

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
08 February 2024 (08.02.2024)



(10) International Publication Number
WO 2024/030606 A1

(51) International Patent Classification:
G16B 20/00 (2019.01) *G16B 40/20* (2019.01)
G16B 25/10 (2019.01)

(21) International Application Number:
PCT/US2023/029477

(22) International Filing Date:
04 August 2023 (04.08.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/395,778 05 August 2022 (05.08.2022) US

(71) Applicant: **ILLUMINA, INC.** [US/US]; 5200 Illumina
Way, San Diego, California 92122 (US).

(72) Inventor: **JAGANATHAN, Kishore**; c/o Illumina, Inc.,
5200 Illumina Way, San Diego, California 92122 (US).

(74) Agent: **RARIDEN, John M.** et al.; P.O. Box 692289, Hous-
ton, Texas 77269 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, 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, KR, 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.

(54) Title: ARTIFICIAL INTELLIGENCE-BASED DETECTION OF GENE CONSERVATION AND EXPRESSION PRESER-
VATION AT BASE RESOLUTION

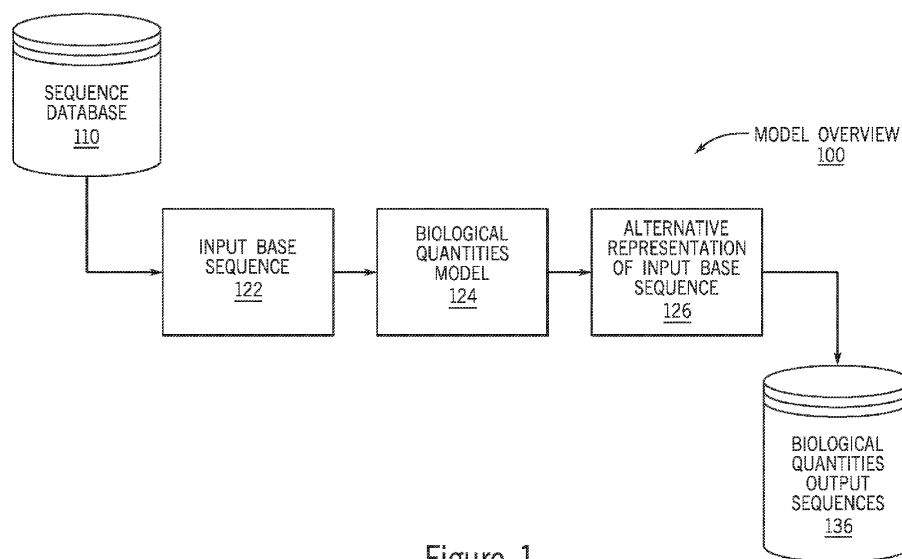


Figure 1

(57) Abstract: The technology disclosed relates to detecting gene conservation and expression preservation. In particular, the technology disclosed relates to detecting gene conservation and epigenetic signals for a reference genetic sequence (302) in comparison to a variant of the reference genetic sequence (302) at base resolution through the generation of a plurality of alternative representations (300) of the sequence in chromatin form which may represent evolutionary conservation, transcription initiation, or epigenetic signals, mapping the plurality of alternative chromatin sequences to a gene expression alterability classifier (2700) to generate a gene expression class prediction for the variant, and mapping the alternative chromatin sequence to a pathogenicity predictor to detect pathogenicity of variants.

[Continued on next page]

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (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).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*