DB (P:79, N:3125)					
		F-Measure	**Precision**	**Recall**	**Accuracy**	**ROC Area (=AUC)**	**PRC Area (=prAUC)**
NetMHCpan (%Rank_EL)	< 0.5	0.285	0.176	0.747	0.908	0.942	0.317
MhcFlurry-2.0 (presentation_score)	> 0.5	0.289	0.175	0.835	0.899	0.949	0.325
PRIME (%Rank)	< 0.5	0.323	0.208	0.722	0.925	0.919	0.328
DeepHLApan (Immunogenic_score)	> 0.5	0.059	0.031	0.468	0.629	0.538	0.032
IEDB tool	> 0.0	0.061	0.032	0.633	0.519	0.598	0.032
DeepImmuno	> 0.5	0.110	0.058	0.892	0.139	0.608	0.099
TESLA 컨소시움 기준	-	0.382	0.385	0.380	0.970	0.682	0.390
본 발명 (NeoPro-Onco)	> 0.5	0.415	0.358	0.494	0.966	0.955	0.429

[114] 본 발명의 경우, 10-fold cross-validation 수행 결과이며, 기존 공개된 프로그램 및 알고리즘의 경우, 전체 데이터를 대상으로 평가 수행한 결과이다. [표 4]를 참조하면, 본 발명에 따른 방법이 기존 공개된 프로그램 및 알고리즘 방법 대비, AUC 및 prAUC, F-measure 등에서 더 좋은 예측 정확도를 보임을 확인할 수 있다. [표 5]를 참조하면, 학습에 사용되지 않은 평가 데이터 세트에 대해서도 본 발명에 따른 방법이 기존 공개된 프로그램 및 알고리즘 방법 대비, AUC 및 prAUC, F-measure 등에서 더 좋은 예측 정확도를 보임을 확인할 수 있다.

[115] [표 6]은 모든 요소를 포함하는 모델과 DIFFERENCE와 RATIO를 제외한 모델의 성능을 비교한 결과를 나타낸 것이다.

[116] [표6]

Algorithm	**Threshold**	**Training DB (P:315, N:4067)**		Independent test DB (P:79, N:3125)	
		ROC Area (=AUC)	**PRC Area (=prAUC)**	**ROC Area (=AUC)**	**PRC Area (=prAUC)**

모든 요소를 고려한 모델	> 0.5	0.724	0.234	0.955	0.429
12개의 differential index를 제외한 모델	> 0.5	0.730	0.226	0.954	0.408

[117] 면역원성/비면역원성 에피토프의 샘플수가 동일하지 않기 때문에, 더 적합한 성능 비교방법인 prAUC 기준으로 비교하였다. [표 6]을 참조하면, DIFFERENCE와 RATIO를 포함한 모델이 DIFFERENCE와 RATIO가 제외된 모델보다 더 높은 성능을 보여 Differential index(DIFFERENCE와 RATIO)가 면역원성 에피토프를 예측하는데 기여함을 확인할 수 있다.

[118]

[119] 이상, 본 발명내용의 특정한 부분을 상세히 기술하였는 바, 당업계의 통상의 지식을 가진 자에게 있어서, 이러한 구체적인 기술은 단지 바람직한 실시양태일 뿐이며, 이에 의해 본 발명의 범위가 제한되는 것이 아닌 점은 명백할 것이다. 따라서 본 발명의 실질적인 범위는 첨부된 청구항들과 그것들의 등가물에 의해 정의된다고 할 것이다.

청구범위

[청구항 1] 돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계;
상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산하는 단계;
상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index를 계산하는 단계;
상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측하는 단계;를 포함하는 돌연변이 에피토프의 면역원성을 예측하는 방법.

[청구항 2] 제1항에 있어서,
상기 학습된 인공 지능 모델에 계산한 값을 입력하기 전, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index에 대해 표준화 및 정규화 중 적어도 하나를 수행하는 단계를 더 포함하는 돌연변이 에피토프의 면역원성을 예측하는 방법.

