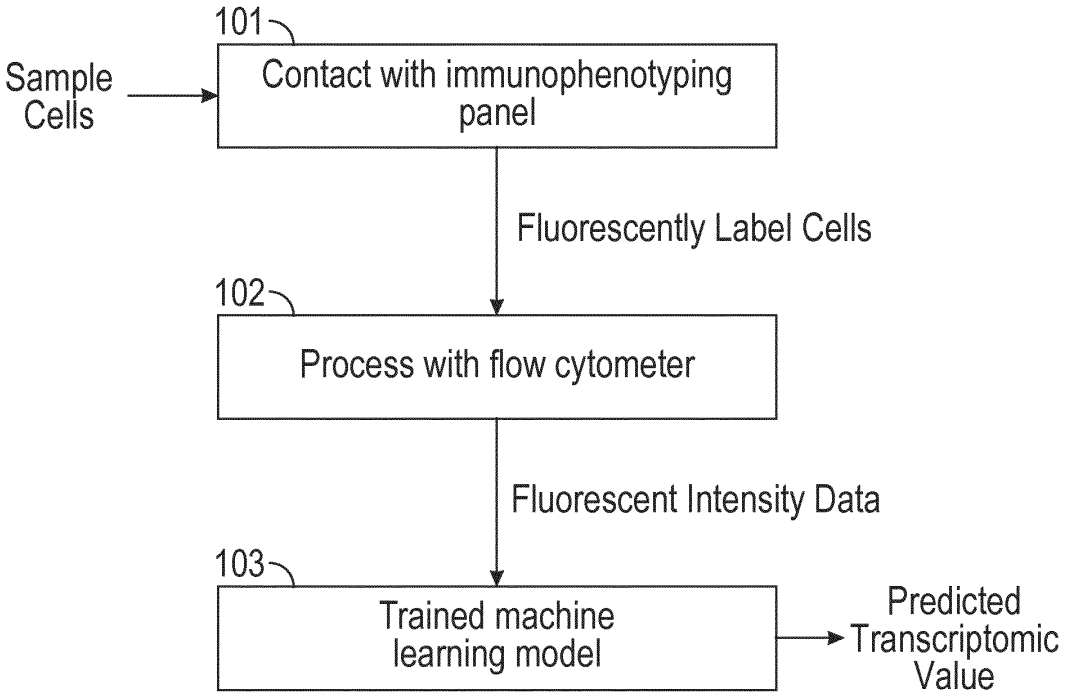

101
Sample Cells
Contact with immunophenotyping panel
Fluorescently Label Cells
102
Process with flow cytometer
Fluorescent Intensity Data
103
Trained machine learning model
Predicted Transcriptomic Value

FIG. 1

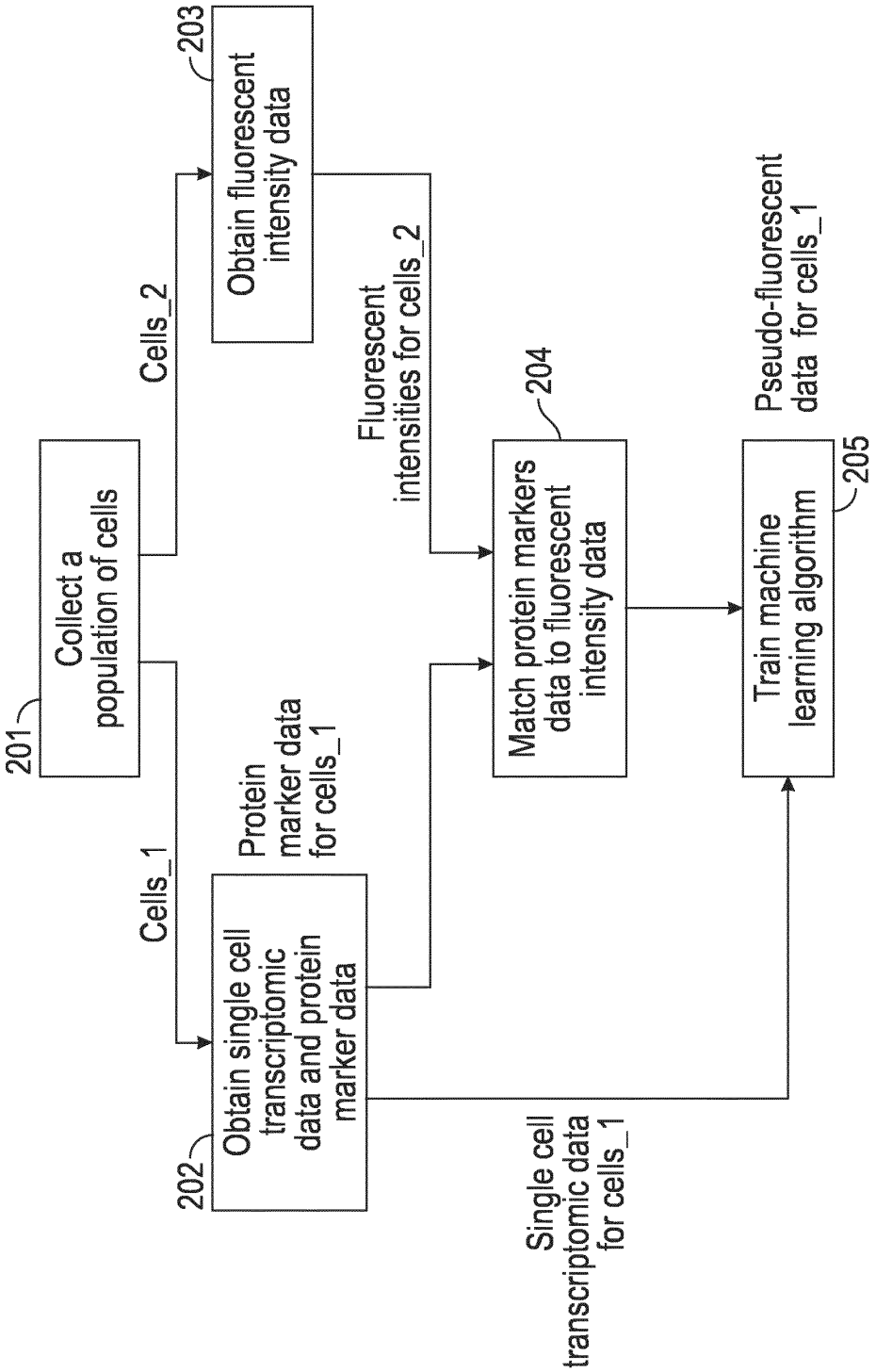

Collect a population of cells
201
Cells_1
Cells_2
Obtain single cell transcriptomic data and protein marker data
202
Protein marker data for cells_1
Obtain fluorescent intensity data
203
Fluorescent intensities for cells_2
Match protein markers data to fluorescent intensity data
204
Single cell transcriptomic data for cells_1
Train machine learning algorithm
205
Pseudo-fluorescent data for cells_1

FIG. 2

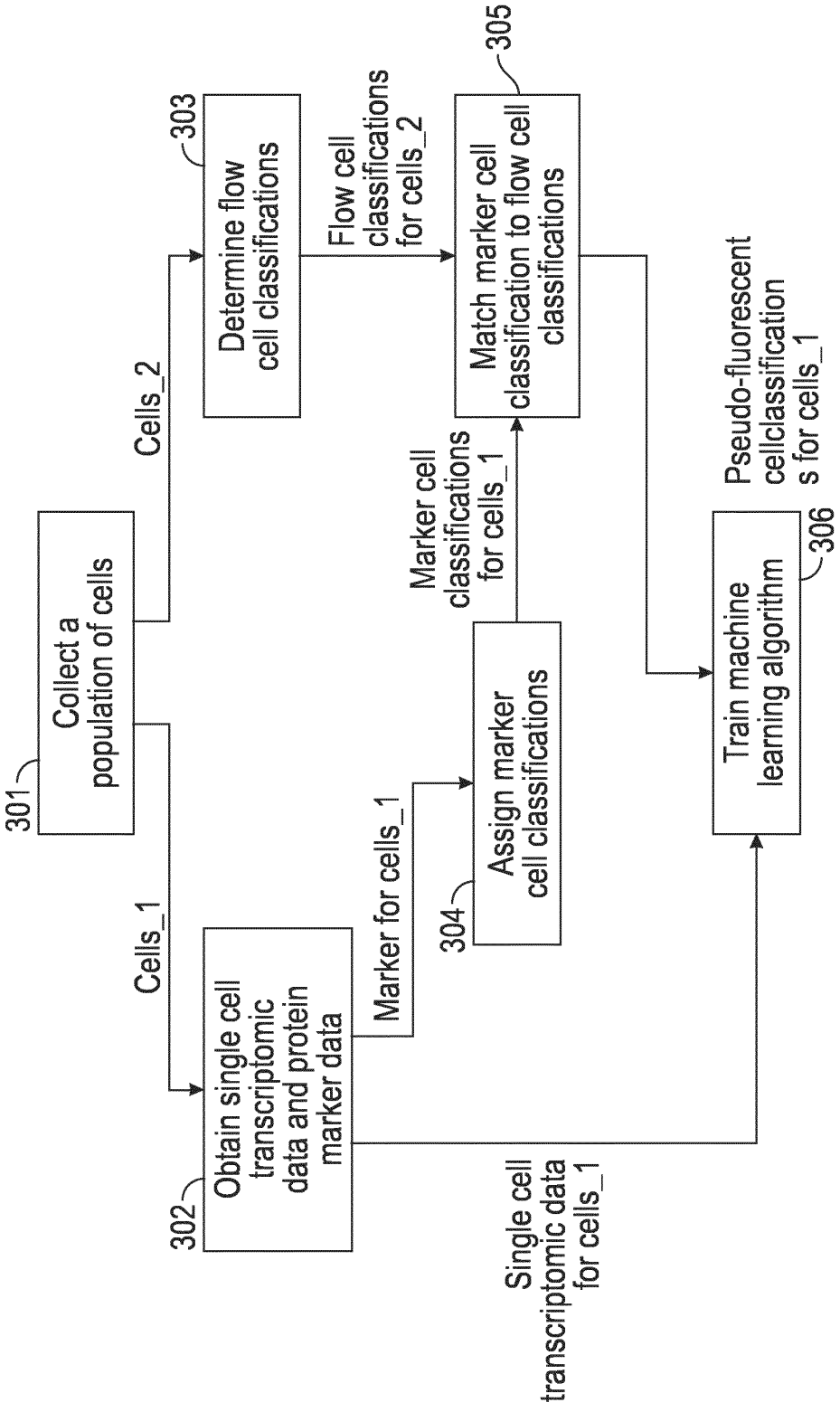

Collect a population of cells
301
Cells_1
Cells_2
Obtain single cell transcriptomic data and protein marker data
302
Determine flow cell classifications
303
Flow cell classifications for cells_2
Marker for cells_1
Assign marker cell classifications
304
Marker cell classifications for cells_1
Match marker cell classification to flow cell classifications
305
Single cell transcriptomic data for cells_1
Train machine learning algorithm
306
Pseudo-fluorescent cellclassification s for cells_1

FIG. 3

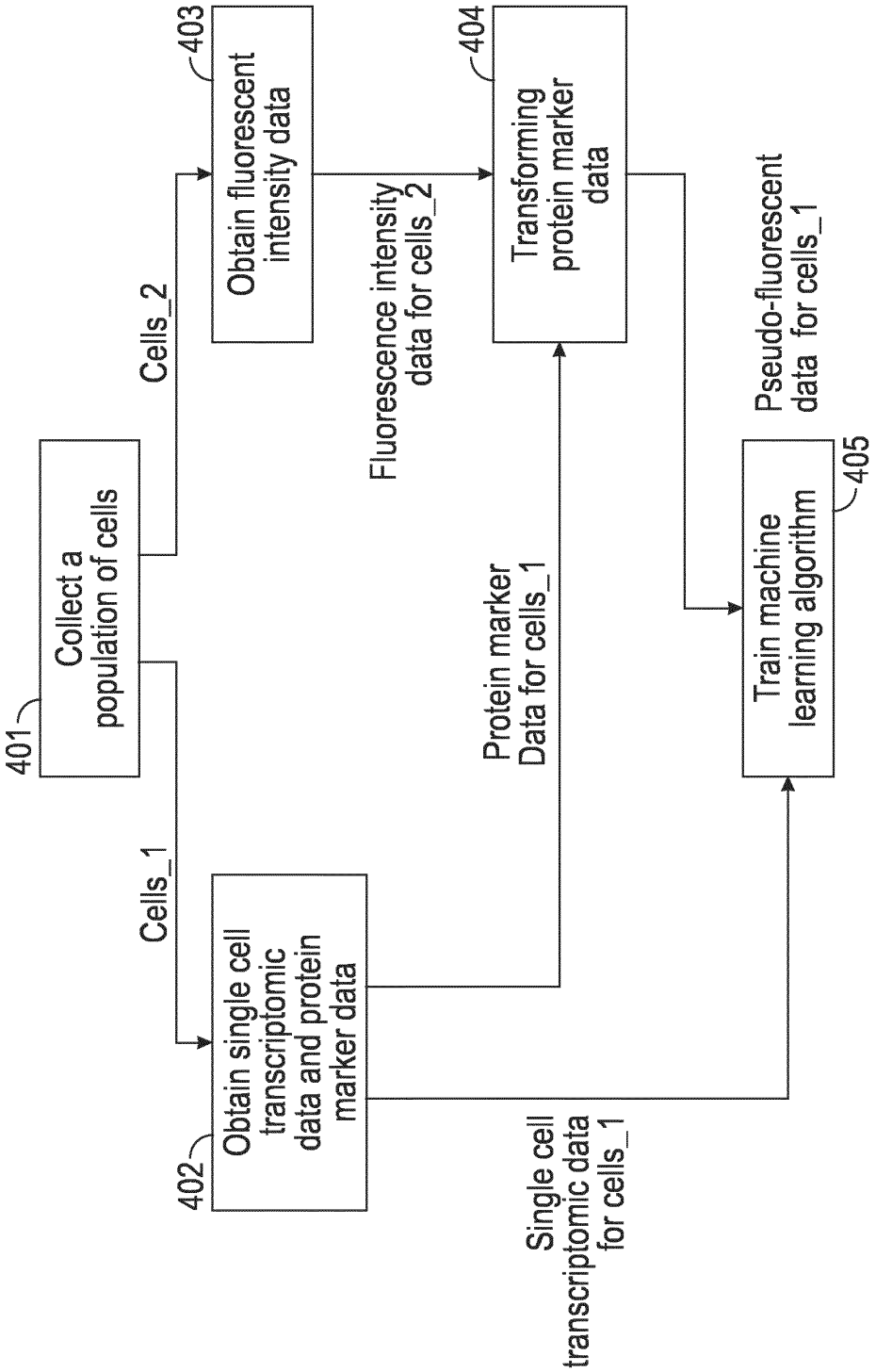

401
Collect a population of cells
Cells_1
Cells_2
402
Obtain single cell transcriptomic data and protein marker data
403
Obtain fluorescent intensity data
Fluorescence intensity data for cells_2
404
Transforming protein marker data
Protein marker Data for cells_1
Single cell transcriptomic data for cells_1
Pseudo-fluorescent data for cells_1
405
Train machine learning algorithm

FIG. 4

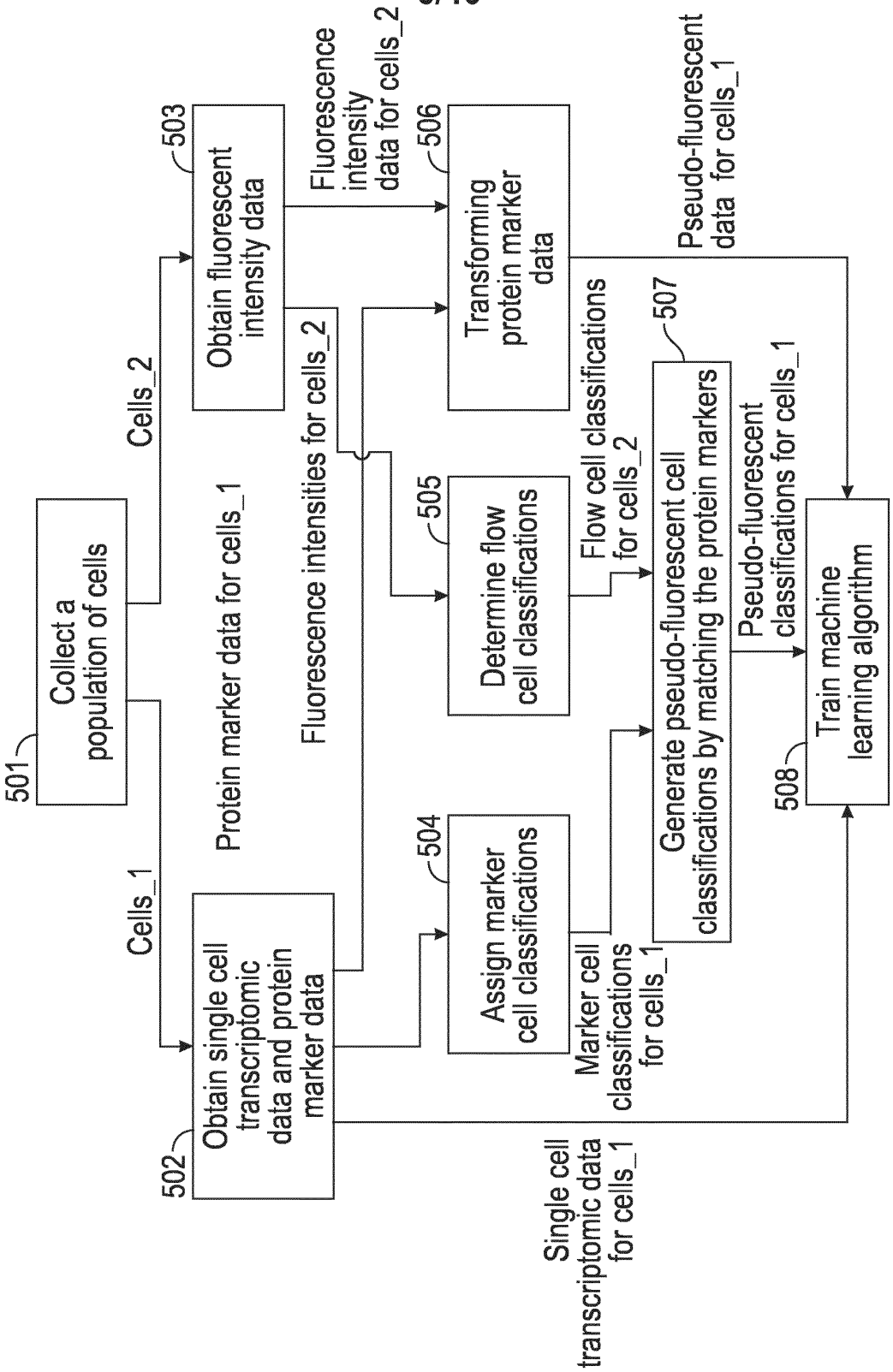

501
Collect a population of cells
Cells_1
Cells_2
502
Obtain single cell transcriptomic data and protein marker data
503
Obtain fluorescent intensity data
Protein marker data for cells_1
Fluorescence intensities for cells_2
Fluorescence intensity data for cells_2
504
Assign marker cell classifications
505
Determine flow cell classifications
506
Transforming protein marker data
Flow cell classifications for cells_2
Pseudo-fluorescent data for cells_1
507
Generate pseudo-fluorescent cell classifications by matching the protein markers
Marker cell classifications for cells_1
Pseudo-fluorescent classifications for cells_1
508
Train machine learning algorithm
Single cell transcriptomic data for cells_1

