

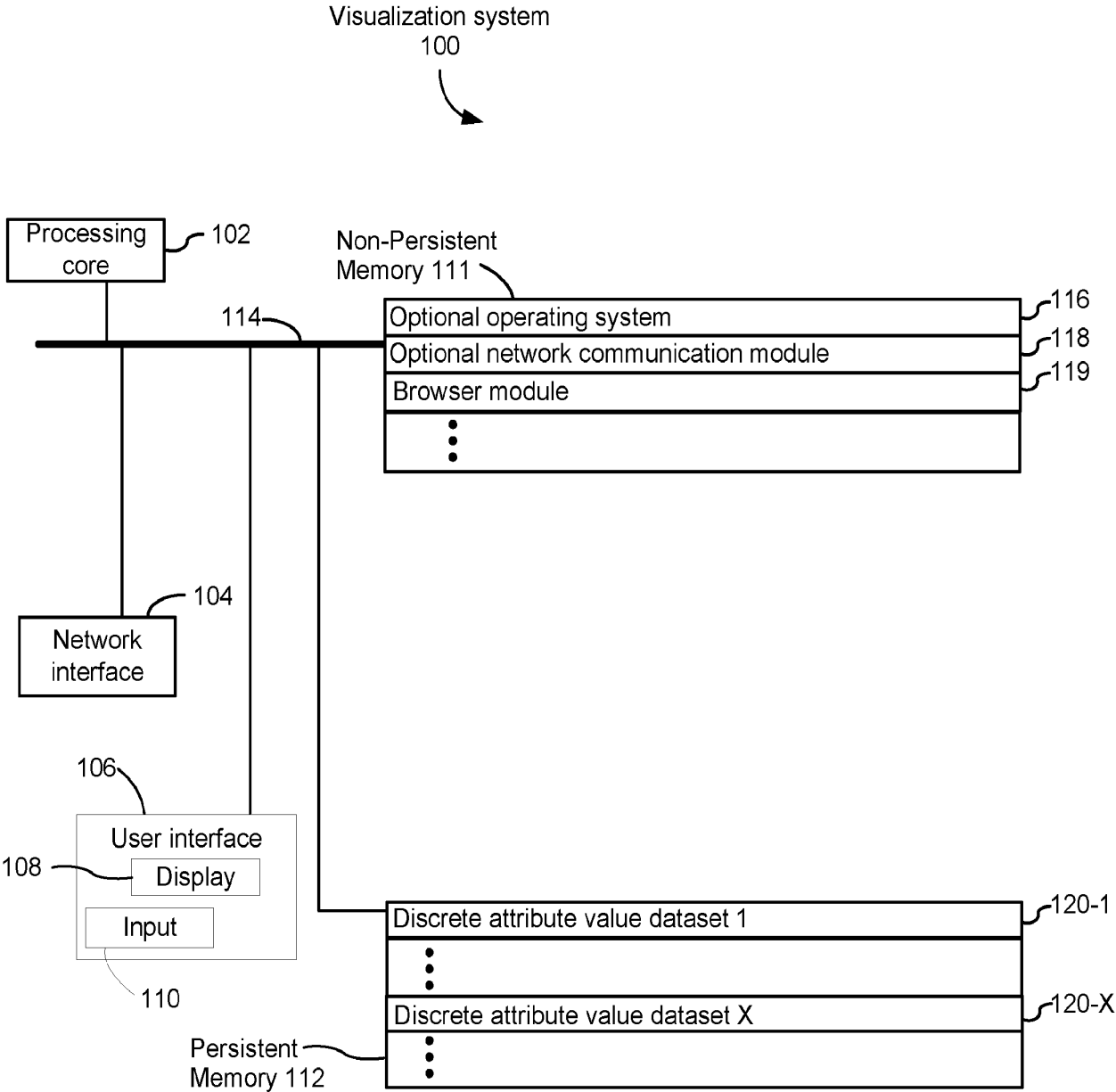


Figure 1A

Discrete attribute value dataset	120-1
Feature-barcode matrix	121-1
Cell 1	126-1
Gene 1 for cell 1	122-1-1
(GEX) Discrete attribute value for gene 1	124-1-1
Gene 2 for Cell 1	122-2-1
(GEX) Discrete attribute value for gene 2	124-2-1
⋮	
Gene M for Cell 1	122-M-1
(GEX) Discrete attribute value for Cell M	125-M-1
ATAC Peak 1 for cell 1	123-1-1
ATAC Fragment count for peak 1	125-1-1
ATAC Peak 2 for Cell 1	123-2-1
ATAC Fragment count for peak 2	125-2-1
⋮	
ATAC Peak L for Cell 1	123-L-1
ATAC Fragment count for peak L	125-L-1
GEX Principal component value 1 for cell 1	164-1-1
⋮	
GEX Principal component value N for cell 1	164-1-N
GEX Two-dimensional datapoint for cell 1	166-1
GEX Cluster assignment for cell 1	158-1
GEX Category assignment 1 for cell 1	170-1-1
Class 1 for GEX category 1	172-1-1-1
⋮	
Class M for GEX category 1	172-1-1-M
⋮	
GEX Category assignment Q for cell 1	170-1-Q
Class 1 for GEX category Q	172-1-Q-1
⋮	
Class Z for GEX category Q	172-1-Q-Z
Continued on Figure 1C...	

Figure 1B
SUBSTITUTE SHEET (RULE 26)

Discrete attribute value dataset 1 - continued	120-1 (continued)
Feature-barcode matrix - continued	121-1 (continued)
Cell 1	126-1 (continued)
ATAC Principal component value 1	165-1-1
⋮	
ATAC Principal component value P	165-1-P
ATAC Two-dimensional datapoint for cell 1	167-1
ATAC Cluster assignment for cell 1	159-1
ATAC Category assignment 1 for cell 1	171-1-1
Class 1 for ATAC category 1	173-1-1-1
⋮	
Class M for ATAC category 1	173-1-1-M
⋮	
ATAC Category assignment P for cell 1	171-1-P
Class 1 for ATAC category P	173-1-P-1
⋮	
Class W for ATAC category P	173-1-P-W
⋮	
Cell Y	126-Y
Gene 1	193-1-1
Nearby gene inferred ATAC count (for each cell)	194-1-1
Promoter sum inferred count (for each cell)	195-1-1
⋮	
Gene M	193-1-M
Nearby gene inferred ATAC count (for each cell)	194-1-M
Promoter sum inferred count (for each cell)	195-1-M
GEX t-SNE projection	196
GEX UMAP projection	197
ATAC t-SNE projection	198
ATAC UMAP projection	199

Figure 1C
SUBSTITUTE SHEET (RULE 26)

Discrete attribute value dataset 1 - continued	120-1 (continued)
Array of chromosome indices	180-1
Chromosome index 1	182-1-1
⋮	
Chromosome index L+M	182-1-L+M
Array of starting positions	183-1
Starting position 1	184-1-1
⋮	
Starting position L+M	184-1-L+M
Array of ending positions	185-1
Ending position 1	186-1-1
⋮	
Ending position L+M	186-1-L+M
Feature-linkage matrix	187-1
Pointer array (P)	188-1
Index 1	189-1-1
⋮	
Index L	189-1-L
Index array (X)	190-1
ATAC peak identity 1	190-1-1
⋮	
ATAC peak identity K	190-1-K
Correlation array (C)	191-1
Correlation 1	191-1-1
⋮	
Correlation K	191-1-K
Significance array (S)	192-1
Significance 1	192-1-1
⋮	
Significance K	192-1-K
⋮	

Figure 1D
SUBSTITUTE SHEET (RULE 26)

A visualization system 100 comprising one or more processing cores, a persistent memory and a non-persistent memory. The persistent memory and the non-persistent memory collectively storing instructions for performing the following method.

Store the dataset 120 in persistent memory. The dataset 120 comprises a corresponding GEX discrete attribute value 124 for each gene 122 in a plurality of genes for each respective cell 126 in a plurality of cells. The dataset 120 further comprises a corresponding ATAC fragment count 125 for each ATAC peak 123 in a plurality of ATAC peaks for each respective cell 126 in the plurality of cells. In some embodiments, the dataset 120 redundantly represents the corresponding discrete attribute value 124 for each gene in the plurality of genes for each respective cell in the plurality of cells in both a compressed sparse row format and a compressed sparse column format in which genes for a respective cell that have a null discrete attribute data value are discarded. In some embodiments the dataset 120 is compressed in accordance with a blocked compression algorithm.

Each gene 122 in the plurality of genes is a respective gene in a plurality of genes. Each discrete attribute value 124 is a count of transcript reads within the cell that map to a respective gene in the plurality of genes. Each cell 126 is a single cell. The GEX data represents a whole transcriptome shotgun sequencing experiment that quantifies gene expression from a single cell in counts of transcript reads mapped to the genes.

Cluster the dataset using the discrete attribute values for each gene in the plurality of genes for each respective cell in the plurality of cells thereby assigning each respective cell in the plurality of cells to a corresponding cluster in a first plurality of clusters. Each respective cluster in the first plurality of clusters consists of a unique different subset of the plurality of cells. The clustering loads less than the entirety of the dataset into the non-persistent memory at any given time during the clustering. Separately cluster the dataset using the ATAC fragment counts for each ATAC peak 123 in the plurality of ATAC peaks for each respective cell in the plurality of cells thereby assigning each respective cell in the plurality of cells to a corresponding cluster in a second plurality of clusters. Each respective cluster in the second plurality of clusters consists of a unique different subset of the plurality of cells.

The clustering is performed on a computer system remote from the visualization system prior to storing the dataset in persistent memory.

The clustering the dataset comprises hierarchical clustering, agglomerative clustering using a nearest-neighbor algorithm, agglomerative clustering using a farthest-neighbor algorithm, agglomerative clustering using an average linkage algorithm, agglomerative clustering using a centroid algorithm, or agglomerative clustering using a sum-of-squares algorithm.

The clustering the dataset comprises k-means clustering, a fuzzy k-means clustering algorithm, or Jarvis-Patrick clustering.

A

Figure 2A
SUBSTITUTE SHEET (RULE 26)

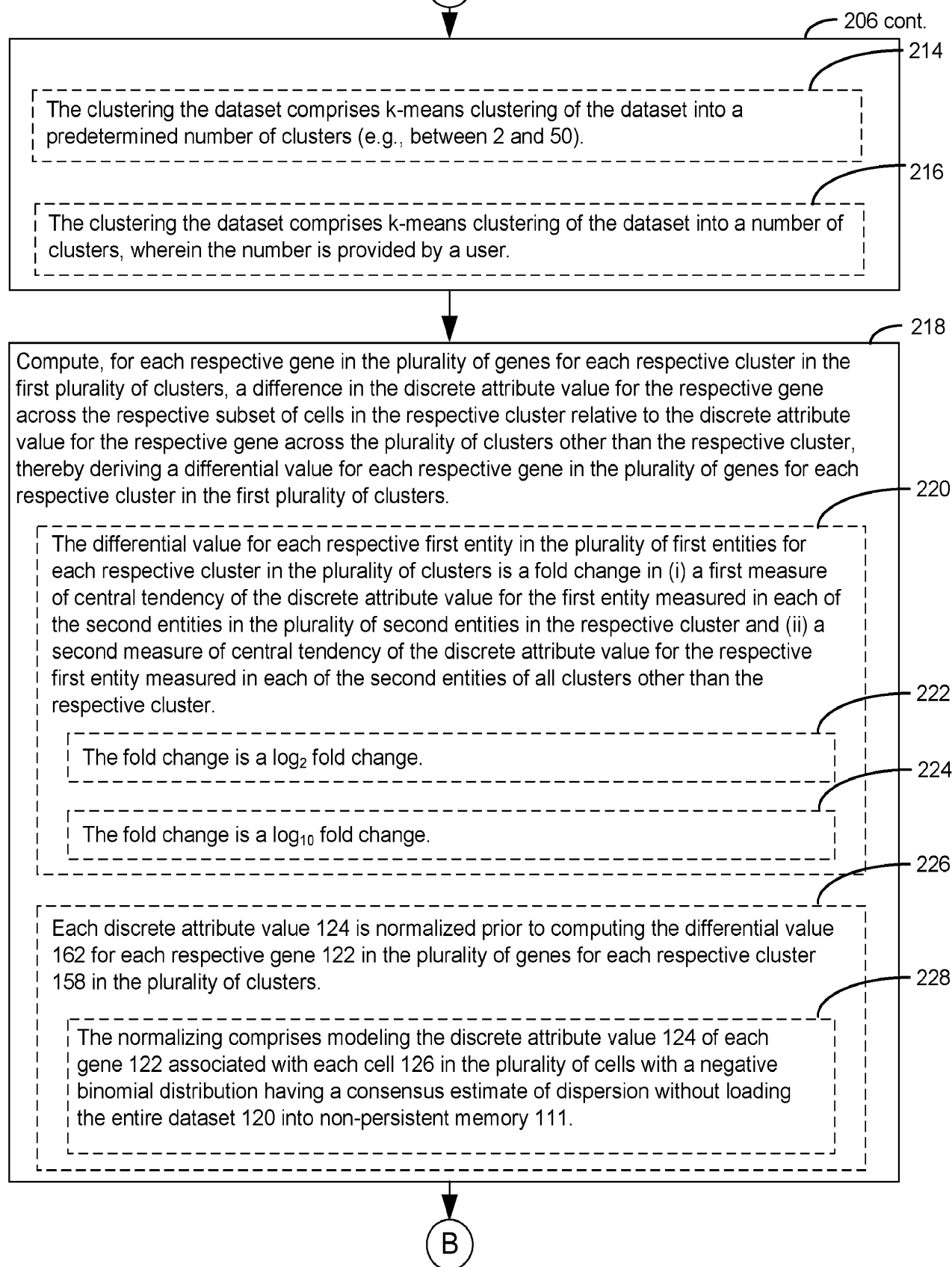


Figure 2B
SUBSTITUTE SHEET (RULE 26)

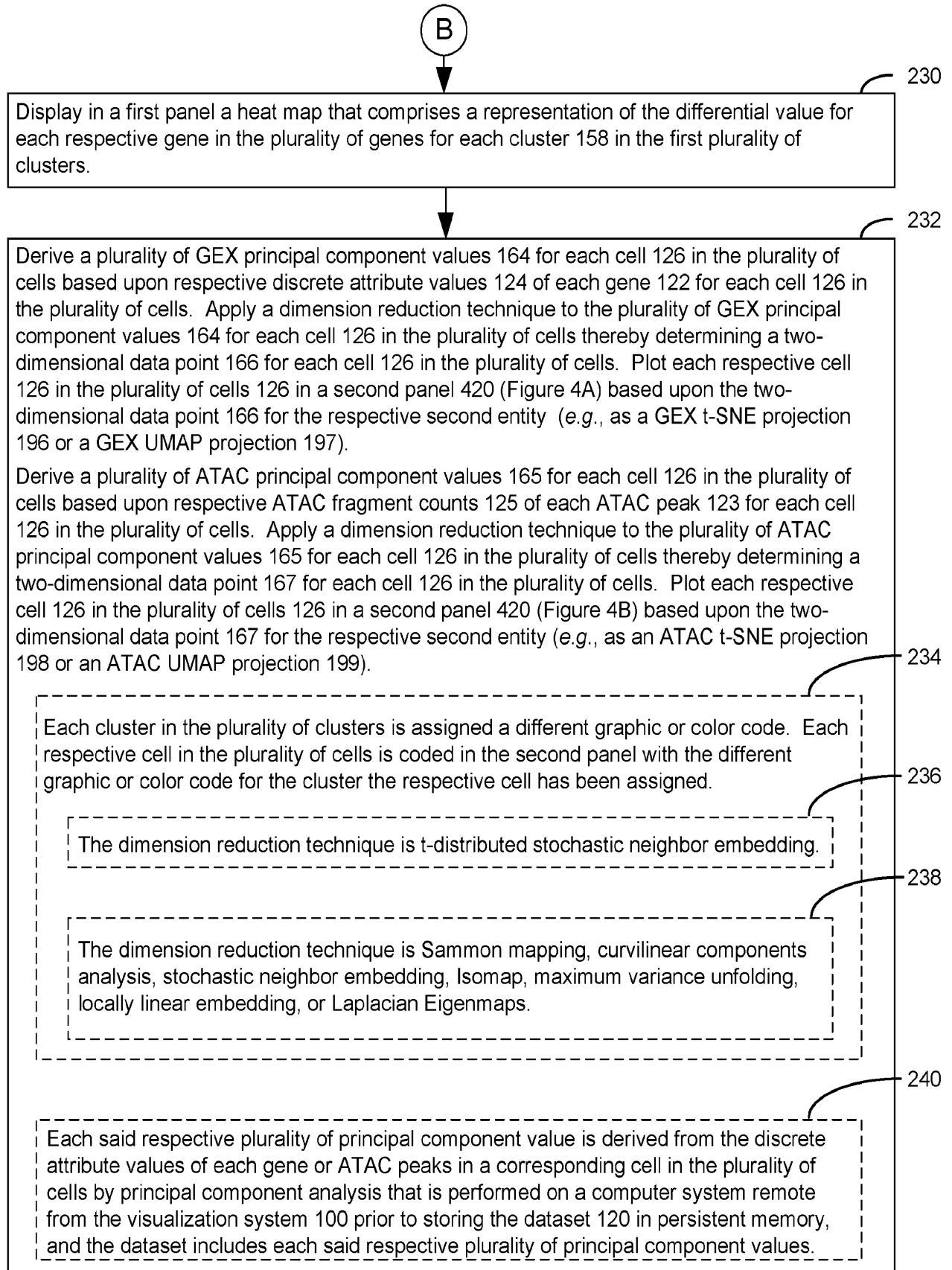
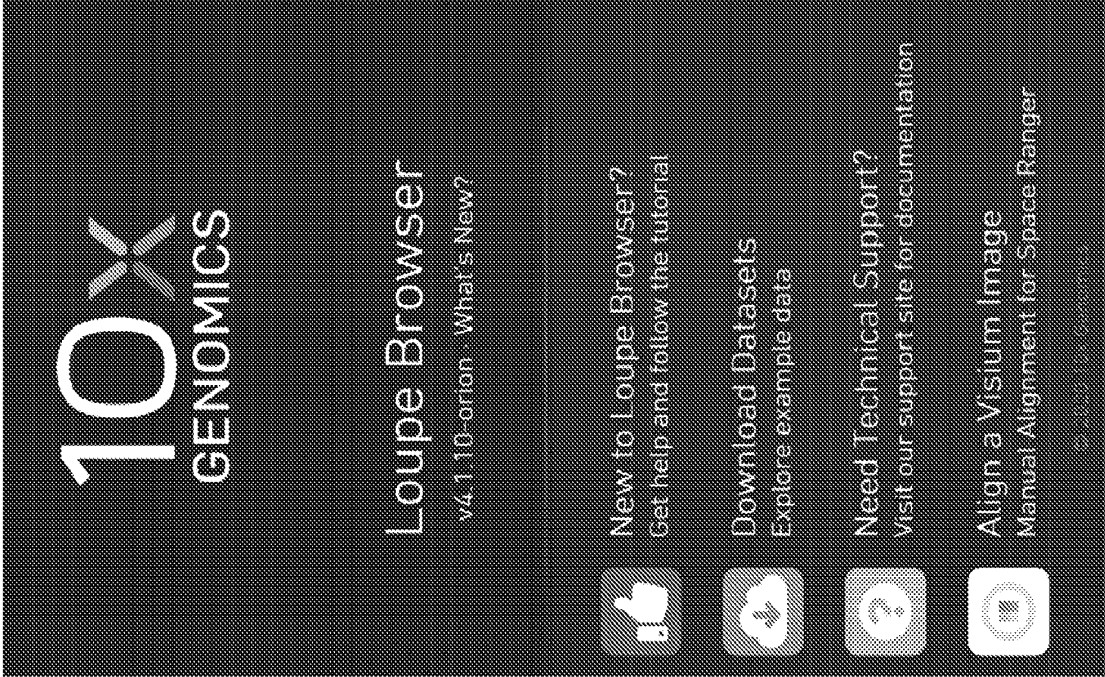


Figure 2C
SUBSTITUTE SHEET (RULE 26)



Recent Files

120

126

Clear Recent

Type	Name	Cell Count	Last Opened
📄	atac_gex_intron	9,461	2 minutes ago
📄	aggr	—	about 2 months ago
📄	fluo	—	about 2 months ago
📄	multichannel-aggr	—	about 2 months ago
📄	AMLTutorial.cloupe	8,414	
📄	ATACTutorial.cloupe	5,335	
📄	SpatialTutorial.cloupe	—	