[청구항 3] 제1항에 있어서, 상기 differential index는,
상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 비율(RATIO), 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값의 차이(DIFFERENCE) 중 적어도 하나를 포함하는, 돌연변이 에피토프의 면역원성을 예측하는 방법.

[청구항 4] 제1항에 있어서, 상기 differential index는,
인산화, 소수성, 알려진 에피토프와의 유사성, 자기-단백체와의 비유사성, 에피토프 안정도, 항원 처리, MHC I 결합 친화도, pMHC 안정도, 면역, 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 중 적어도 하나에 대한 differential index인, 돌연변이 에피토프의 면역원성을 예측하는 방법.

[청구항 5] 제1항에 있어서, 상기 돌연변이 에피토프와 상기 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 면역원성에 영향을 주는 정도를 계산하는 단계는,
알고리즘, 프로그램, 또는 툴을 이용해 계산하는 단계인, 돌연변이 에피토프의 면역원성을 예측하는 방법.

[청구항 6] 제1항에 있어서, 상기 생물학적 프로세스는,
항원 처리 단계, 항원 제시 단계, 면역 단계, 및 종양 미세 환경으로 구성되고,

상기 종양 미세 환경에 관여하는 요소는 염증 반응, B-세포 선형 에피토프, B-세포 형태 이상 에피토프 중 적어도 하나인, 돌연변이 에피토프의 면역원성을 예측하는 방법.

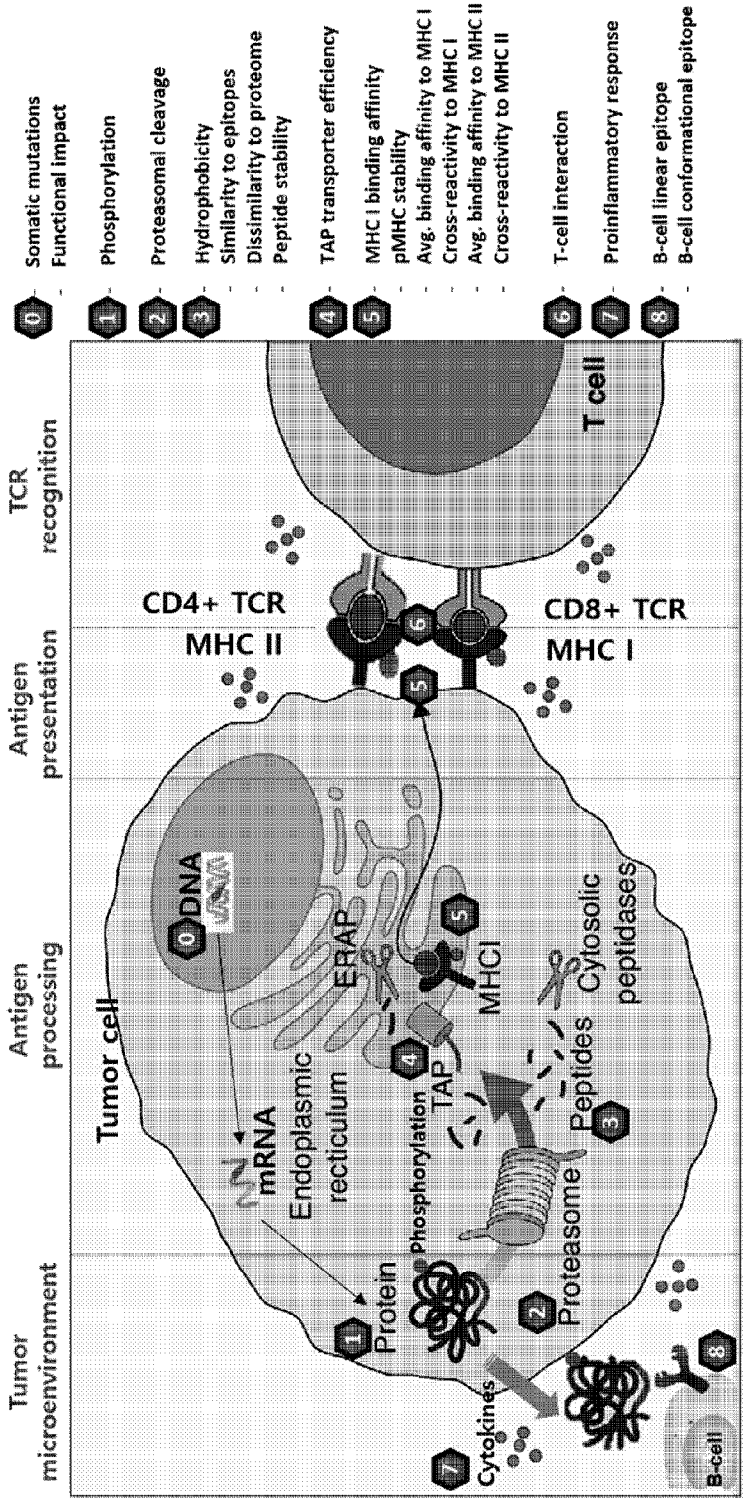
[청구항 7] 제1항 내지 제6항 중 어느 한 항에 따른 방법을 실행하기 위한 컴퓨터 프로그램을 기록하는 컴퓨터 판독 가능한 기록 매체.