FIG. 5

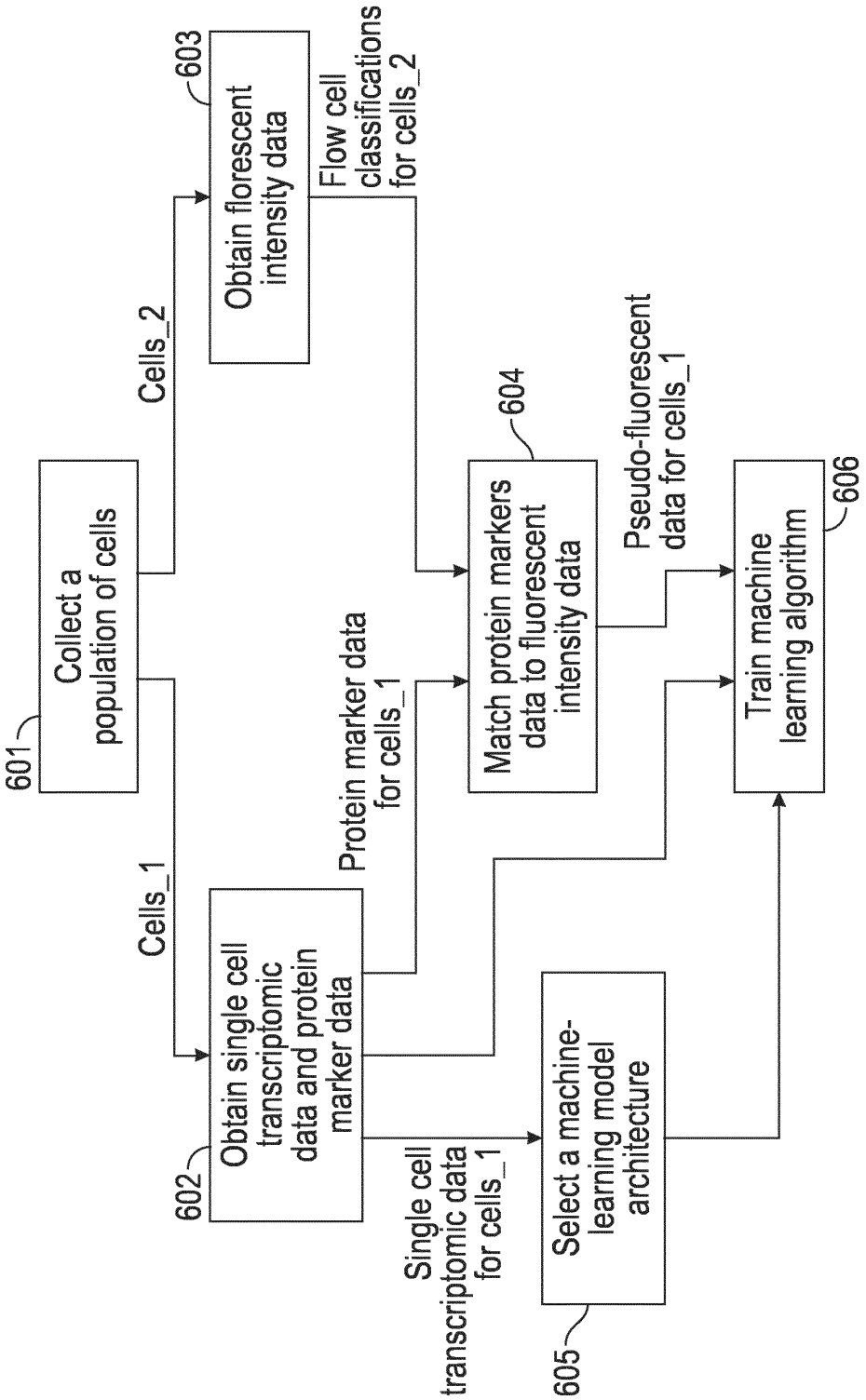

601
Collect a population of cells
Cells_1
Cells_2
602
Obtain single cell transcriptomic data and protein marker data
603
Obtain florescent intensity data
Flow cell classifications for cells_2
Protein marker data for cells_1
Single cell transcriptomic data for cells_1
604
Match protein markers data to fluorescent intensity data
Pseudo-fluorescent data for cells_1
605
Select a machine-learning model architecture
606
Train machine learning algorithm

FIG. 6

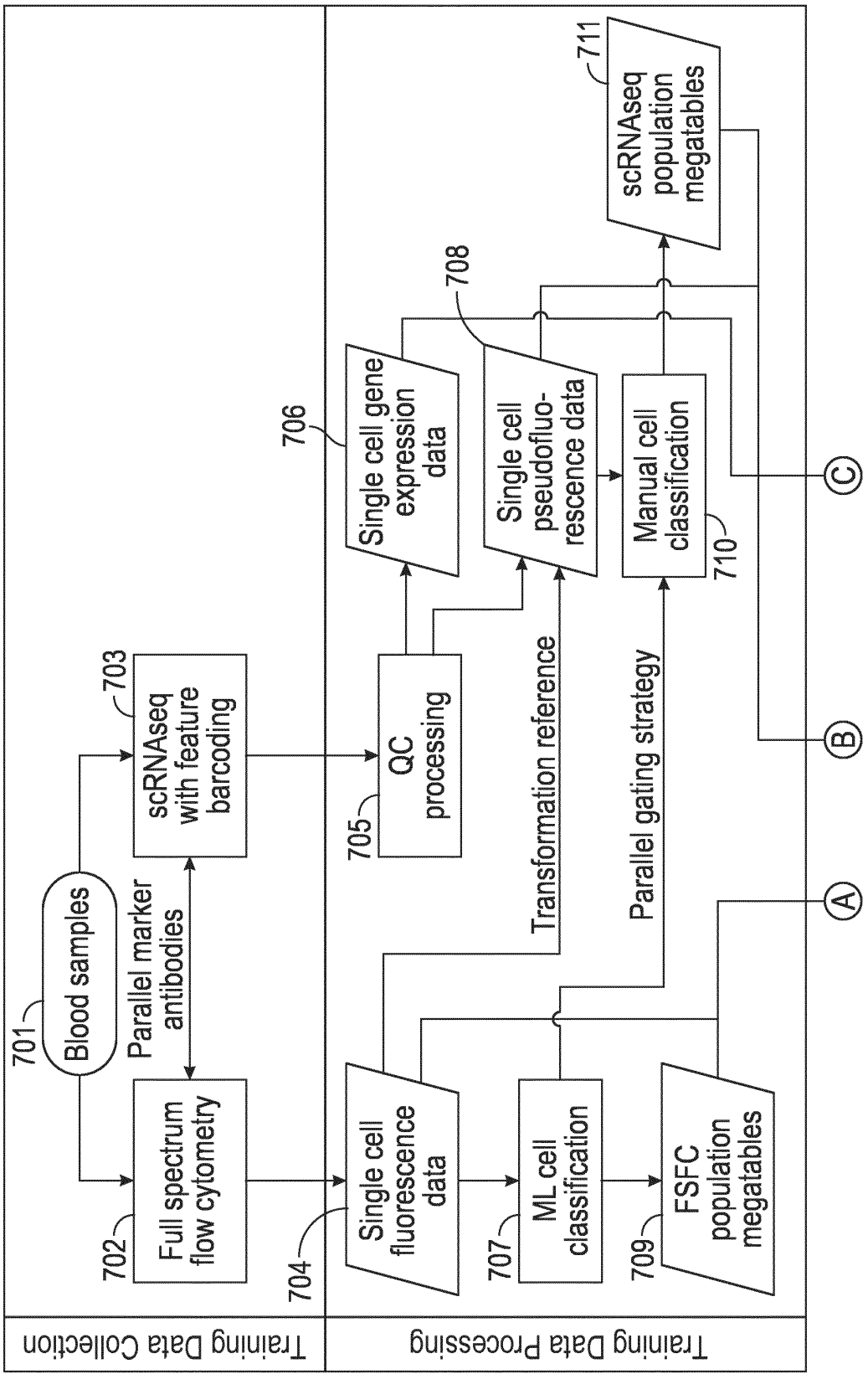

Training Data Collection
Training Data Processing
Blood samples
701
Parallel marker antibodies
Full spectrum flow cytometry
702
scRNAseq with feature barcoding
703
Single cell fluorescence data
704
QC processing
705
Single cell gene expression data
706
ML cell classification
707
Single cell pseudofluo-rescence data
708
FSFC population megatables
709
Manual cell classification
710
scRNAseq population megatables
711
Transformation reference
Parallel gating strategy
A
B
C

FIG. 7A

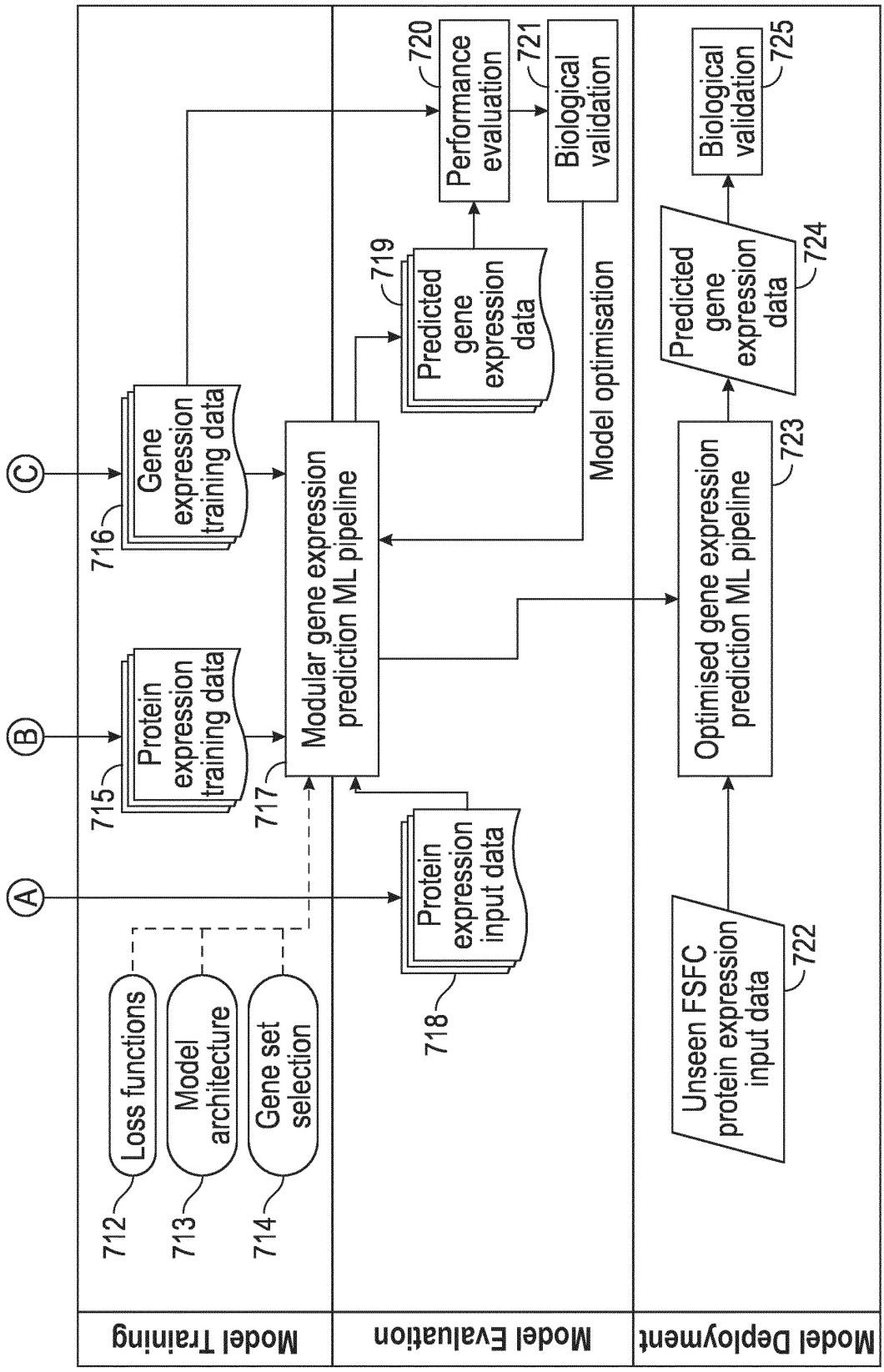

Model Training
Model Evaluation
Model Deployment
A
B
C
712
Loss functions
713
Model architecture
714
Gene set selection
715
Protein expression training data
716
Gene expression training data
717
Modular gene expression prediction ML pipeline
718
Protein expression input data
719
Predicted gene expression data
720
Performance evaluation
721
Biological validation
Model optimisation
722
Unseen FSFC protein expression input data
723
Optimised gene expression prediction ML pipeline
724
Predicted gene expression data
725
Biological validation

FIG. 7B

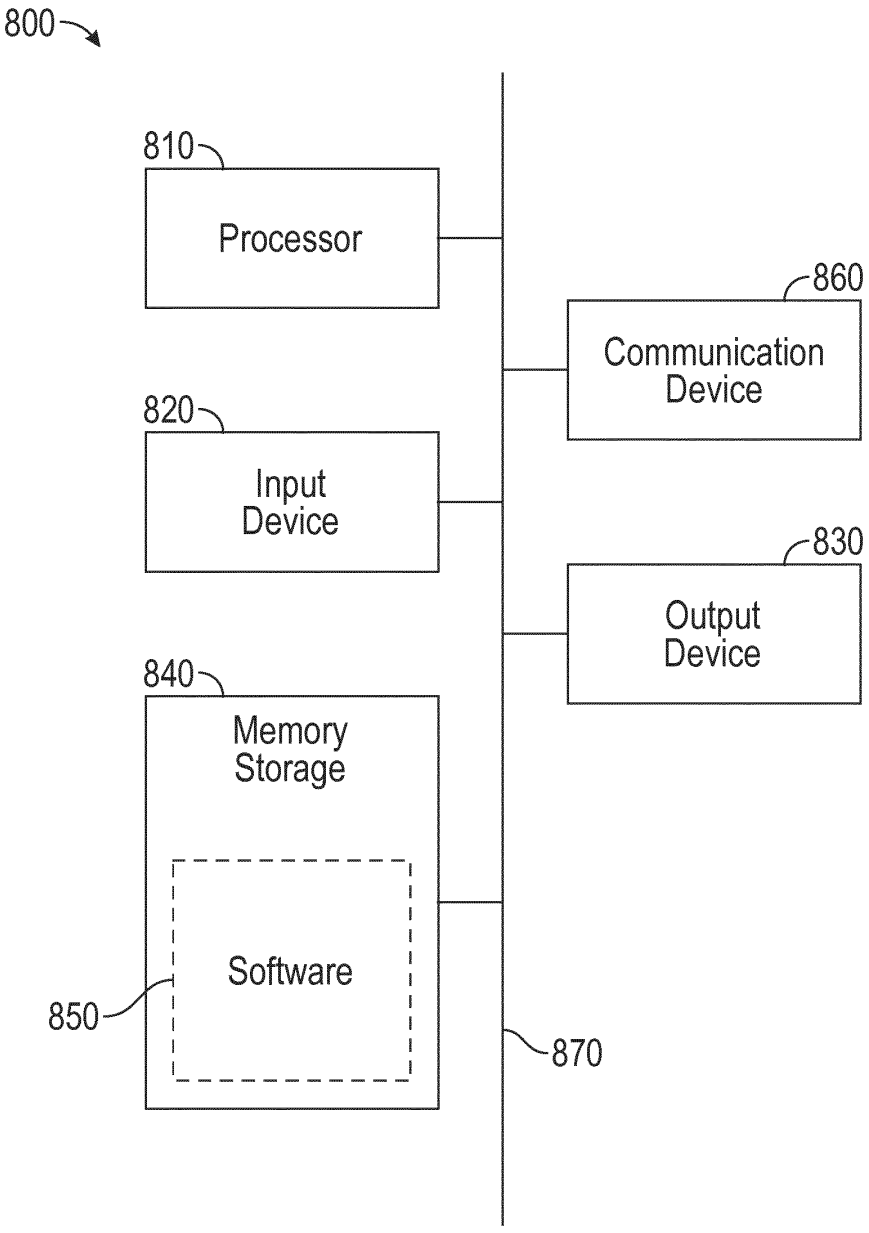

800
810
Processor
860
Communication Device
820
Input Device
830
Output Device
840
Memory Storage
Software
850
870

FIG. 8

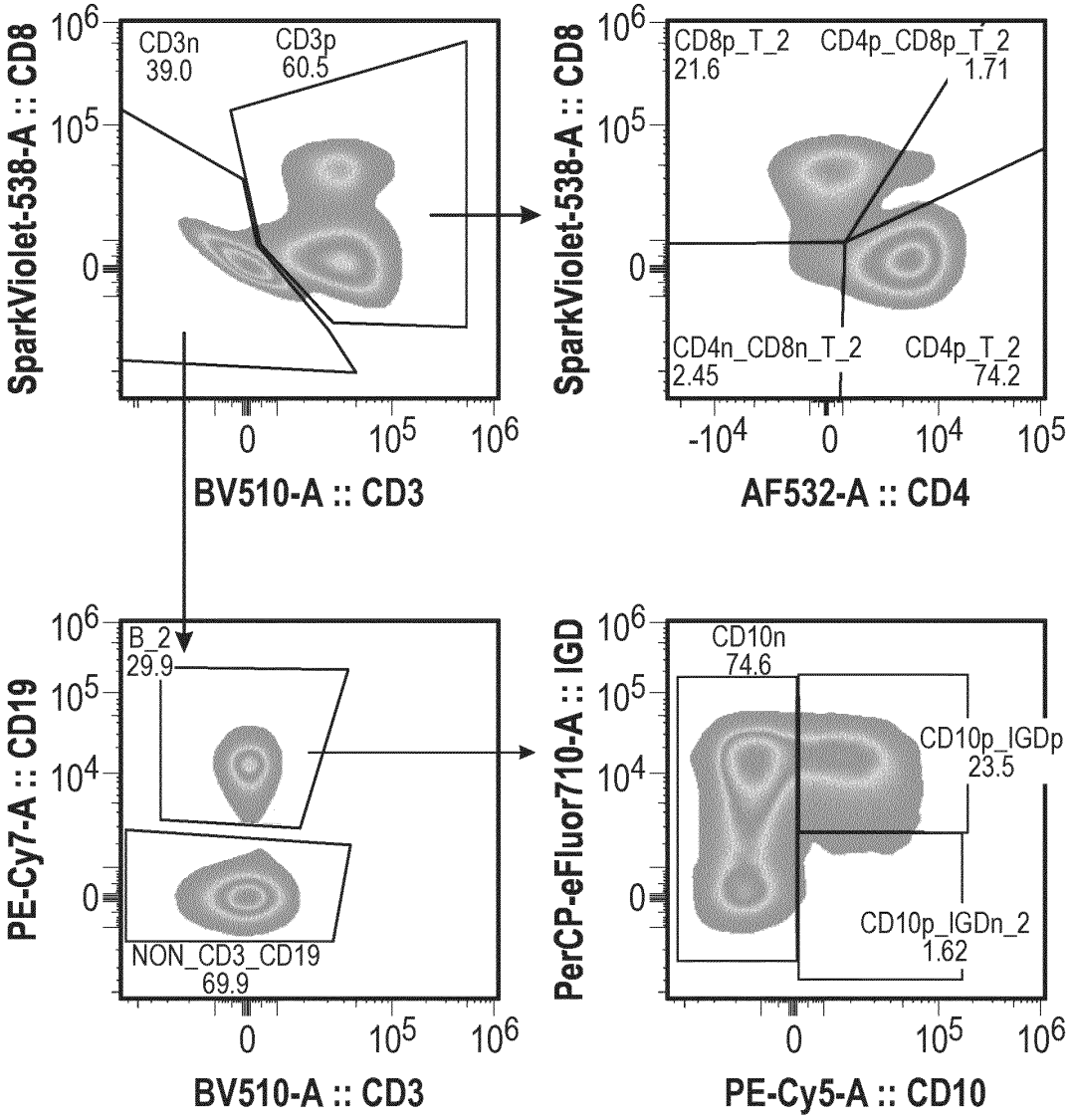

SparkViolet-538-A :: CD8
CD3n
39.0
CD3p
60.5
BV510-A :: CD3
SparkViolet-538-A :: CD8
CD8p_T_2
21.6
CD4p_CD8p_T_2
1.71
CD4n_CD8n_T_2
2.45
CD4p_T_2
74.2
AF532-A :: CD4
PE-Cy7-A :: CD19
B_2
29.9
NON_CD3_CD19
69.9
BV510-A :: CD3
PerCP-eFluor710-A :: IGD
CD10n
74.6
CD10p_IGDp
23.5
CD10p_IGDn_2
1.62
PE-Cy5-A :: CD10

