

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

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
10 March 2022 (10.03.2022)



(10) International Publication Number
WO 2022/051528 A1

(51) International Patent Classification:

C12N 15/86 (2006.01) C12N 5/074 (2010.01)
C12N 15/90 (2006.01)

(21) International Application Number:

PCT/US2021/048905

(22) International Filing Date:

02 September 2021 (02.09.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/074,987 04 September 2020 (04.09.2020) US

(71) Applicant: **10X GENOMICS, INC.** [US/US]; 6230
Stoneridge Mall Road, Pleasanton, California 94588-3260
(US).

(72) Inventors: **MAHESHWARI, Arundhati Shamoni**; c/o
10x Genomics, Inc., 6230 Stoneridge Mall Road, Pleasanton,
California 94588-3260 (US). **GOPALAN, Vijay Kumar Sreenivasa**; c/o 10x Genomics, Inc., 6230 Stoneridge
Mall Road, Pleasanton, California 94588-3260 (US). **OLSEN, Brett**; c/o 10x Genomics, Inc., 6230 Stoneridge
Mall Road, Pleasanton, California 94588-3260 (US). **HEP-
LER, Nicolaus, Lance**; c/o 10x Genomics, Inc., 6230
Stoneridge Mall Road, Pleasanton, California 94588-3260
(US).

(74) Agent: **NOVAK, Jason, J.**; Haynes and Boone, LLP, 2323
Victory Avenue, Suite 700, Dallas, Texas 75219 (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,

(54) Title: SYSTEMS AND METHODS FOR IDENTIFYING CELL-ASSOCIATED BARCODES IN MULTI-GENOMIC FEATURE DATA FROM SINGLE-CELL PARTITIONS

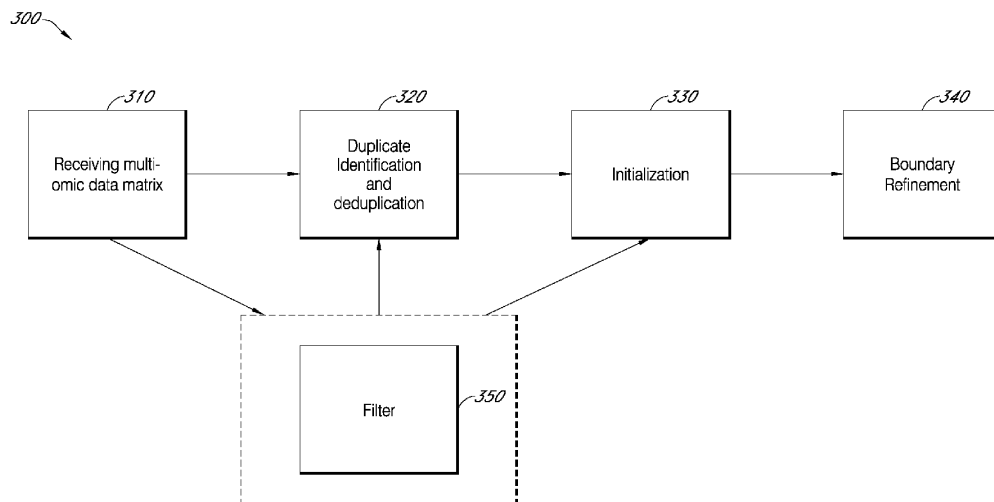


FIG. 3

(57) Abstract: Methods and systems may be provided for distinguishing cell populations from non-cell populations within a data set, the method comprising receiving a data set at least associated with a plurality of cells, wherein the data set comprises molecule counts of at least two genomic features for each cell; identifying duplicate subsets of data points from the data set; generating deduplicated data by condensing data points from each duplicate subset into a single data point; applying a pre-set threshold to divide the deduplicated data into an initial cell population and a non-cell population, wherein the pre-set threshold is determined using the molecule counts; and generating a refined cell population and a non-cell population by adjusting boundaries of the initial cell population and non-cell population using clustering.

[Continued on next page]

WO 2022/051528 A1

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.

- (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))*