[청구항 8] 돌연변이 에피토프를 입력받거나, 상기 돌연변이 에피토프의 면역원성을 예측한 결과를 출력하는 입출력 장치;

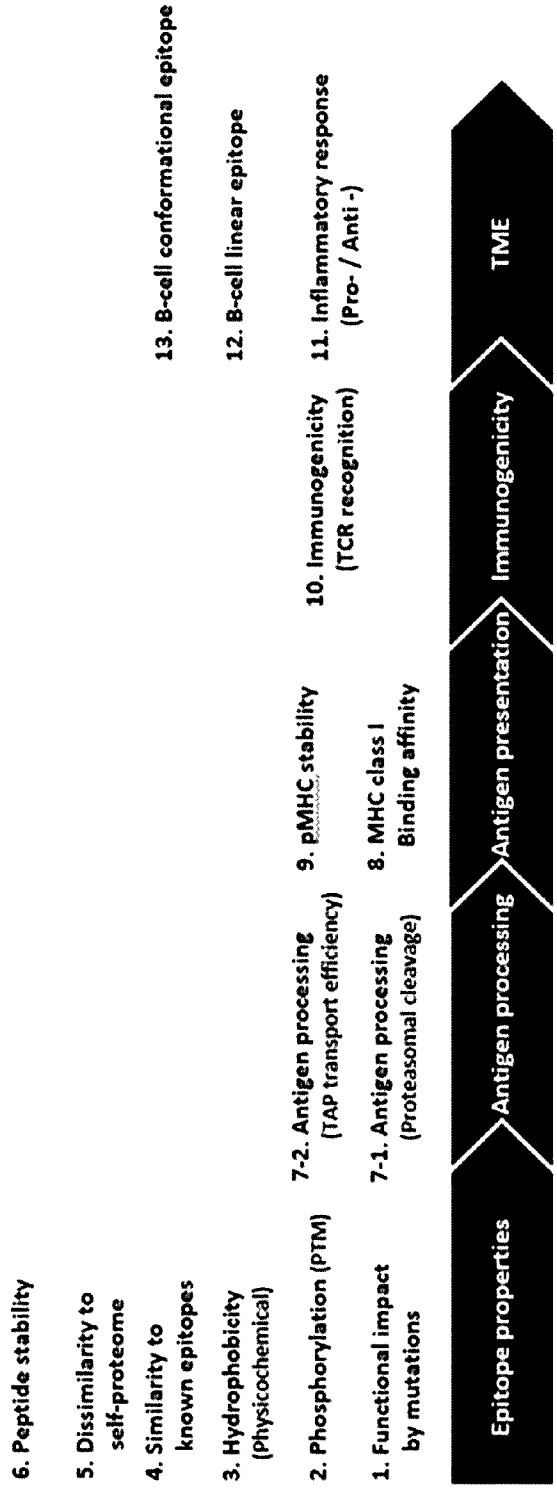
돌연변이 에피토프와 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성을 기초로 계산된 differential index와 돌연변이 에피토프에 대해 생물학적 프로세스에 관여하는 요소와 에피토프의 특성을 기초로 계산된 값을 이용하여 돌연변이 에피토프의 면역원성 예측하도록 학습된 인공 지능 모델을 저장하는 저장 장치; 및

돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산하고, 상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산하고, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index를 계산하고, 상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측하는 연산 장치를 포함하는, 돌연변이 에피토프의 면역원성을 예측하는 장치.

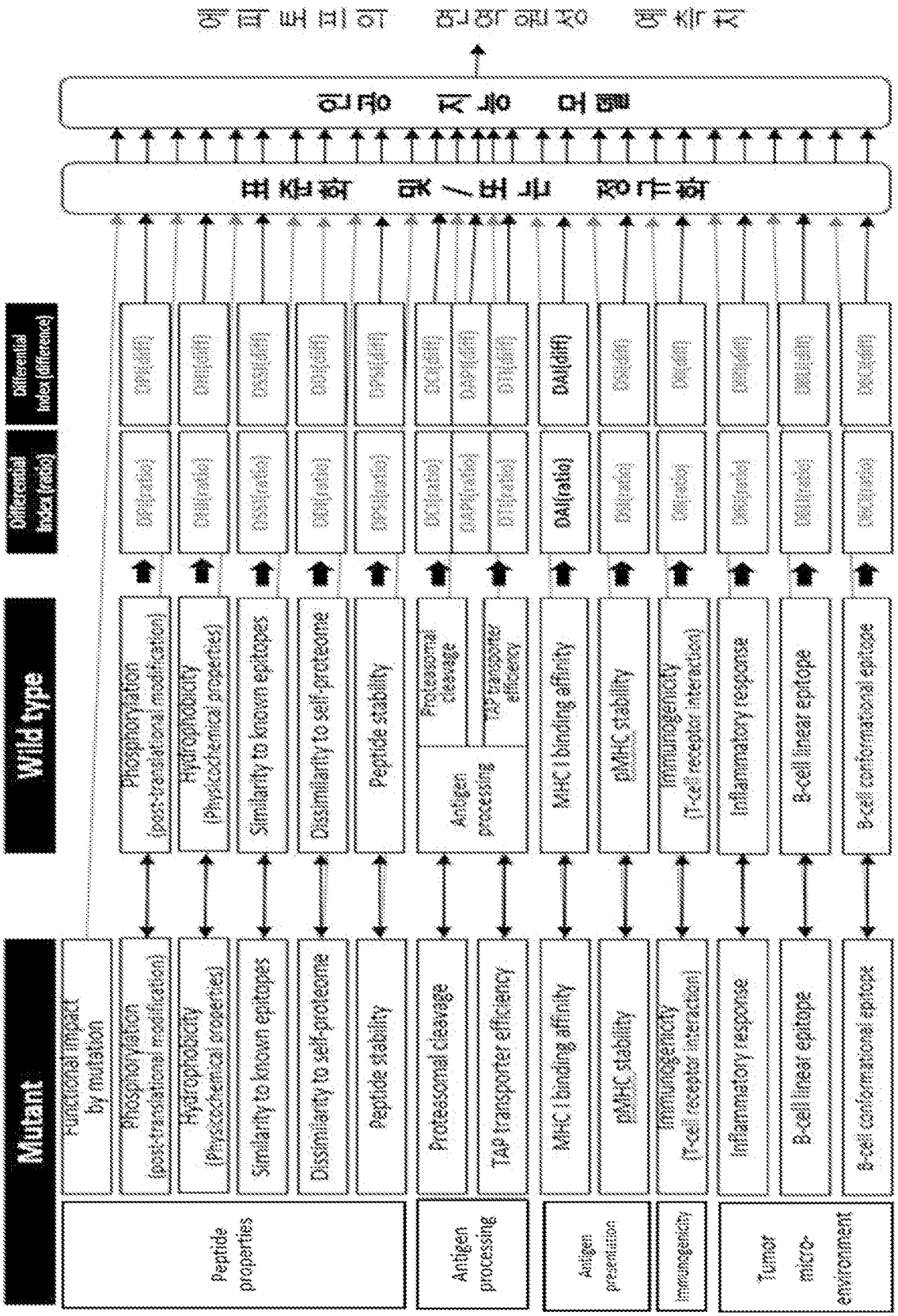
[도1]