FIG. 9A

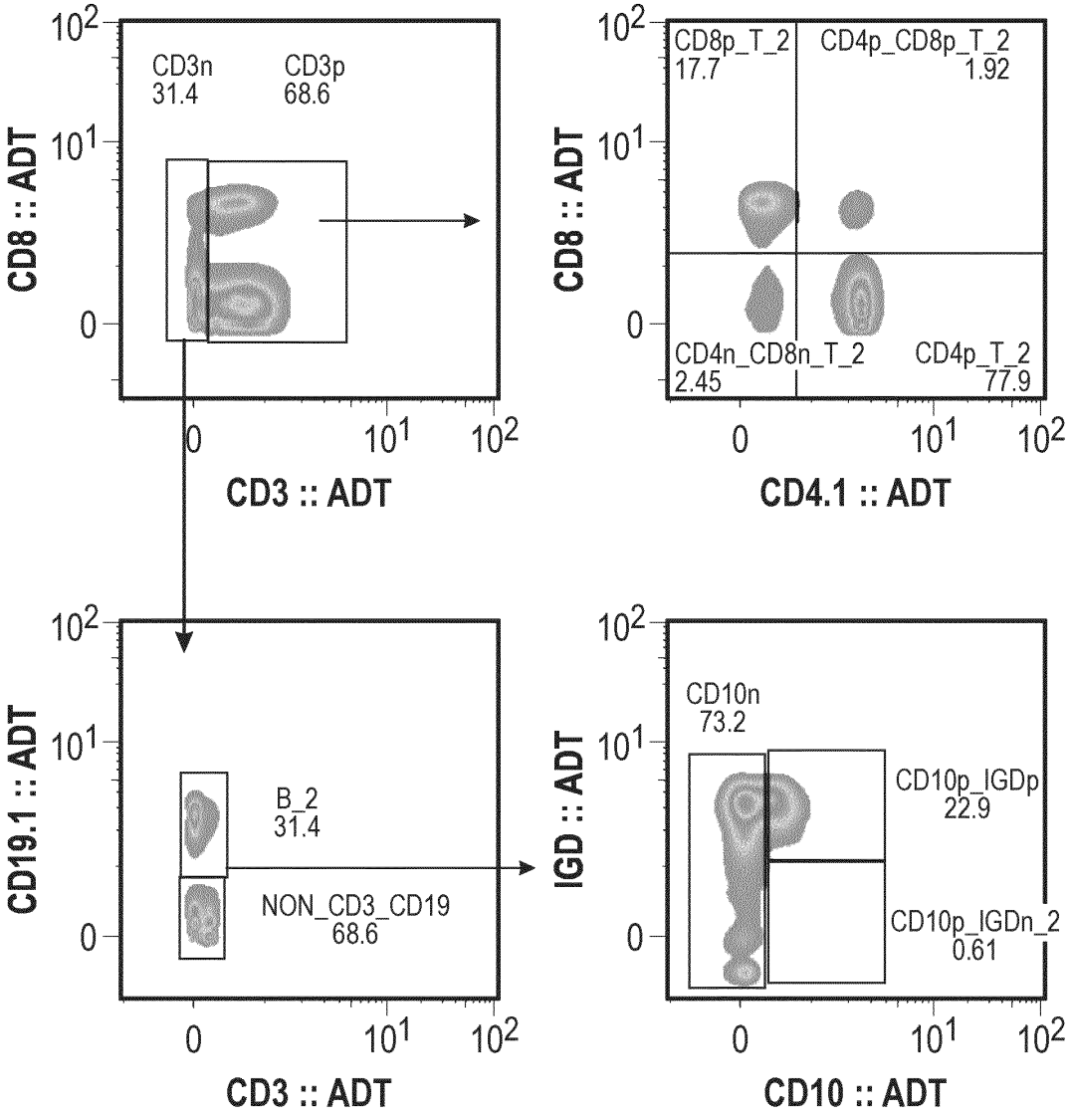

CD3n
31.4
CD3p
68.6
CD8 :: ADT
CD3 :: ADT
CD8p_T_2
17.7
CD4p_CD8p_T_2
1.92
CD4n_CD8n_T_2
2.45
CD4p_T_2
77.9
CD4.1 :: ADT
B_2
31.4
NON_CD3_CD19
68.6
CD19.1 :: ADT
CD10n
73.2
CD10p_IGDp
22.9
CD10p_IGDn_2
0.61
IGD :: ADT
CD10 :: ADT

FIG. 9B

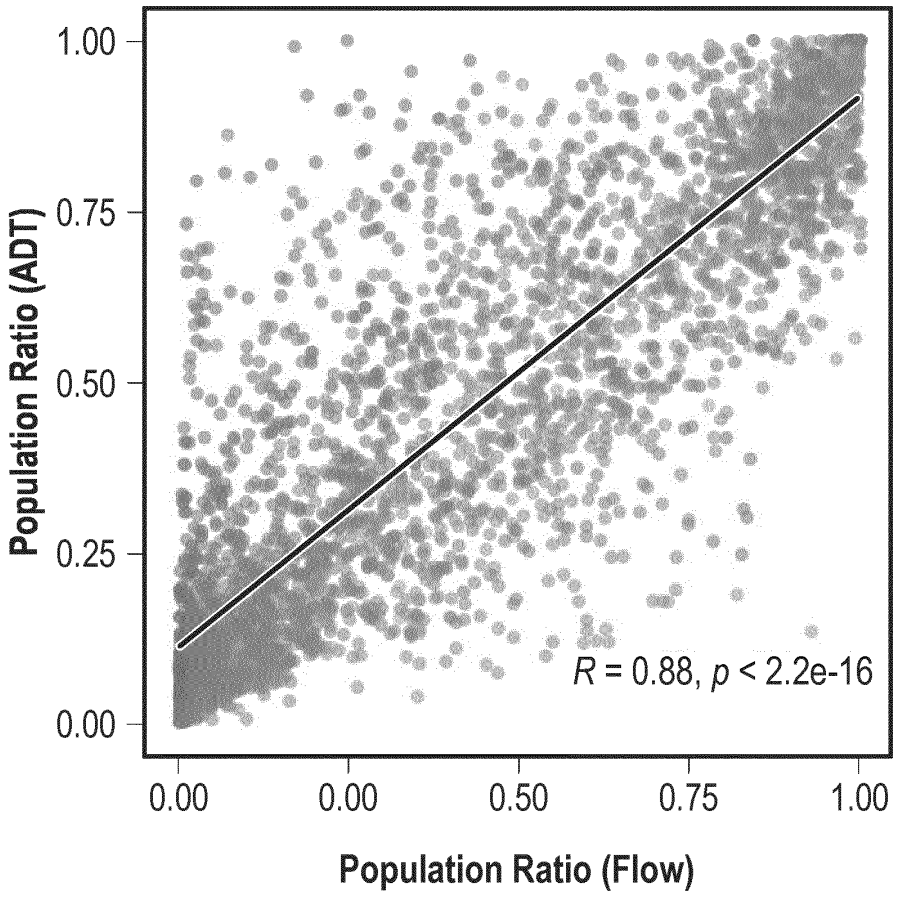

1.00
0.75
0.50
0.25
0.00
Population Ratio (ADT)
R = 0.88, p < 2.2e-16
0.00
0.00
0.50
0.75
1.00
Population Ratio (Flow)

FIG. 10

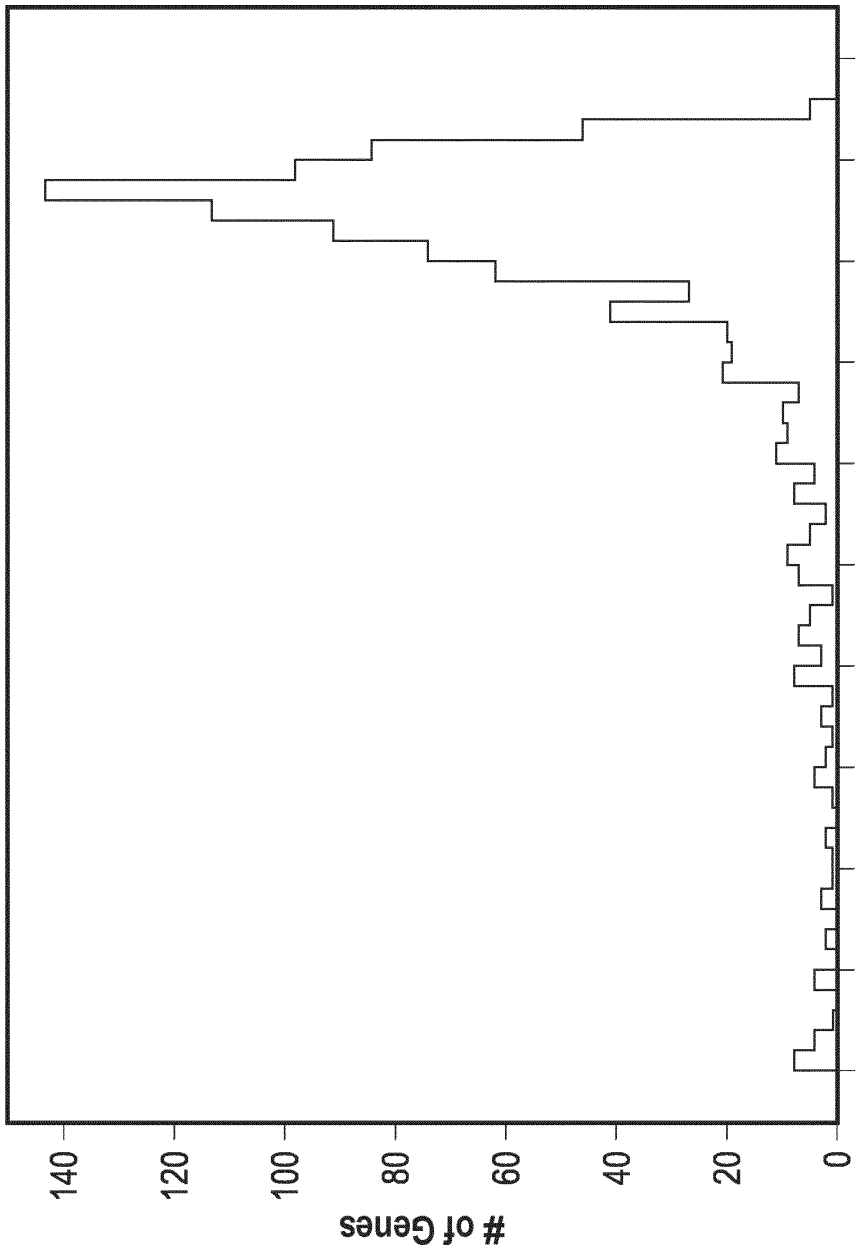


FIG. 11

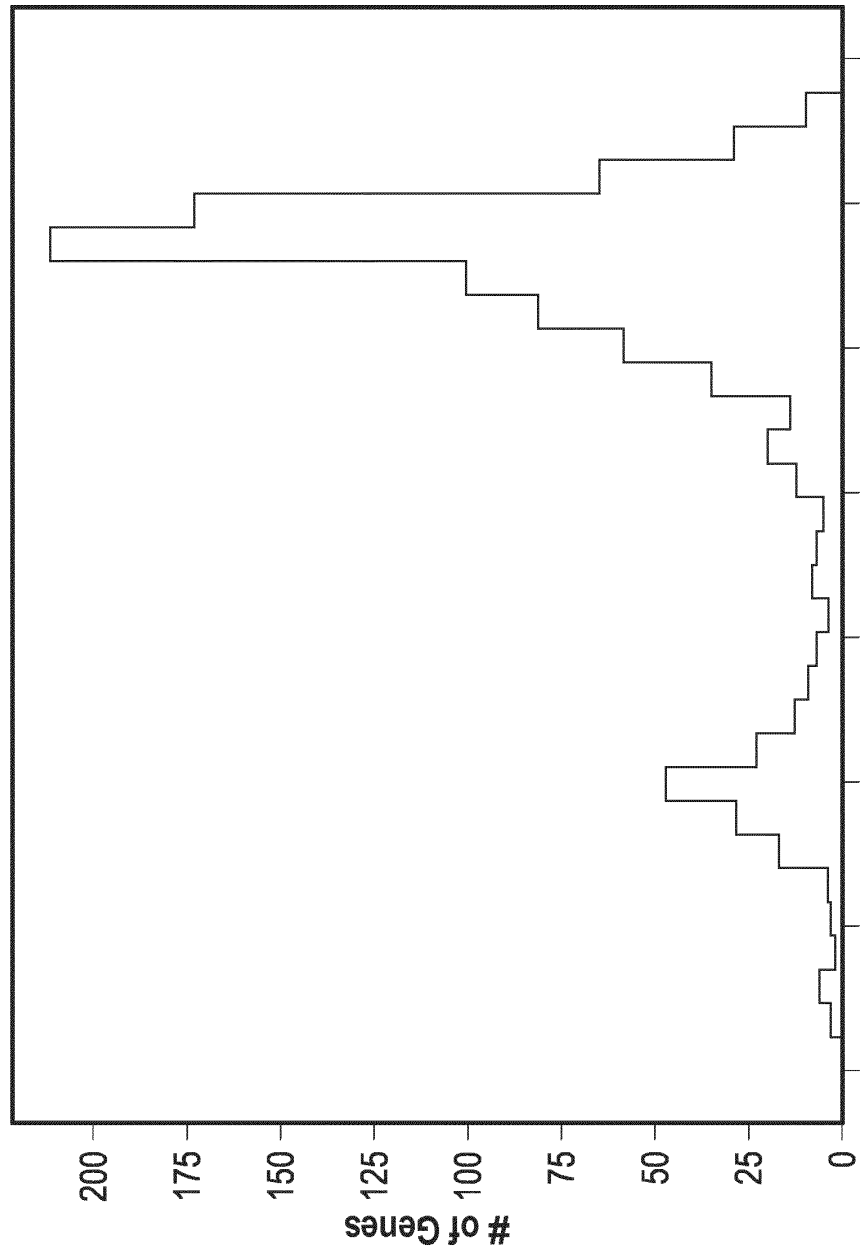


FIG. 12

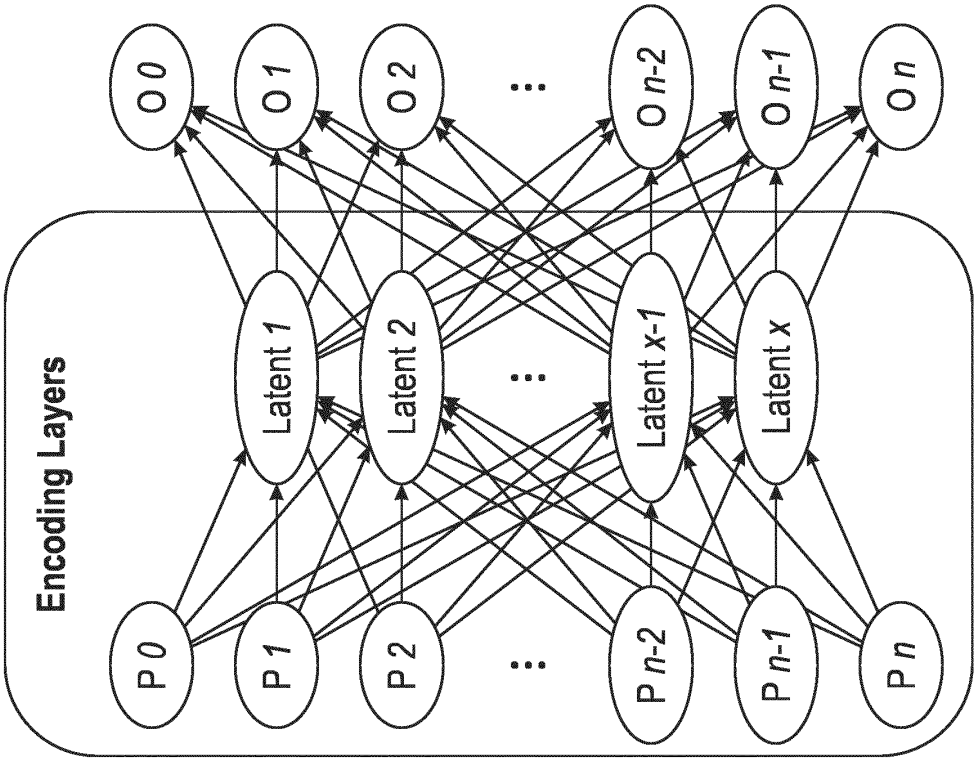

Encoding Layers
O 0
O 1
O 2
...
O n-2
O n-1
O n
Latent 1
Latent 2
...
Latent x-1
Latent x
P 0
P 1
P 2
...
P n-2
P n-1
P n

FIG. 13A

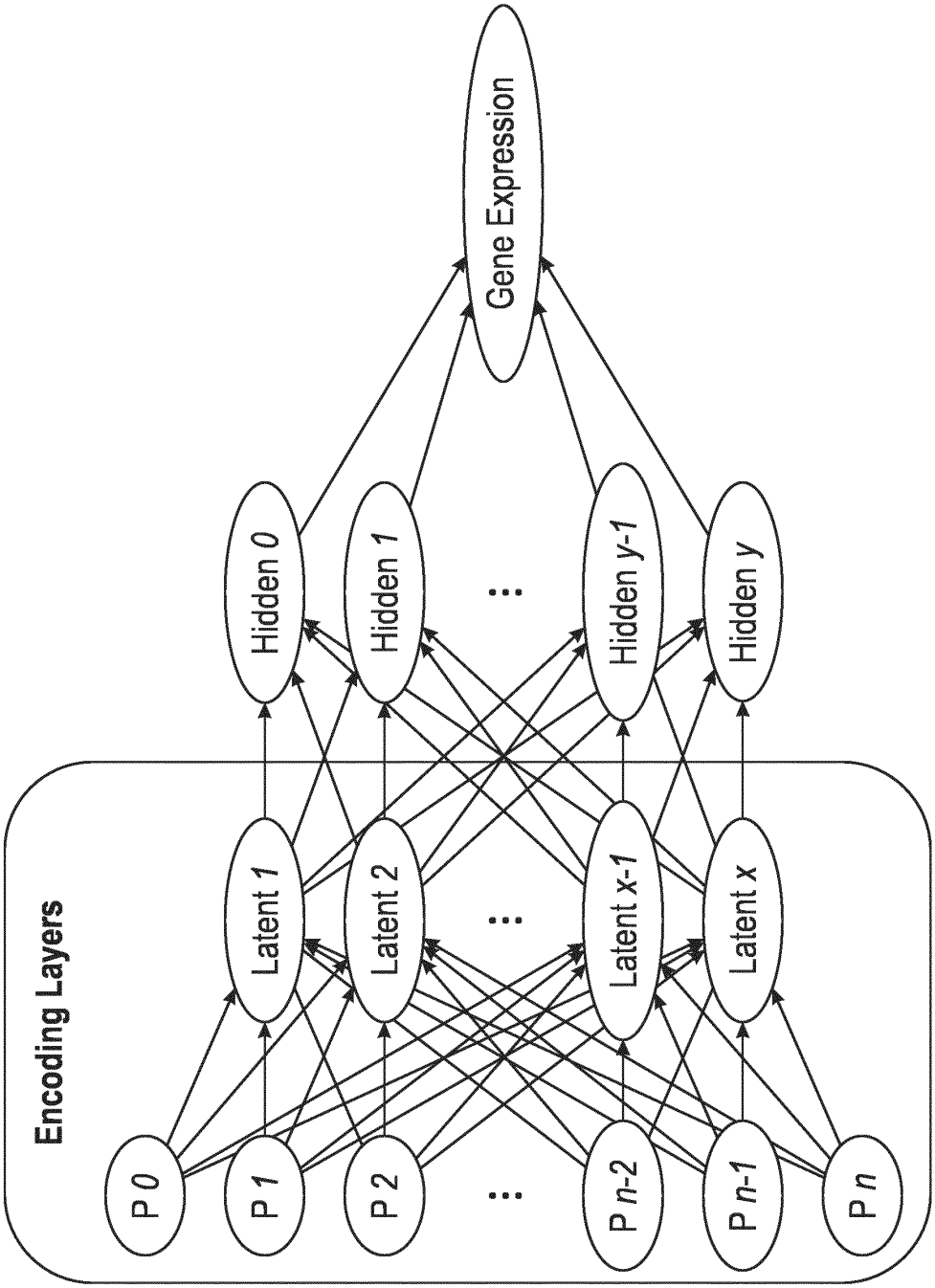

Encoding Layers
P 0
P 1
P 2
P n-2
P n-1
P n
Latent 1
Latent 2
Latent x-1
Latent x
Hidden 0
Hidden 1
Hidden y-1
Hidden y
Gene Expression

FIG. 13B

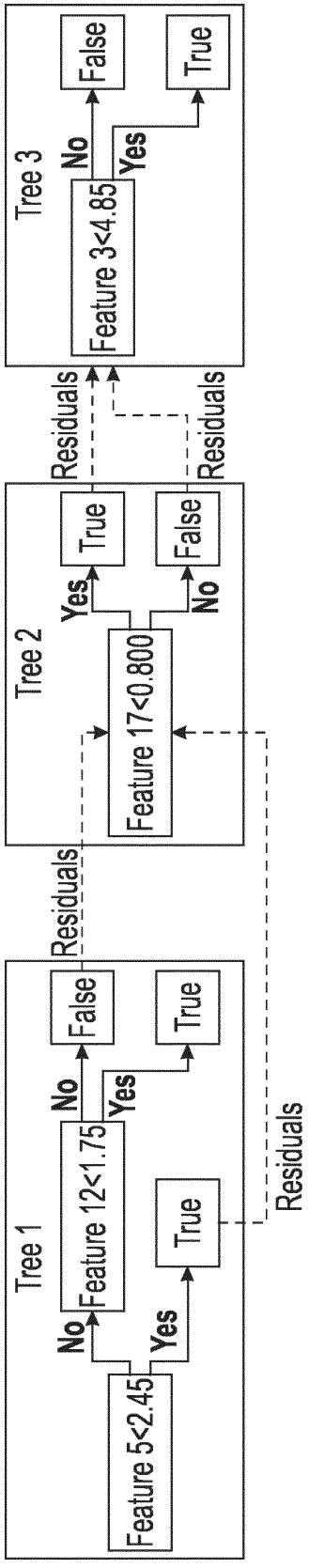

Tree 1
Feature 5<2.45
No
Yes
Feature 12<1.75
No
Yes
False
True
True
Residuals
Residuals
Tree 2
Feature 17<0.800
Yes
No
True
False
Residuals
Residuals
Tree 3
Feature 3<4.85
No
Yes
False
True

