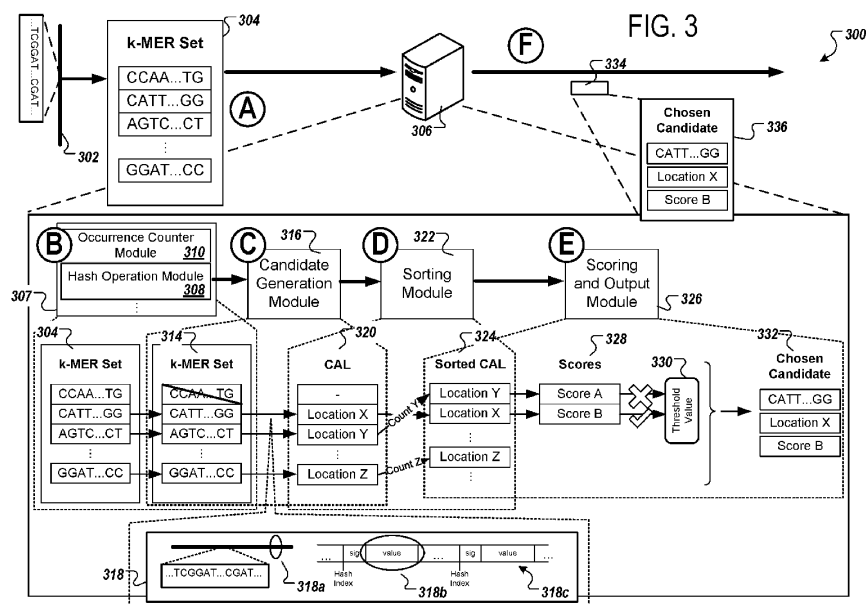




- (51) International Patent Classification:
G16B 30/00 (2019.01) *G16B 30/10* (2019.01)
- (21) International Application Number:
PCT/US2021/050557
- (22) International Filing Date:
15 September 2021 (15.09.2021)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
63/078,890 15 September 2020 (15.09.2020) US
- (71) Applicant: **ILLUMINA, INC.** [US/US]; 5200 Illumina Way, San Diego, California 92122 (US).
- (72) Inventor: **RIZK, Guillaume Alexandre Pascal**; 5200 Illumina Way, San Diego, California 92122 (US).
- (74) Agent: **DARNO, Patrick** et al.; Fish & Richardson P.C., P. O. Box 1022, Minneapolis, Minnesota 55440-1022 (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, 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.

(54) Title: SOFTWARE ACCELERATED GENOMIC READ MAPPING



(57) Abstract: Methods, systems, apparatus, and computer programs are disclosed for software-accelerated genomic data read mapping. In one aspect, the method can include actions of obtaining a k-mer seed from a genomic data read, generating a genomic signature based on the obtained k-mer seed, determining a reference sequence location that match at least a portion of the k-mer seed using a hash data structure, wherein the hash data structure comprises N data cells comprising a first portion storing a predetermined genomic signature and a second portion storing a value that corresponds to a first occurrence of a reference sequence location that match at least a portion of the k-mer seed from which the predetermined genomic signature was derived, and selecting the determined reference sequence location as an actual alignment for the obtained k-mer seed based on one or more alignment scores.

[Continued on next page]

(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))*