Somatic mutations
Functional impact
Phosphorylation
Proteasomal cleavage
Hydrophobicity
Similarity to epitopes
Dissimilarity to proteome
Peptide stability
TAP transporter efficiency
MHC I binding affinity
pMHC stability
Avg. binding affinity to MHC I
Cross-reactivity to MHC I
Avg. binding affinity to MHC II
Cross-reactivity to MHC II
T-cell interaction
Proinflammatory response
B-cell linear epitope
B-cell conformational epitope
TCR recognition
Antigen presentation
Antigen processing
Tumor microenvironment
T cell
CD4+ TCR
MHC II
CD8+ TCR
MHC I
Tumor cell
DNA
mRNA
Endoplasmic reticulum
ERAP
Phosphorylation
TAP
MHCI
Peptides
Cytosolic peptidases
Protein
Proteasome
Cytokines
B-cell

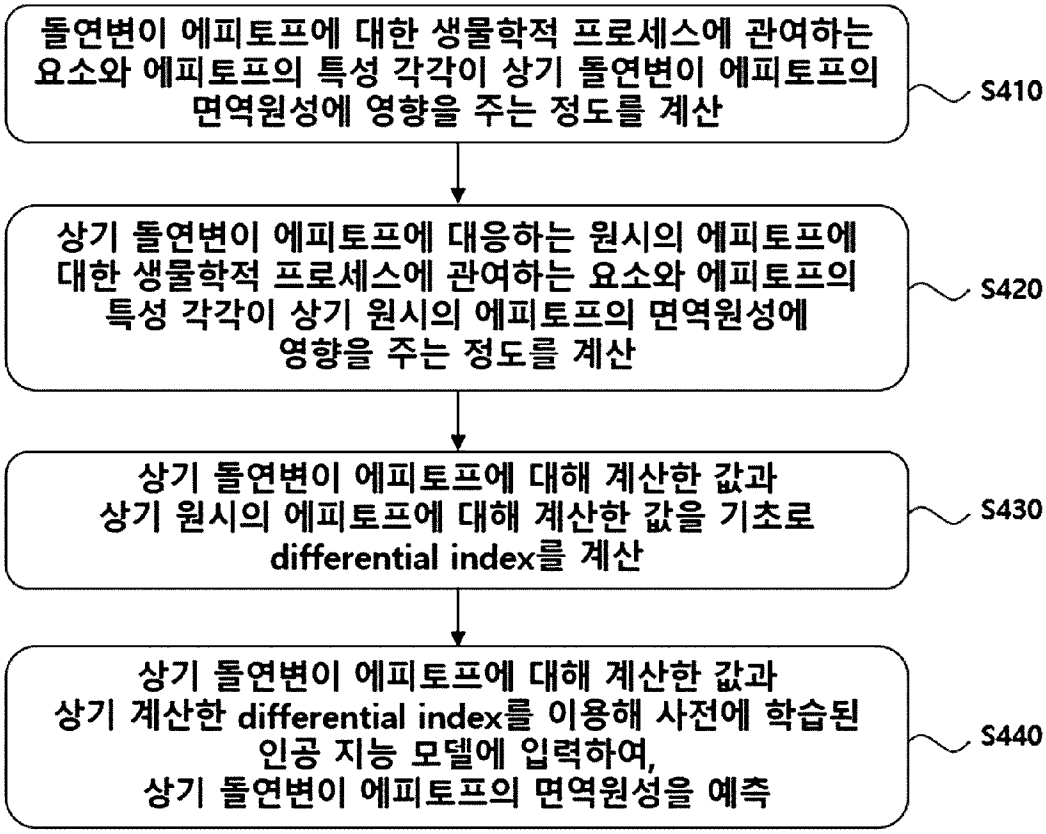
[도2]