Browse for a Loupe Browser File

Figure 3

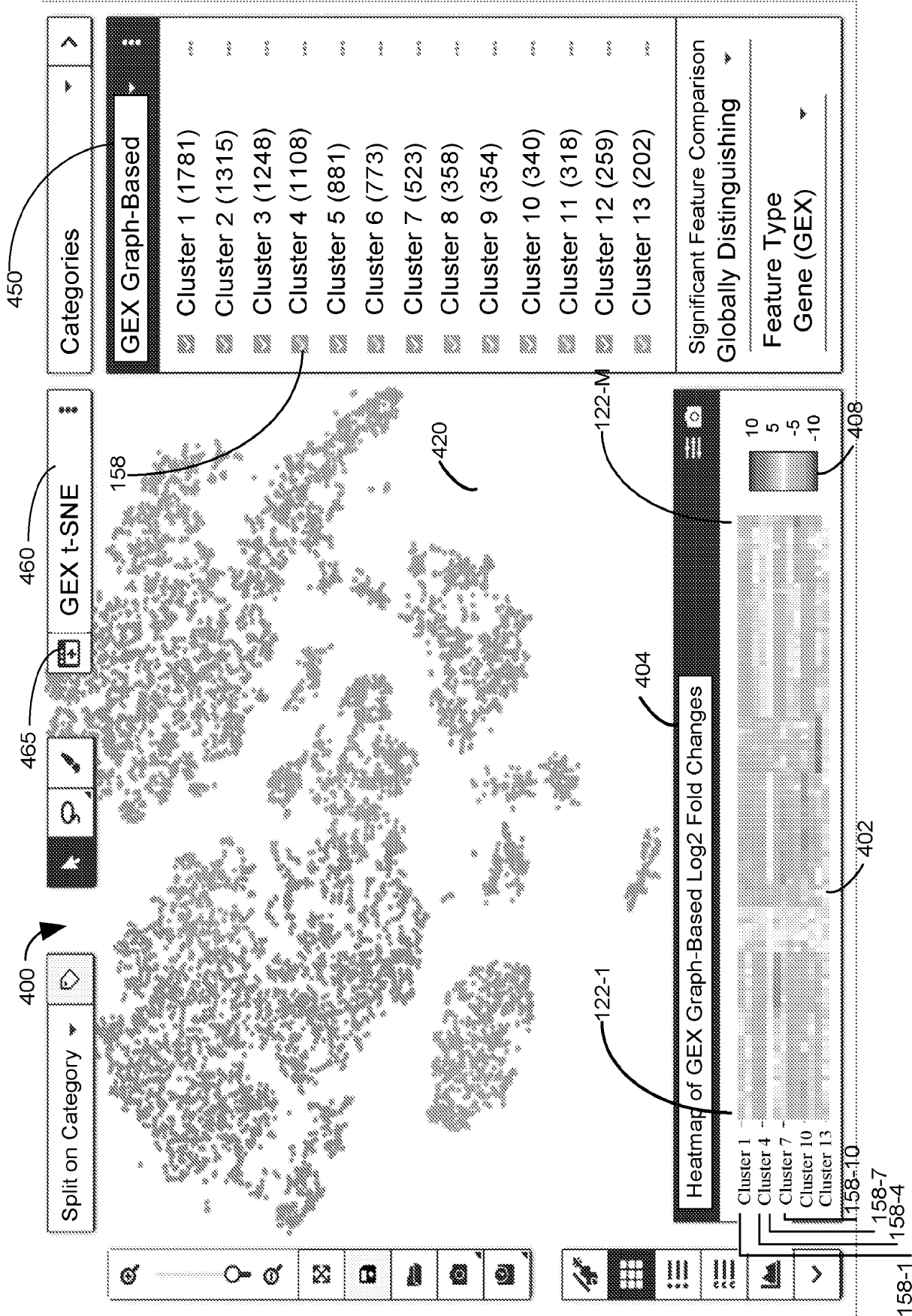


Figure 4A

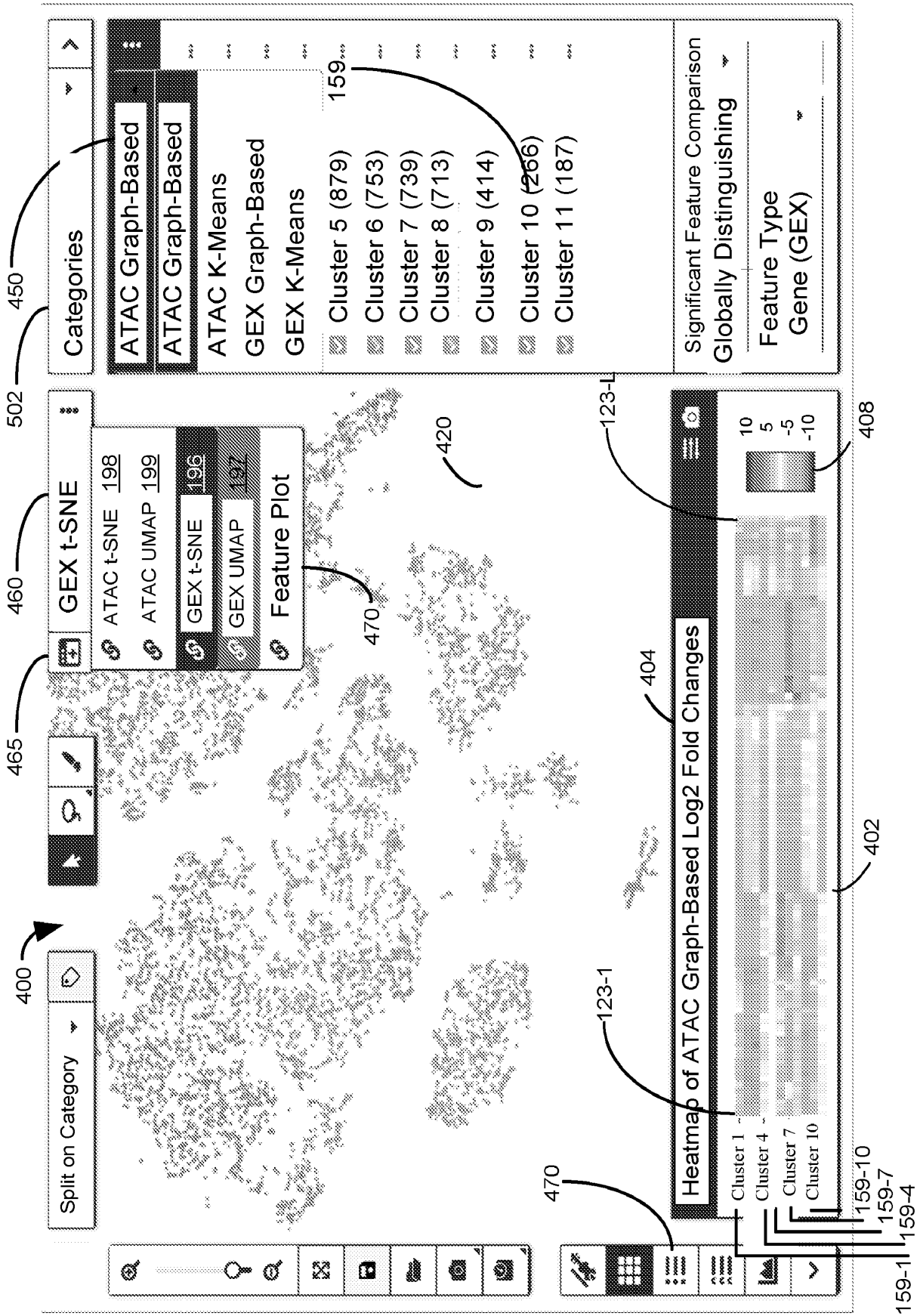


Figure 4B

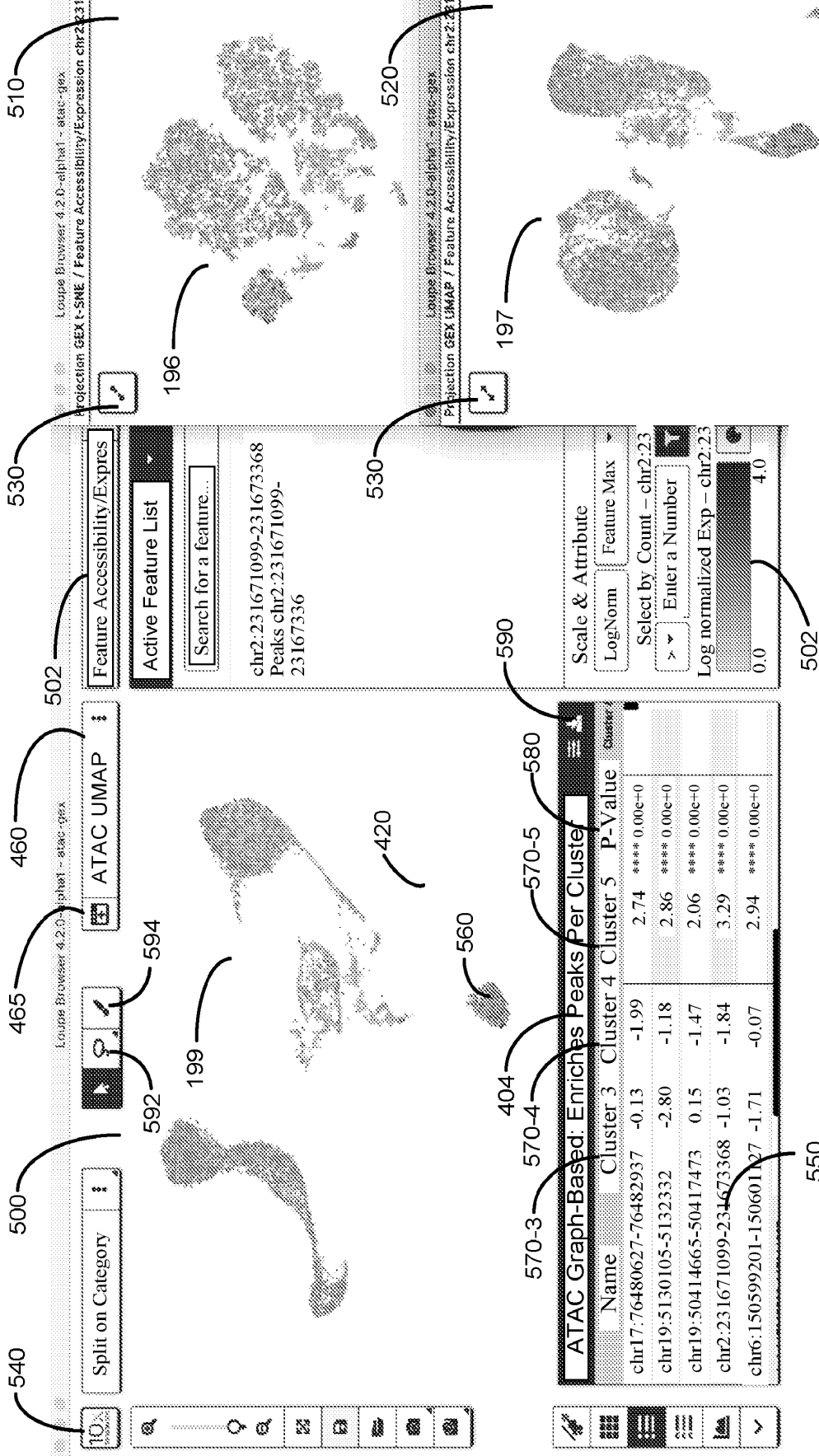


Figure 5

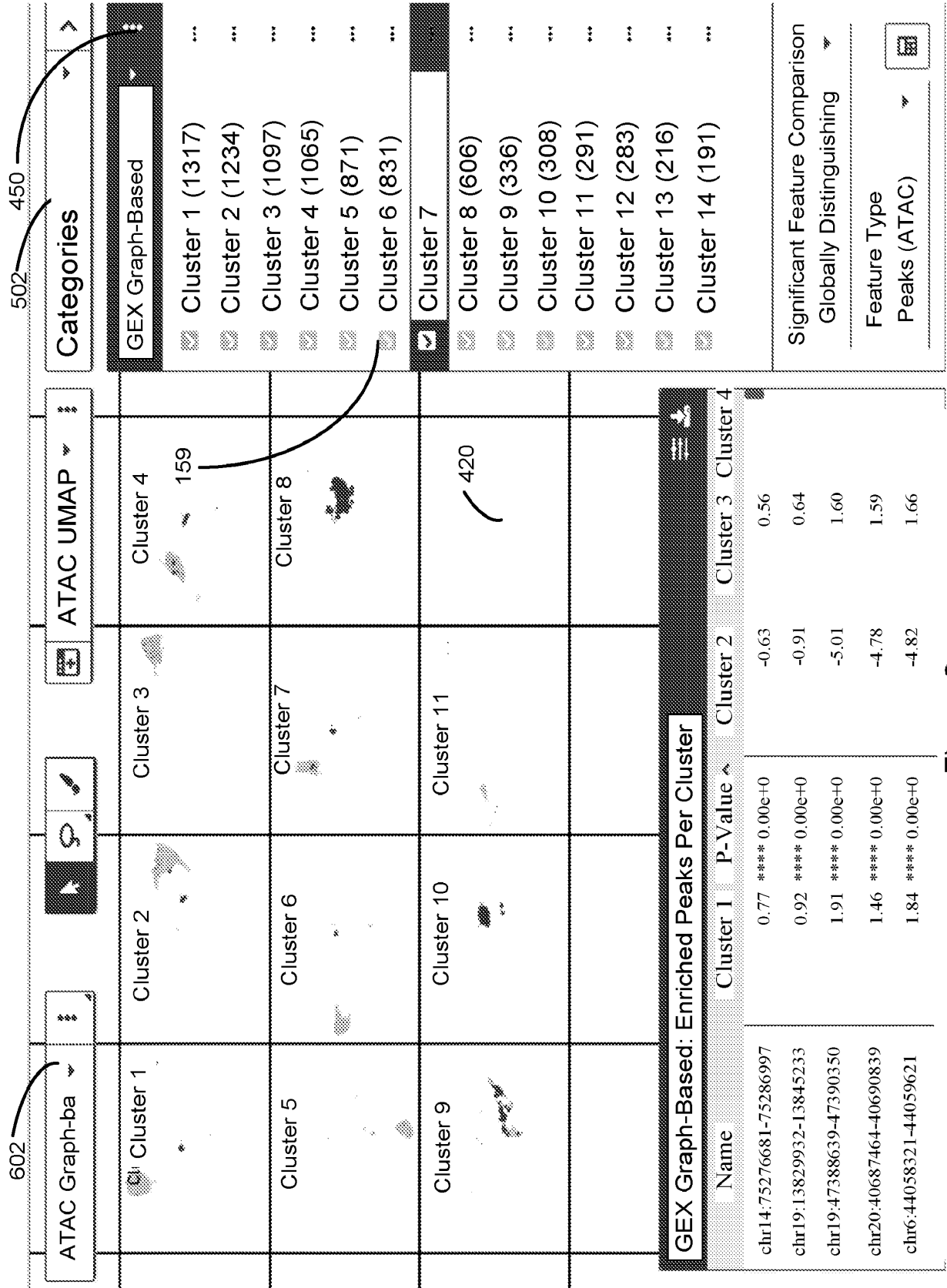


Figure 6

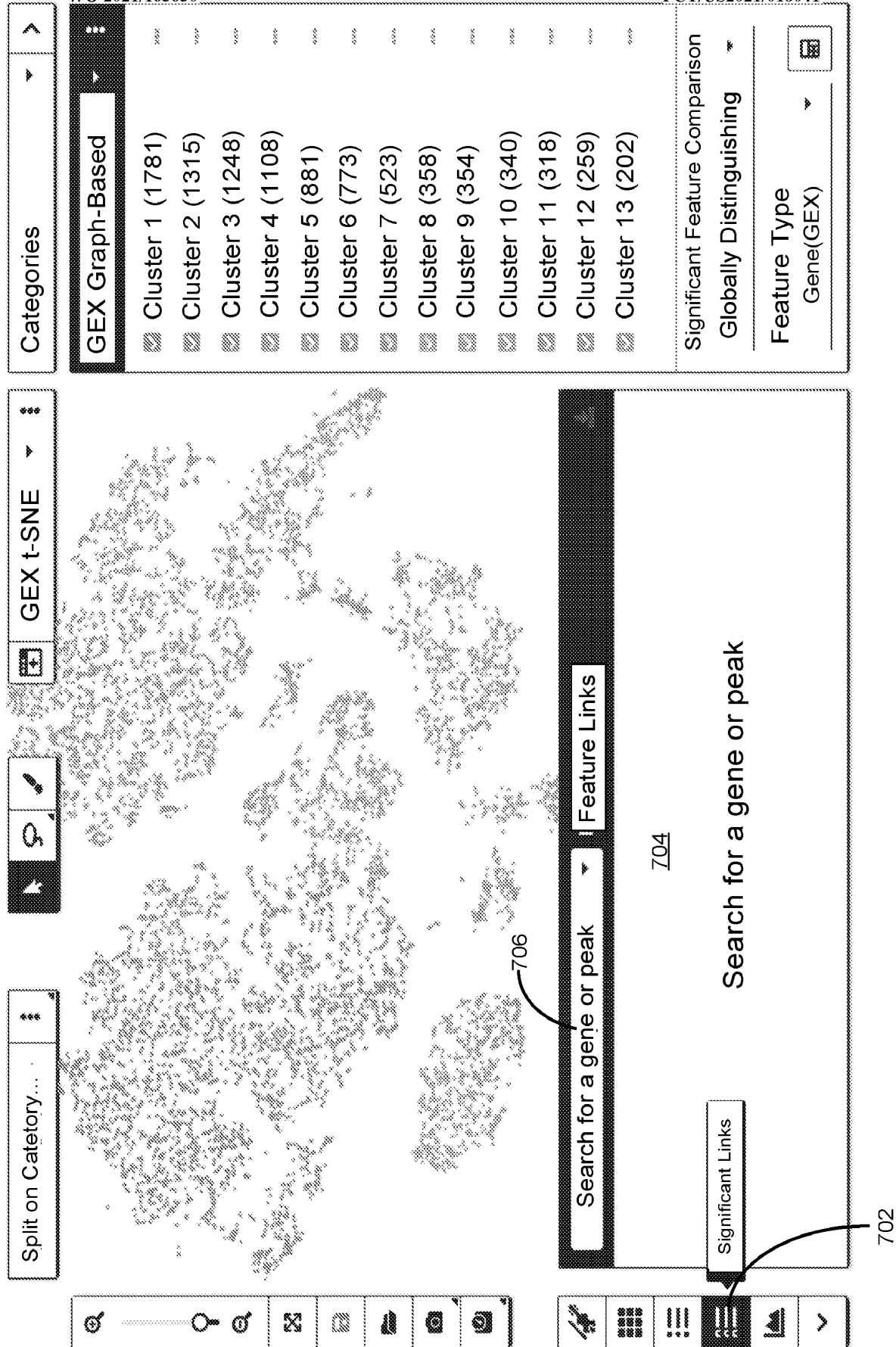


Figure 7

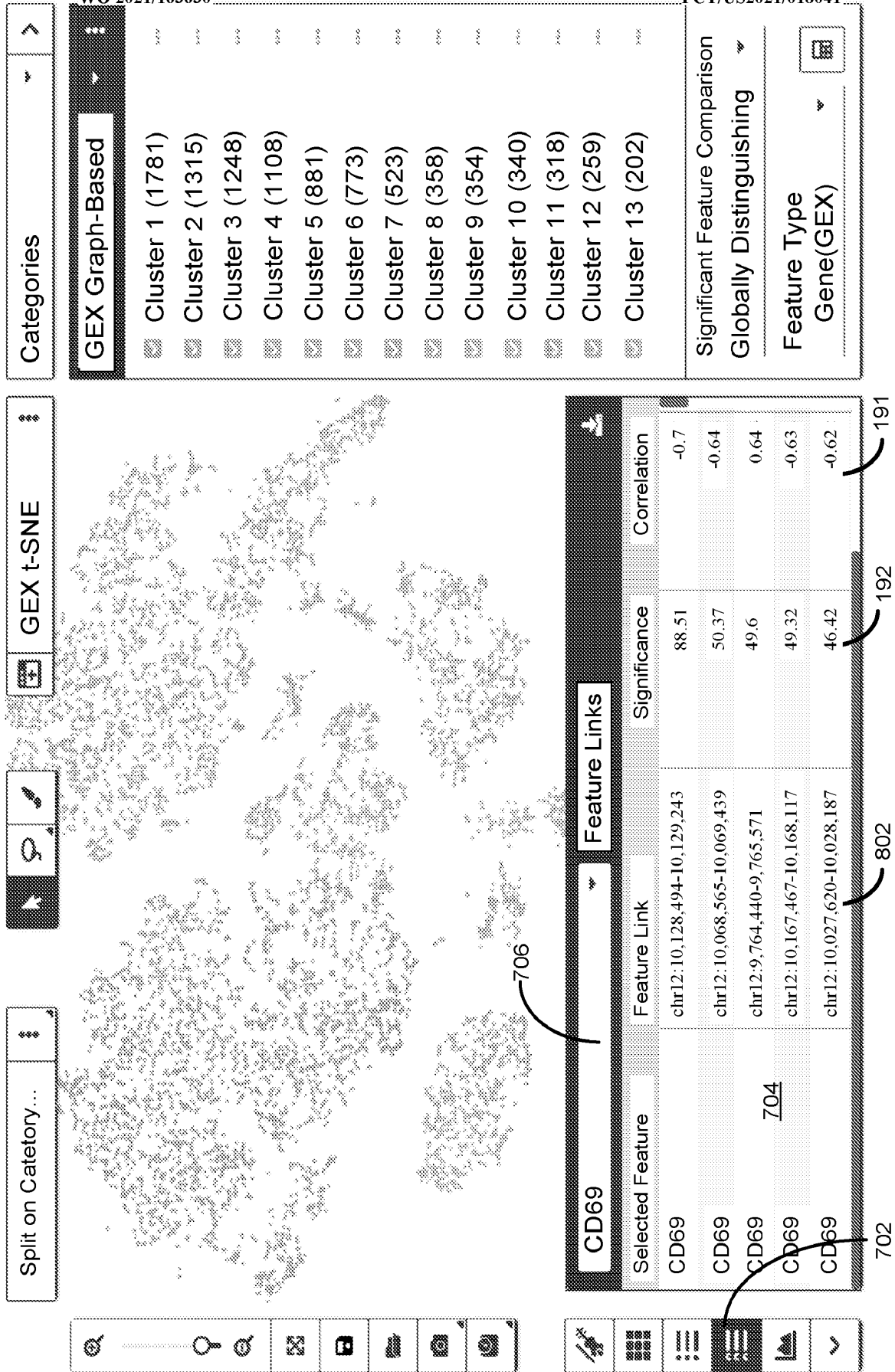


Figure 8

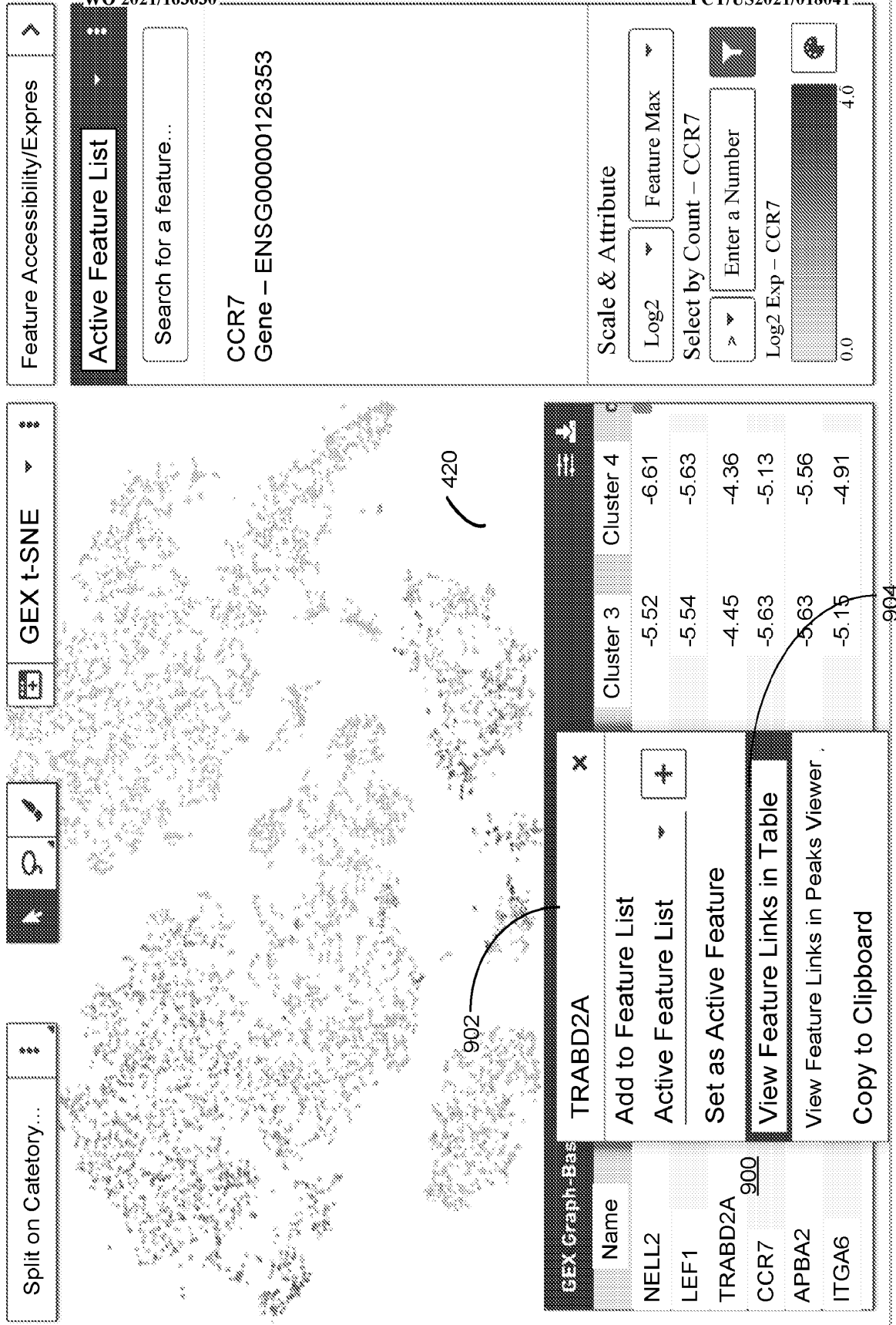


Figure 9