FIG. 14

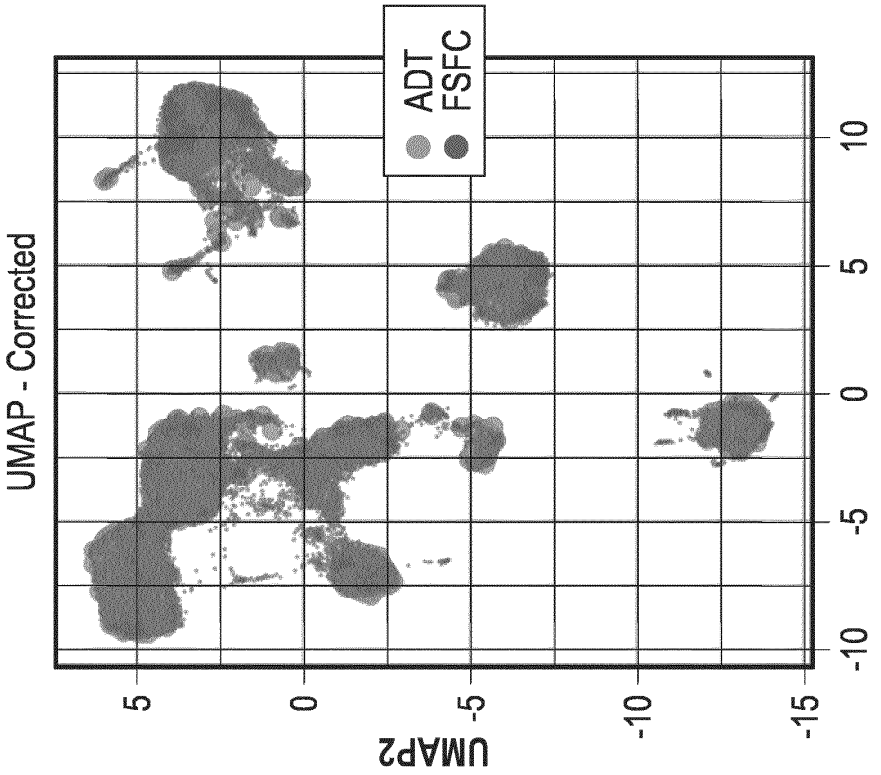


FIG. 15B

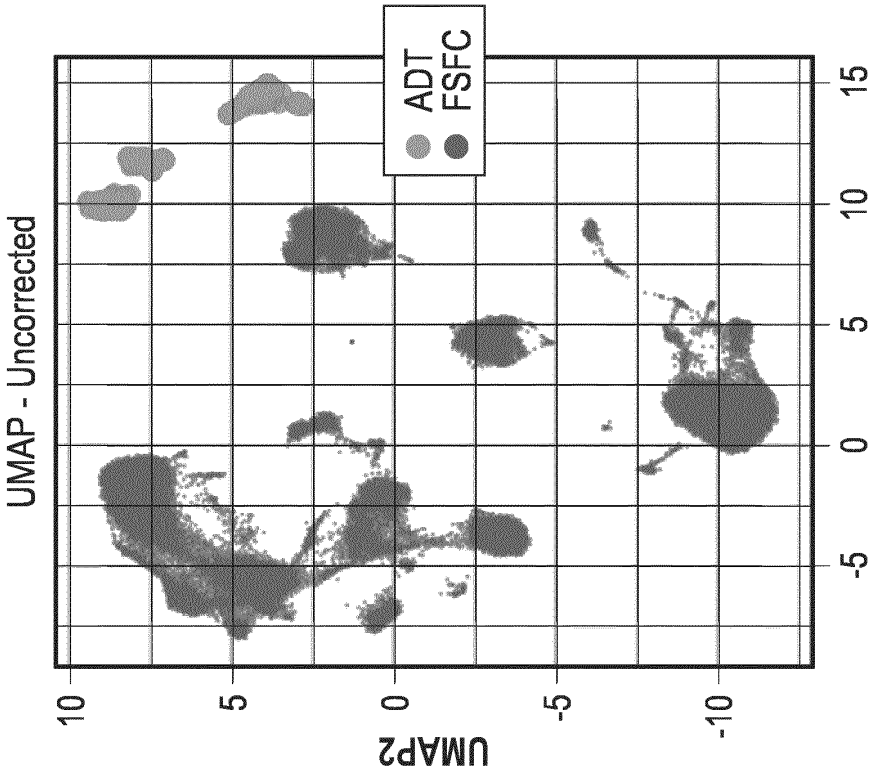


FIG. 15A

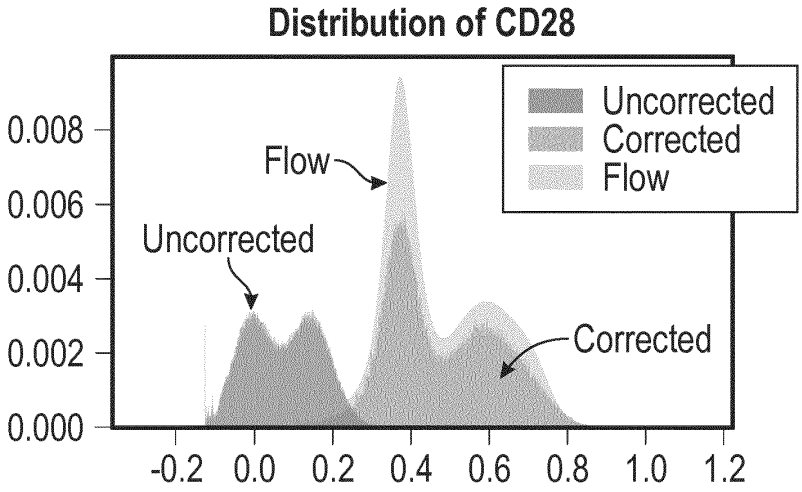


FIG. 16A

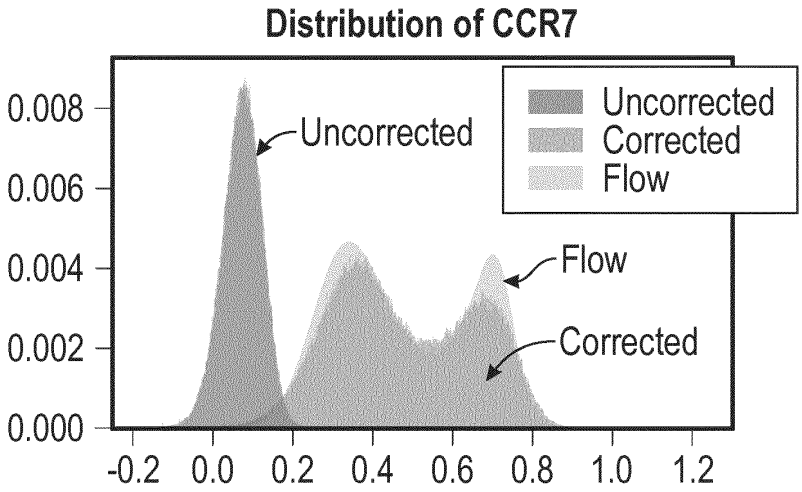


FIG. 16B

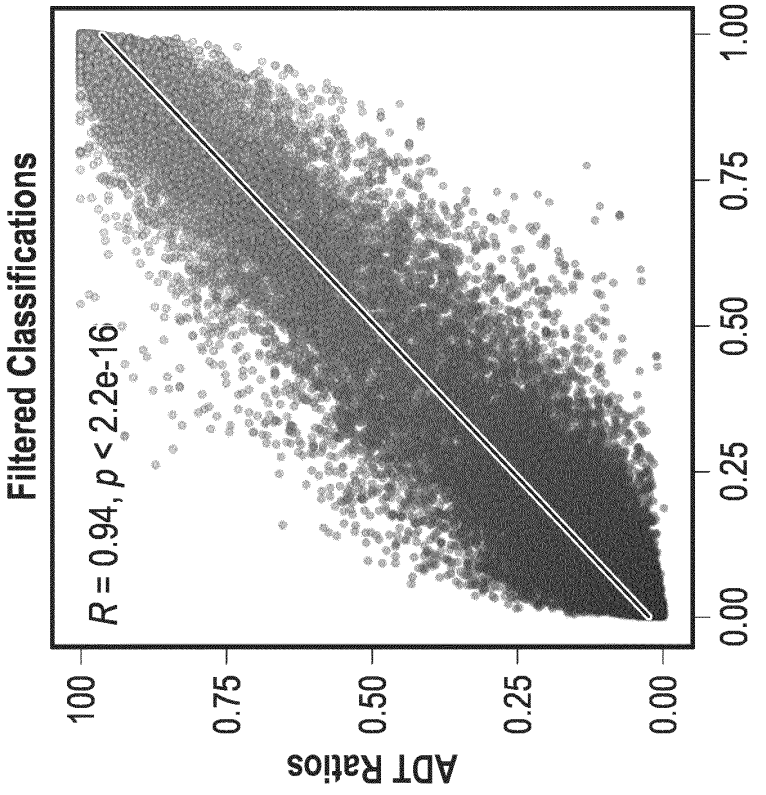


FIG. 17B

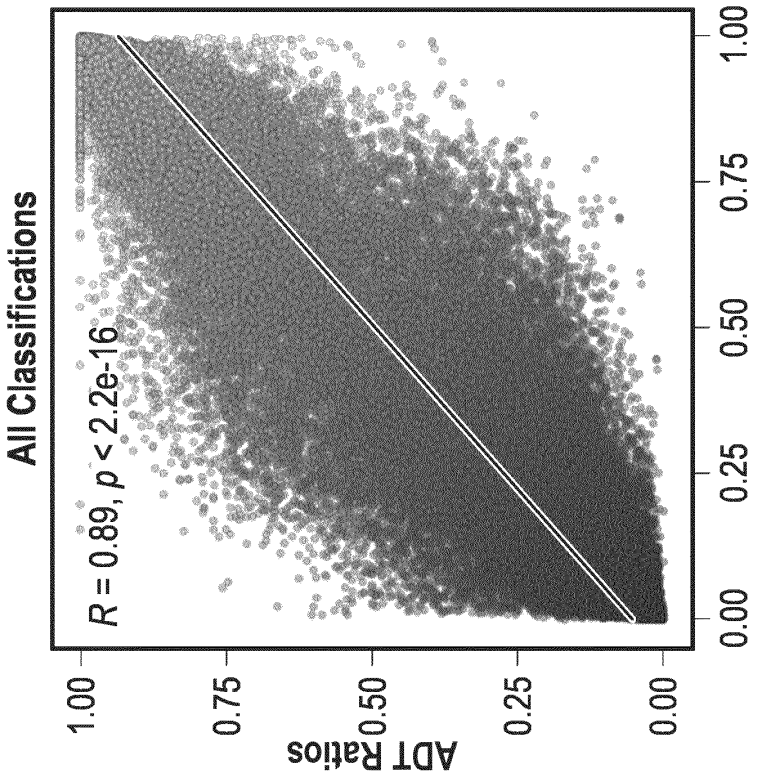


FIG. 17A

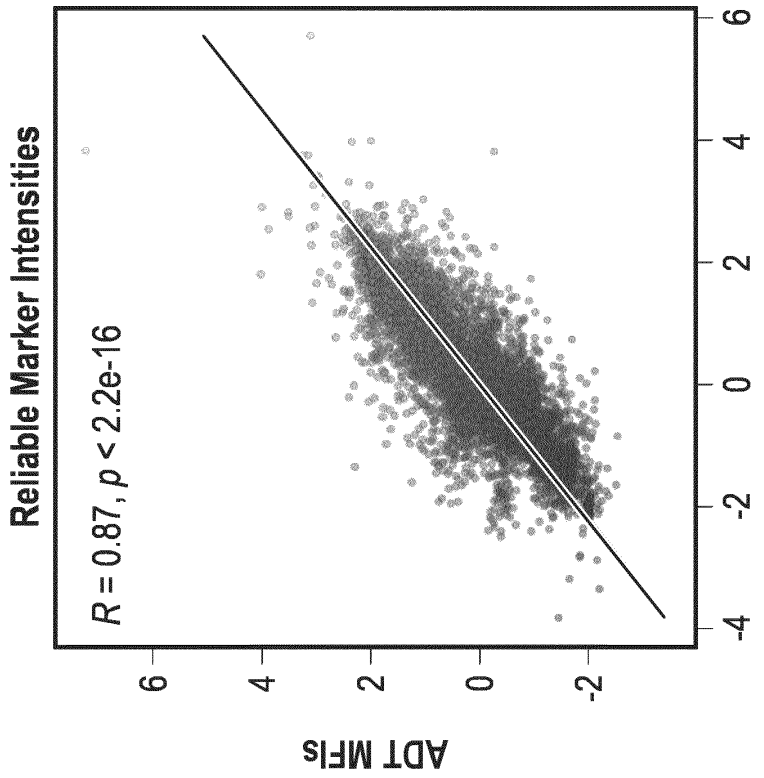


FIG. 18B

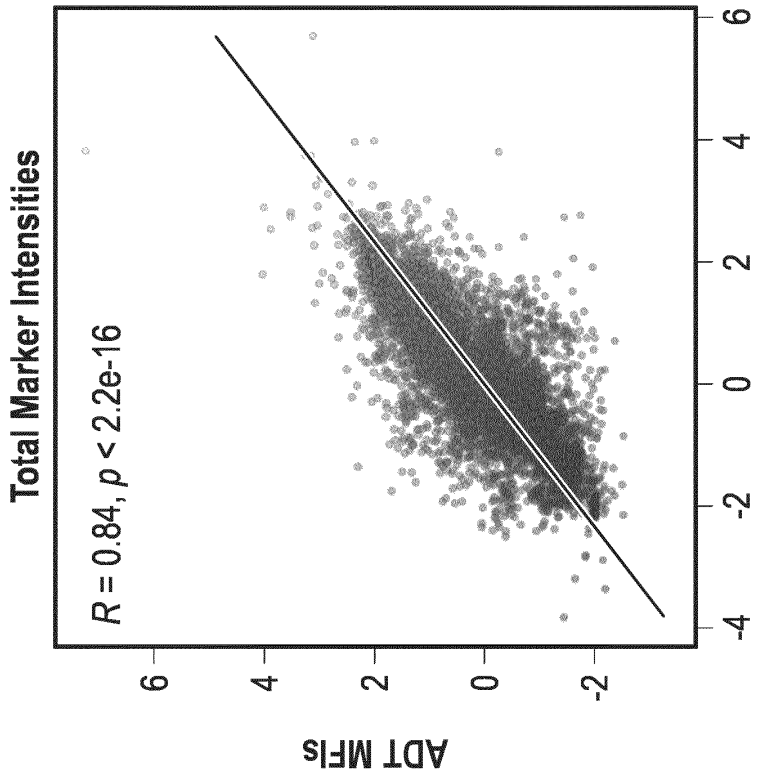


FIG. 18A

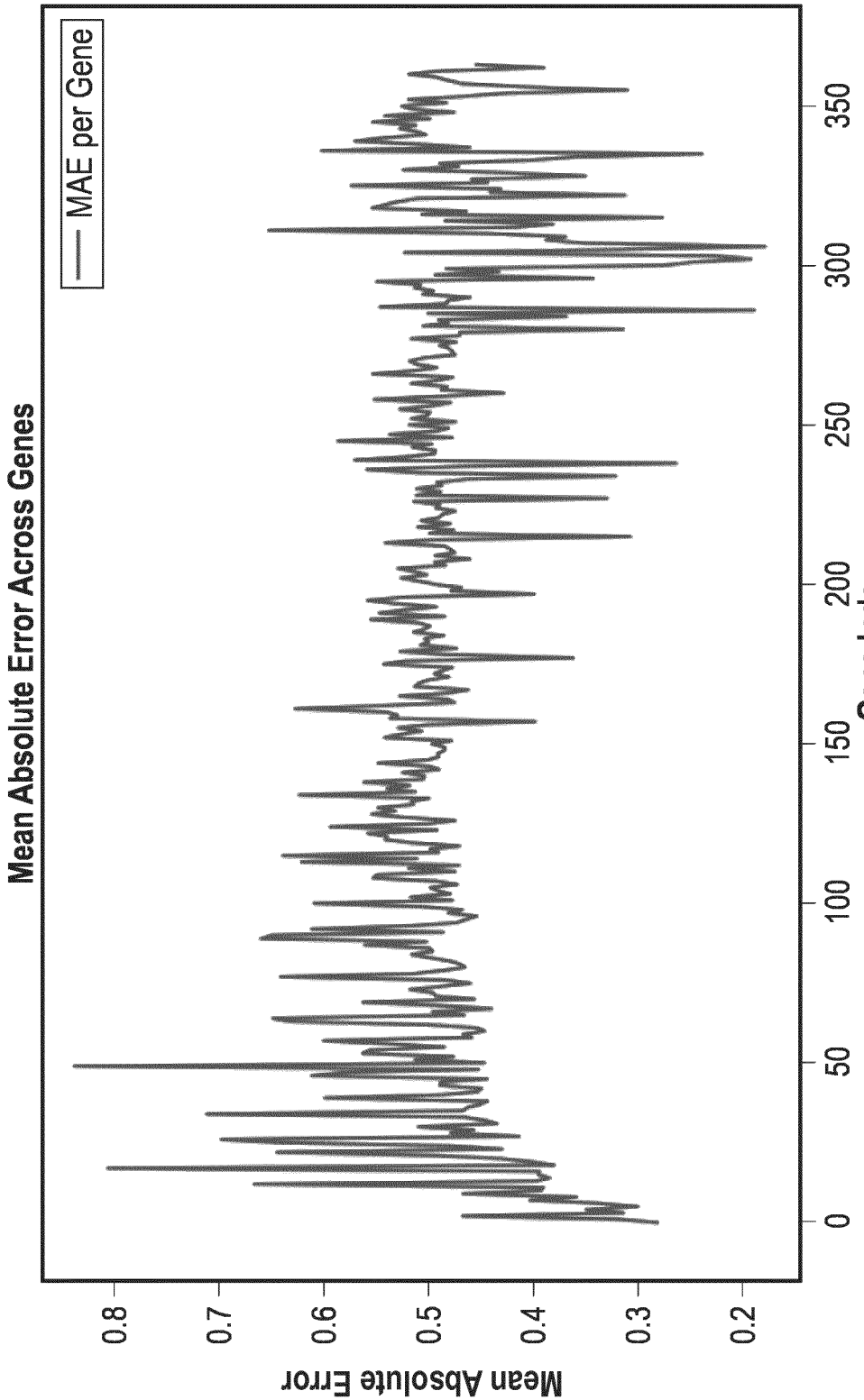

Mean Absolute Error Across Genes
MAE per Gene
Mean Absolute Error
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Gene Index
0
50
100
150
200
250
300
350

