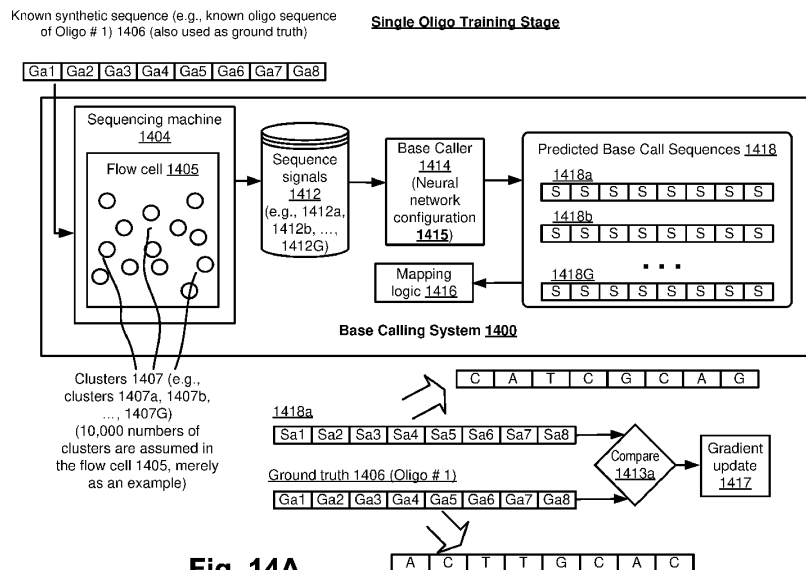




- (51) **International Patent Classification:**
G16B 50/30 (2019.01) *G16B 30/00* (2019.01)
G16B 40/20 (2019.01)
- (21) **International Application Number:**
 PCT/US2022/035564
- (22) **International Filing Date:**
 29 June 2022 (29.06.2022)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
- | | | |
|------------|---------------------------|----|
| 63/216,404 | 29 June 2021 (29.06.2021) | US |
| 63/216,419 | 29 June 2021 (29.06.2021) | US |
| 17/830,287 | 01 June 2022 (01.06.2022) | US |
| 17/830,316 | 01 June 2022 (01.06.2022) | US |
- (71) **Applicant:** ILLUMINA, INC. [US/US]; 5200 Illumina Way, San Diego, California 92122 (US).
- (72) **Inventors:** KIA, Amirali; 811 NEI Camino Real, Apt. 302, San Mateo, Alaska 94401 (US). DUTTA, Anindita; 5200 Illumina Way, San Diego, California 92122 (US).
- (74) **Agent:** FUNK, Christopher et al.; Keller Preece PLLC, 1010 N 500 E., Suite 210, North Salt Lake, UT 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, 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,

(54) **Title:** SELF-LEARNED BASE CALLER, TRAINED USING OLIGO SEQUENCES**Fig. 14A**

(57) **Abstract:** A method of progressively training a base caller is disclosed. The method includes iteratively initially training a base caller with analyte comprising a single-oligo base sequence, and generating labelled training data using the initially trained base caller. At operations (i), the base caller is further trained with analyte comprising multi-oligo base sequences, and labelled training data is generated using the further trained base caller. Operations (i) are iteratively repeated to further train the base caller. In an example, during at least one iteration, a complexity of neural network configuration loaded within the base caller is increased. In an example, labelled training data generated during an iteration is used to train the base caller during an immediate subsequent iteration.

[Continued on next page]

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))*