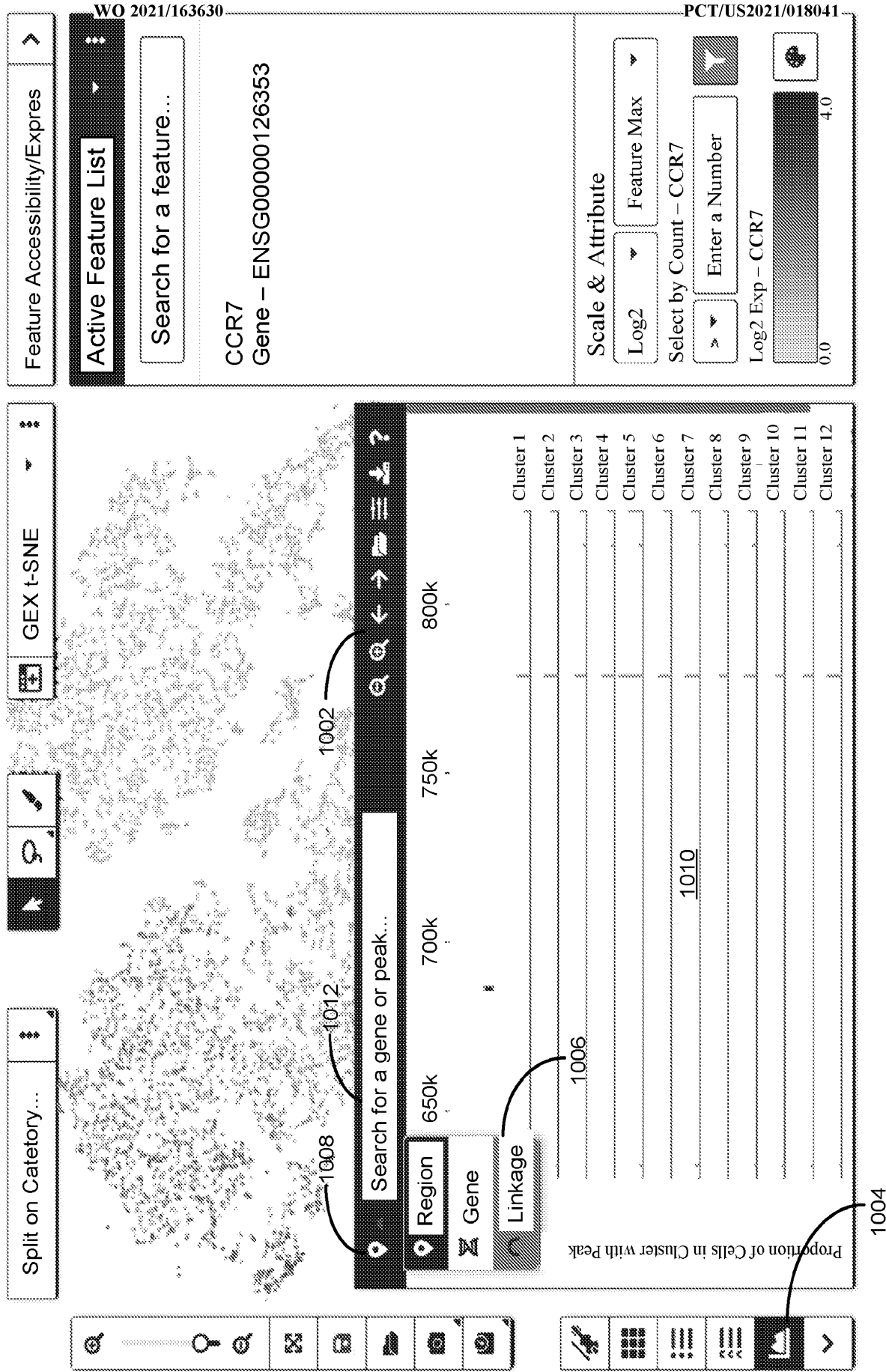


Figure 10

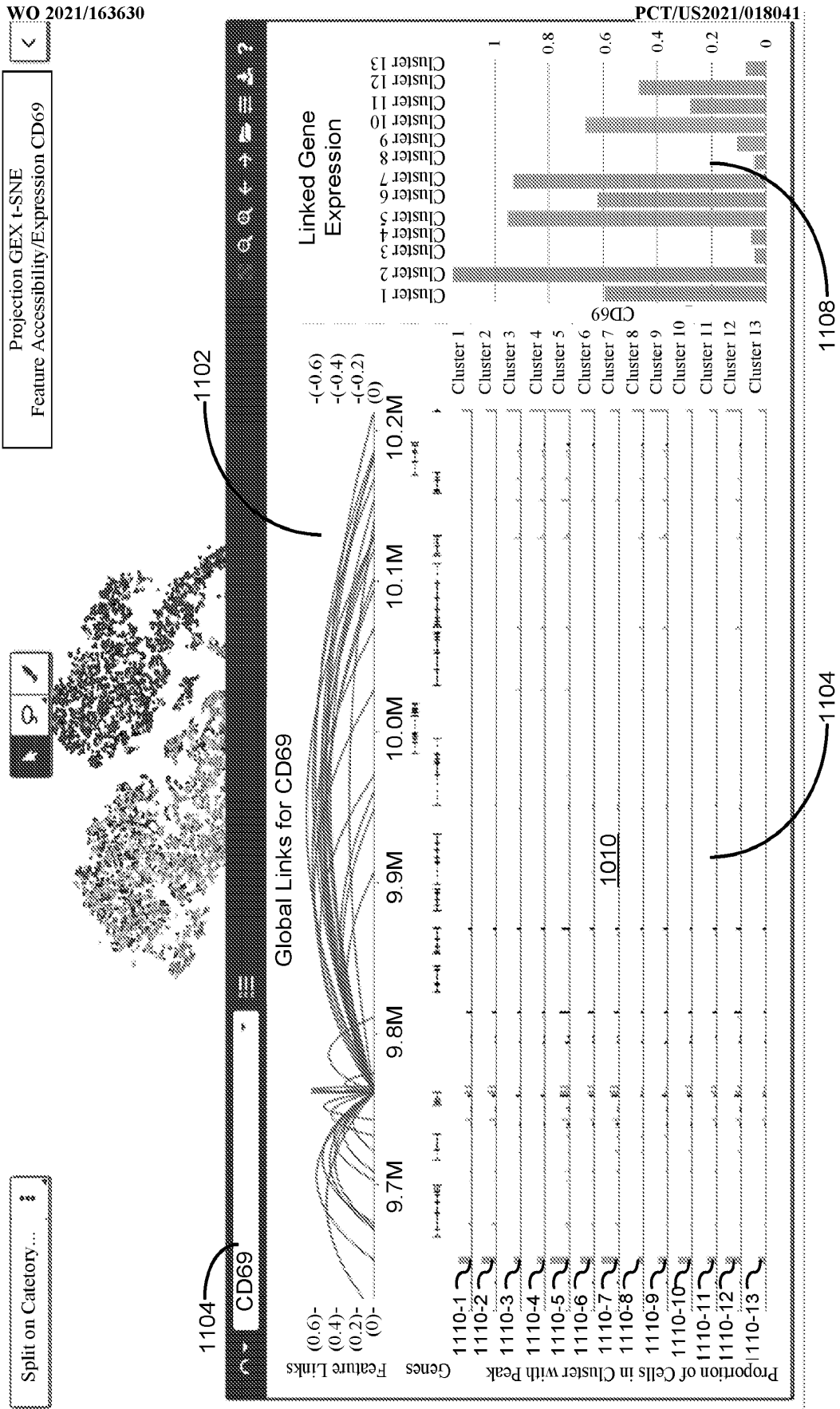


Figure 11

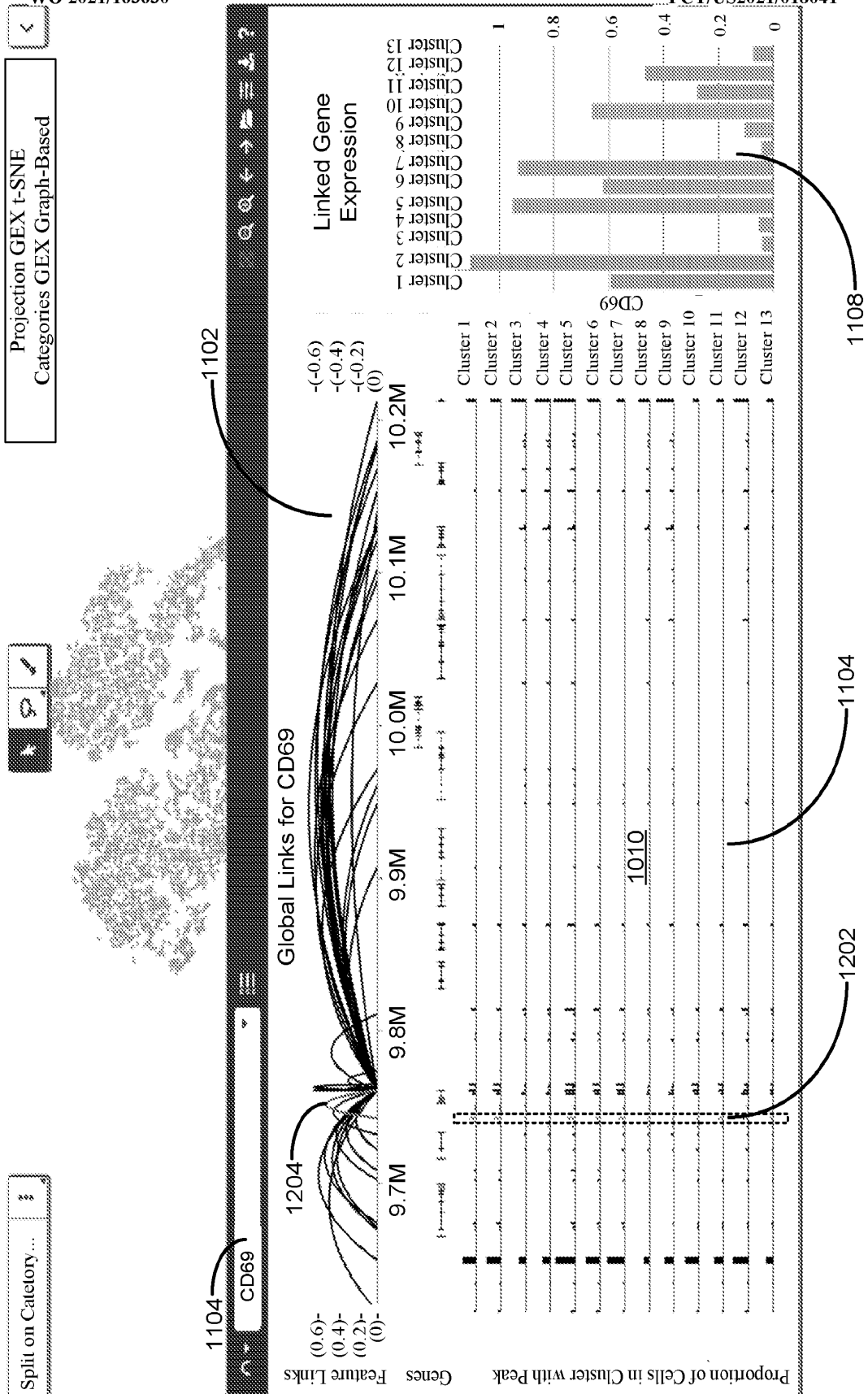


Figure 12

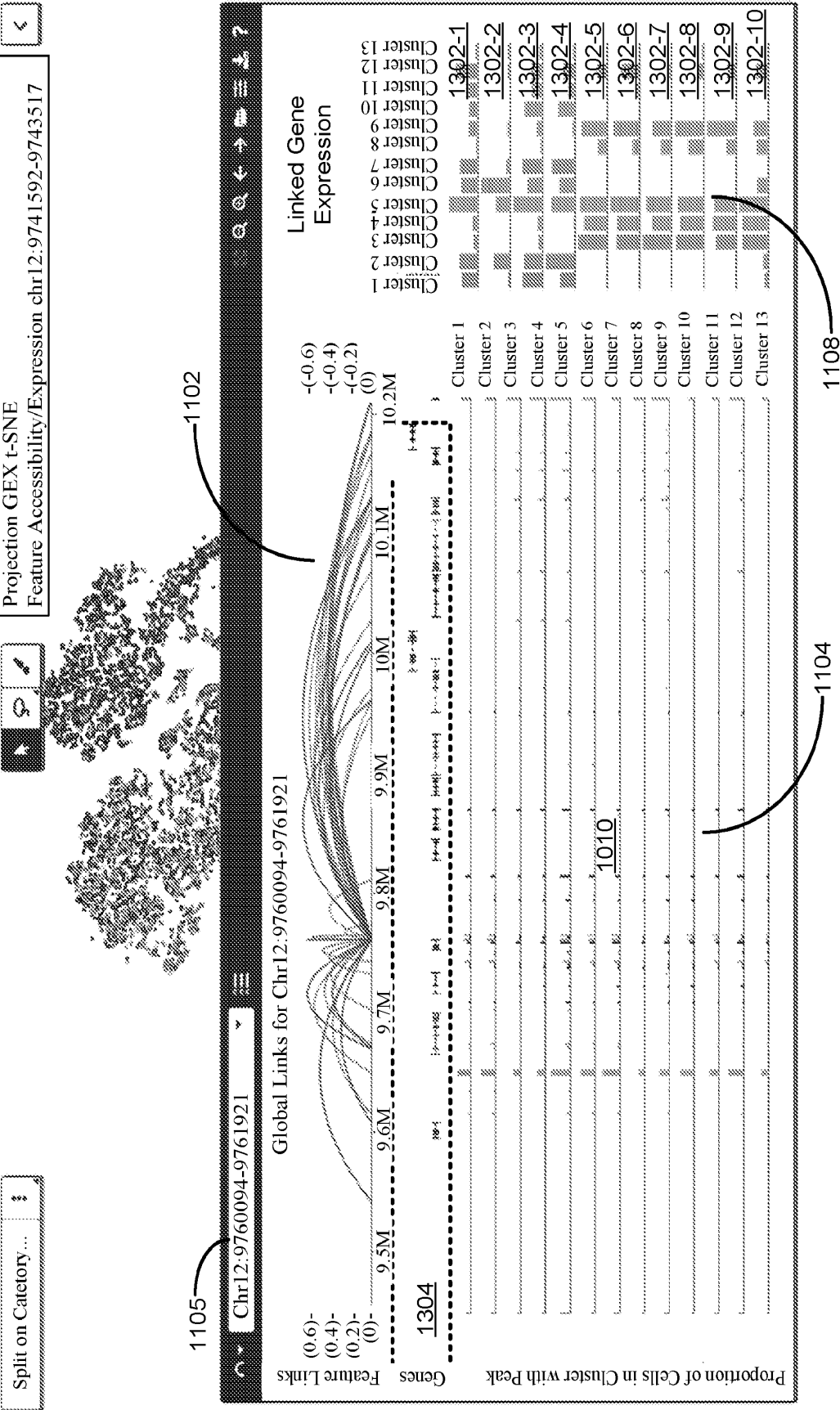


Figure 13A

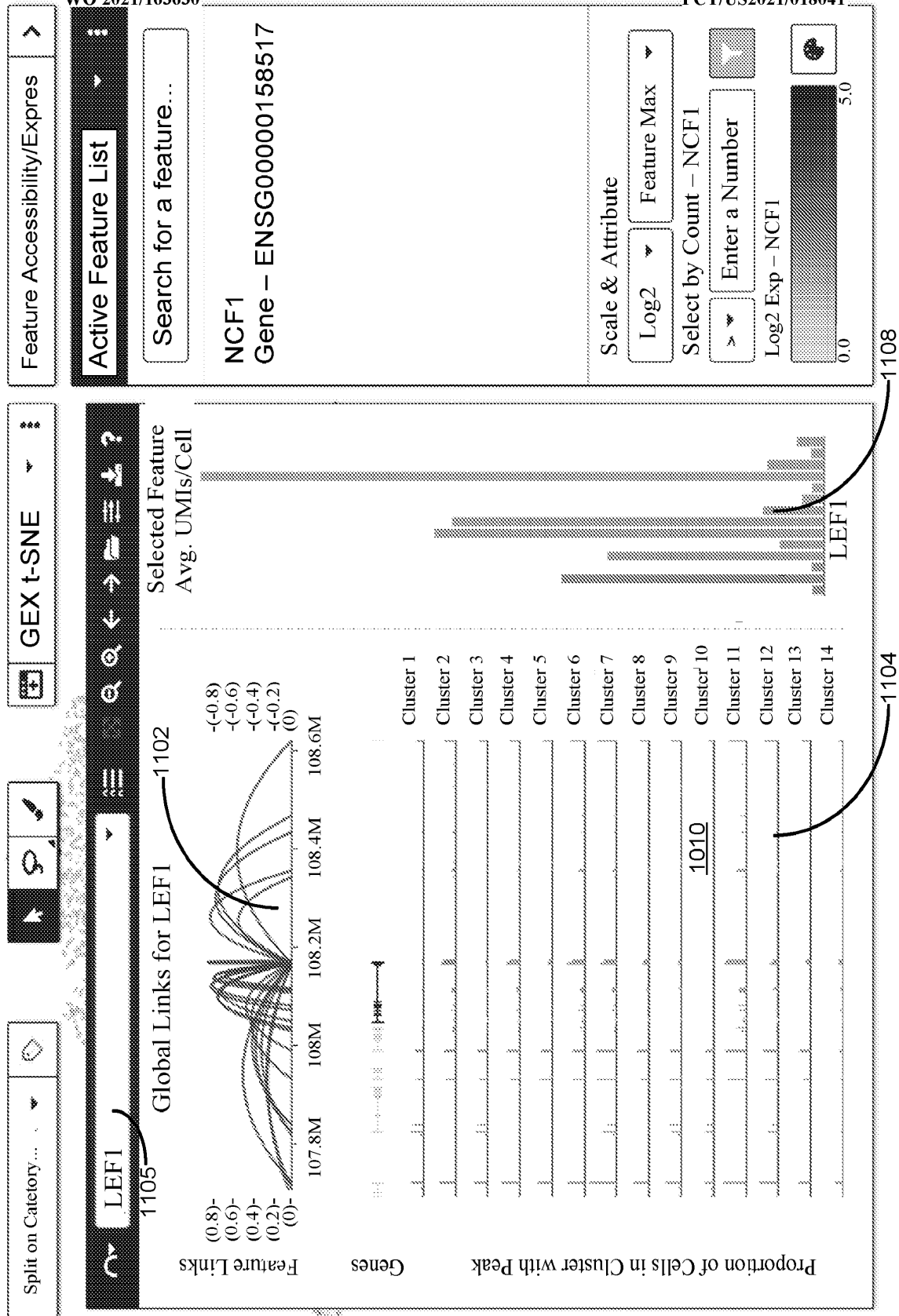


Figure 13B

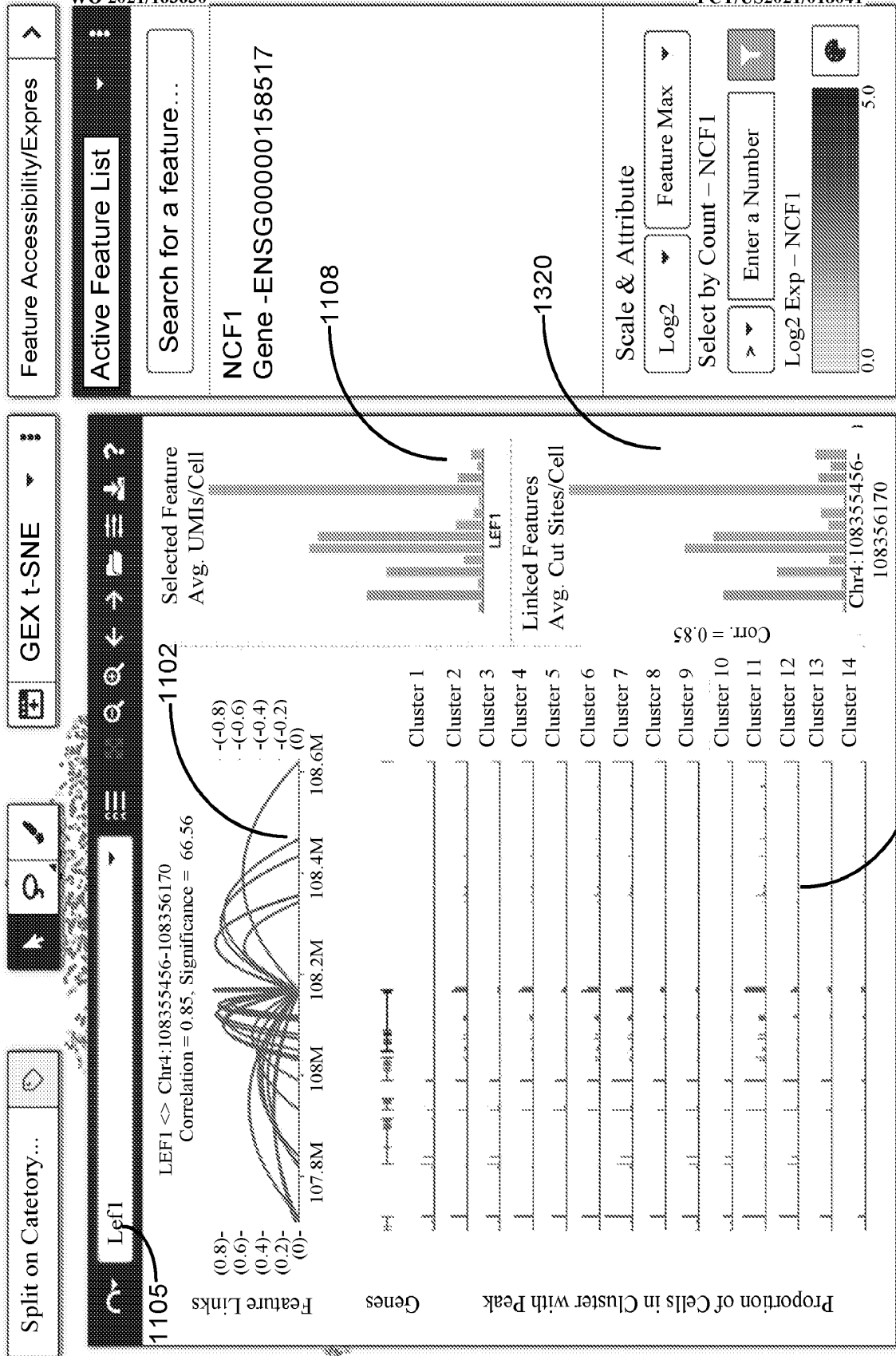


Figure 13C

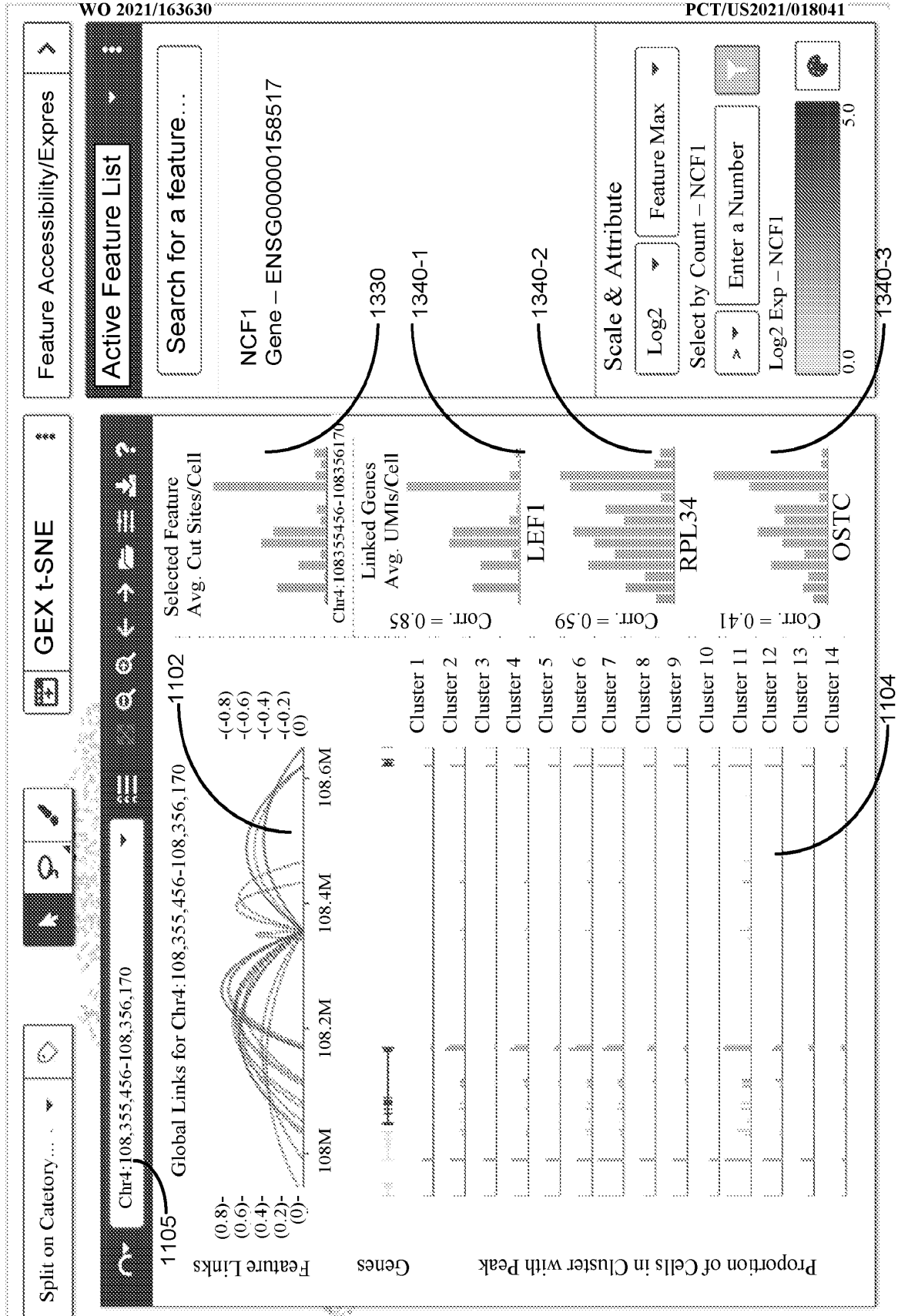


Figure 13D

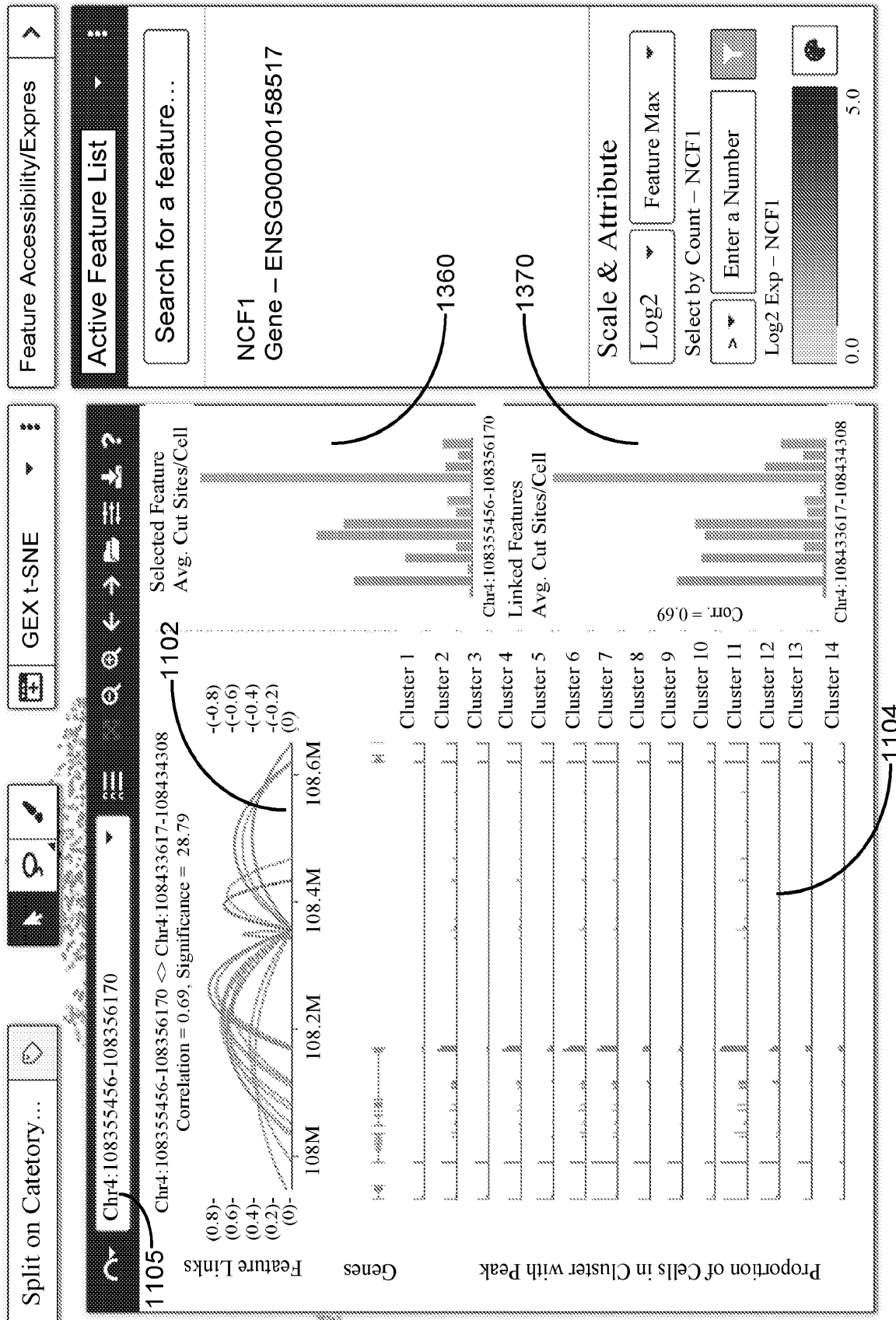


Figure 13E