FIG. 19

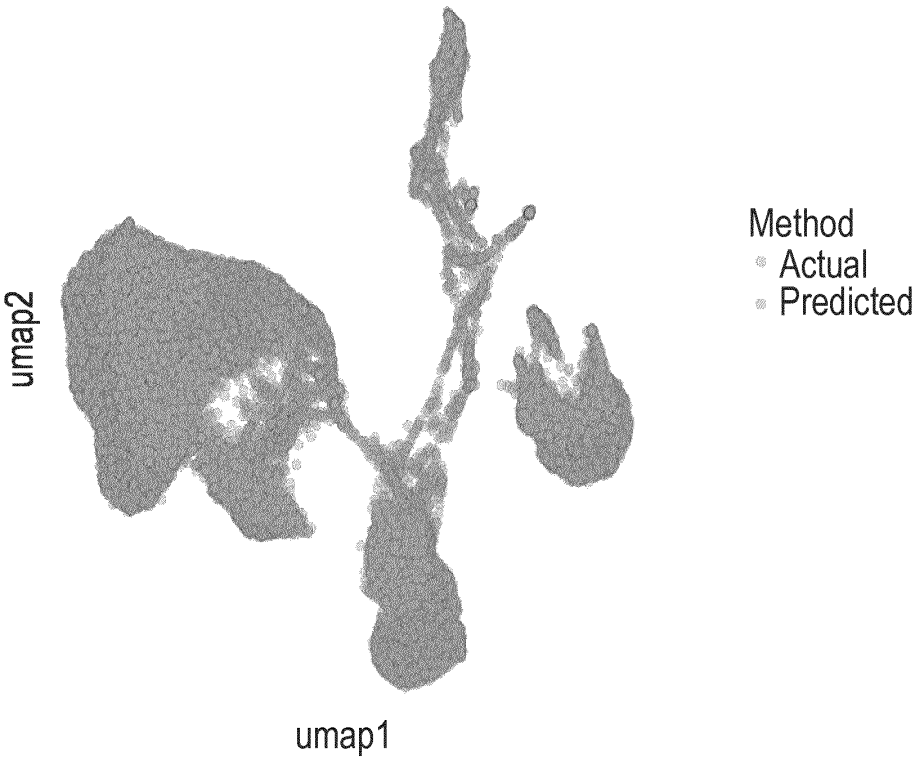


FIG. 20A

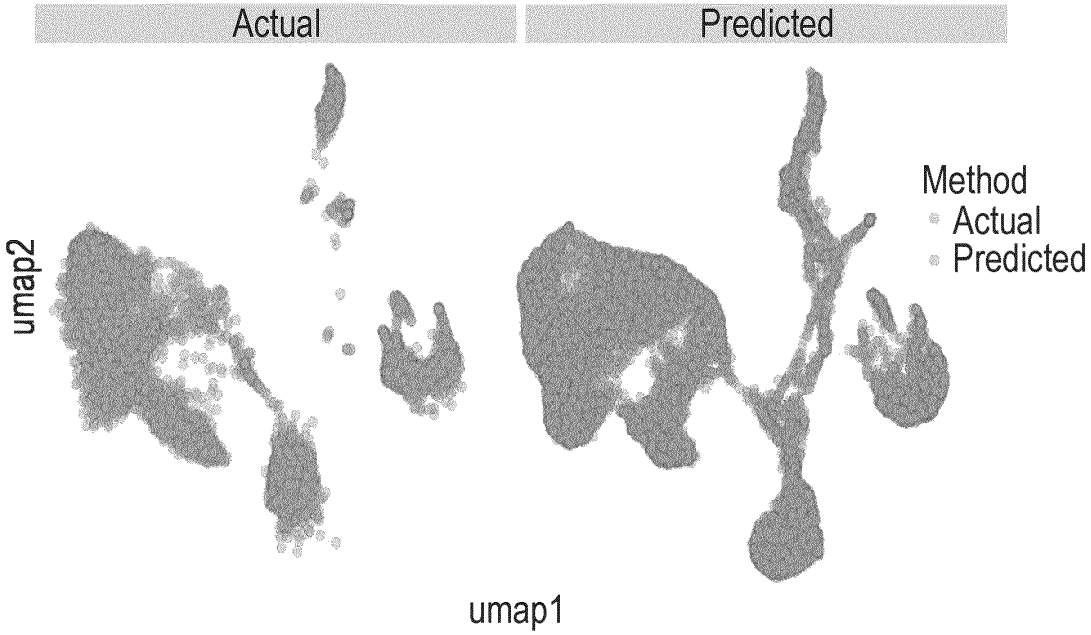


FIG. 20B

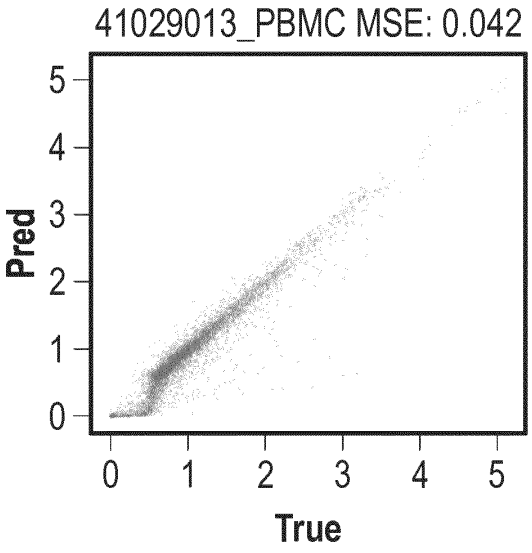


FIG. 21A

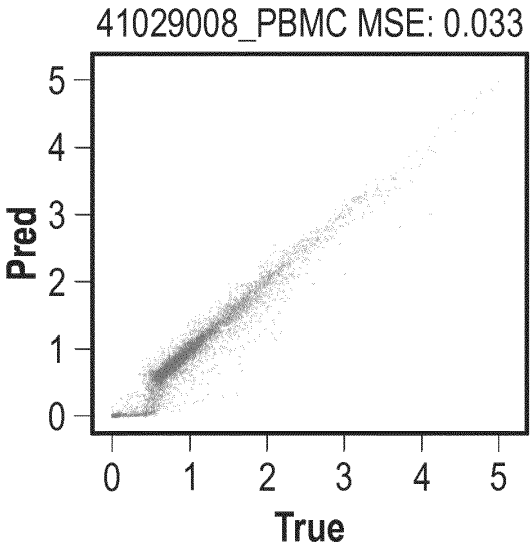


FIG. 21B

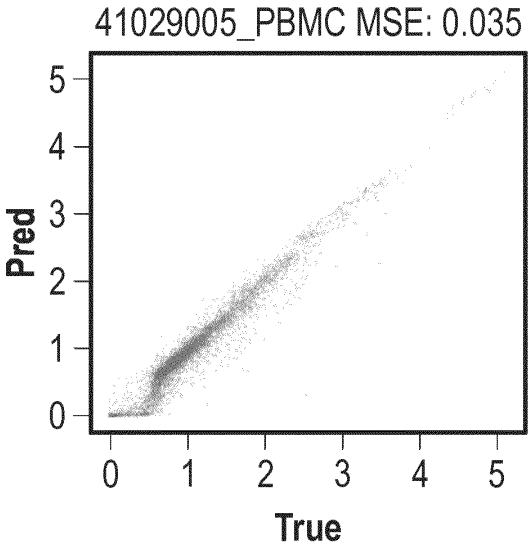


FIG. 21C

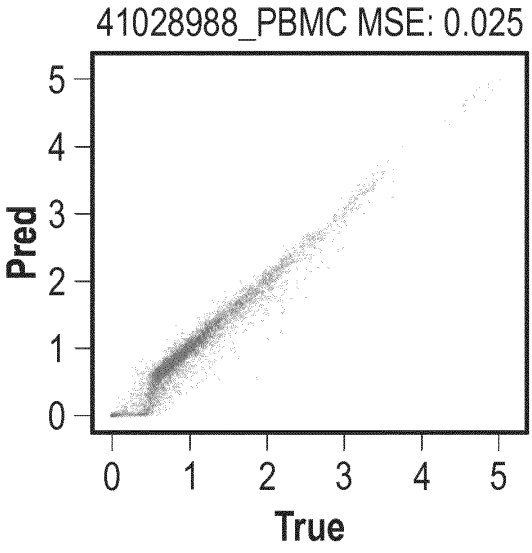


FIG. 21D

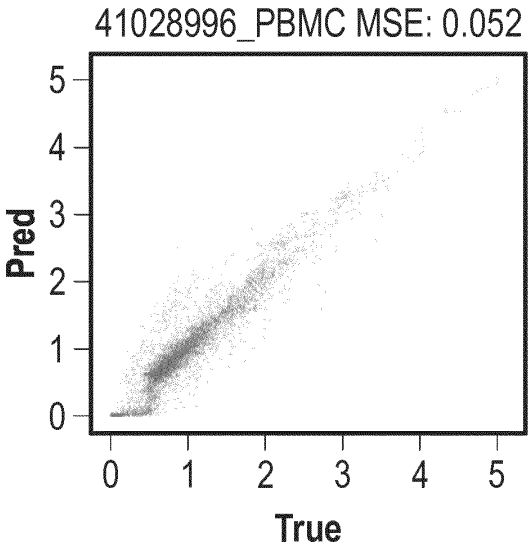


FIG. 21E

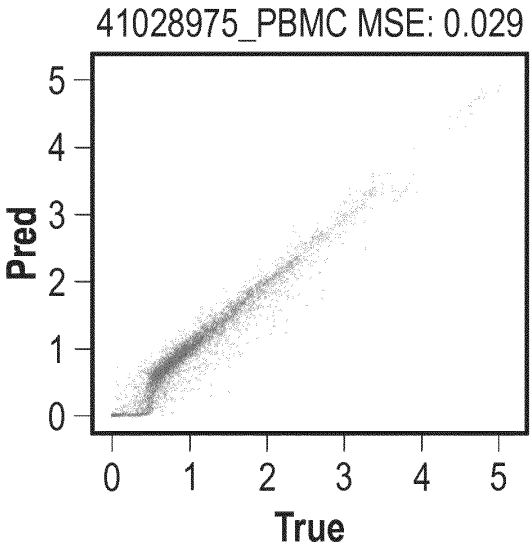


FIG. 21F

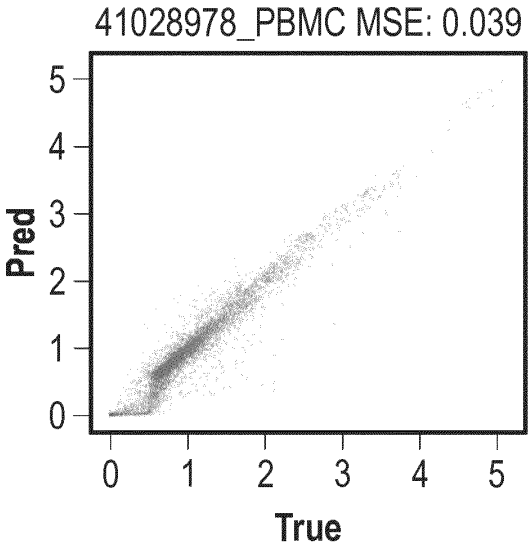


FIG. 21G

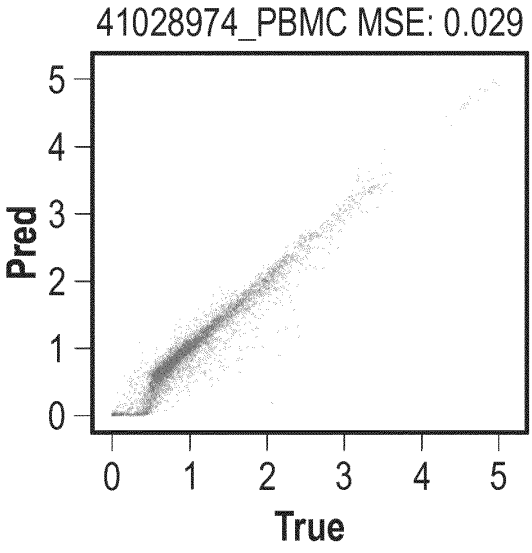


FIG. 21H

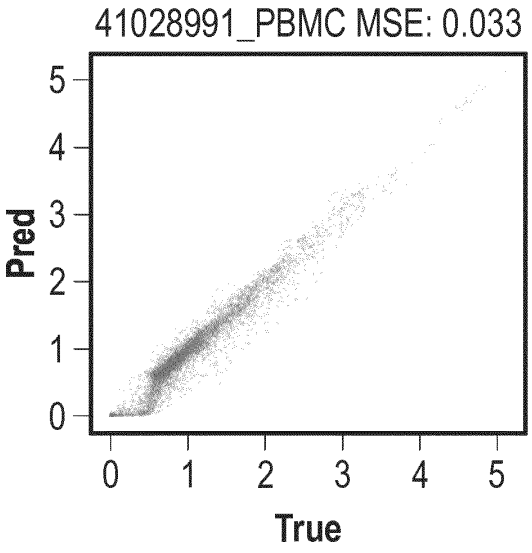


FIG. 21I

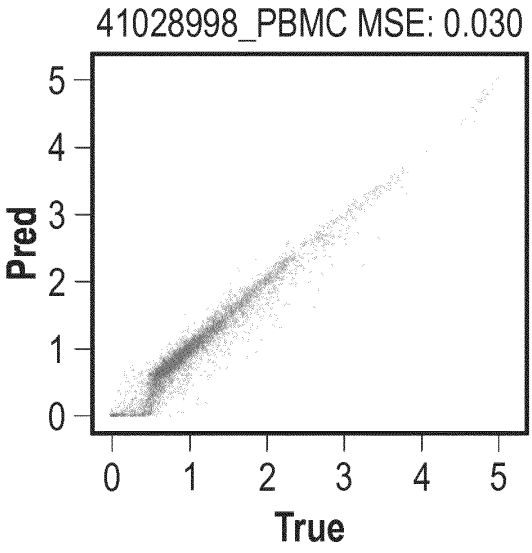


FIG. 21J

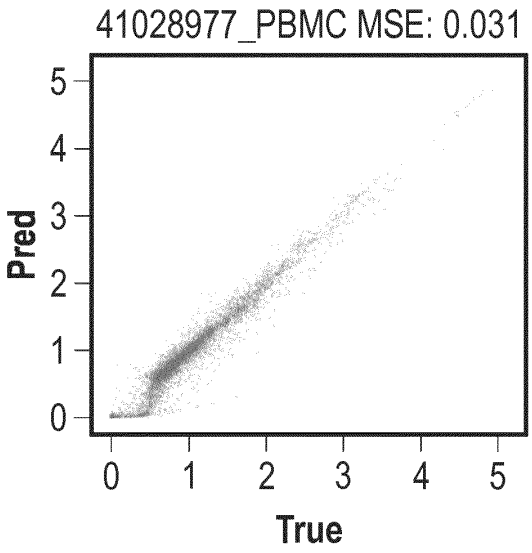


FIG. 21K

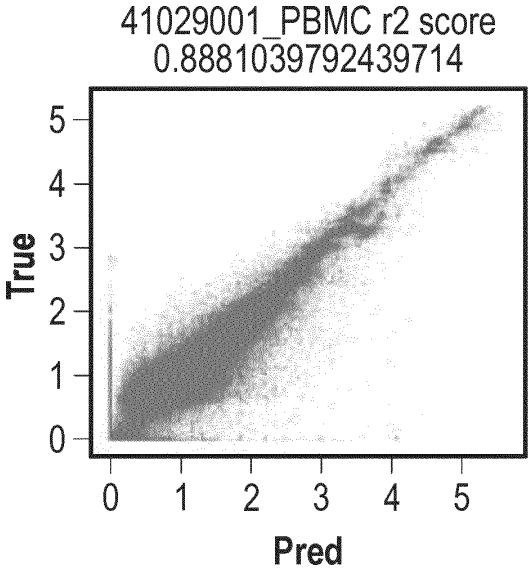


FIG. 22A

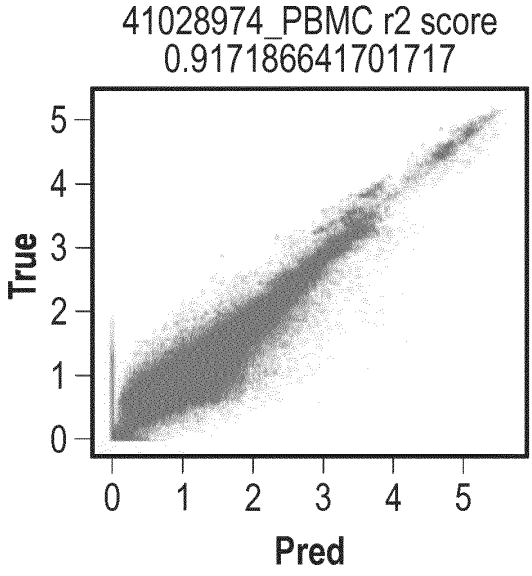


FIG. 22B

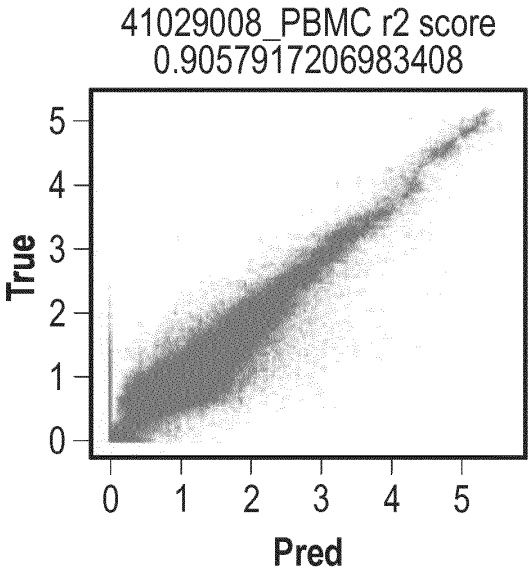


FIG. 22C

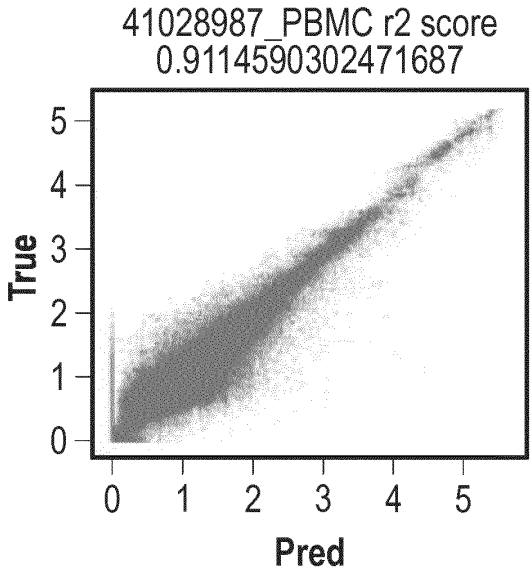


FIG. 22D

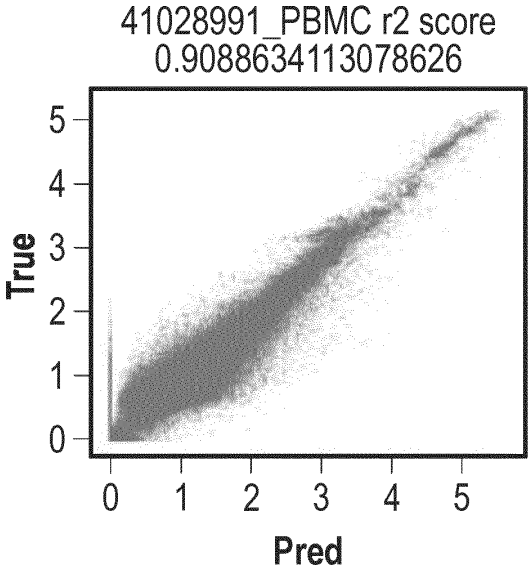


FIG. 22E

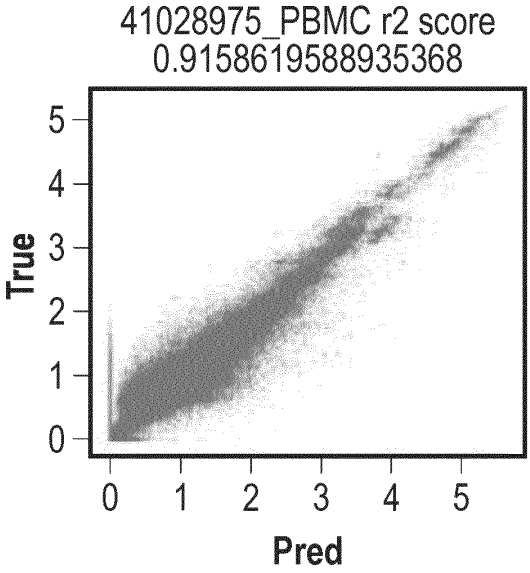


FIG. 22F

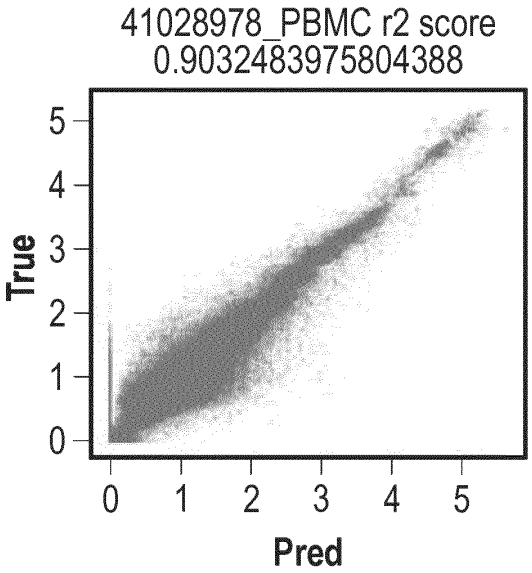


FIG. 22G

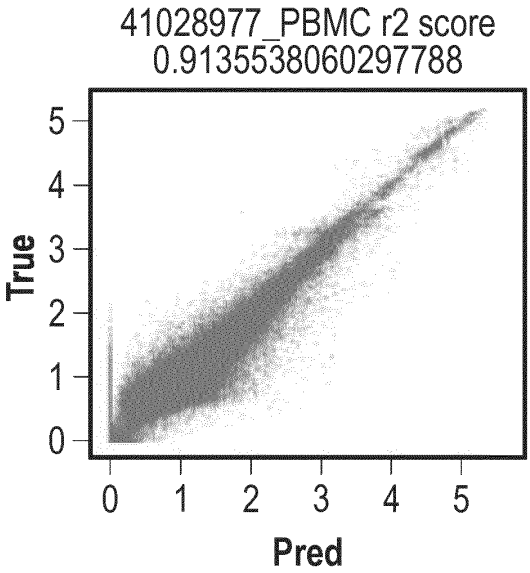


FIG. 22H

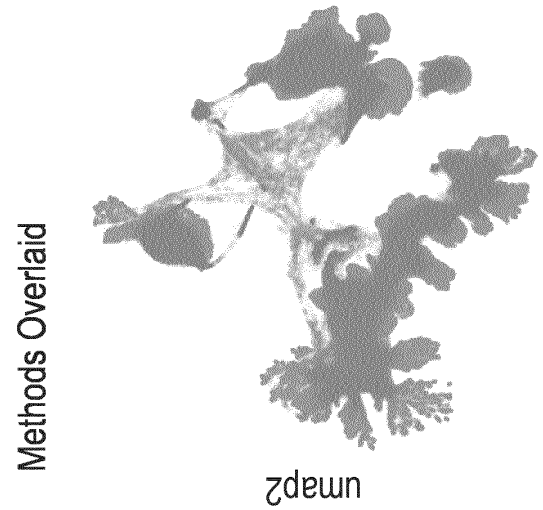


FIG. 23B

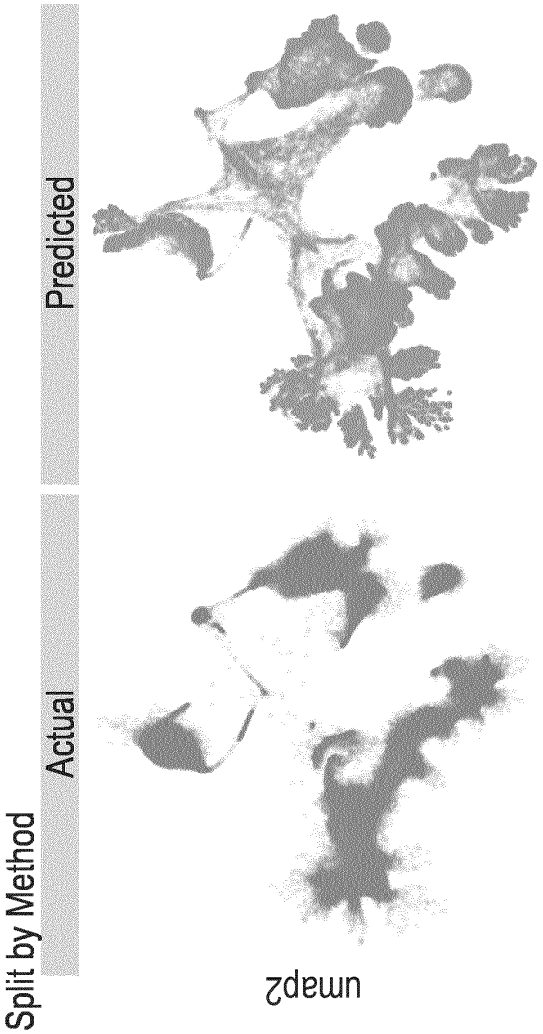


FIG. 23A

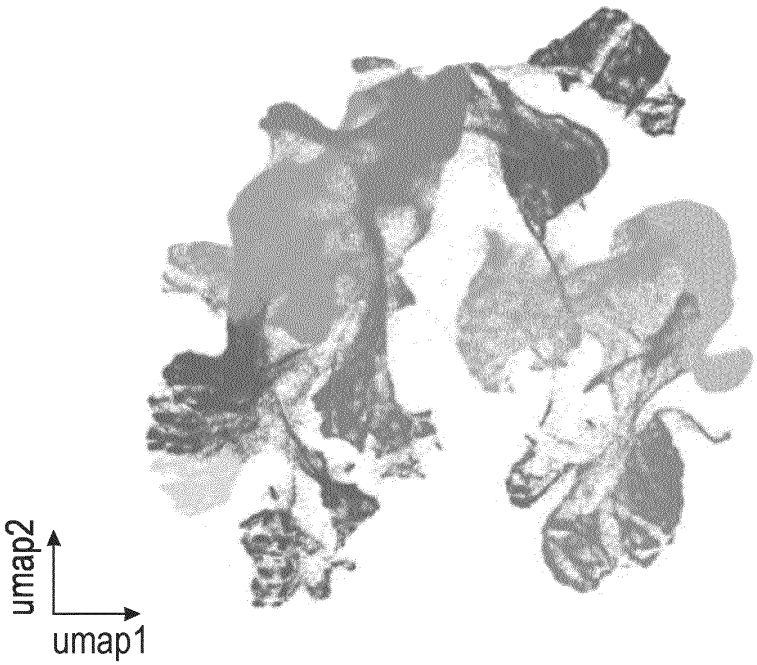


FIG. 24A

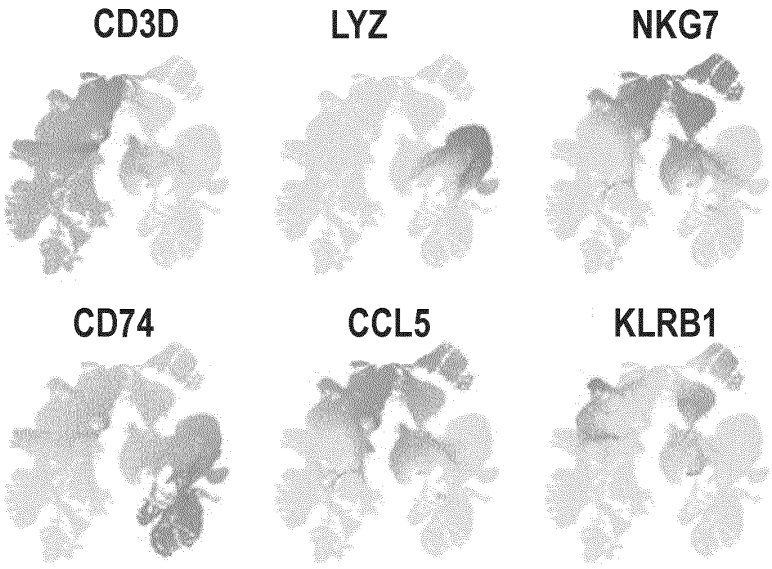
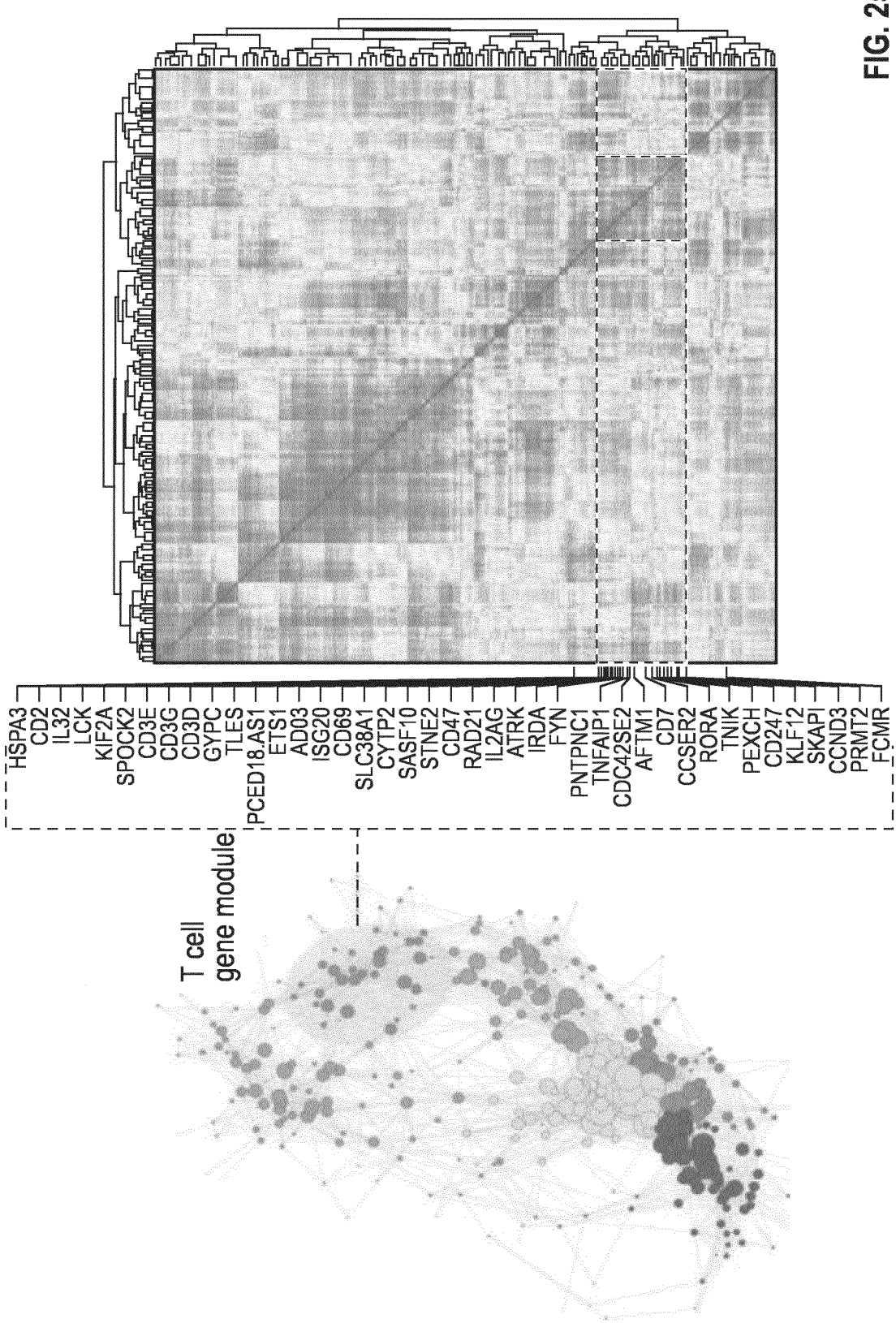


FIG. 24B

FIG. 25