6. Peptide stability
5. Dissimilarity to self-proteome
4. Similarity to known epitopes
3. Hydrophobicity (Physicochemical)
2. Phosphorylation (PTM)
1. Functional impact by mutations
Epitope properties
7-2. Antigen processing (TAP transport efficiency)
7-1. Antigen processing (Proteasomal cleavage)
Antigen processing
9. pMHC stability
8. MHC class I Binding affinity
Antigen presentation
13. B-cell conformational epitope
12. B-cell linear epitope
10. Immunogenicity (TCR recognition)
Immunogenicity
11. Inflammatory response (Pro- / Anti -)
TME

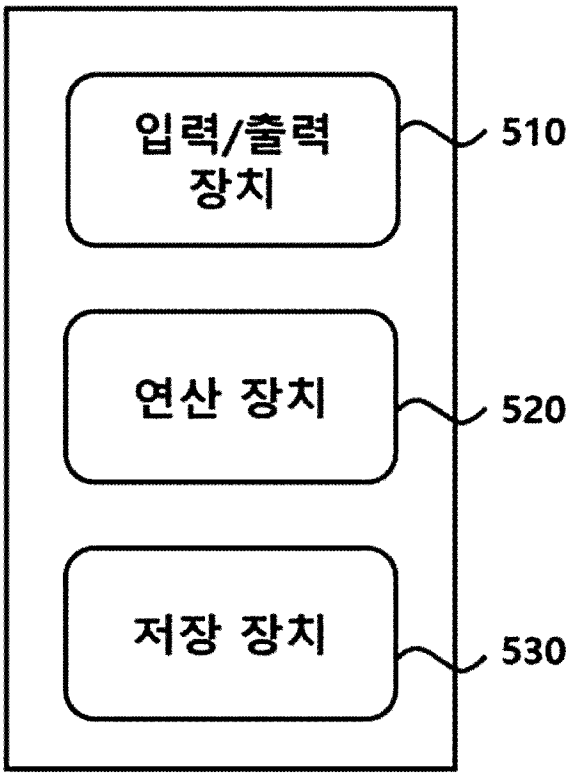
[도 3]


Mutant
Wild type
Differential index (ratio)
Differential index (difference)
Functional impact by mutation
Phosphorylation (post-translational modification)
Hydrophobicity (Physicochemical properties)
Similarity to known epitopes
Dissimilarity to self-proteome
Peptide stability
Proteasomal cleavage
TAP transporter efficiency
MHC I binding affinity
pMHC stability
Immunogenicity (T-cell receptor interaction)
Inflammatory response
B-cell linear epitope
B-cell conformational epitope
Peptide properties
Antigen processing
Antigen presentation
Immunogenicity
Tumor micro-environment

[도4]