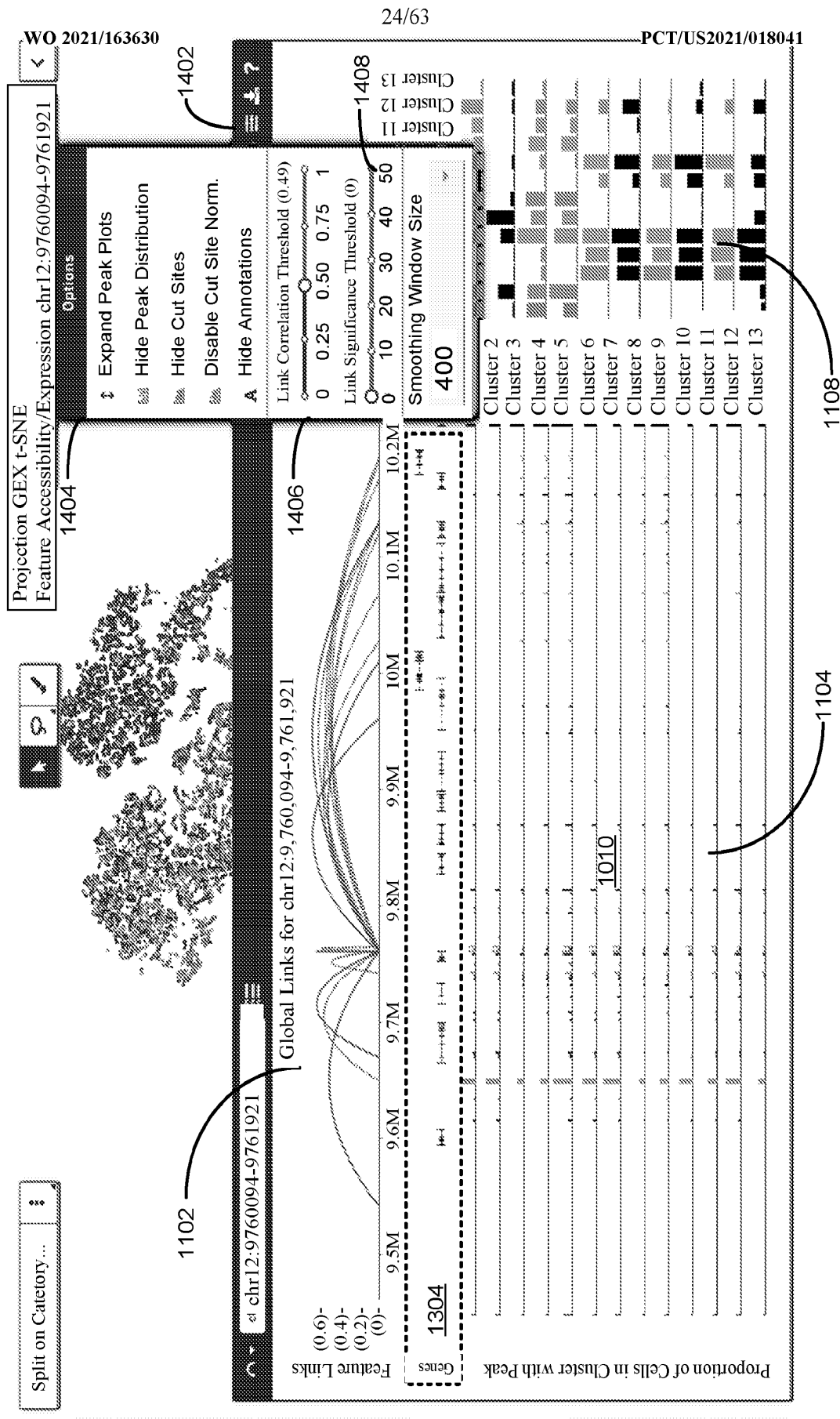


Figure 14

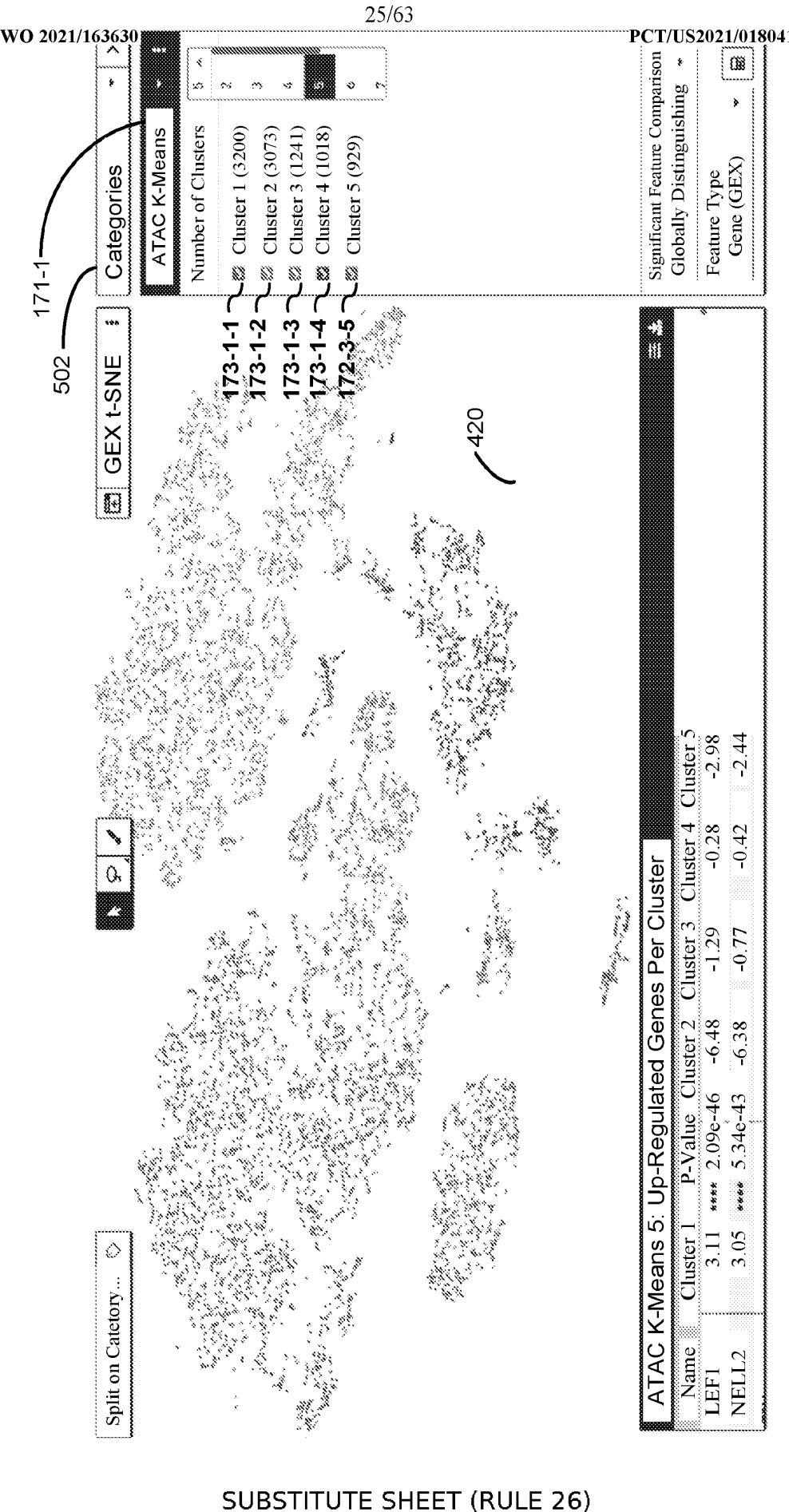


Figure 15

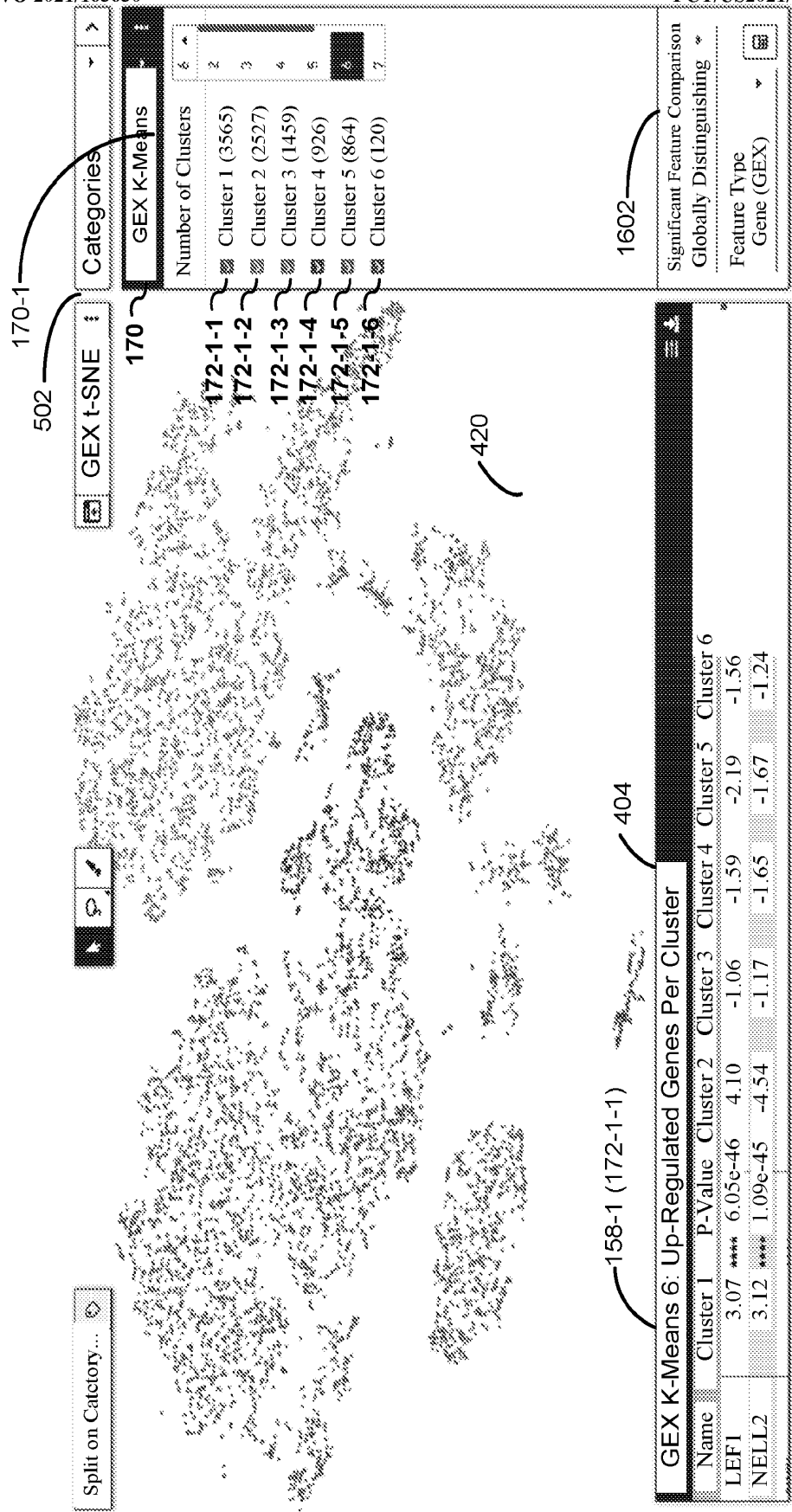
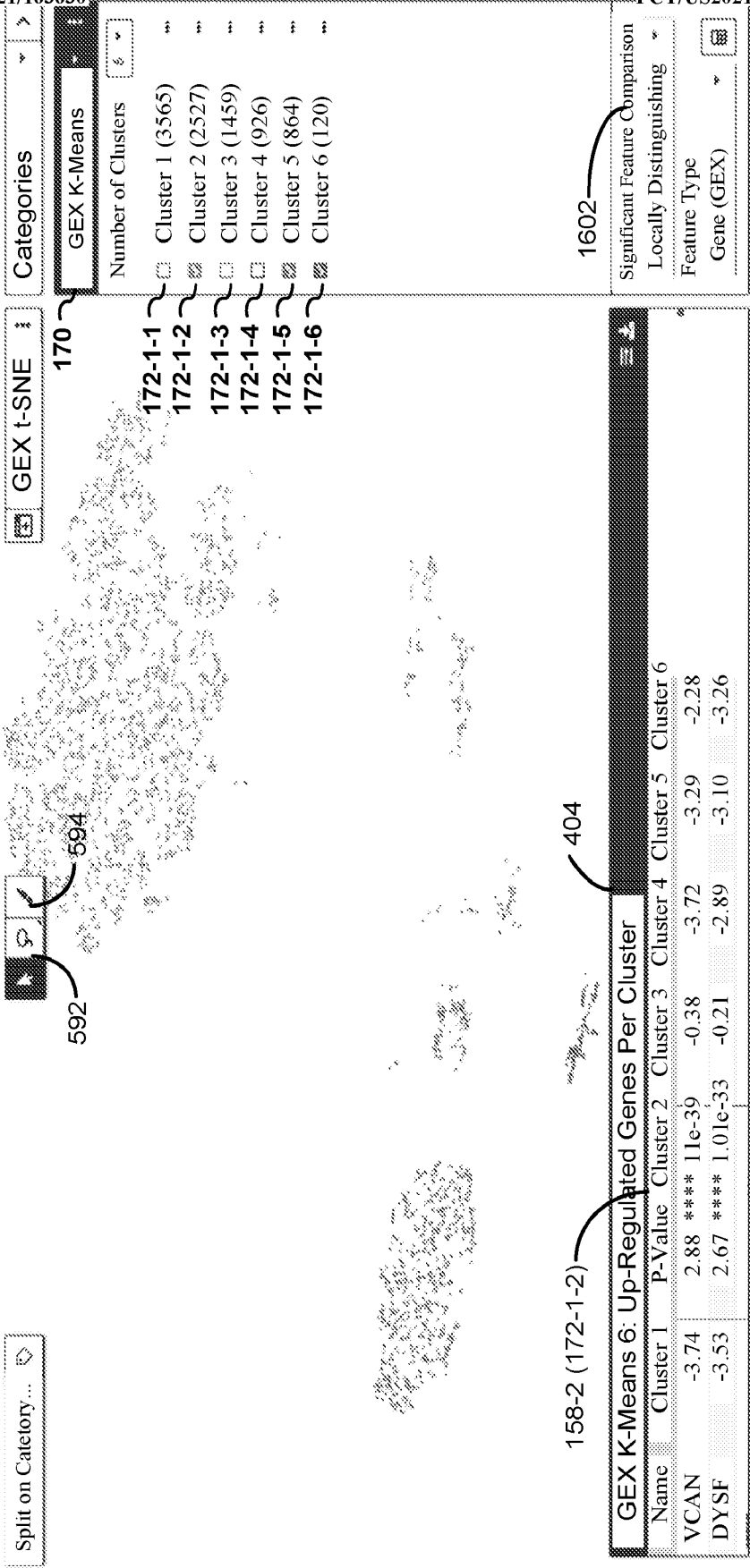


Figure 16



158-2 (172-1-2)

404

GEX K-Means 6: Up-Regulated Genes Per Cluster

Name	Cluster 1	P-Value	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
VCAN	-3.74	2.88 ***	11e-39	-0.38	-3.72	-3.29	-2.28
DYSF	-3.53	2.67 ***	1.01e-33	-0.21	-2.89	-3.10	-3.26