HSPA3
CD2
IL32
LCK
KIF2A
SPOCK2
CD3E
CD3G
CD3D
GYPC
TLES
PCED18.AS1
ETS1
AD03
ISG20
CD69
SLC38A1
CYTP2
SASF10
STNE2
CD47
RAD21
IL2AG
ATRK
IRDA
FYN
PNTPNC1
TNFAIP1
CDC42SE2
AFTM1
CD7
CCSER2
RORA
TNIK
PEXCH
CD247
KLF12
SKAP1
CCND3
PRMT2
FCMR
T cell
gene module

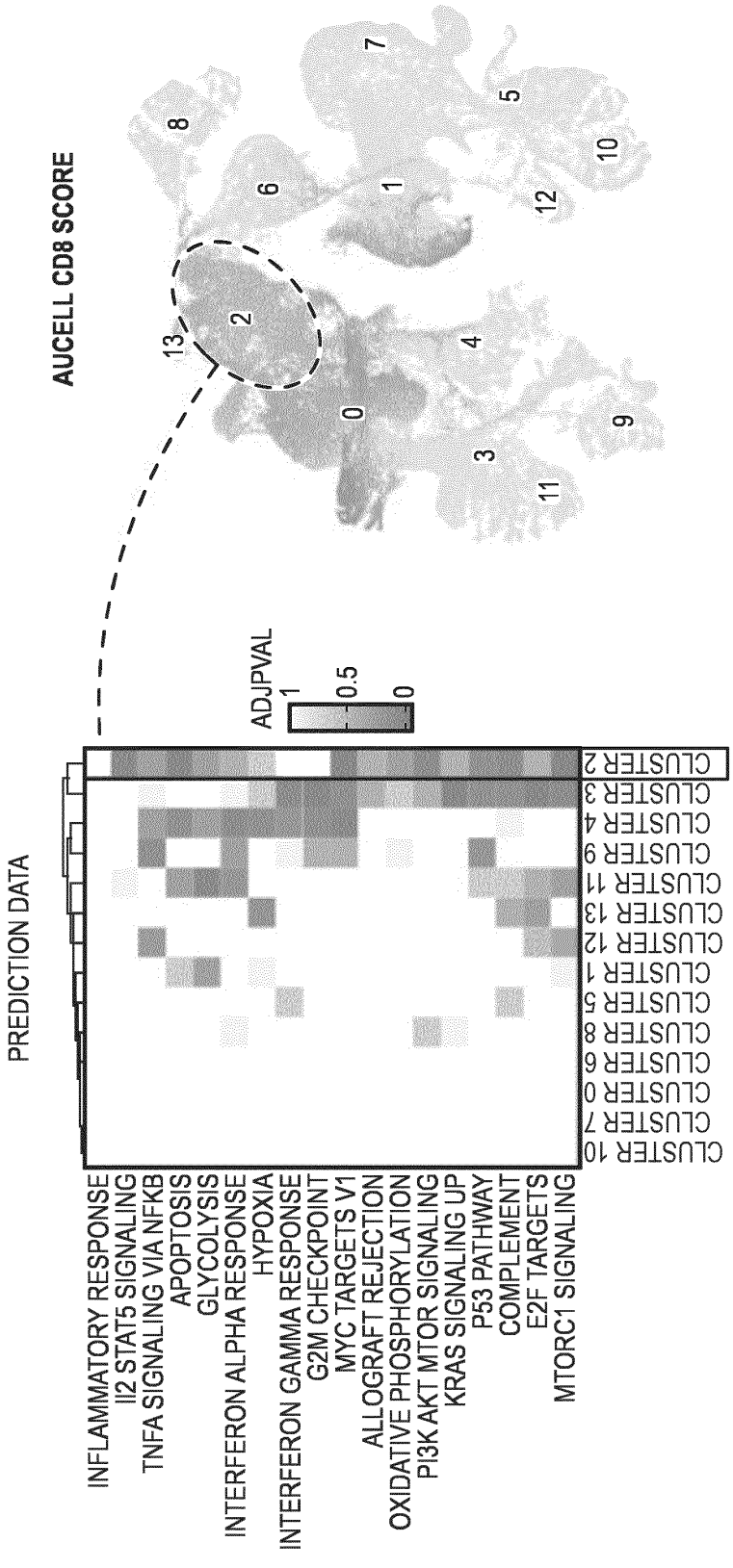

AUCELL CD8 SCORE
0
1
2
3
4
5
6
7
8
9
10
11
12
13
ADJPVAL
1
0.5
0
PREDICTION DATA
CLUSTER 2
CLUSTER 3
CLUSTER 4
CLUSTER 9
CLUSTER 11
CLUSTER 13
CLUSTER 12
CLUSTER 1
CLUSTER 5
CLUSTER 8
CLUSTER 6
CLUSTER 0
CLUSTER 7
CLUSTER 10
INFLAMMATORY RESPONSE
Il2 STAT5 SIGNALING
TNFA SIGNALING VIA NFKB
APOPTOSIS
GLYCOLYSIS
INTERFERON ALPHA RESPONSE
HYPOXIA
INTERFERON GAMMA RESPONSE
G2M CHECKPOINT
MYC TARGETS V1
ALLOGRAFT REJECTION
OXIDATIVE PHOSPHORYLATION
PI3K AKT MTOR SIGNALING
KRAS SIGNALING UP
P53 PATHWAY
COMPLEMENT
E2F TARGETS
MTORC1 SIGNALING

FIG. 26A

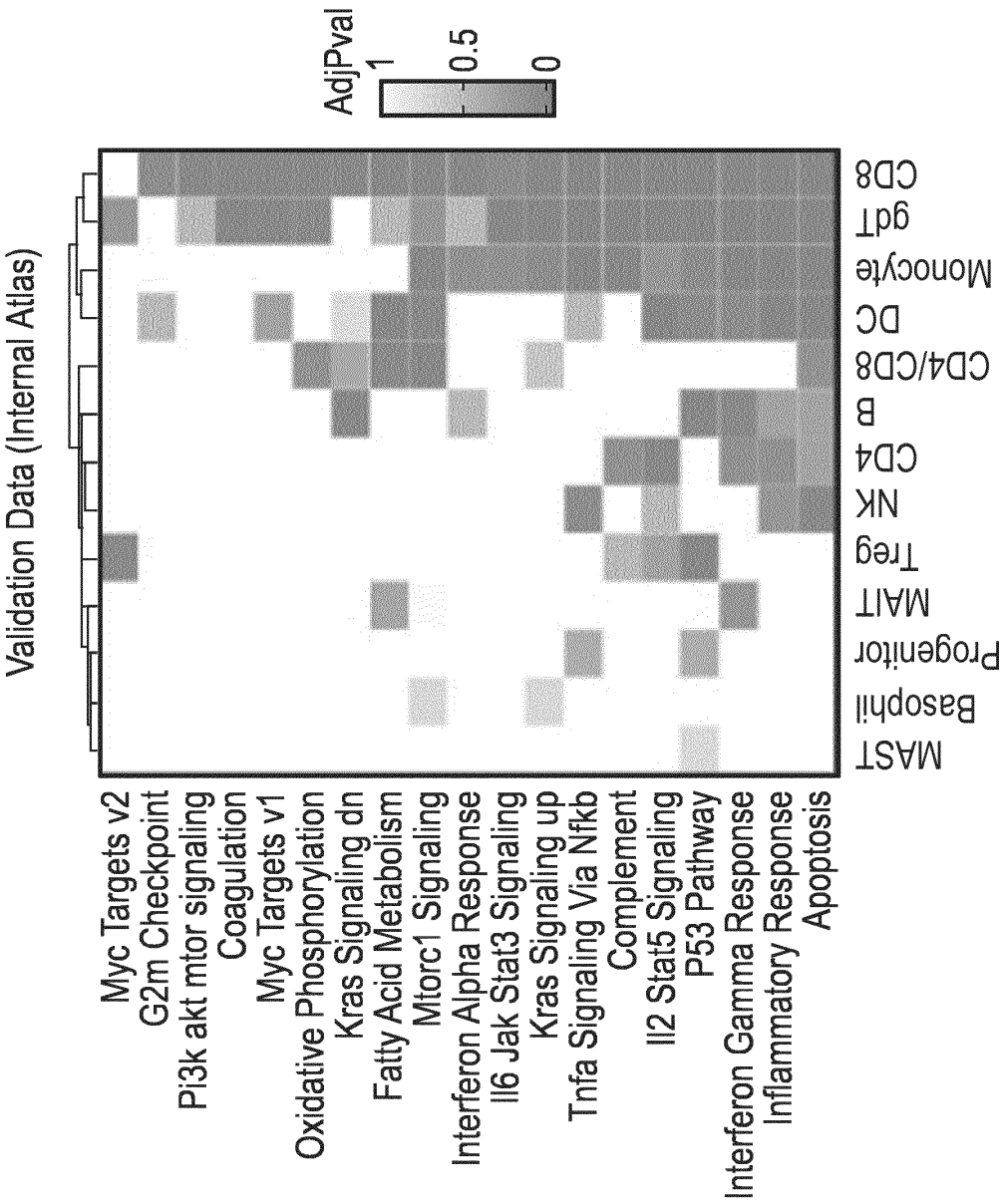

AdjPval
1
0.5
0
Validation Data (Internal Atlas)
Myc Targets v2
G2m Checkpoint
Pi3k akt mtor signaling
Coagulation
Myc Targets v1
Oxidative Phosphorylation
Kras Signaling dn
Fatty Acid Metabolism
Mtorc1 Signaling
Interferon Alpha Response
Il6 Jak Stat3 Signaling
Kras Signaling up
Tnfa Signaling Via Nfkb
Complement
Il2 Stat5 Signaling
P53 Pathway
Interferon Gamma Response
Inflammatory Response
Apoptosis
MAST
Basophil
Progenitor
MAIT
Treg
NK
CD4
B
CD4/CD8
DC
Monocyte
gdT
CD8

