

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일
2022년 12월 1일 (01.12.2022)



(10) 국제공개번호
WO 2022/250513 A1

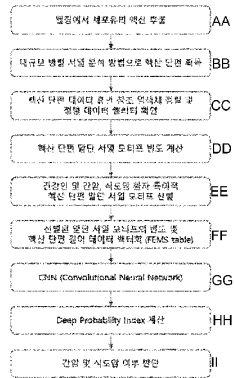
- (51) 국제특허분류:
G16H 50/20 (2018.01) G16H 30/40 (2018.01)
G16H 50/50 (2018.01) G06N 3/08 (2006.01)
G16B 30/10 (2019.01) G06N 3/04 (2006.01)
G16B 40/20 (2019.01)
- (21) 국제출원번호: PCT/KR2022/007651
- (22) 국제출원일: 2022년 5월 30일 (30.05.2022)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2021-0068891 2021년 5월 28일 (28.05.2021) KR
- (71) 출원인: 주식회사 지씨지놈 (GC GENOME CORPORATION) [KR/KR]; 16924 경기도 용인시 기흥구 이현로30번길 107, Gyeonggi-do (KR). 재단법인 아산사회복지재단 (THE ASAN FOUNDATION) [KR/KR]; 05505 서울특별시 송파구 올림픽로43길 88, Seoul (KR). 울산대학교 산학협력단 (UNIVERSITY OF ULSAN FOUN-

DATION FOR INDUSTRY COOPERATION) [KR/KR]; 44610 울산광역시 남구 대학로 93, Ulsan (KR).

- (72) 발명자: 조은혜 (CHO, Eun Hae); 16924 경기도 용인시 기흥구 이현로30번길 107, Gyeonggi-do (KR). 이태림 (LEE, Tae-Rim); 16924 경기도 용인시 기흥구 이현로30번길 107, Gyeonggi-do (KR). 박숙련 (PARK, Sook Ryun); 05649 서울특별시 송파구 양재대로 1218, #204-2102, Seoul (KR).
- (74) 대리인: 이치영 등 (LEE, Cheo Young et al.); 06210 서울특별시 강남구 언주로 430, 16층, 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, CL, CN, CO, CR, CU, 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, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,

(54) Title: METHOD FOR DIAGNOSING CANCER AND PREDICTING CANCER TYPE BY USING TERMINAL SEQUENCE MOTIF FREQUENCY AND SIZE OF CELL-FREE NUCLEIC ACID FRAGMENT

(54) 발명의 명칭: 세포유리 핵산단편 말단 서열 모티프 빈도 및 크기를 이용한 암 진단 및 암 종 예측방법



- AA ... Extract cell-free nucleic acids from plasma
BB ... Obtain nucleic acid fragments by using massive parallel sequencing method
CC ... Align nucleic acid fragment data human reference chromosomes and identify alignment data quality
DD ... Calculate nucleic acid fragment terminal sequence motif frequency
EE ... Select nucleic acid fragment terminal sequence motifs specific to healthy people and patients with liver cancer and esophageal cancer
FF ... Vectorize frequency of selected terminal sequence motifs and nucleic acid fragment length data (FEMS table)
GG ... CNN (Convolutional Neural Network)
HH ... Calculate Deep Probability Index
II ... Determine presence of liver cancer and esophageal cancer

(57) Abstract: The present invention relates to a method for diagnosing cancer and predicting a cancer type, by using a terminal sequence motif frequency and the size of a cell-free nucleic acid fragment and, more particularly, to a method for diagnosing cancer and predicting a cancer type, by using a method of extracting nucleic acids from a biological sample, deriving a terminal sequence motif frequency of a nucleic acid fragment and the size of the nucleic acid fragment on the basis of a read obtained by obtaining and aligning sequence information, generating the same into vectorized data, inputting the data into a trained artificial intelligence model, and analyzing a calculated value. Because the method for diagnosing cancer and predicting a cancer type by using a terminal sequence motif frequency and the size of a cell-free nucleic acid fragment, according to the present invention, generates vectorized data and analyzes the data by using an AI algorithm, the method shows high sensitivity and accuracy even if read coverage is low, and thus is useful.

(57) 요약서: 본 발명은 세포유리 핵산단편 말단 서열 모티프 빈도 및 크기를 이용한 암 진단 및 암 종 예측방법에 관한 것으로, 보다 구체적으로는 생체시료에서 핵산을 추출하여, 서열정보를 획득하여 정렬한 리드를 기반으로 핵산단편의 말단 서열 모티프 빈도와 핵산단편의 크기를 도출한 다음, 이를 벡터화된 데이터로 생성한 후, 학습된 인공지능 모델에 입력하여 계산된 값을 분석하는 방법을 이용한 암 진단 및 암 종 예측방법에 관한 것이다. 본 발명에 따른 세포유리 핵산단편 말단 서열 모티프 빈도 및 크기를 이용한 암 진단 및 암 종 예측방법은 벡터화된 데이터를 생성하여 AI 알고리즘을 이용하여 분석하기 때문에 리드 커버리지가 낮더라도 높은 민감도와 정확도를 나타내어 유용하다.

[다음 쪽 계속]

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 역
내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE,
LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM,
ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유
럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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))
- 명세서의 서열목록 부분과 함께 (규칙 5.2(a))