돌연변이 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 돌연변이 에피토프의 면역원성에 영향을 주는 정도를 계산
S410
상기 돌연변이 에피토프에 대응하는 원시의 에피토프에 대한 생물학적 프로세스에 관여하는 요소와 에피토프의 특성 각각이 상기 원시의 에피토프의 면역원성에 영향을 주는 정도를 계산
S420
상기 돌연변이 에피토프에 대해 계산한 값과 상기 원시의 에피토프에 대해 계산한 값을 기초로 differential index를 계산
S430
상기 돌연변이 에피토프에 대해 계산한 값과 상기 계산한 differential index를 이용해 사전에 학습된 인공 지능 모델에 입력하여, 상기 돌연변이 에피토프의 면역원성을 예측
S440

[도5]


입력/출력 장치
510
연산 장치
520
저장 장치
530

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2023/017786

A. CLASSIFICATION OF SUBJECT MATTER

G16B 5/00(2019.01)i; **G16B 20/20**(2019.01)i; **G16B 40/00**(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G16B 5/00(2019.01); G16B 20/00(2019.01); G16B 20/20(2019.01); G16B 30/00(2019.01); G16B 40/00(2019.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean utility models and applications for utility models: IPC as above
Japanese utility models and applications for utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & keywords: 신생항원(neoantigen), 에피토프(epitope), 면역원성(immunogenicity), 돌연변이(mutation), 비율(ratio), 차이(difference), 인덱스(index), 원시(wild), 인공지능 모델(artificial intelligence model)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SCHMIDT, J. et al. Prediction of neo-epitope immunogenicity reveals TCR recognition determinants and provides insight into immunoediting. Cell Reports Medicine. 16 February 2021, vol. 2, thesis no. 100194, pp. 1-14, Star Methods (e1-e7). See entire document.	1-8
A	SMITH, C. C. et al. Machine-learning prediction of tumor antigen immunogenicity in the selection of therapeutic epitopes. Cancer Immunology Research. 2019, vol. 7, no. 10, pp. 1591-1604. See entire document.	1-8
A	KR 10-2399419 B1 (BIONTECH SE et al.) 19 May 2022 (2022-05-19) See entire document.	1-8
A	GARTNER, J. J. et al. A machine learning model for ranking candidate HLA class I neoantigens based on known neoepitopes from multiple human tumor types. Nature Cancer. May 2021, vol. 2, pp. 563-574. See entire document.	1-8

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
08 February 2024	**08 February 2024**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office **Government Complex-Daejeon Building 4, 189 Cheongsa-ro, Seo-gu, Daejeon 35208**	
Facsimile No. **+82-42-481-8578**	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2023/017786

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2020-0109269 A (SYNTEKABIO CO., LTD.) 22 September 2020 (2020-09-22) See entire document.	1-8

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/KR2023/017786

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
KR 10-2399419 B1	19 May 2022	AU 2014-264943 A1	26 November 2015
		AU 2014-264943 B2	12 September 2019
		AU 2019-275637 A1	02 January 2020
		AU 2022-200863 A1	03 March 2022
		CA 2911945 A1	13 November 2014
		CA 2911945 C	17 October 2023
		CN 105451759 A	30 March 2016
		CN 105451759 B	05 February 2021
		CN 113219179 A	06 August 2021
		EP 2994159 A1	16 March 2016
		EP 2994159 B1	10 March 2021
		EP 3895727 A1	20 October 2021
		JP 2016-521128 A	21 July 2016
		JP 2020-103297 A	09 July 2020
		JP 2021-121801 A	26 August 2021
		JP 6710634 B2	17 June 2020
		JP 6882550 B2	02 June 2021
		JP 7405792 B2	26 December 2023
		KR 10-2016-0030101 A	16 March 2016
		US 11222711 B2	11 January 2022
		US 2016-0125129 A1	05 May 2016
		US 2022-0093209 A1	24 March 2022
		WO 2014-180490 A1	13 November 2014
		WO 2014-180569 A1	13 November 2014
		ZA 201508048 B	31 May 2017
KR 10-2020-0109269 A	22 September 2020	KR 10-2406699 B1	08 June 2022
		US 2022-0130489 A1	28 April 2022
		WO 2020-185010 A1	17 September 2020