FIG. 26B



FIG. 27A



FIG. 27B

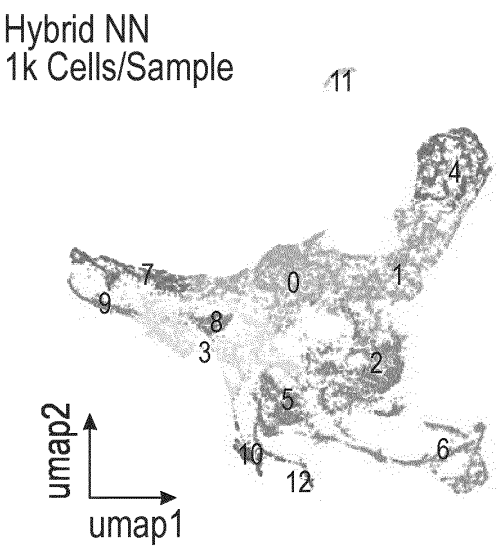


FIG. 28A

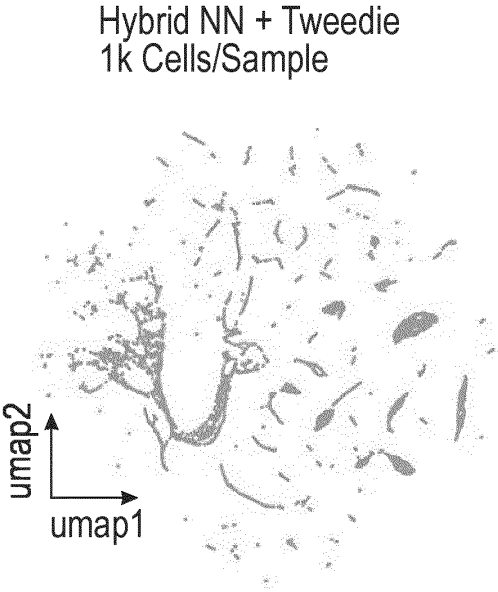


FIG. 28B

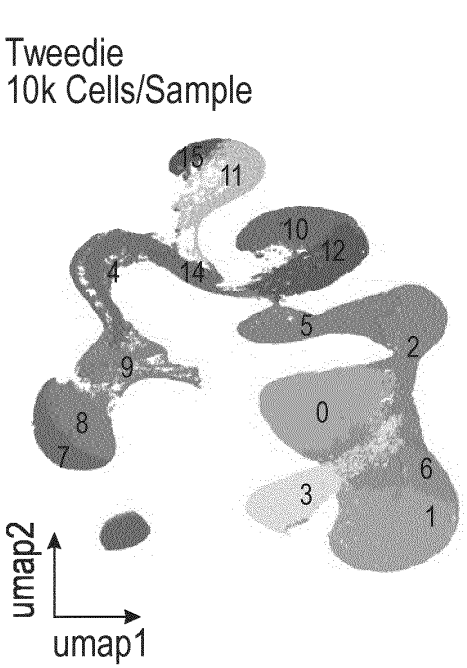


FIG. 28C

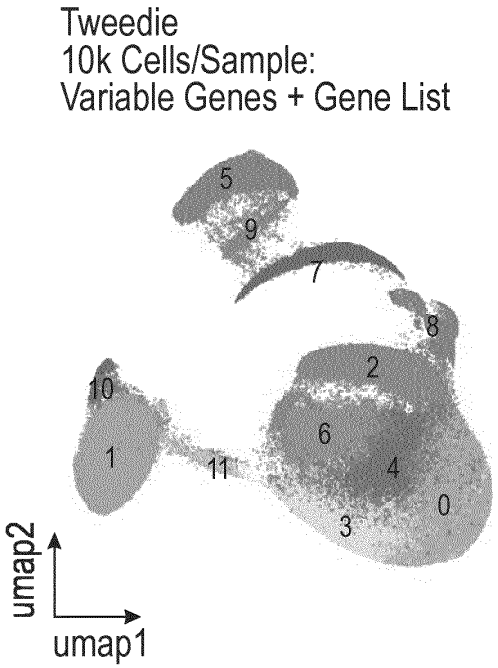


FIG. 28D

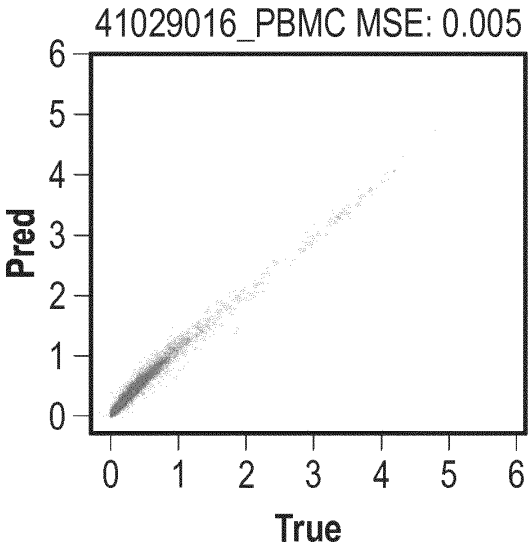


FIG. 29A

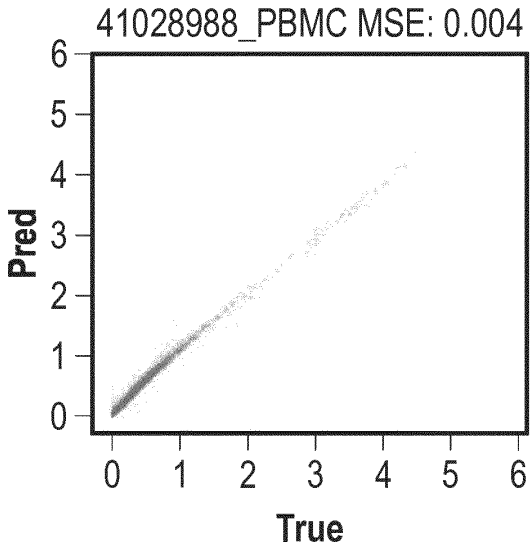


FIG. 29B

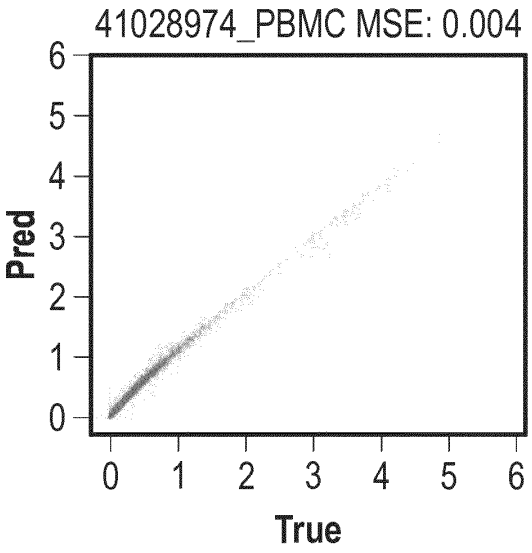


FIG. 29C

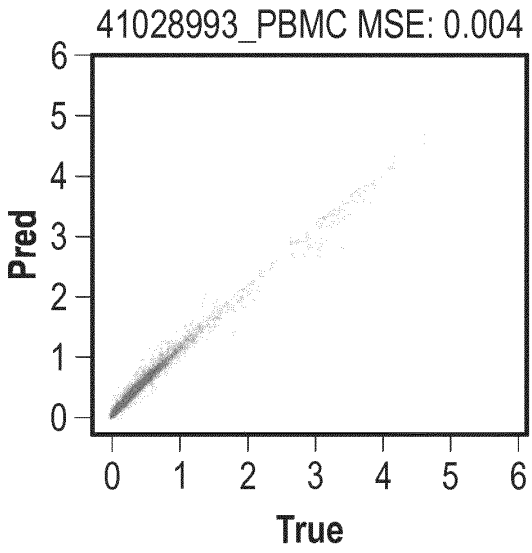


FIG. 29D

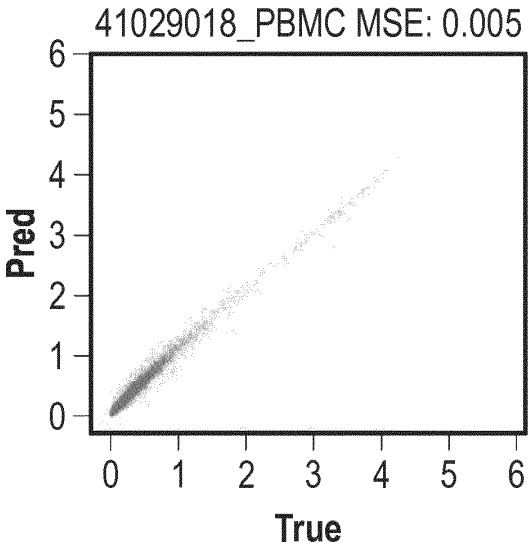


FIG. 29E

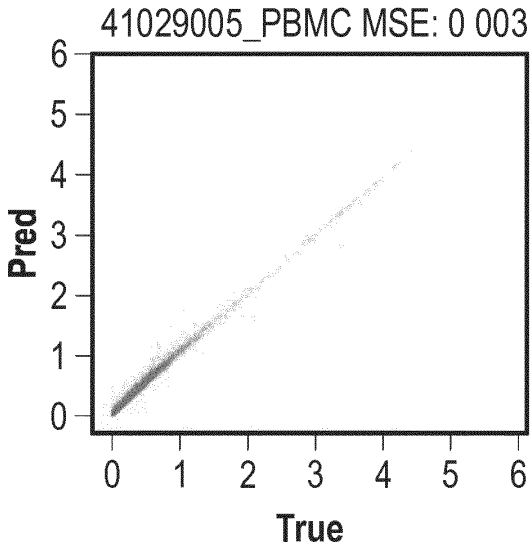


FIG. 29F

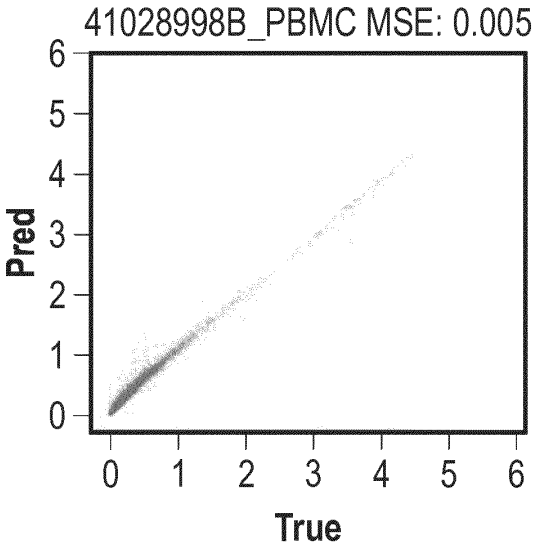


FIG. 29G

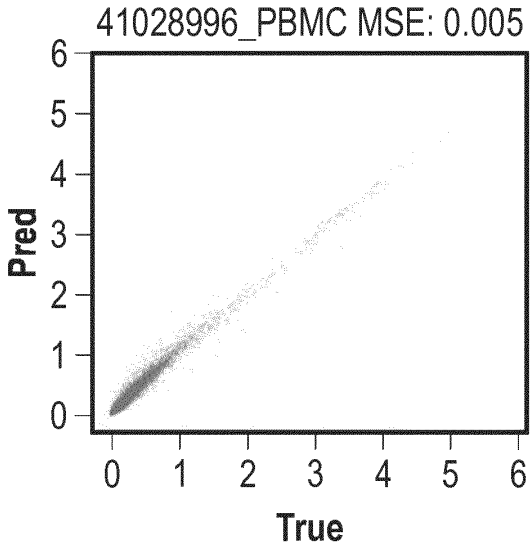


FIG. 29H

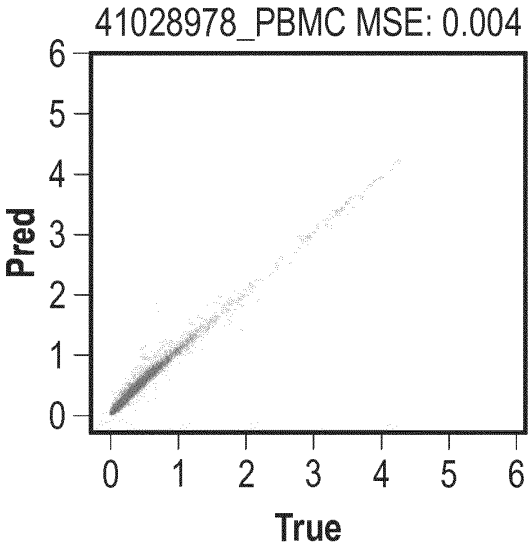


FIG. 29I

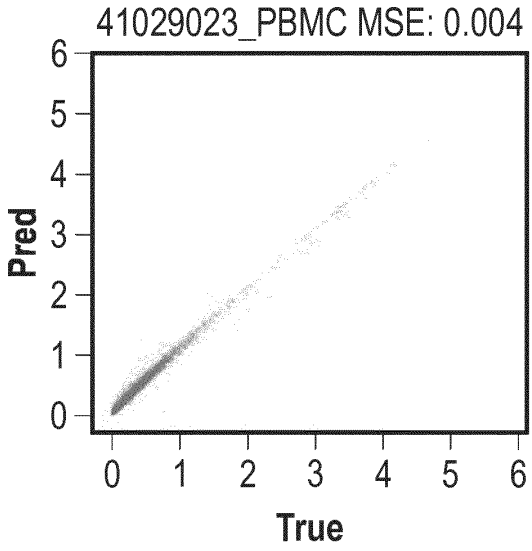


FIG. 29J

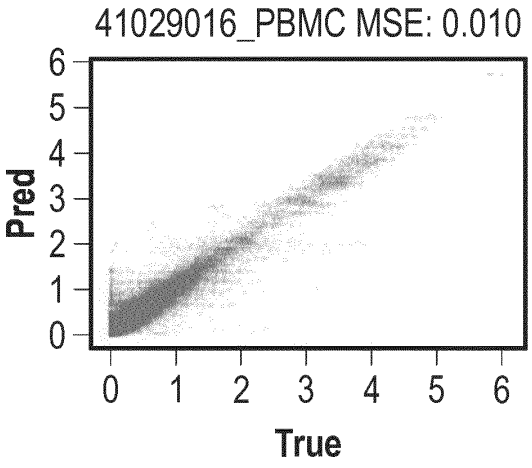


FIG. 30A

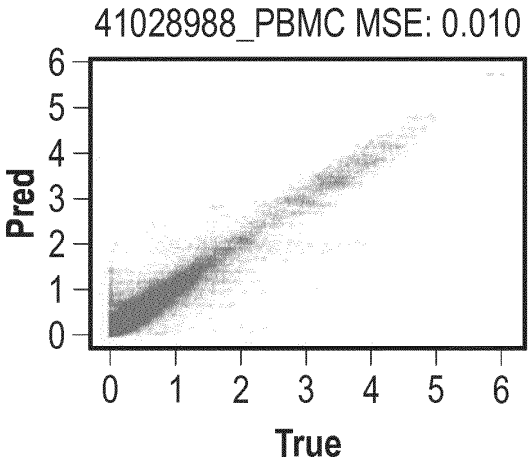


FIG. 30B

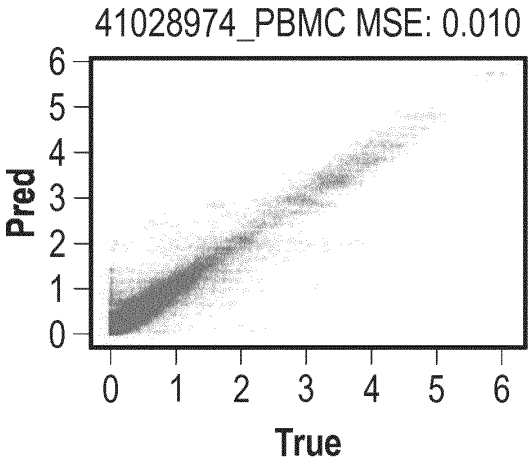


FIG. 30C

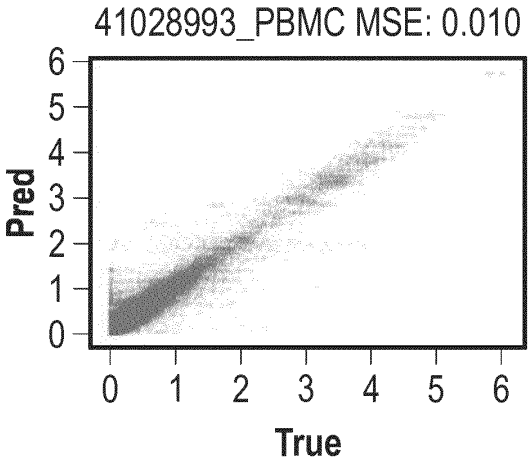


FIG. 30D

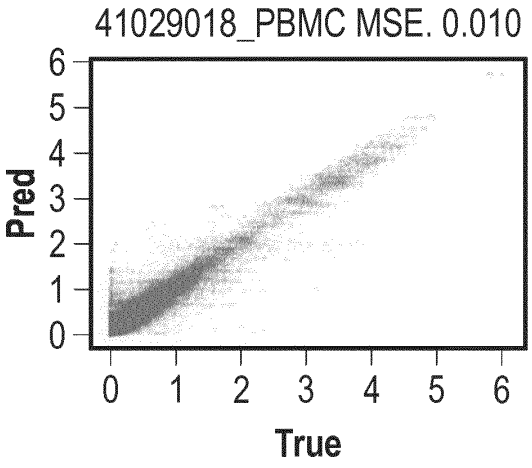


FIG. 30D

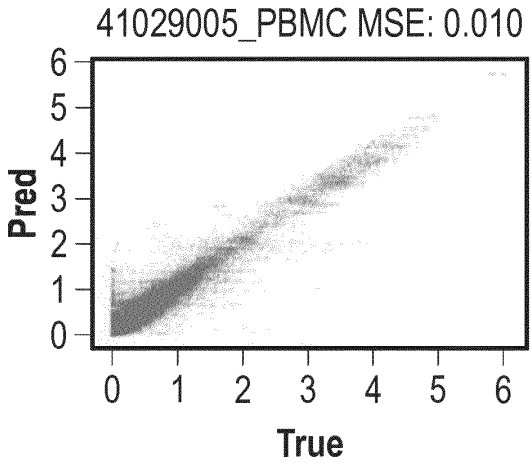


FIG. 30E

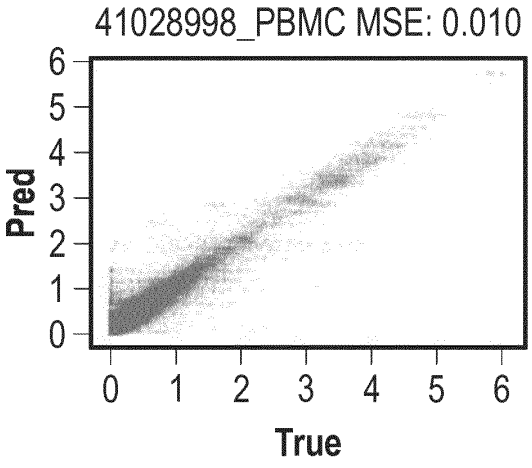


FIG. 30F

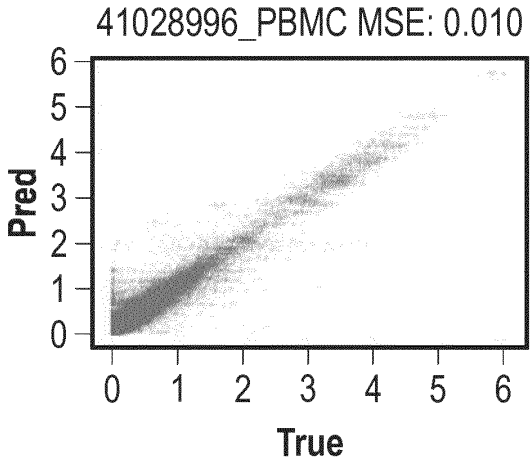


FIG. 30G