Figure 17

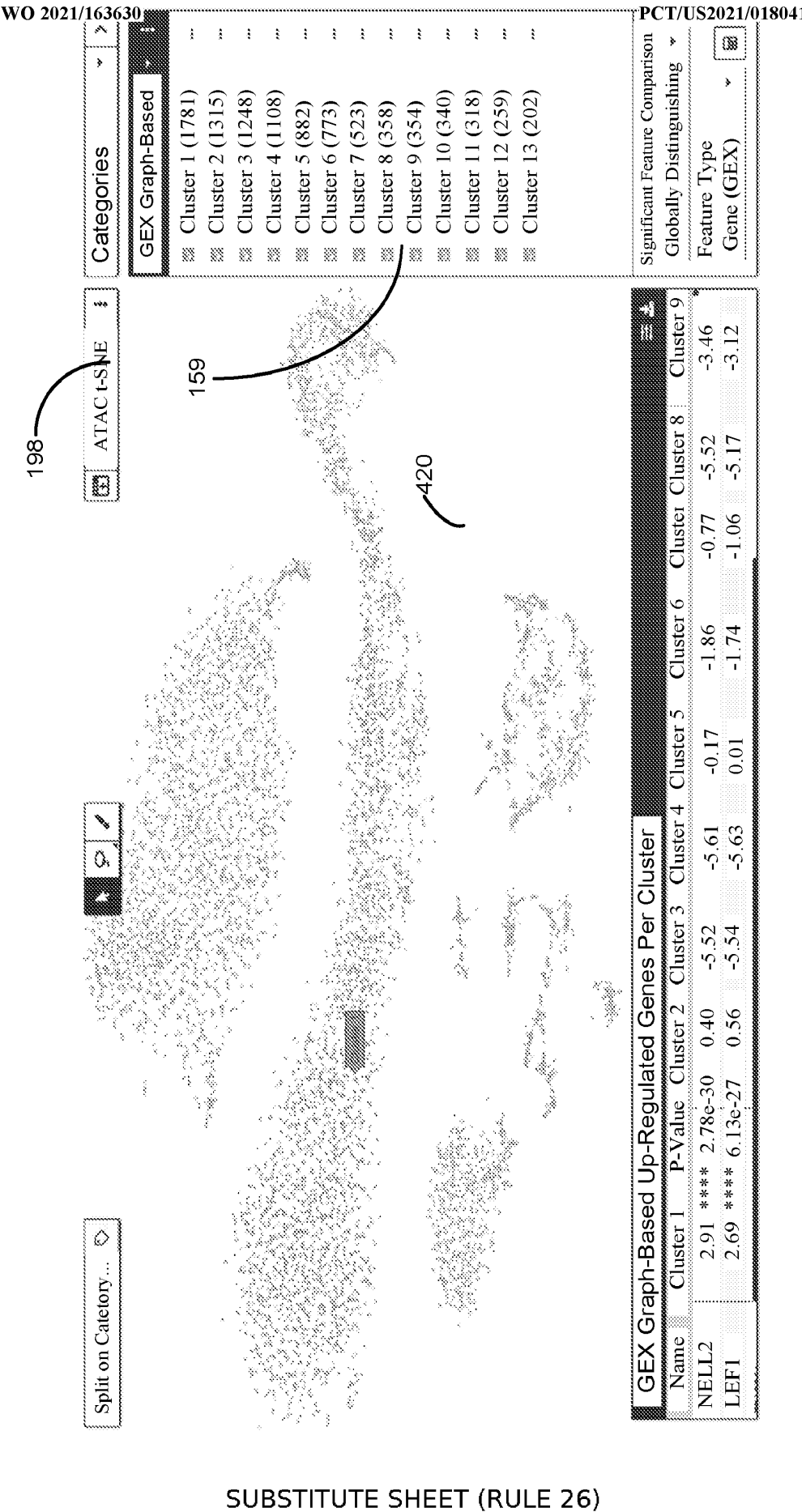


Figure 18

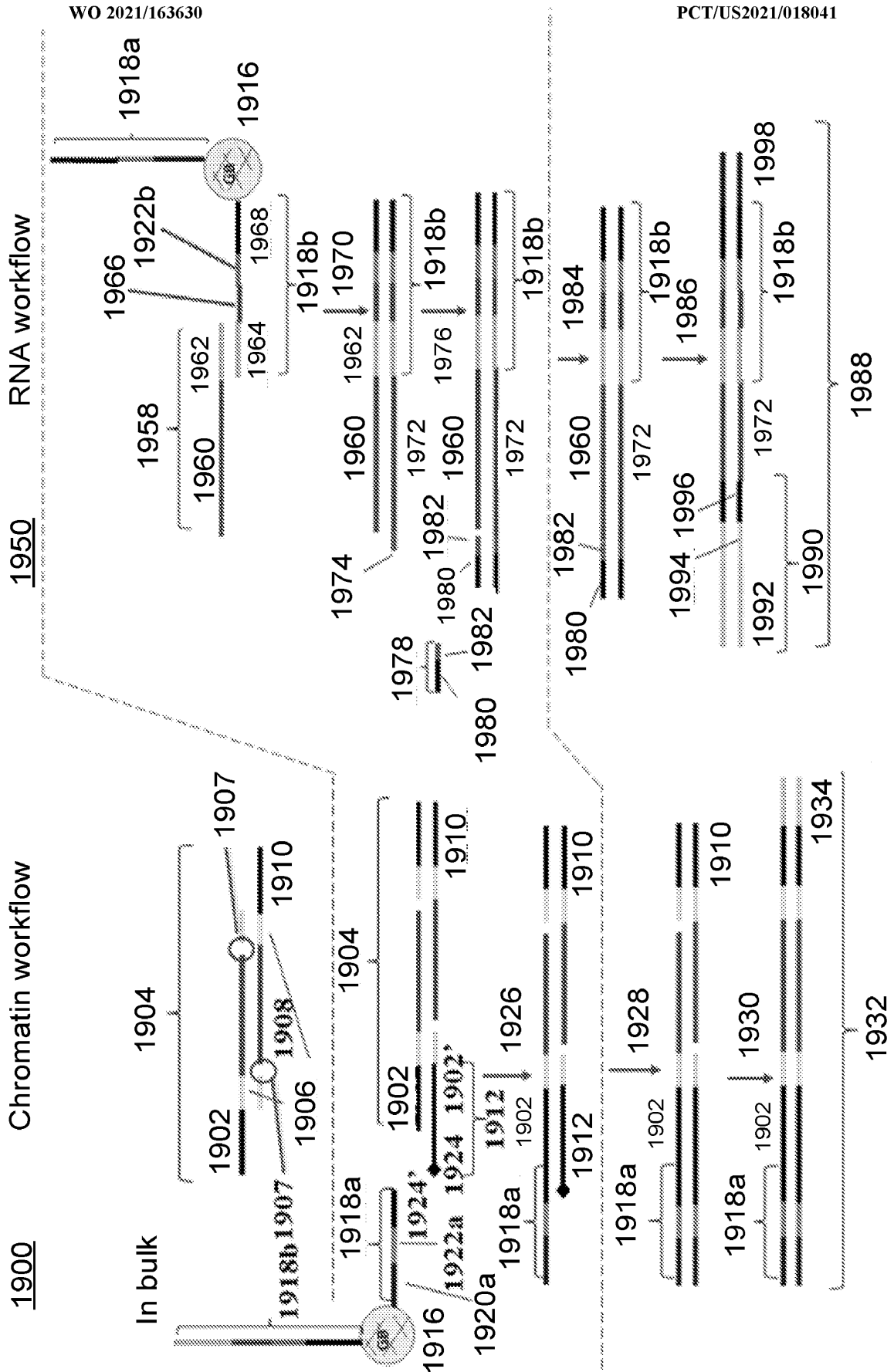


Figure 19

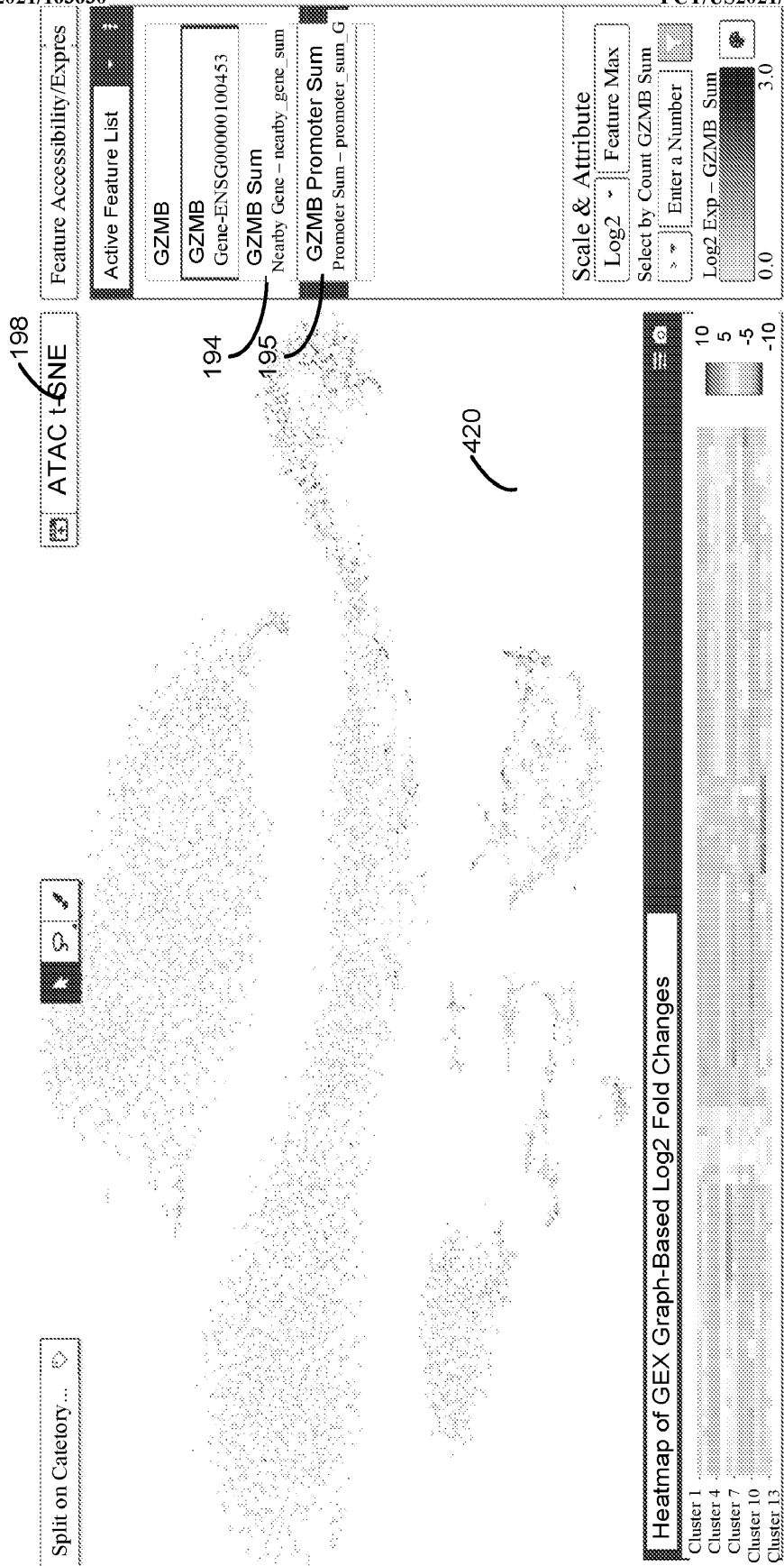


Figure 20

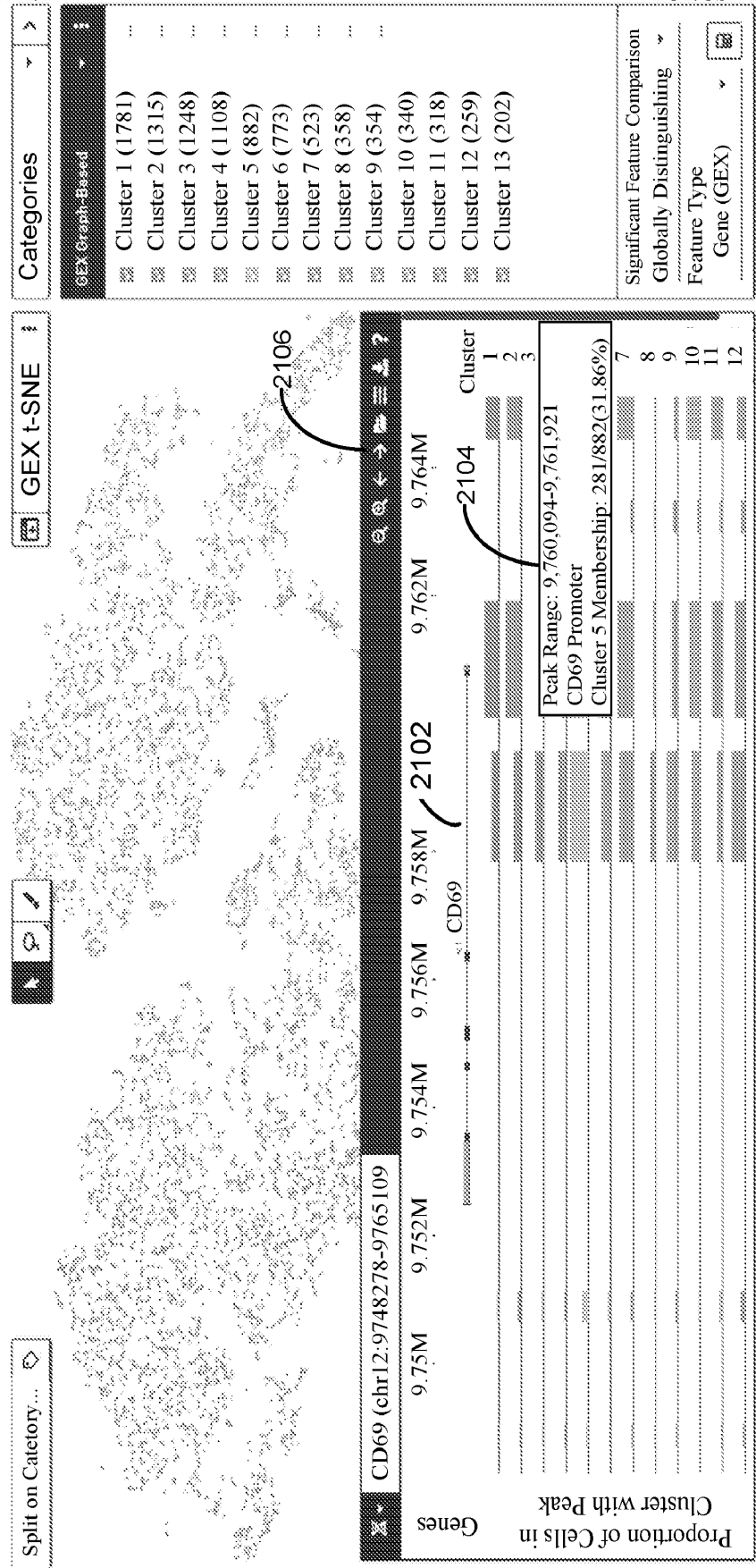


Figure 21

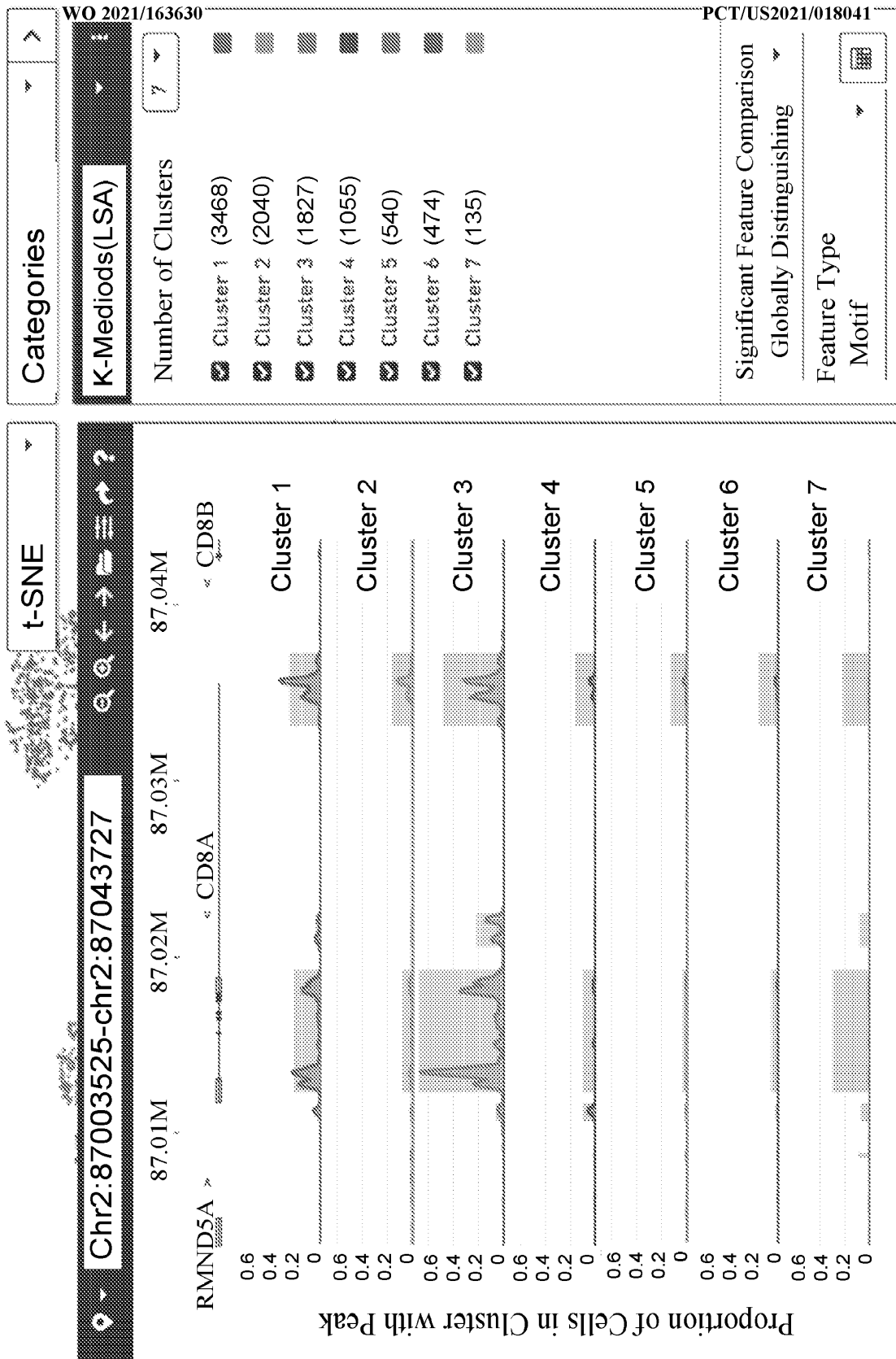


Figure 22

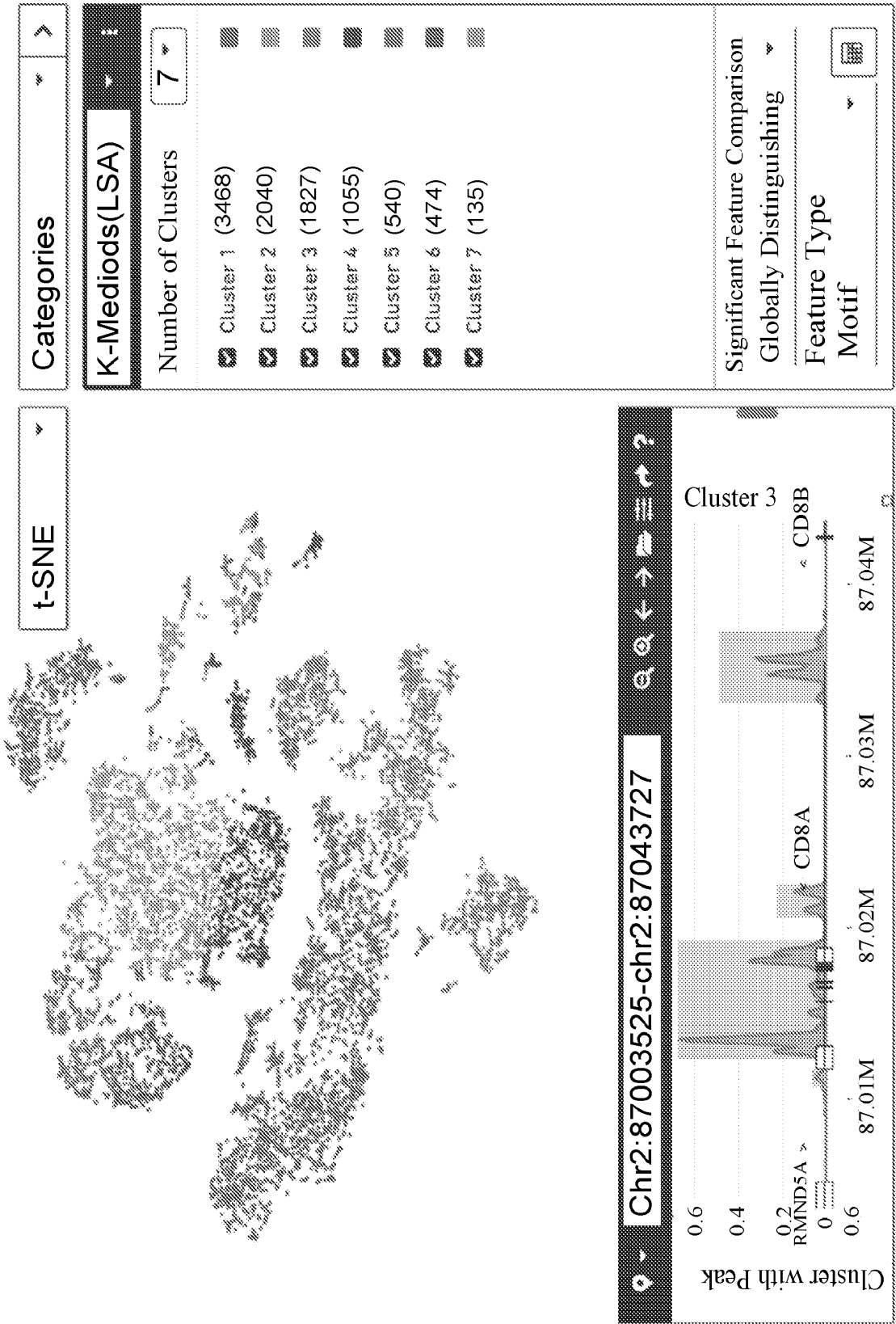


Figure 23

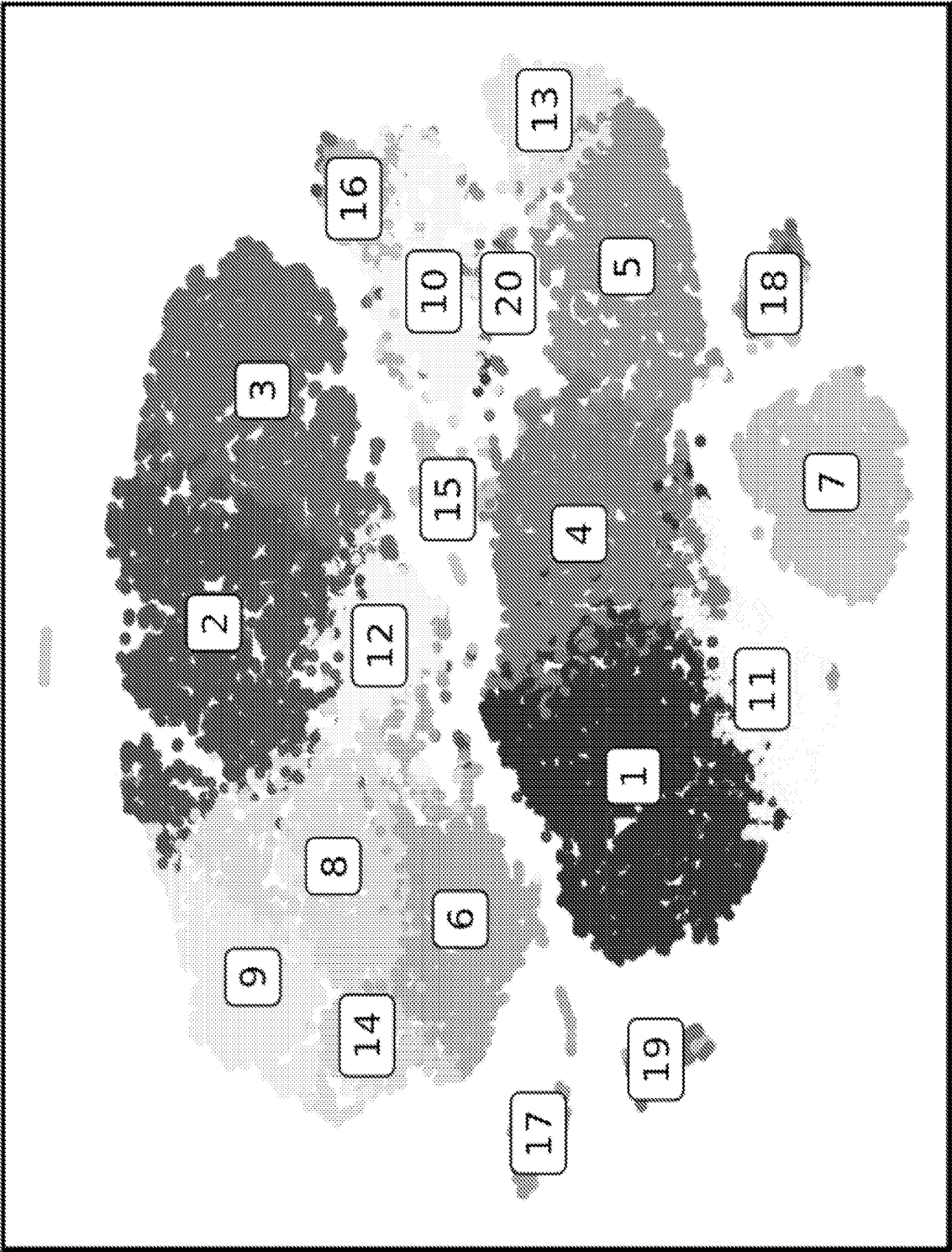


Figure 24A

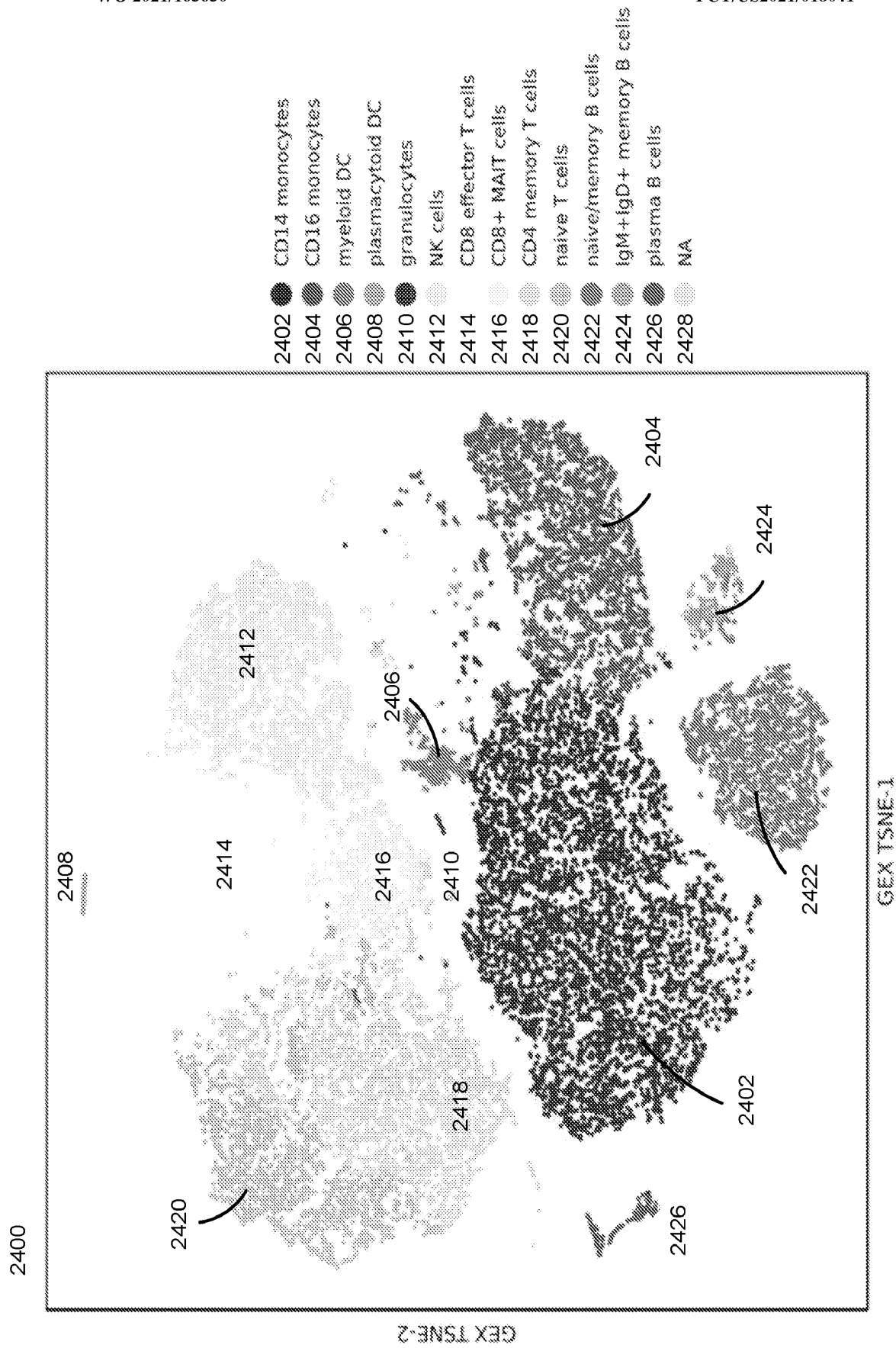


Figure 24B

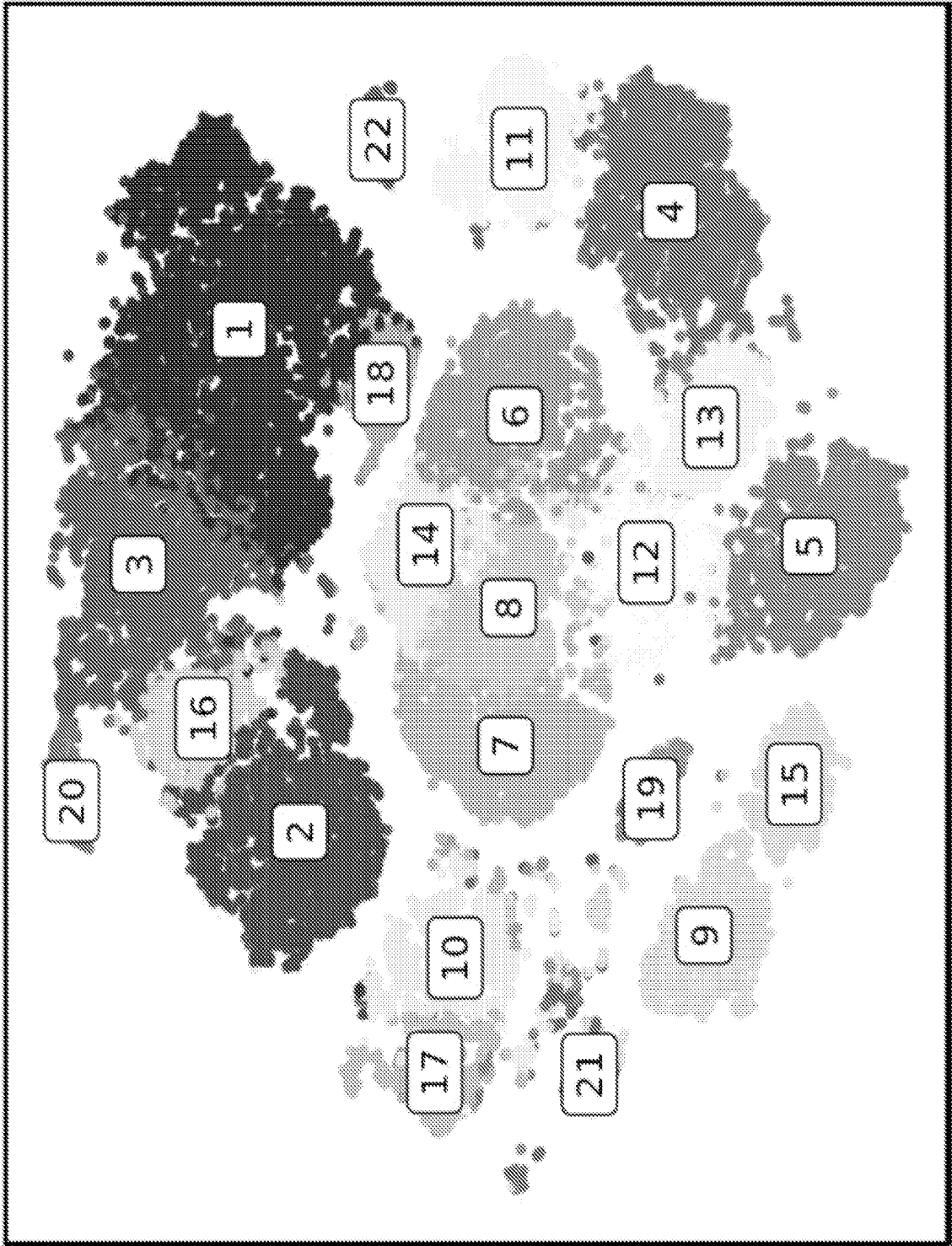


Figure 25A

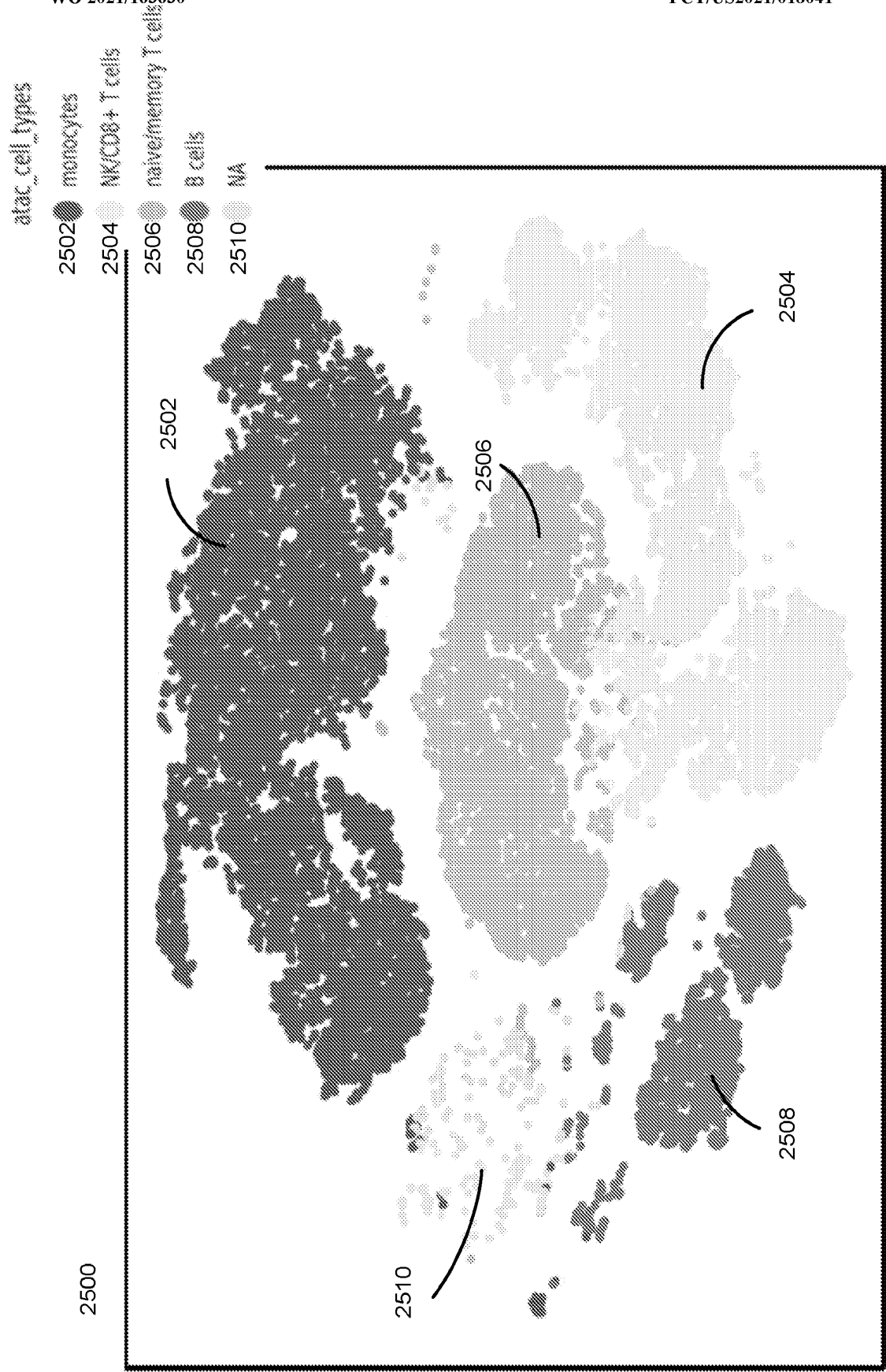


Figure 25B

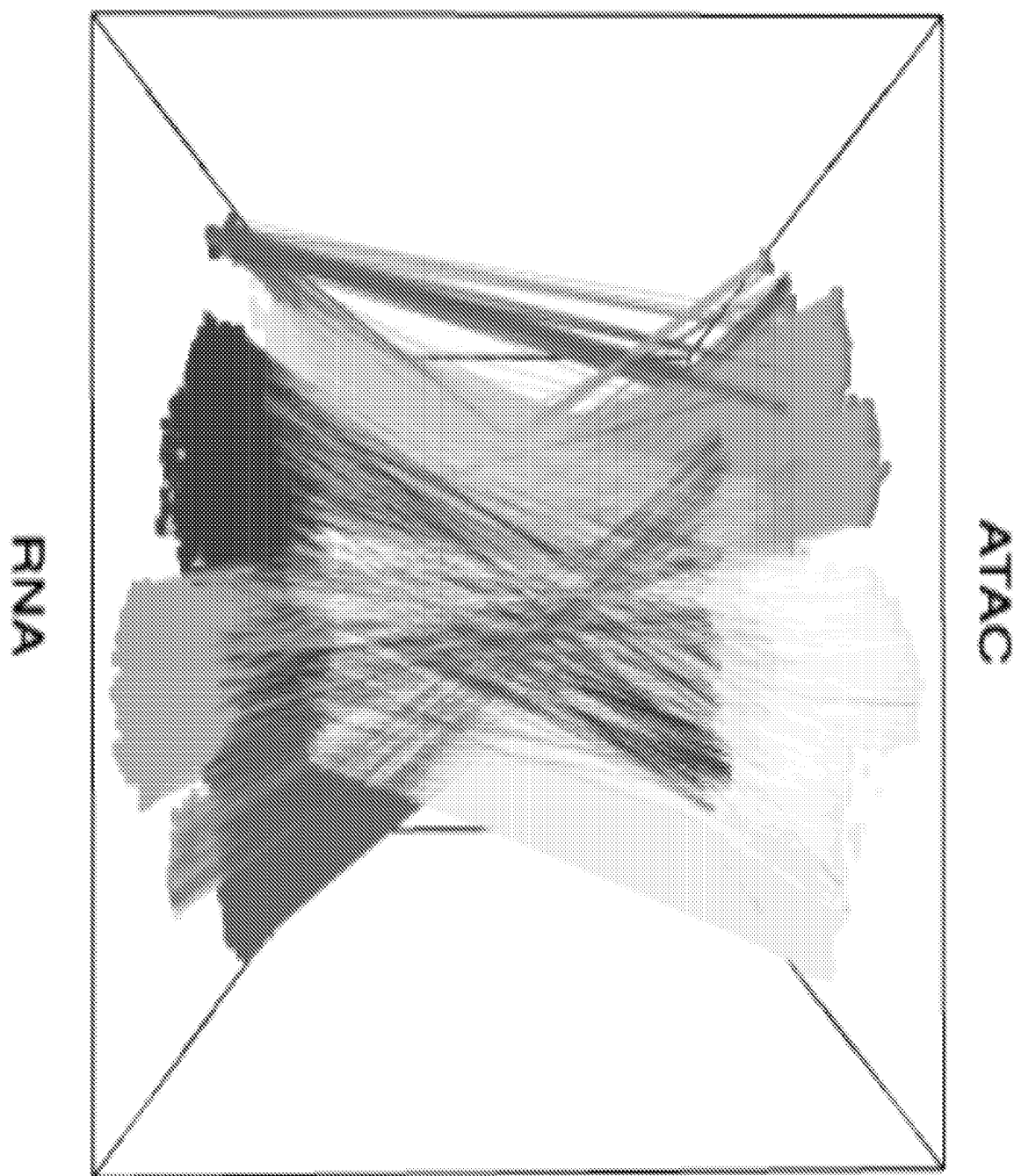


Figure 26A

SUBSTITUTE SHEET (RULE 26)

