

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
26 January 2023 (26.01.2023)



(10) International Publication Number
WO 2023/004323 A1

(51) International Patent Classification:

G16B 30/00 (2019.01) G16B 40/20 (2019.01)
G06N 20/00 (2019.01)

(21) International Application Number:

PCT/US2022/073899

(22) International Filing Date:

19 July 2022 (19.07.2022)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

17/384,423 23 July 2021 (23.07.2021) US

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

(72) Inventors: **PARNABY, Gavin Derek**; 5200 Illumina Way,
San Diego, California 92122 (US). **VISVANATH, Arun**;
5200 Illumina Way, San Diego, California 92122 (US). **DE-
JONG, Antoine Jean**; 5200 Illumina Way, San Diego, Cal-
ifornia 92122 (US).

(74) Agent: **FUNK, Christopher** et al.; 1010 North 500 East,
Suite 210, North Salt Lake, Utah 84054 (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, 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, 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) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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).

Published:

— with international search report (Art. 21(3))

(54) Title: MACHINE-LEARNING MODEL FOR RECALIBRATING NUCLEOTIDE-BASE CALLS

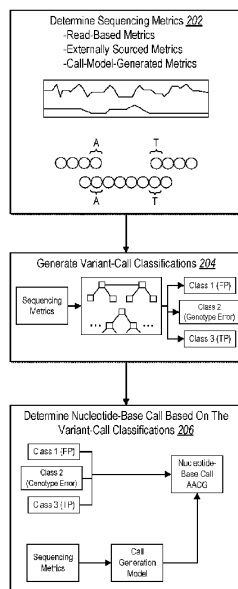


Fig. 2

(57) Abstract: This disclosure describes methods, non-transitory computer readable media, and systems that can utilize a machine learning model to recalibrate nucleotide-base calls (e.g., variant calls) of a call-generation model. For instance, the disclosed systems can train and utilize a call-recalibration-machine-learning model to generate a set of predicted variant-call classifications based on sequencing metrics associated with a sample nucleotide sequence. Leveraging the set of variant-call classifications, the disclosed systems can further update or modify nucleotide-base calls (e.g., variant calls) corresponding to genomic coordinates. Indeed, the disclosed systems can generate an initial nucleotide-base call based on sequencing metrics for nucleotide reads of a sample sequence utilizing a call-generation model and further utilize a call-recalibration-machine-learning model to generate classification predictions for updating or recalibrating the initial nucleotide-base call from a subset of the same sequencing metrics or other sequencing metrics.

WO 2023/004323 A1