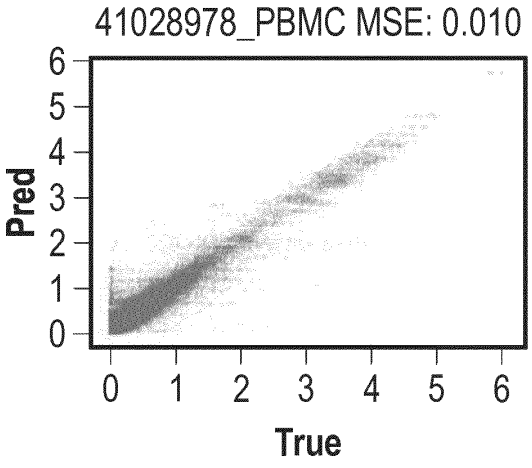


FIG. 30H

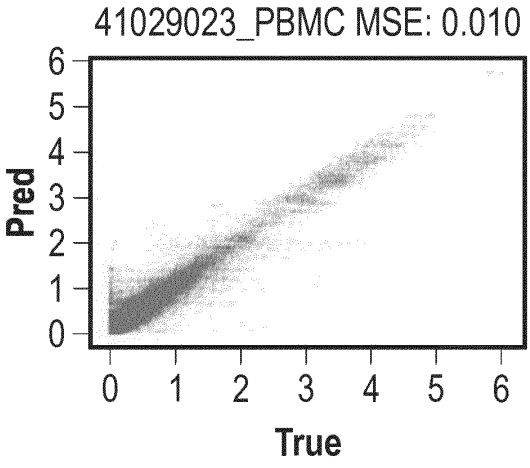


FIG. 30I

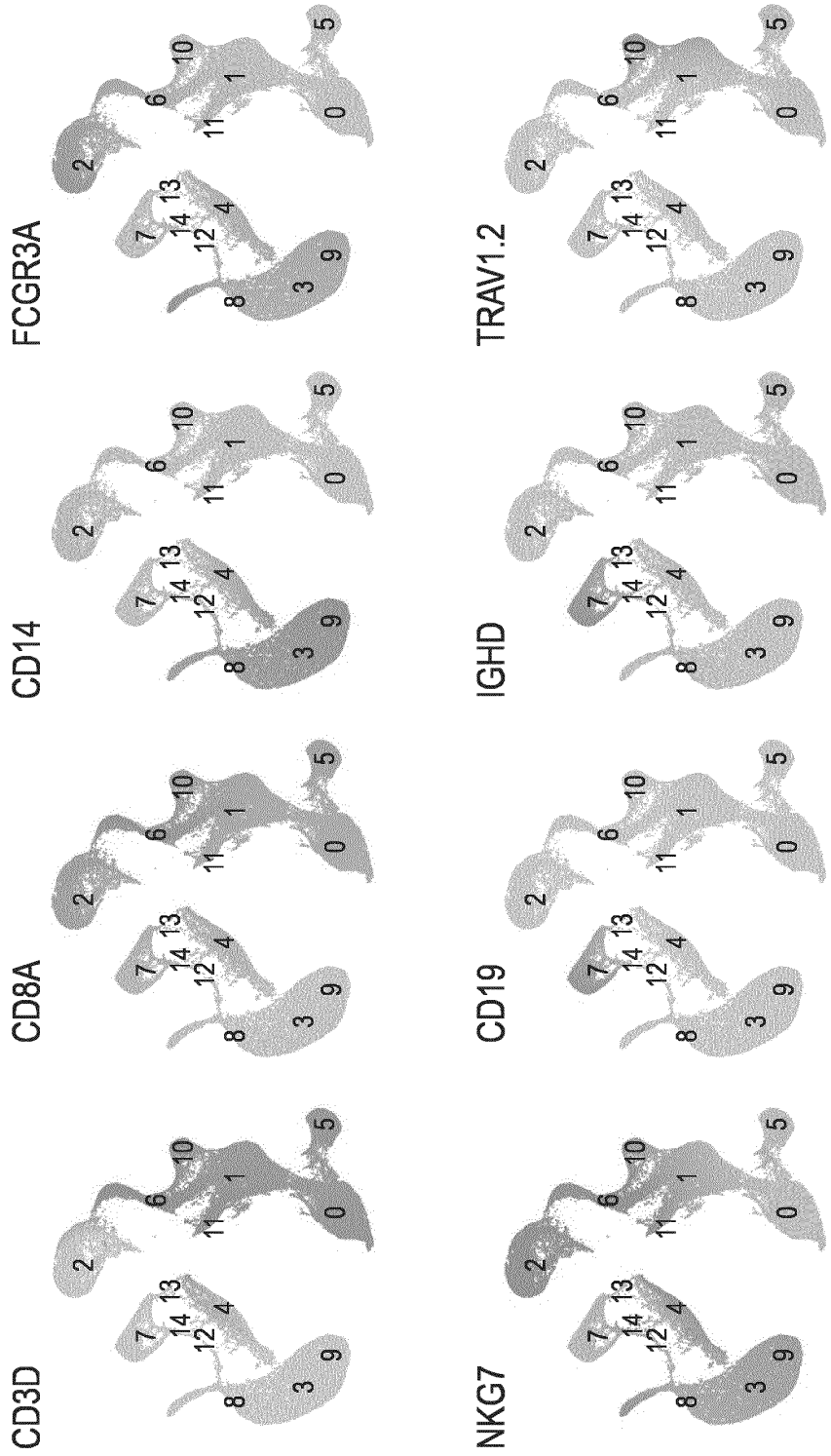


FIG. 31

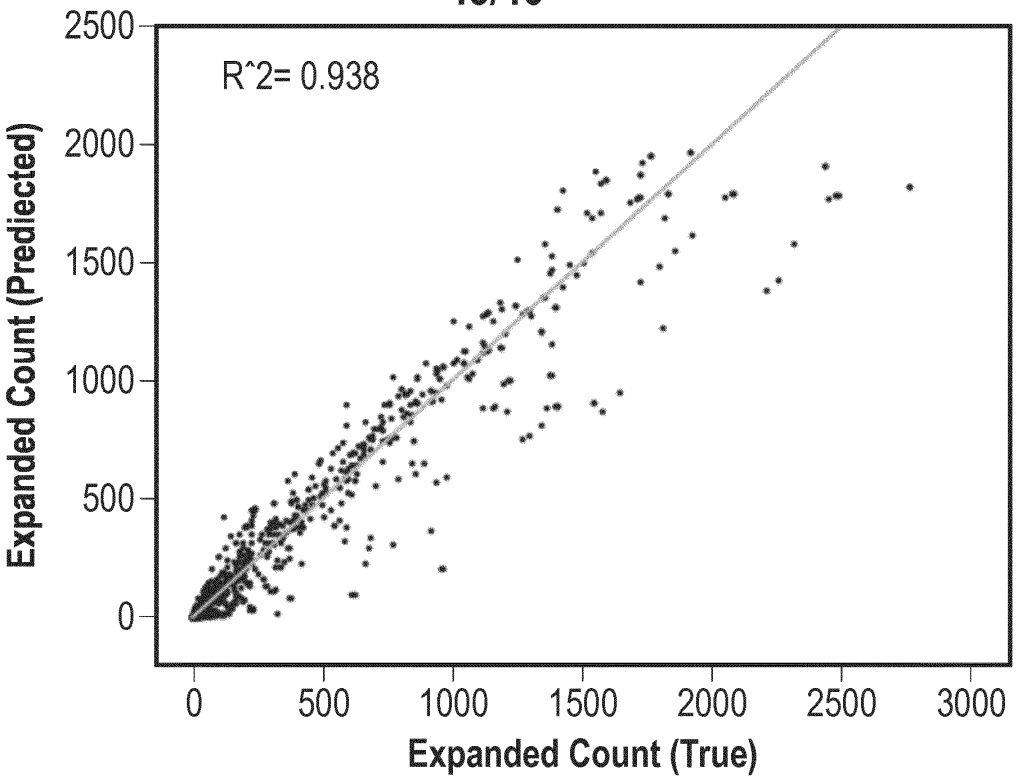


FIG. 32A

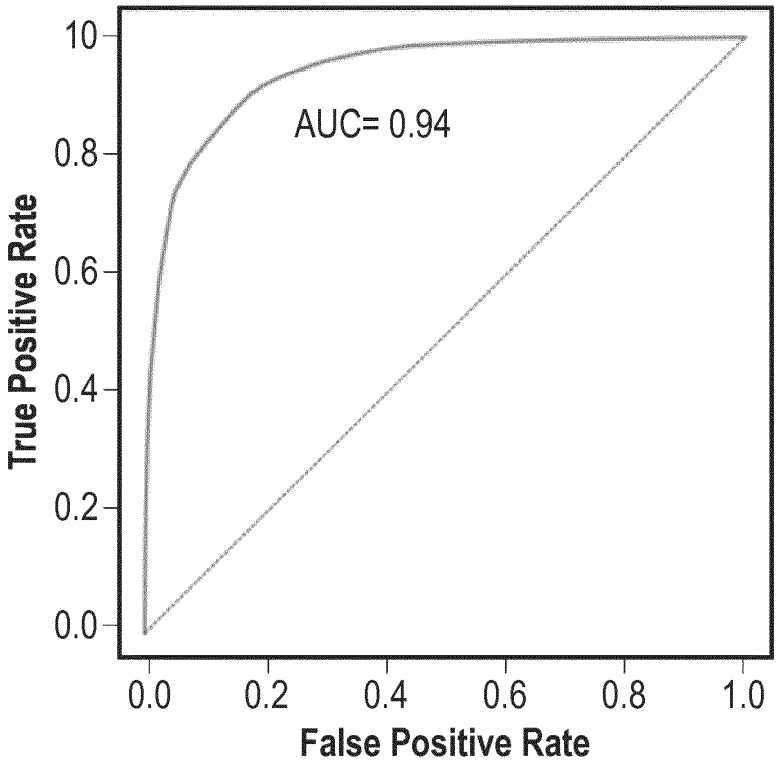


FIG. 32B

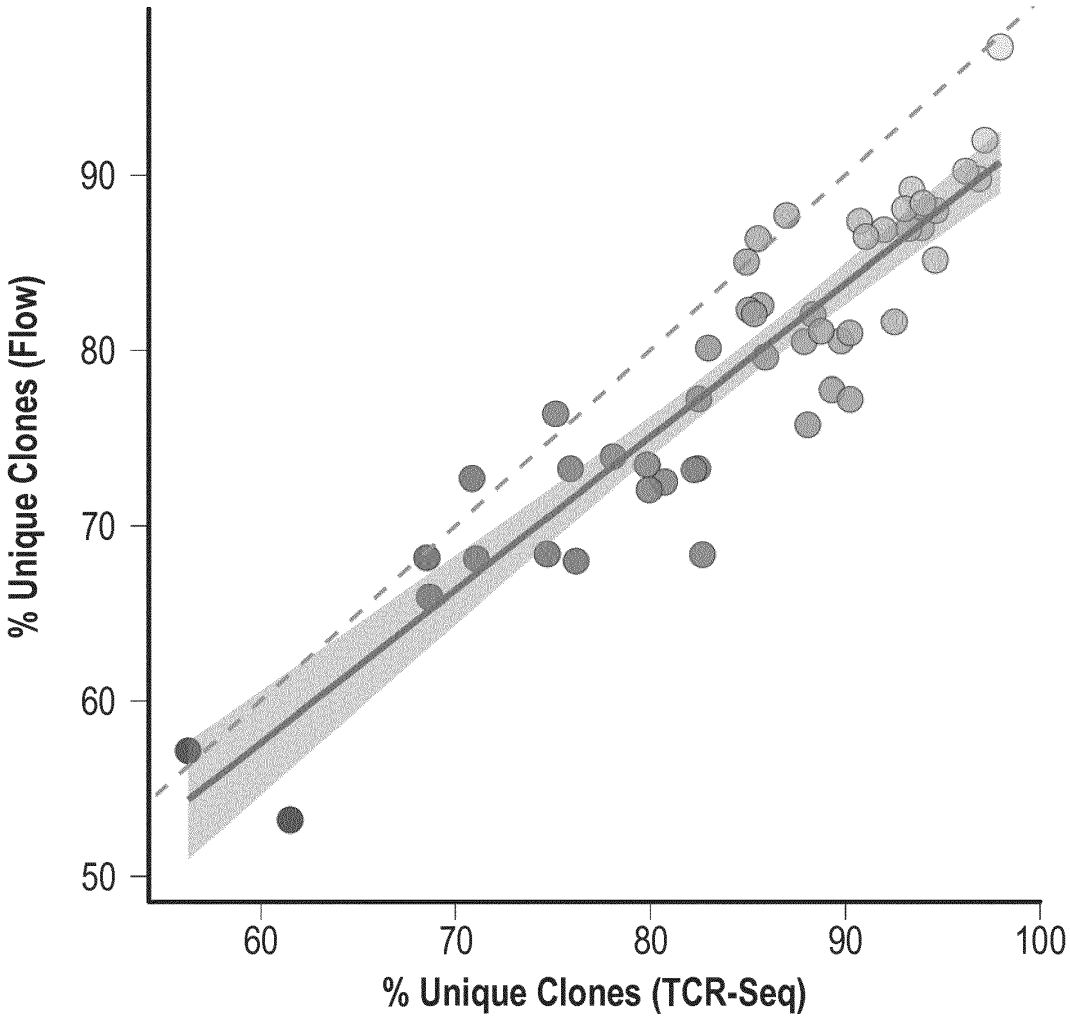

% Unique Clones (Flow)
90
80
70
60
50
60
70
80
90
100
% Unique Clones (TCR-Seq)

FIG. 33

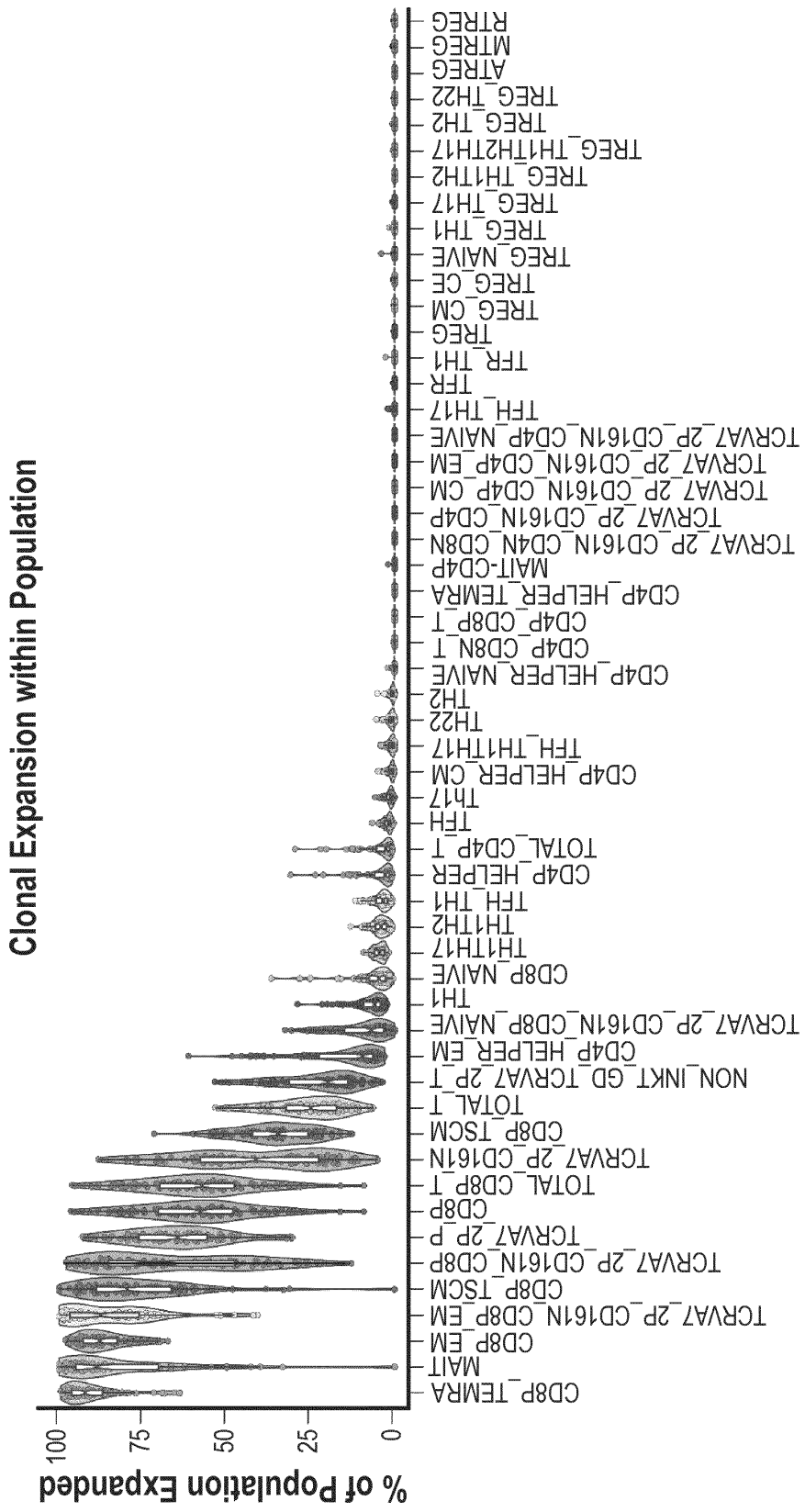

Clonal Expansion within Population
% of Population Expanded
100
75
50
25
0
CD8P_TEMRA
MAIT
CD8P_EM
TCRVA7_2P_CD161N_CD8P_EM
CD8P_TSCM
TCRVA7_2P_CD161N_CD8P
TCRVA7_2P_P
CD8P
TOTAL_CD8P_T
TCRVA7_2P_CD161N
CD8P_TSCM
TOTAL_T
NON_INKT_GD_TCRVA7_2P_T
CD4P_HELPER_EM
TCRVA7_2P_CD161N_CD8P_NAIVE
TH1
CD8P_NAIVE
TH1TH17
TH1TH2
TFH_TH1
CD4P_HELPER
TOTAL_CD4P_T
TFH
Th17
CD4P_HELPER_CM
TFH_TH1TH17
TH22
TH2
CD4P_HELPER_NAIVE
CD4P_CD8N_T
CD4P_CD8P_T
CD4P_HELPER_TEMRA
MAIT-CD4P
TCRVA7_2P_CD161N_CD4N_CD8N
TCRVA7_2P_CD161N_CD4P
TCRVA7_2P_CD161N_CD4P_CM
TCRVA7_2P_CD161N_CD4P_EM
TCRVA7_2P_CD161N_CD4P_NAIVE
TFH_TH17
TFR
TFR_TH1
TREG
TREG_CM
TREG_CE
TREG_NAIVE
TREG_TH1
TREG_TH17
TREG_TH1TH2
TREG_TH1TH2TH17
TREG_TH2
TREG_TH22
ATREG
MTREG
RTREG

FIG. 34

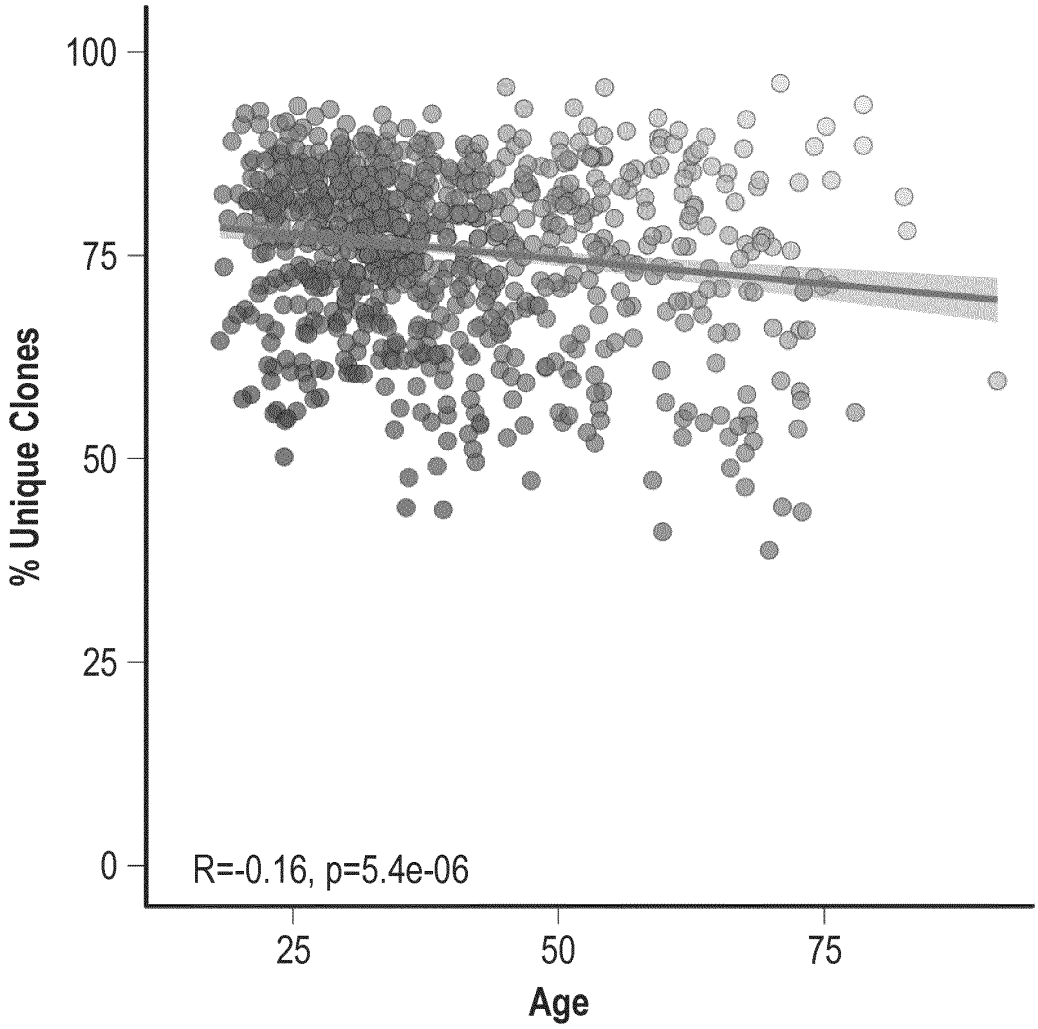


FIG. 35