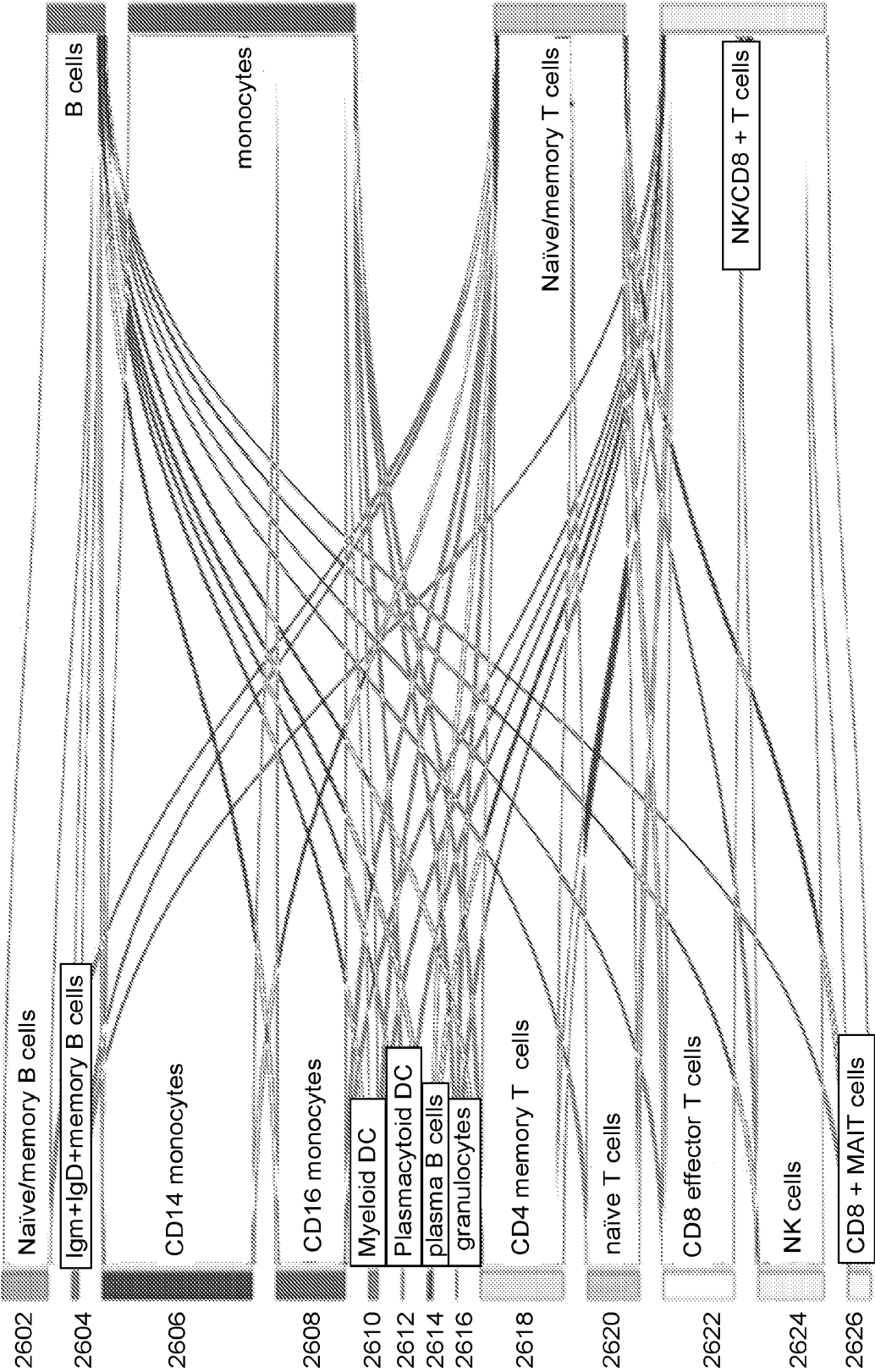


Figure 26B

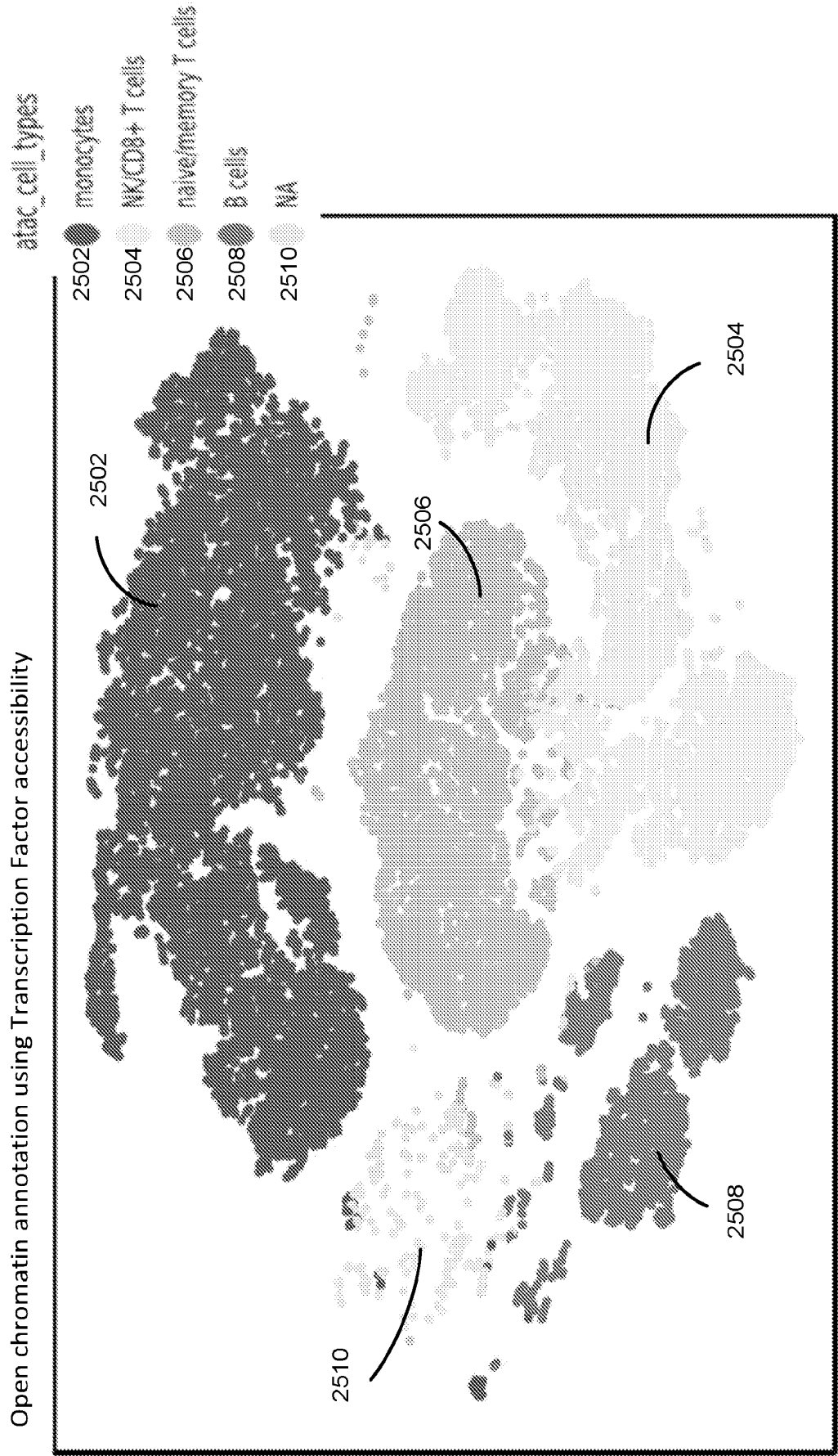


Figure 27A

Open chromatin annotation using Gene Expression markers

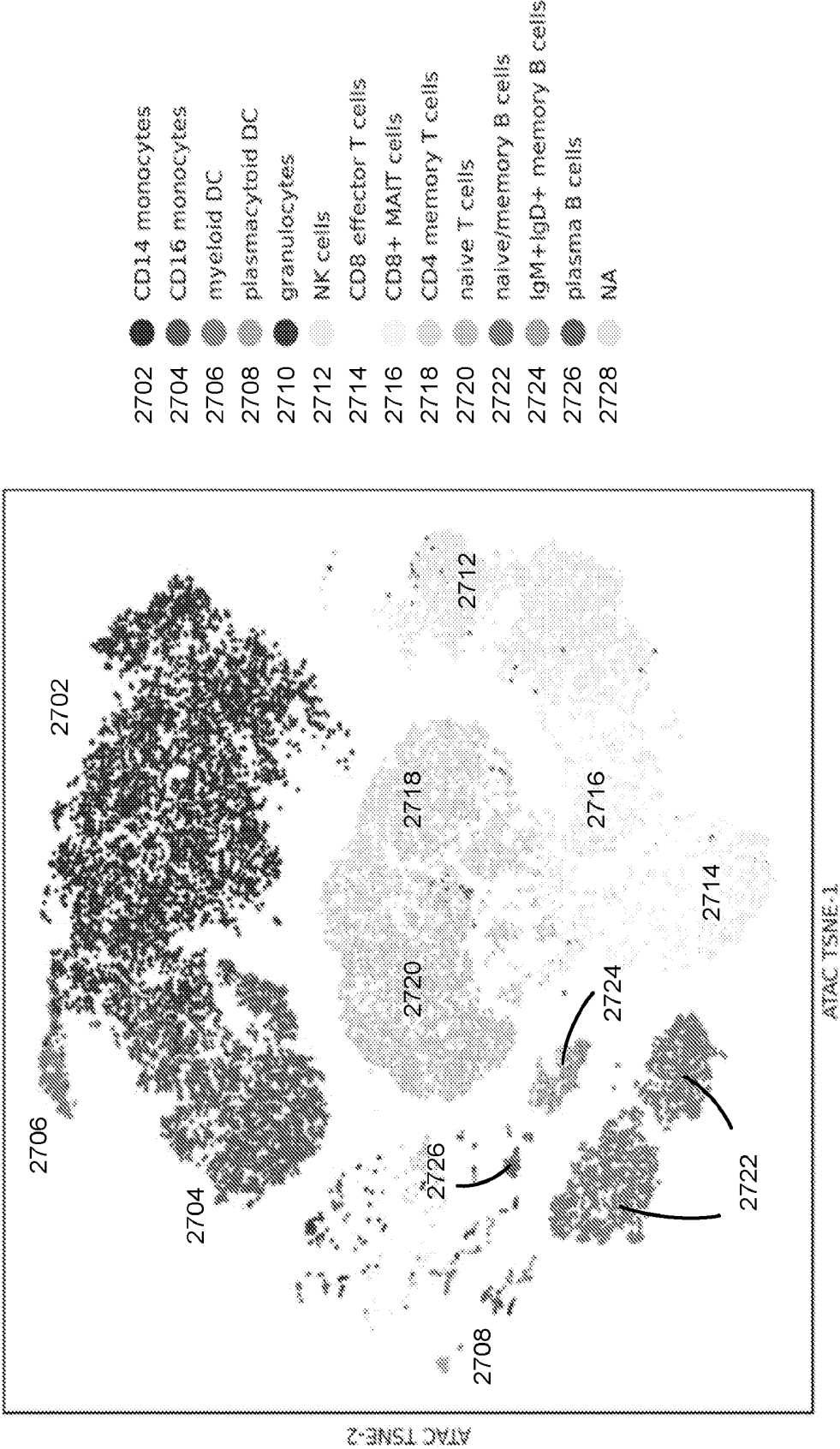


Figure 27B

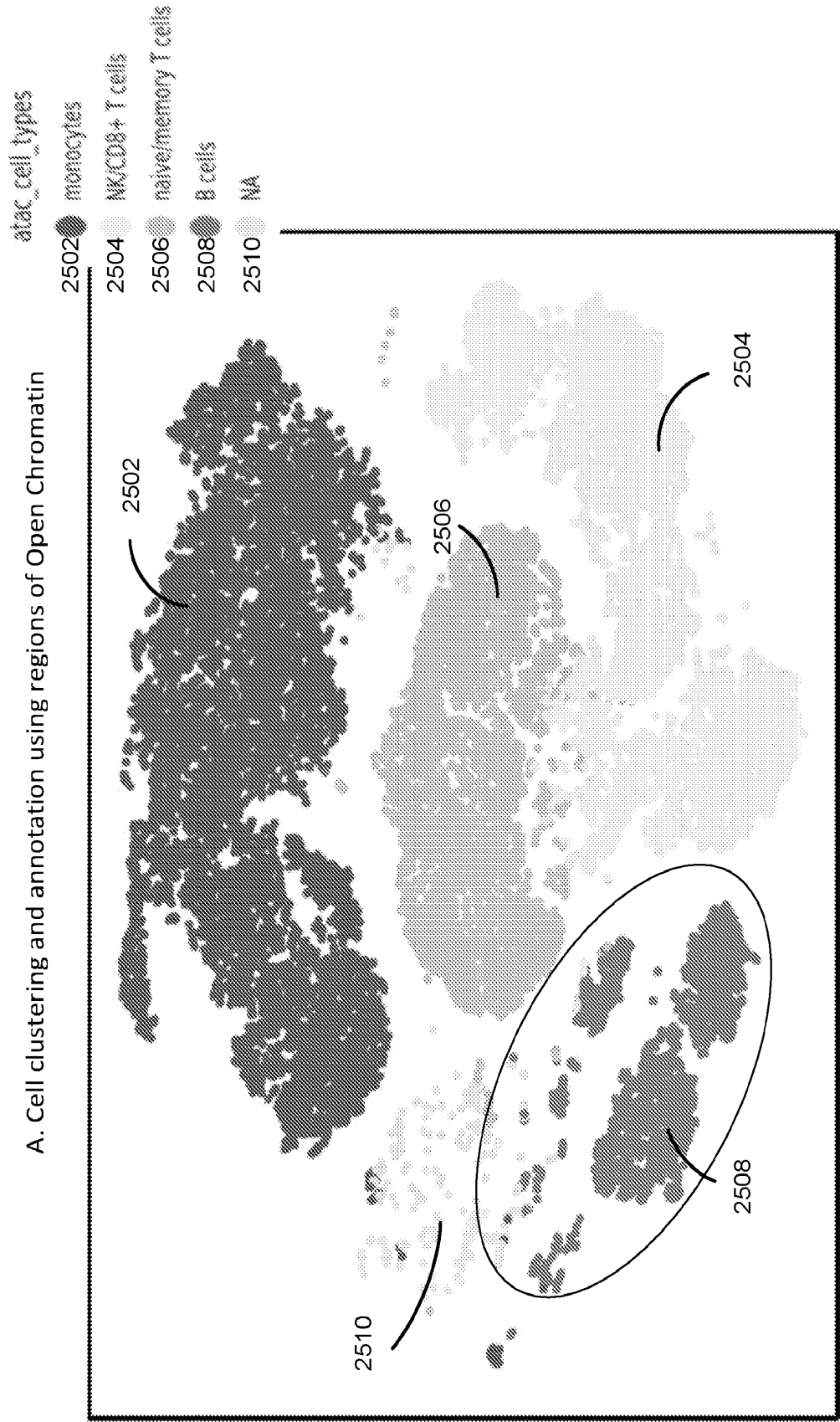


Figure 28A

B. Cell clustering and annotation using Gene Expression markers

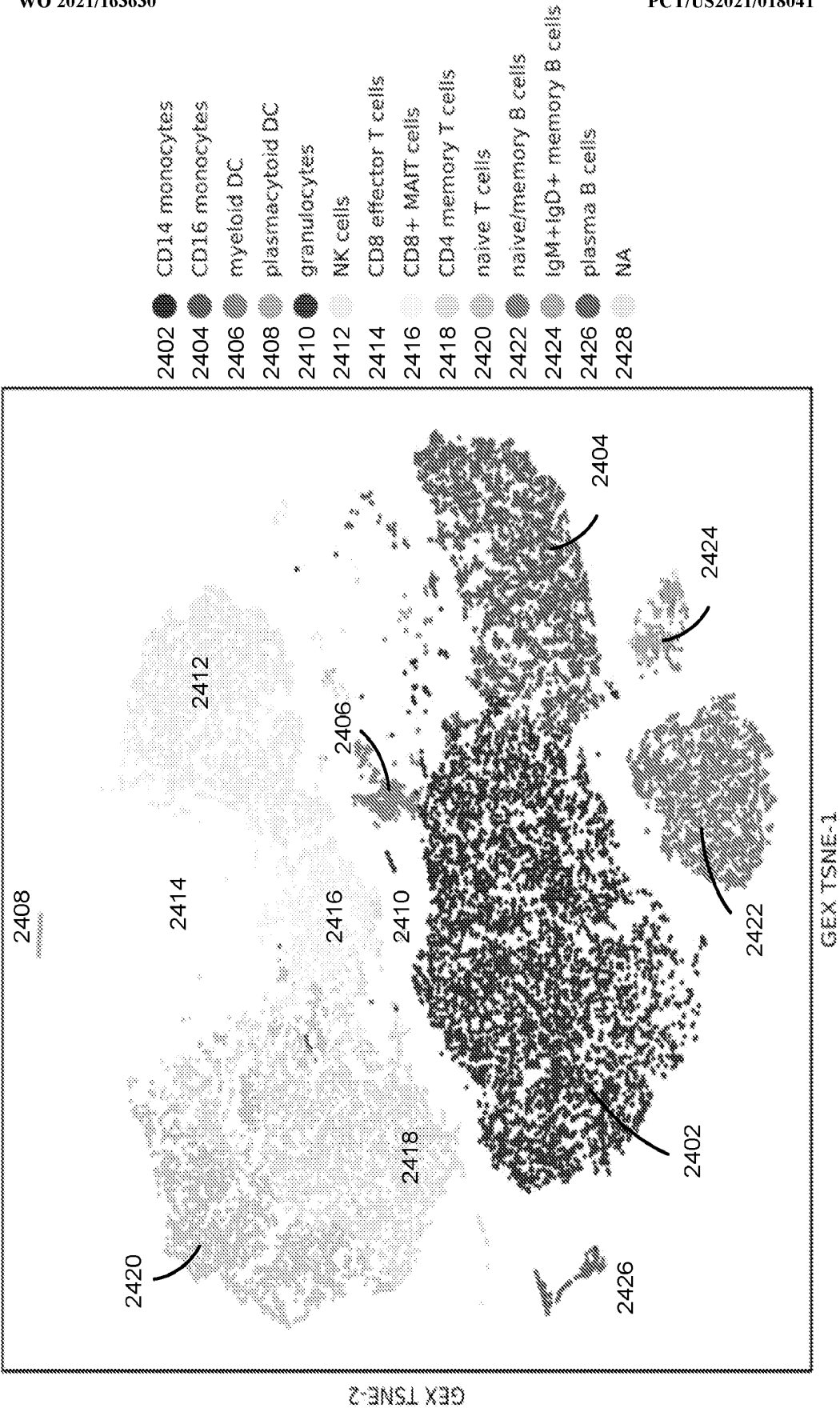


Figure 28B

C. Open Chromatin annotation using Gene Expression markers

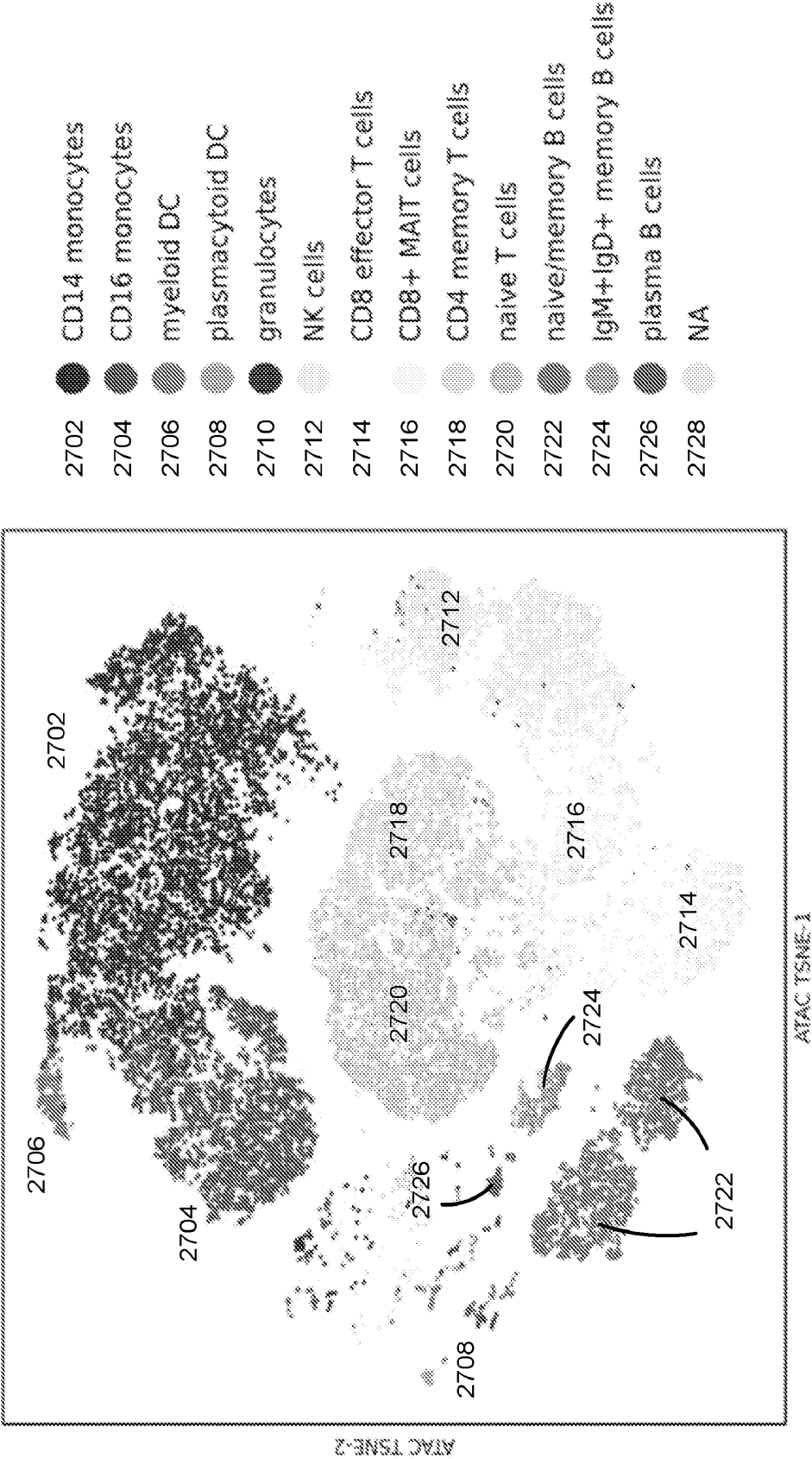
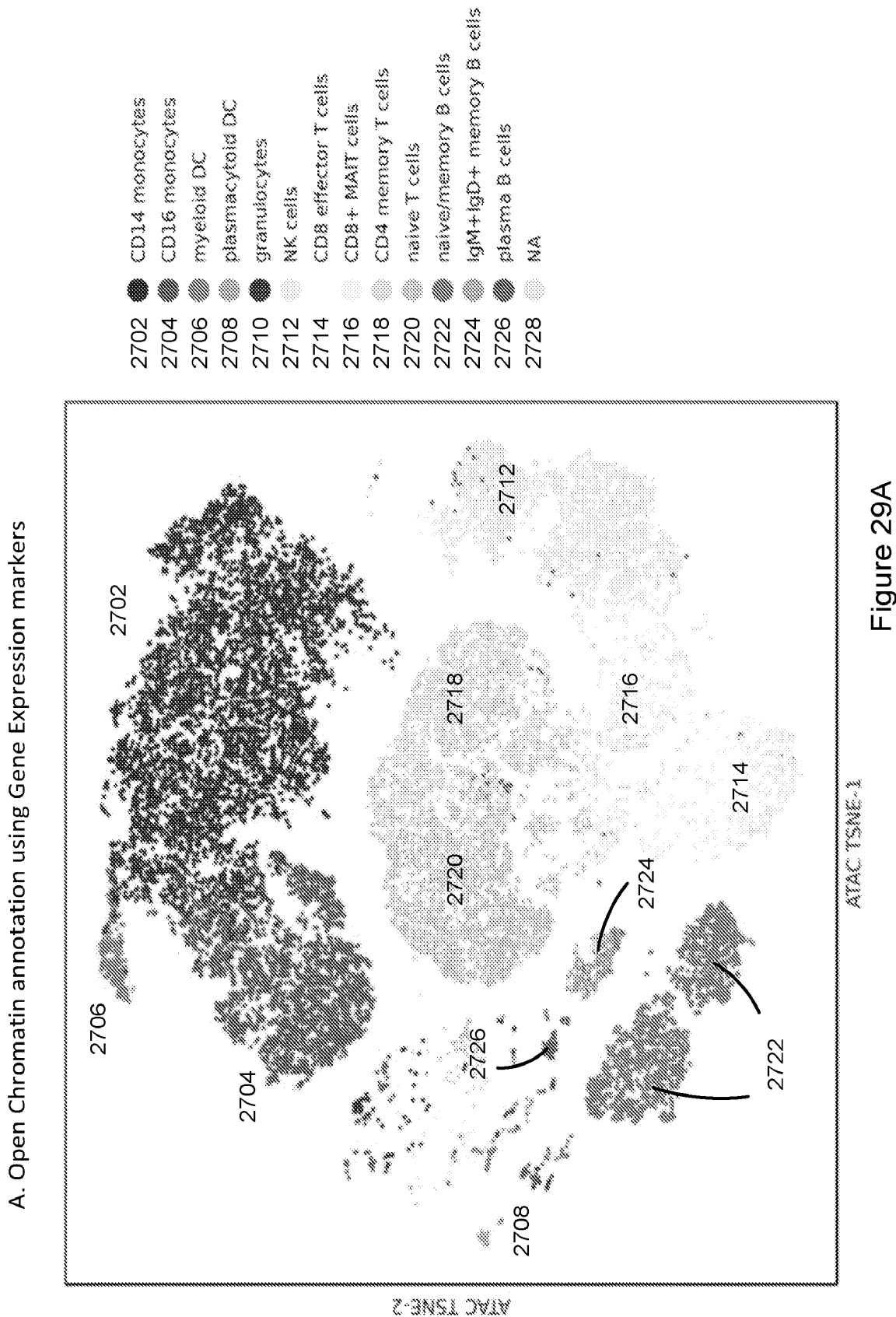


Figure 28C



B. Cell annotation using Gene Expression markers

