



- (51) International Patent Classification:
G16B 20/20 (2019.01) *G16B 20/50* (2019.01)
- (21) International Application Number:
PCT/US2022/033020
- (22) International Filing Date:
10 June 2022 (10.06.2022)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
63/209,608 11 June 2021 (11.06.2021) US
- (71) Applicant: **ILLUMINA, INC.** [US/US]; 5200 Illumina Way, San Diego, CA 92122 (US).
- (72) Inventors: **CHEN, Sai**; 5200 Illumina Way, San Diego, CA 92122 (US). **EBERLE, Michael, A.**; 5200 Illumina Way, San Diego, CA 92122 (US).
- (74) Agent: **CHANG, Y., Paul**; Sheppard Mullin Richter & Hampton LLP, 12275 El Camino Real, Suite 100, San Diego, CA 92130 (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, 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,

(54) Title: DETERMINING PATHOGENIC RFC1 EXPANSIONS FROM SEQUENCING DATA

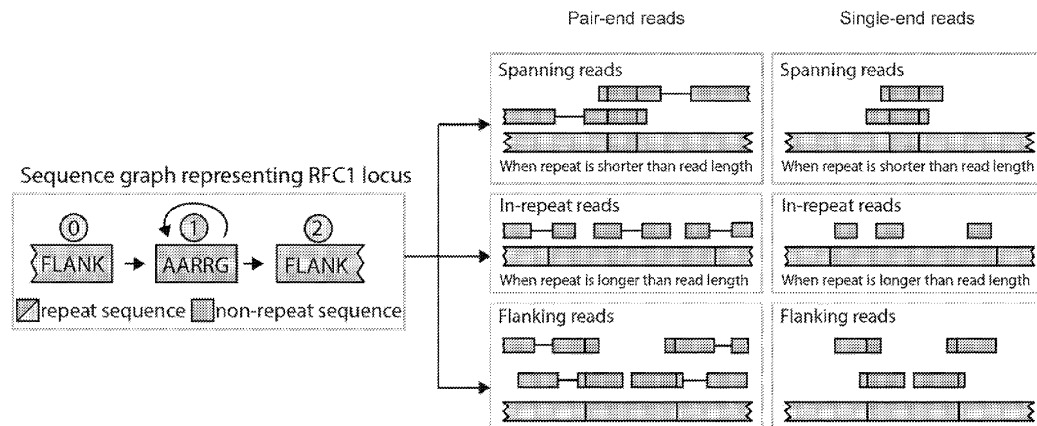


FIG. 1

(57) Abstract: Disclosed herein include systems, devices, and methods for determining repeat expansion status (e.g., pathogenic, carrier, and benign) of a locus of a gene of interest (e.g., at, or at about, chr4:39348424 of hg38 for RFC1). After aligning sequence reads to a sequence graph, the number of occurrences of repeat sequences satisfying predetermined criteria and the frequency of a pathogenic repeat sequence can be determined, which are in turn used to determine a repeat expansion status.

[Continued on next page]

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))*
- *with sequence listing part of description (Rule 5.2(a))*