



(51) International Patent Classification:
G16B 20/20 (2019.01)

(21) International Application Number:
PCT/US2023/028394

(22) International Filing Date:
21 July 2023 (21.07.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/395,774 05 August 2022 (05.08.2022) US

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

(72) Inventors: **JAGANATHAN, Kishore**; c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US). **AGHAMIRZAIE, Delasa**; 6686 Wooster Court, Castro Valley, California 94552 (US). **PANAGIOTOPOULOU, Sofia Kyriazopoulou**; 370 Hillview Ave, Redwood City, California 94062 (US). **FARH, Kai-How**; c/o Illumina, Inc., 5200 Illumina Way, San Diego, California 92122 (US).

(74) Agent: **RARIDEN, John M.** et al.; P.O. Box 692289, Houston, Texas 77269 (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, CV, 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, MG, MK, MN, MU, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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).

(54) Title: COMPUTER-IMPLEMENTED METHODS OF IDENTIFYING RARE VARIANTS THAT CAUSE EXTREME LEVELS OF GENE EXPRESSION

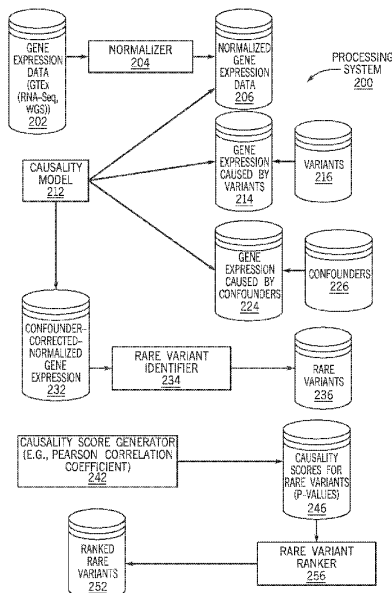


Figure 2

(57) Abstract: The technology disclosed relates to reliably identifying variants that cause extreme levels of gene expression. Extreme levels of gene expression include under expression and over expression. Then, these variants may be used to train artificial intelligence based models for a variety of prediction tasks. One example of the prediction tasks is to produce per-base resolution for chromatin sequences. Another example of the chromatin task is to produce gene expression changes caused by the reliably identified variants.

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))*