

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
03 March 2022 (03.03.2022)



(10) International Publication Number
WO 2022/047038 A1

(51) International Patent Classification:

G16B 30/00 (2019.01) G16B 40/10 (2019.01)
G16B 40/20 (2019.01) G16B 40/30 (2019.01)
G16B 50/30 (2019.01)

(21) International Application Number:

PCT/US2021/047763

(22) International Filing Date:

26 August 2021 (26.08.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/072,032 28 August 2020 (28.08.2020) US
17/411,980 25 August 2021 (25.08.2021) US

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

(72) Inventors: **KASHEFHAGHIGHI, Dorna**; 499 Illinois Street, Suite 210, San Francisco, California 94158 (US).
PARNABY, Gavin Derek; 5200 Illumina Way, San Diego, California 92122 (US).

(74) Agent: **KHAN, Sikander** et al.; Haynes Beffel & Wolfeld LLP, P.O. Box 366, 637 Main Street, Half Moon Bay, California 94019 (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD,

(54) Title: DETECTING AND FILTERING CLUSTERS BASED ON ARTIFICIAL INTELLIGENCE-PREDICTED BASE CALLS

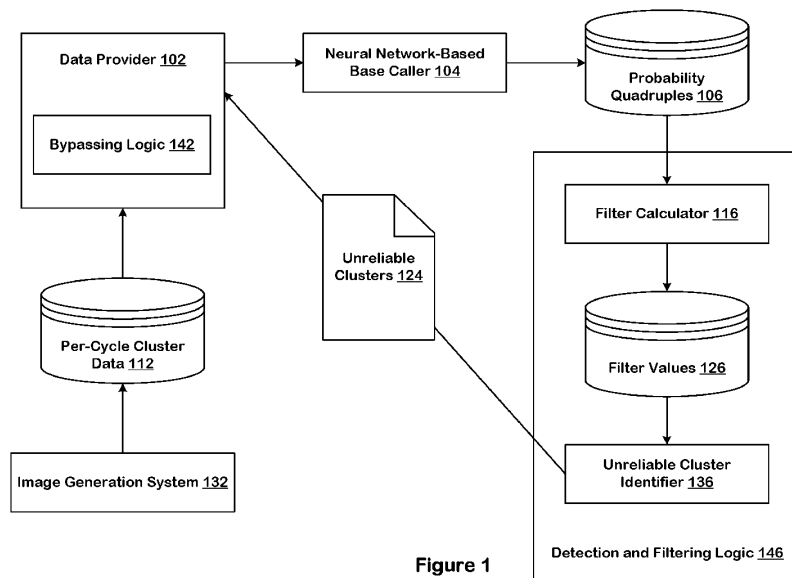


Figure 1

(57) **Abstract:** The technology disclosed relates to identifying unreliable clusters to improve accuracy and efficiency of base calling. The technology disclosed includes accessing per-cycle cluster data for a plurality of clusters and for a first subset of sequencing cycles of a sequencing run, and base calling each cluster in the plurality of clusters at each sequencing cycle in the first subset of sequencing cycles, including generating per-cycle probability quadruple for each cluster and for each sequencing cycle. The technology disclosed includes determining a filter value for each per-cluster, per-cycle probability quadruple based on the probabilities it identifies, identifying those clusters in the plurality of clusters as unreliable clusters whose sequences of filter values contain at least "N" number of filter values below a threshold "M", and bypassing base calling the unreliable clusters at a remainder of sequencing cycles of the sequencing run.

[Continued on next page]

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).

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))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*