

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
11 November 2021 (11.11.2021)



(10) International Publication Number
WO 2021/226523 A3

(51) International Patent Classification:
G16B 30/10 (2019.01) *G16H 50/80* (2018.01)
G16B 20/20 (2019.01)

(74) Agent: **BAKKER, Jila** et al.; P.O. Box 692289, Houston, Texas 77269 (US).

(21) International Application Number:
PCT/US2021/031390

(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, 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.

(22) International Filing Date:
07 May 2021 (07.05.2021)

(25) Filing Language: English

(26) Publication Language: English

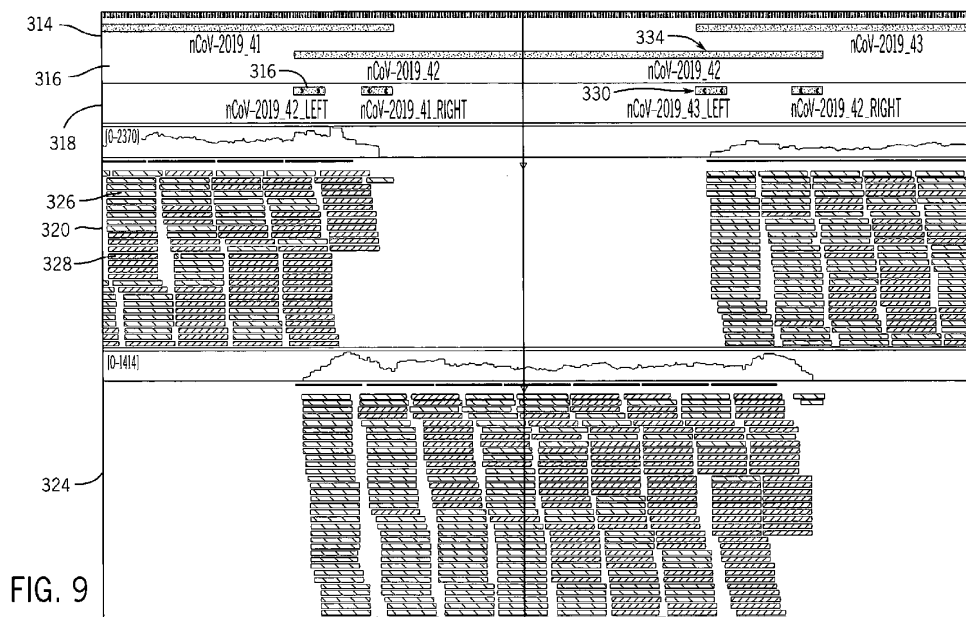
(30) Priority Data:
63/022,296 08 May 2020 (08.05.2020) US

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

(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,

(72) Inventors: **BILKE, Sven**; 5200 Illumina Way, San Diego, California 92122 (US). **SCHLESINGER, Johann Felix Wilhelm**; 5200 Illumina Way, San Diego, California 92122 (US).

(54) Title: GENOME SEQUENCING AND DETECTION TECHNIQUES



(57) Abstract: A nucleic acid sequencing technique is described. Sequence data, e.g., generated by a sequencing device, may be analyzed to scan k-mers of a fixed size n in individual reads in the sequence data. Exact matches of the k-mers in the sequence data with reference k-mers are identified. K-mer matching may be used to identify alternative alleles in sequence data with anomalous distribution associated with contamination or other quality issues and to determine a quality metric in real-time.

[Continued on next page]

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))*
- *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))*

(88) Date of publication of the international search report:

23 December 2021 (23.12.2021)