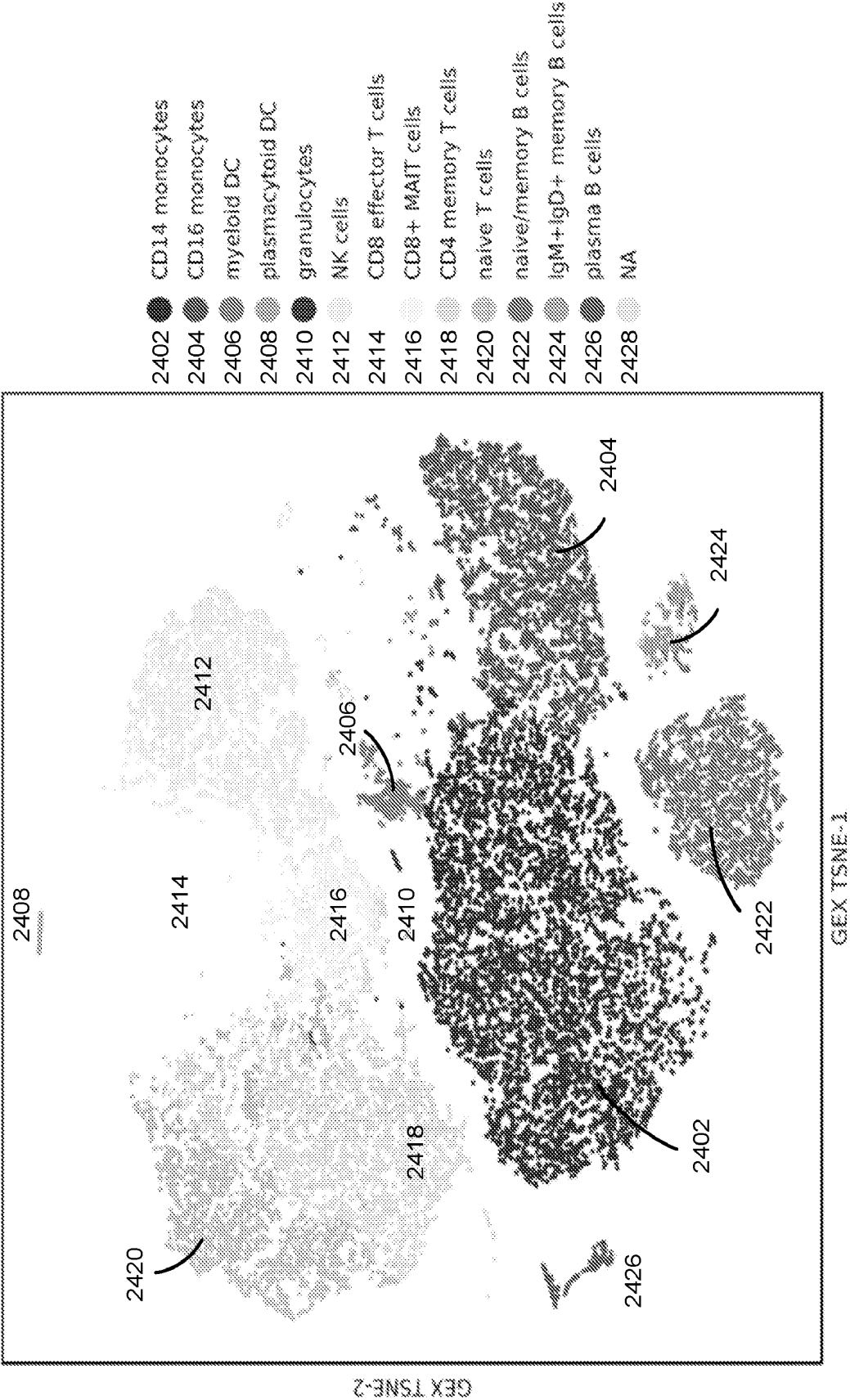
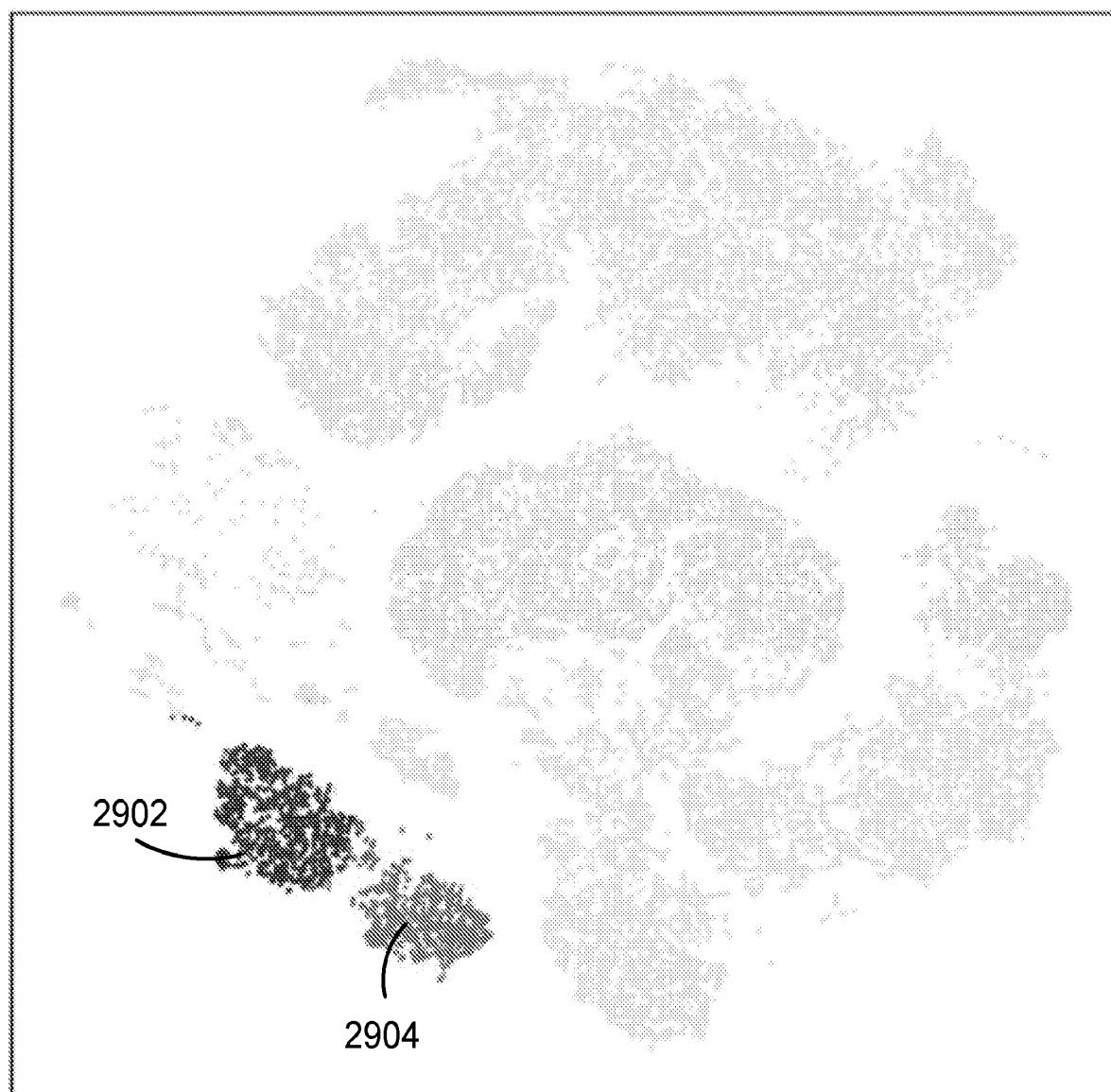


Figure 29B

C. Gene Expression of Open Chromatin clusters

ATAC



2902 Subcluster1
(Memory B cells)

2904 Subcluster 2
(Naïve B cells)

Figure 29C

SUBSTITUTE SHEET (RULE 26)

D. Open chromatin regions of Gene Expression clusters

GEX



2902 Subcluster1
(Memory B cells)

2904 Subcluster 2
(Naïve B cells)

Figure 29D

SUBSTITUTE SHEET (RULE 26)

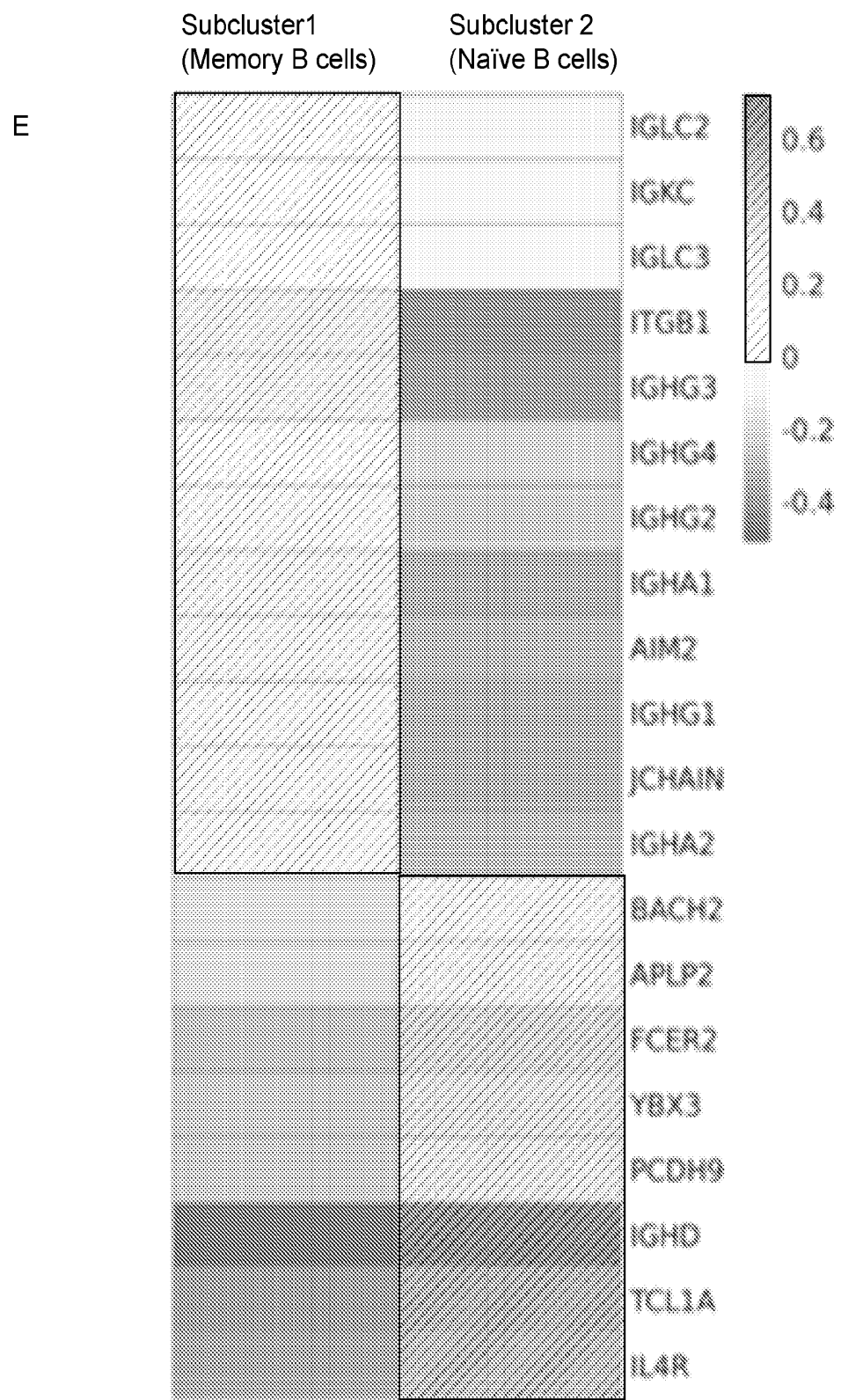


Figure 29E
SUBSTITUTE SHEET (RULE 26)

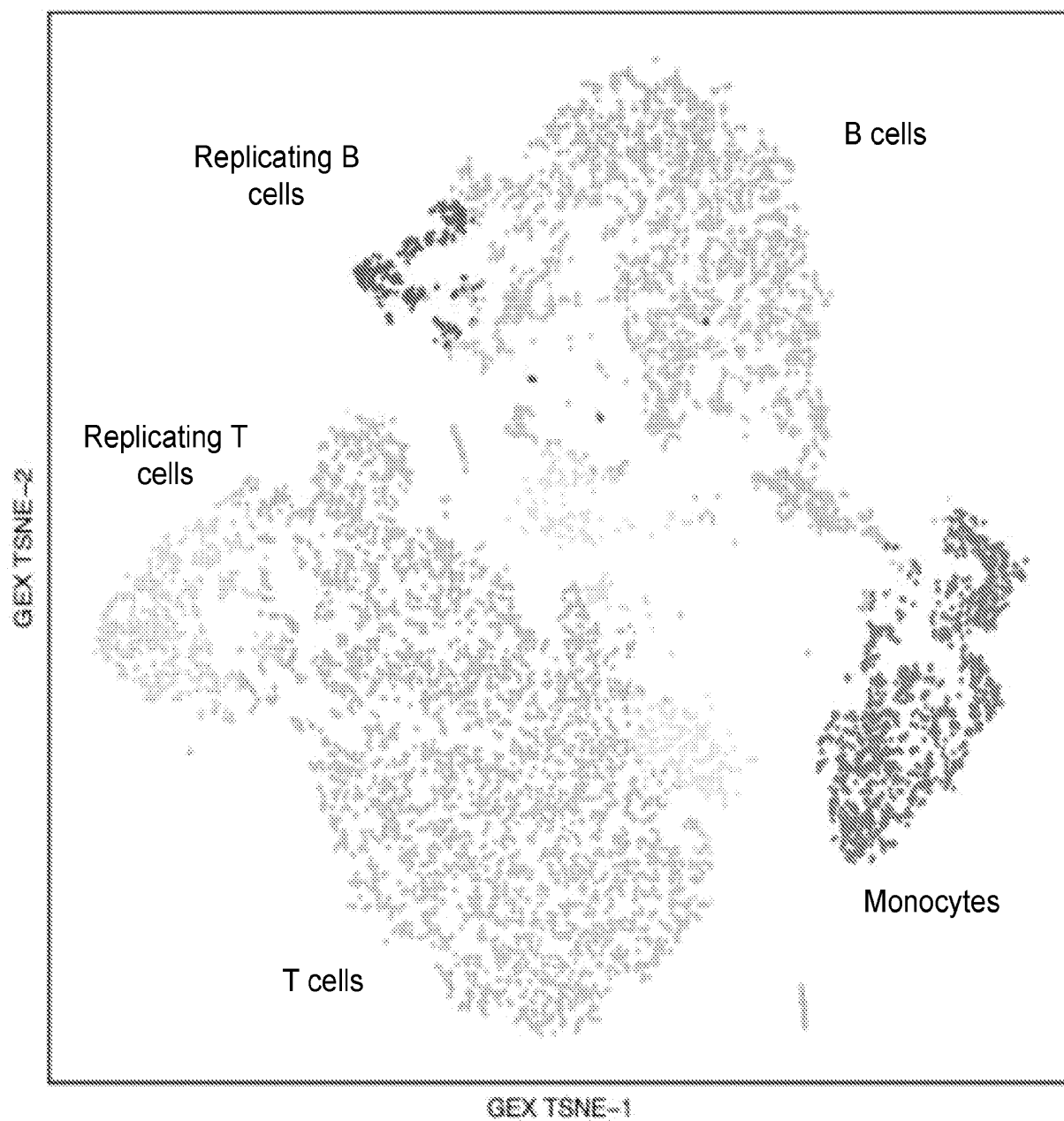


Figure 30A

SUBSTITUTE SHEET (RULE 26)

