



- (51) International Patent Classification:
C12Q 1/68 (2006.01)
- (21) International Application Number:
PCT/SG2014/000172
- (22) International Filing Date:
17 April 2014 (17.04.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
201302940-0 17 April 2013 (17.04.2013) SG
- (71) Applicant: AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis, Singapore 138632 (SG).
- (72) Inventors: **QUAKE, Stephen R.**; c/o Stanford University, 450 Serra Mall, Stanford, California 94305 (US). **BURKHOLDER, William F.**; c/o Institute of Molecular and Cell Biology, Proteos Building, 61 Biopolis Drive, Singapore 138673 (SG). **HONG, Lewis Z.**; c/o Institute of Molecular and Cell Biology, Proteos Building, 61 Biopolis Drive, Singapore 138673 (SG).
- (74) Agent: YUSARN AUDREY; 24 Raffles Place, #27-01 Clifford Centre, Singapore 048621 (SG).
- (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, 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).

[Continued on next page]

(54) Title: METHOD FOR GENERATING EXTENDED SEQUENCE READS

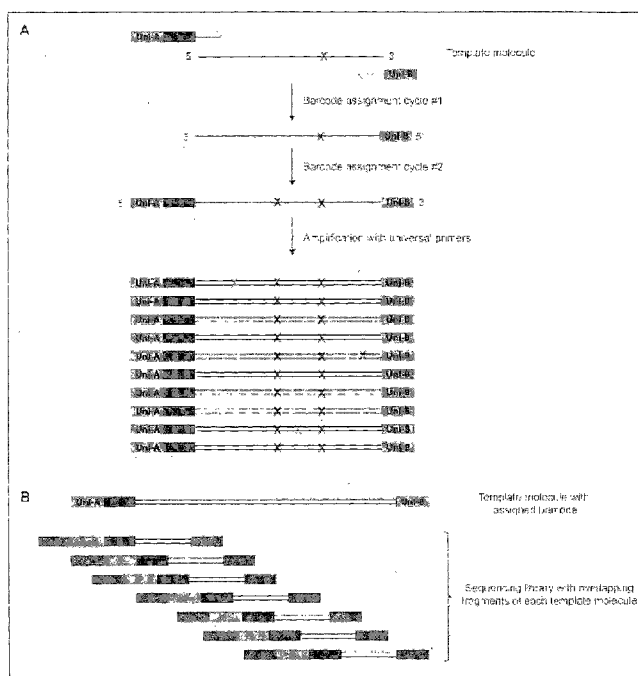


FIGURE 1

(57) Abstract: The present invention provides an approach to increase the effective read length of commercially available sequencing platforms to several kilobases and be broadly applied to obtain long sequence reads from mixed template populations. A method for generating extended sequence reads of long DNA molecules in a sample, comprising the steps of: assigning a specific barcode sequence to each template DNA molecule in a sample to obtain barcode-tagged molecules; amplifying the barcode-tagged molecules; fragmenting the amplified barcode-tagged molecules to obtain barcode-containing fragments; juxtaposing the barcode-containing fragments to random short segments of the original DNA template molecule during the process of generating a sequencing library to obtain demultiplexed reads; and assembling the demultiplexed reads to obtain extended sequence reads for each DNA template molecule, is disclosed.

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *of inventorship (Rule 4.17(iv))*

Published:

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*