국 제 조 사 보 고 서

국제출원번호
PCT/KR2023/017786

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G16B 5/00(2019.01)i; **G16B 20/20**(2019.01)i; **G16B 40/00**(2019.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G16B 5/00(2019.01); G16B 20/00(2019.01); G16B 20/20(2019.01); G16B 30/00(2019.01); G16B 40/00(2019.01)

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC
일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 신생항원(neoantigen), 에피토프(epitope), 면역원성(immunogenicity), 돌연변이(mutation), 비율(ratio), 차이(difference), 인덱스(index), 원시(wild), 인공지능 모델(artificial intelligence model)

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	SCHMIDT, J. 등, 'Prediction of neo-epitope immunogenicity reveals TCR recognition determinants and provides insight into immunoediting', Cell Reports Medicine, 2021.02.16, 2권, 논문번호 100194, 1-14 페이지, Star Methods(e1-e7) 전체 문헌	1-8
A	SMITH, C. C. 등, 'Machine-learning prediction of tumor antigen immunogenicity in the selection of therapeutic epitopes', Cancer Immunology Research, 2019, 7권, 10호, 1591-1604 페이지 전체 문헌	1-8
A	KR 10-2399419 B1 (비온테크 에스이 등) 2022.05.19 전체 문헌	1-8
A	GARTNER, J. J. 등, 'A machine learning model for ranking candidate HLA class I neoantigens based on known neoepitopes from multiple human tumor types', Nature Cancer, 2021.05, 2권, 563-574 페이지 전체 문헌	1-8

☑ 추가 문헌이 C(계속)에 기재되어 있습니다.

☑ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

"A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌
"D" 본 국제출원에서 출원인이 인용한 문헌
"E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌
"L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌
"O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌
"P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌
"T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌
"X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.
"Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.
"&" 동일한 대응특허문헌에 속하는 문헌

국제조사의 실제 완료일	국제조사보고서 발송일
2024년02월08일(08.02.2024)	**2024년02월08일(08.02.2024)**

ISA/KR의 명칭 및 우편주소

대한민국 특허청
(35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사)

팩스 번호 **+82-42-481-8578**

심사관

허주형

전화번호 **+82-42-481-5373**

국 제 조 사 보 고 서

국제출원번호
PCT/KR2023/017786

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	KR 10-2020-0109269 A ((주)신테카바이오) 2020.09.22 전체 문헌	1-8

국 제 조 사 보 고 서
대응특허에 관한 정보

국제출원번호
PCT/KR2023/017786

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2399419 B1	2022/05/19	AU 2014-264943 A1	2015/11/26
		AU 2014-264943 B2	2019/09/12
		AU 2019-275637 A1	2020/01/02
		AU 2022-200863 A1	2022/03/03
		CA 2911945 A1	2014/11/13
		CA 2911945 C	2023/10/17
		CN 105451759 A	2016/03/30
		CN 105451759 B	2021/02/05
		CN 113219179 A	2021/08/06
		EP 2994159 A1	2016/03/16
		EP 2994159 B1	2021/03/10
		EP 3895727 A1	2021/10/20
		JP 2016-521128 A	2016/07/21
		JP 2020-103297 A	2020/07/09
		JP 2021-121801 A	2021/08/26
		JP 6710634 B2	2020/06/17
		JP 6882550 B2	2021/06/02
		JP 7405792 B2	2023/12/26
		KR 10-2016-0030101 A	2016/03/16
		US 11222711 B2	2022/01/11
		US 2016-0125129 A1	2016/05/05
		US 2022-0093209 A1	2022/03/24
		WO 2014-180490 A1	2014/11/13
		WO 2014-180569 A1	2014/11/13
		ZA 201508048 B	2017/05/31
KR 10-2020-0109269 A	2020/09/22	KR 10-2406699 B1	2022/06/08
		US 2022-0130489 A1	2022/04/28
		WO 2020-185010 A1	2020/09/17