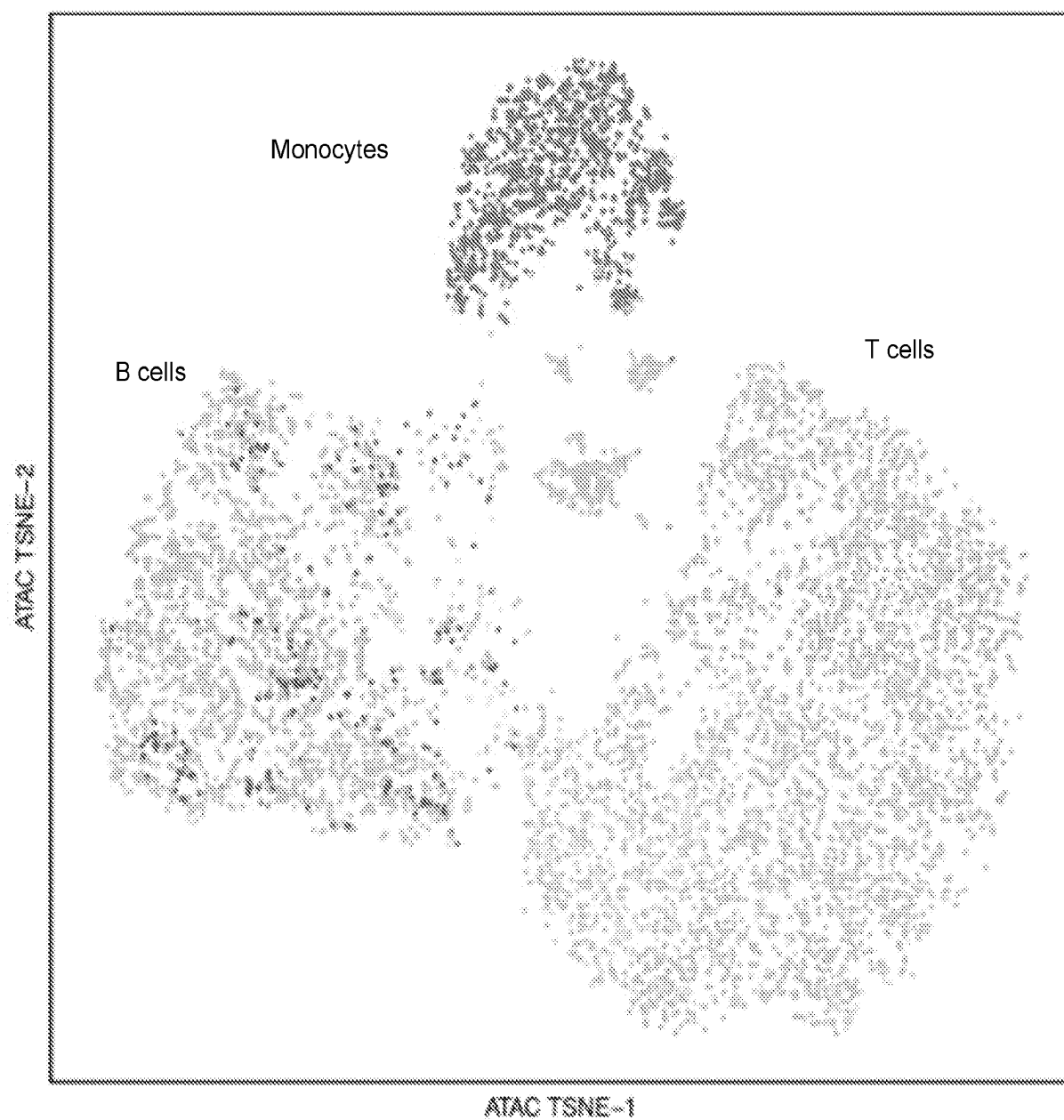


Figure 30B

SNVs

GEX TSNE-2



GEX TSNE-1

Figure 31A

SUBSTITUTE SHEET (RULE 26)

BANK1

GEX TSNE-2



GEX TSNE-1

Figure 31B

SUBSTITUTE SHEET (RULE 26)

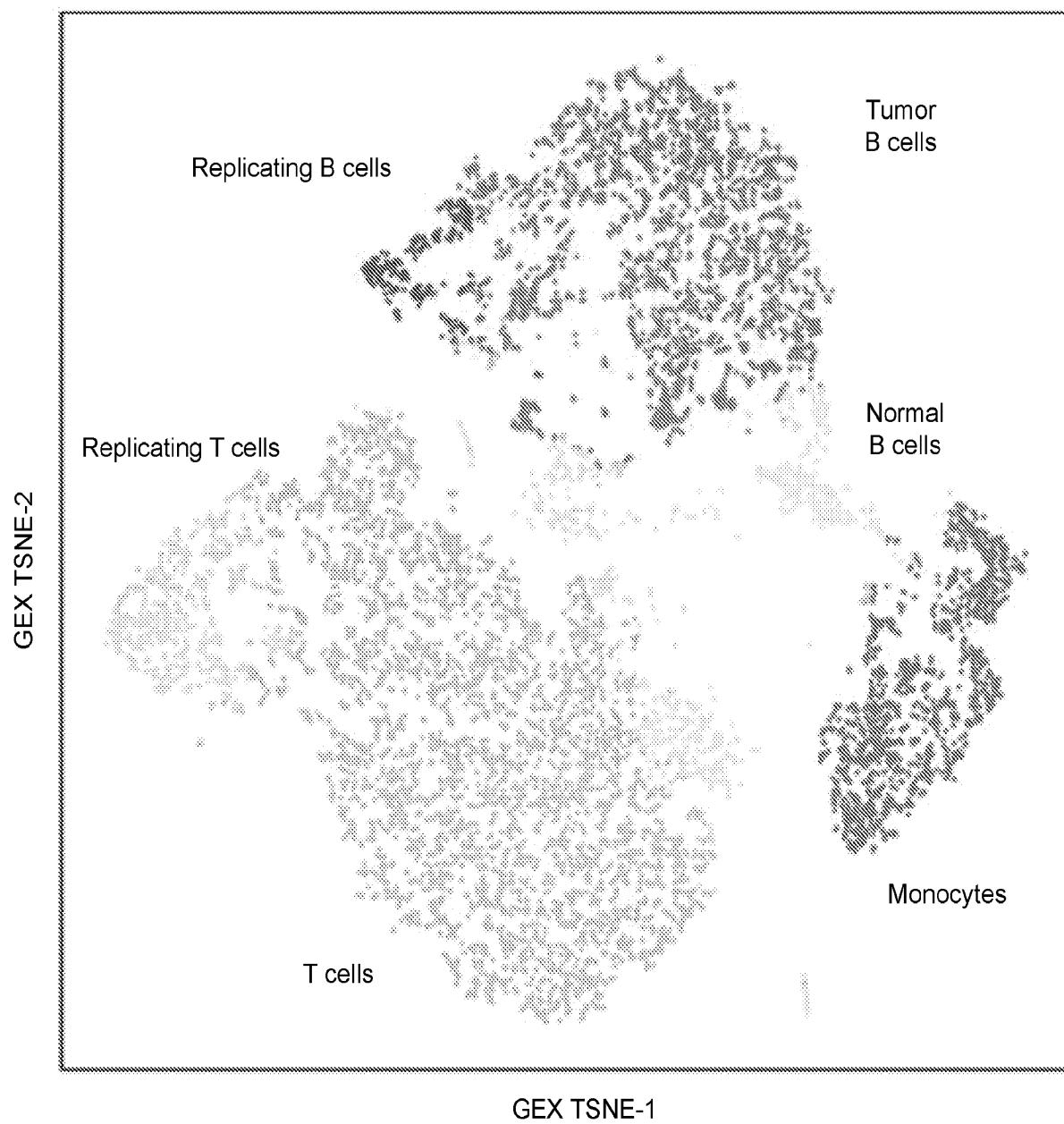


Figure 31C

SUBSTITUTE SHEET (RULE 26)

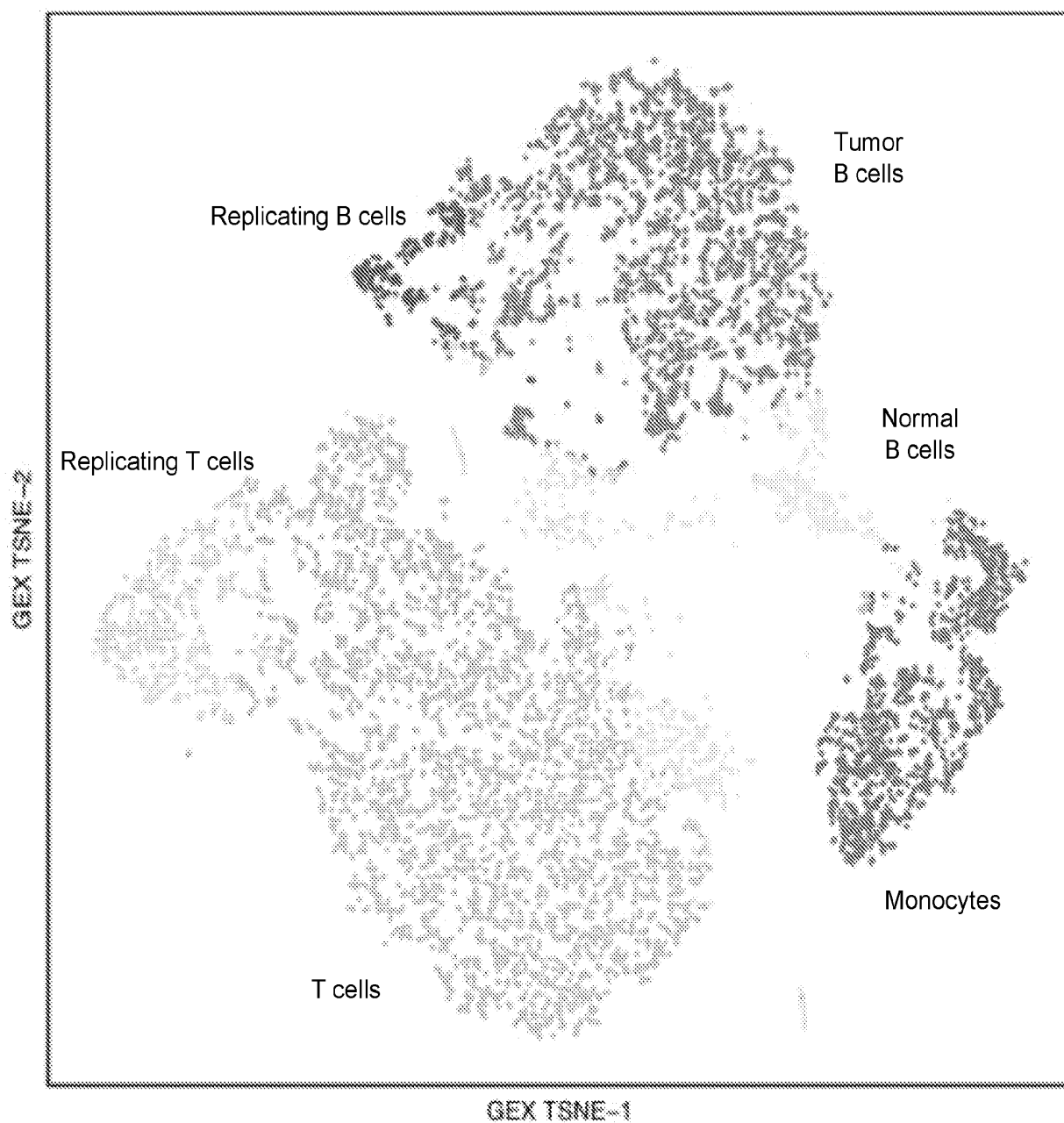


Figure 32A

SUBSTITUTE SHEET (RULE 26)

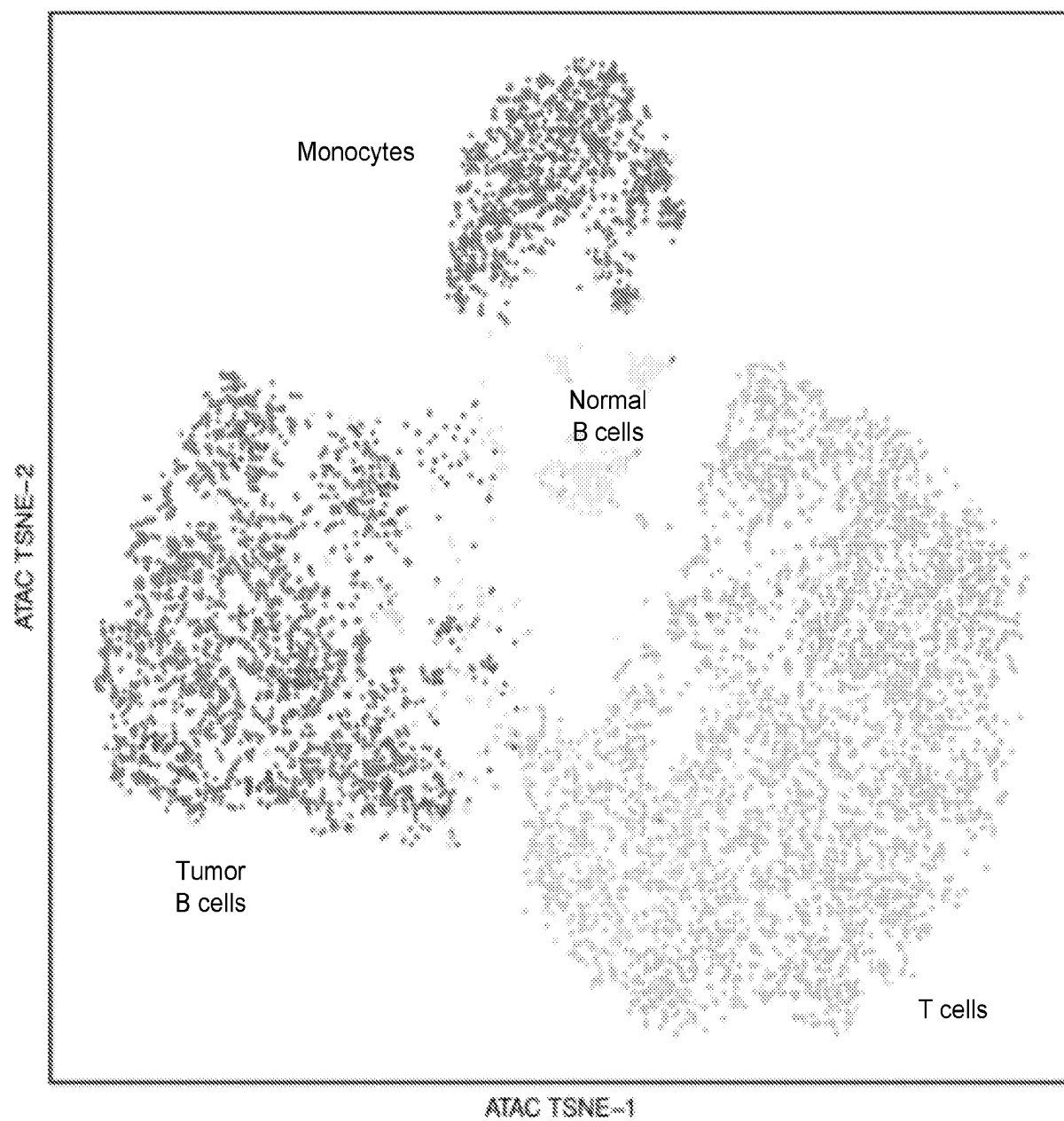


Figure 32B

C

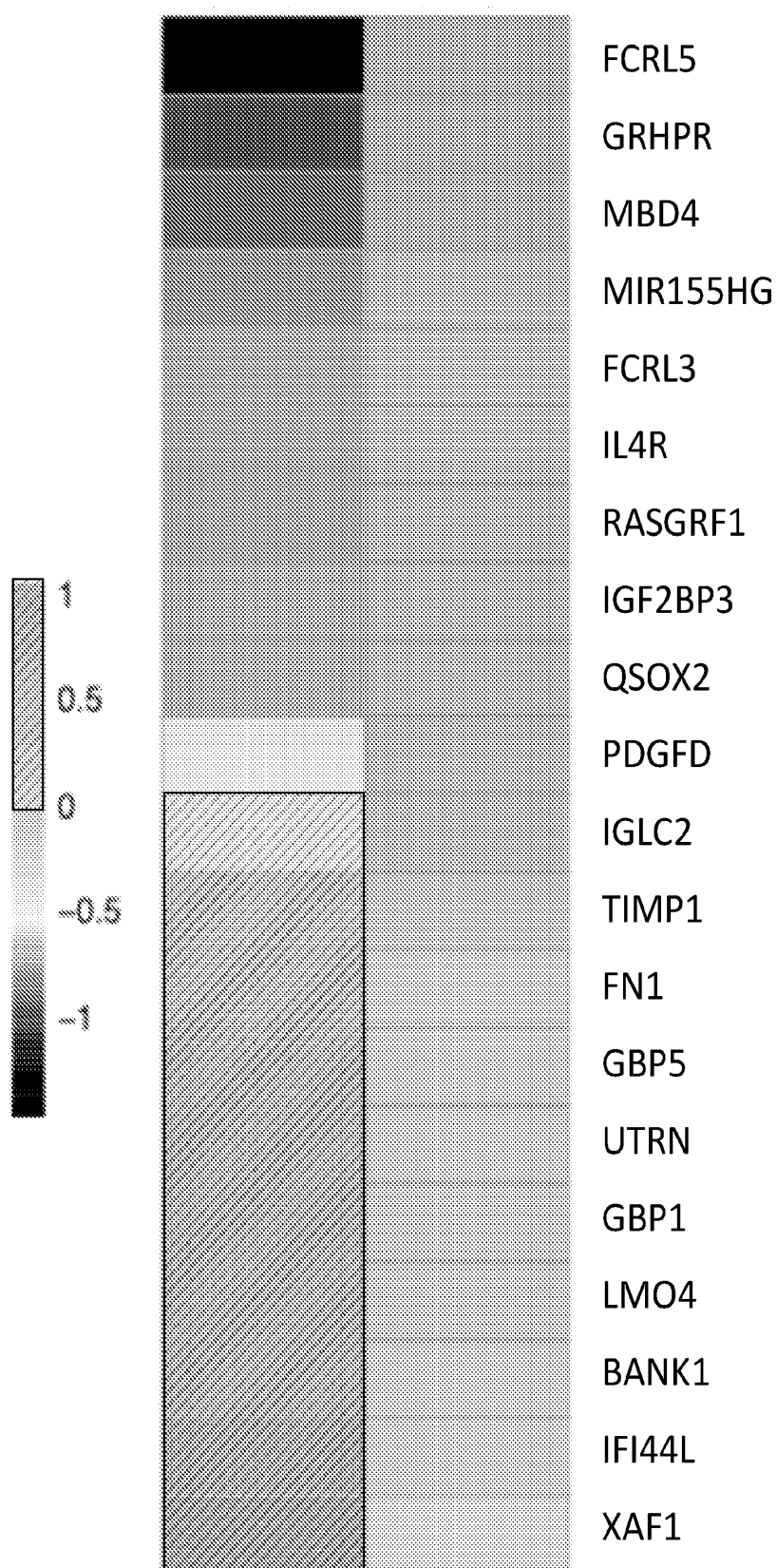


Figure 32C
SUBSTITUTE SHEET (RULE 26)

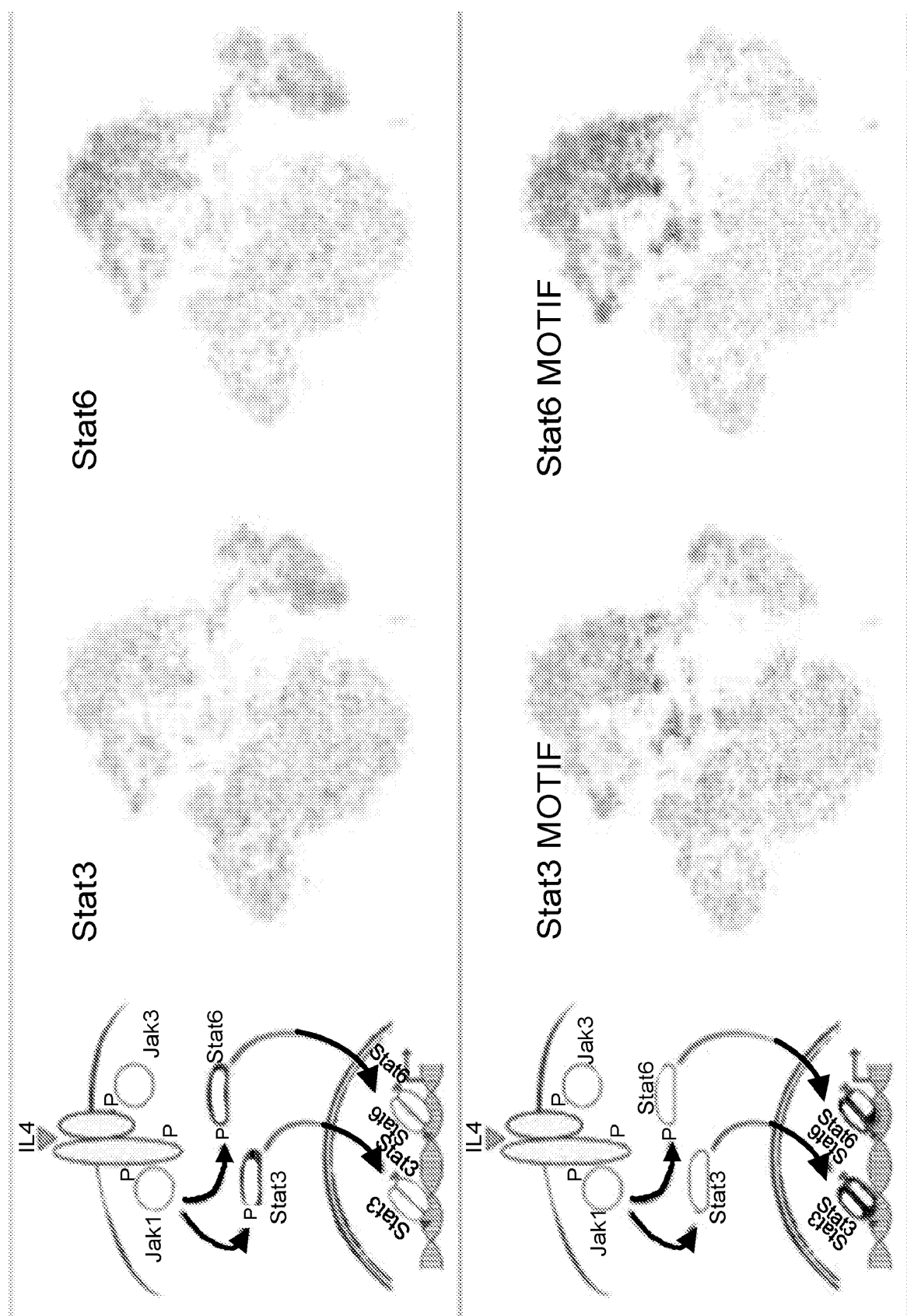


Figure 33A

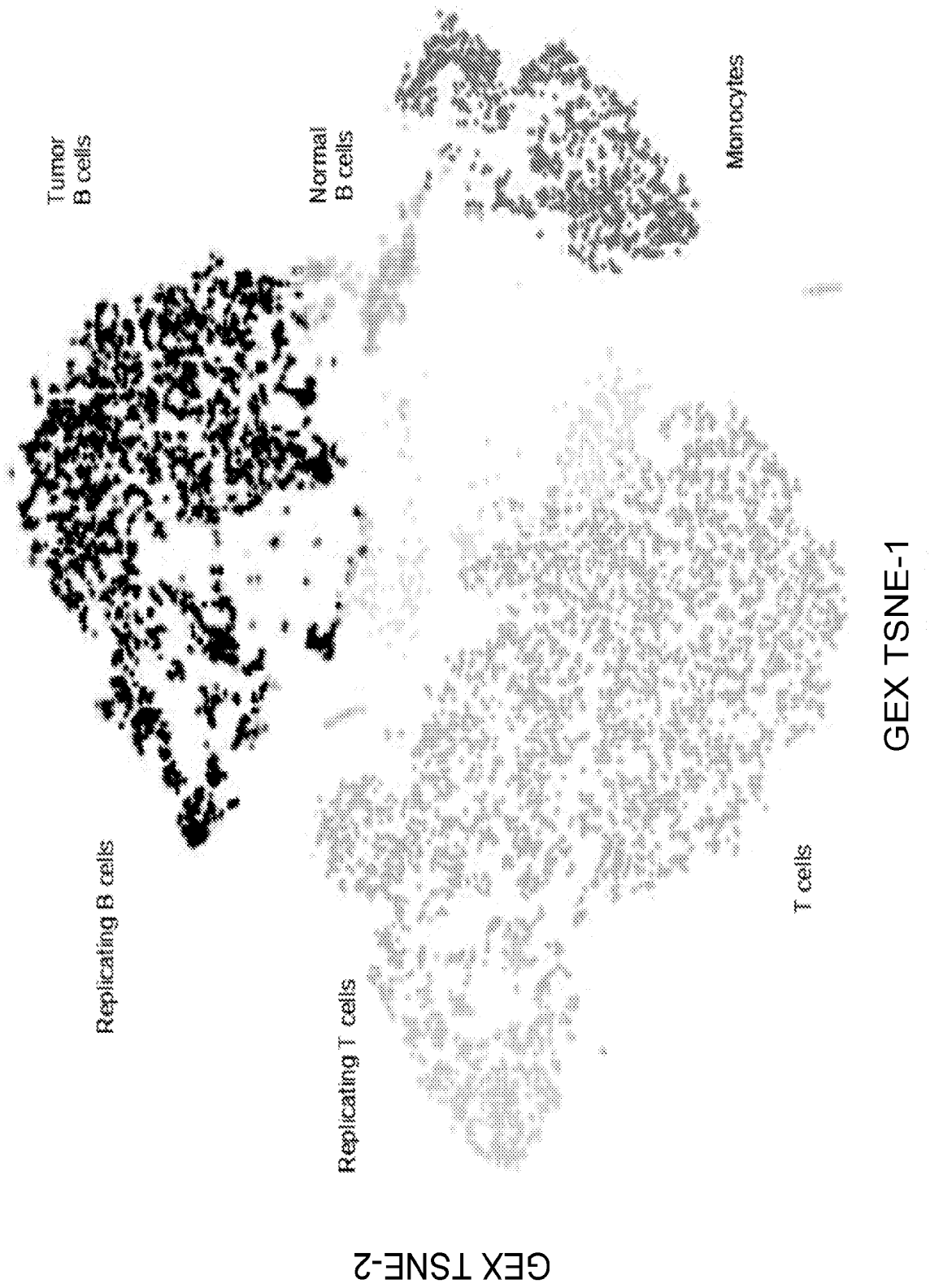


Figure 33B

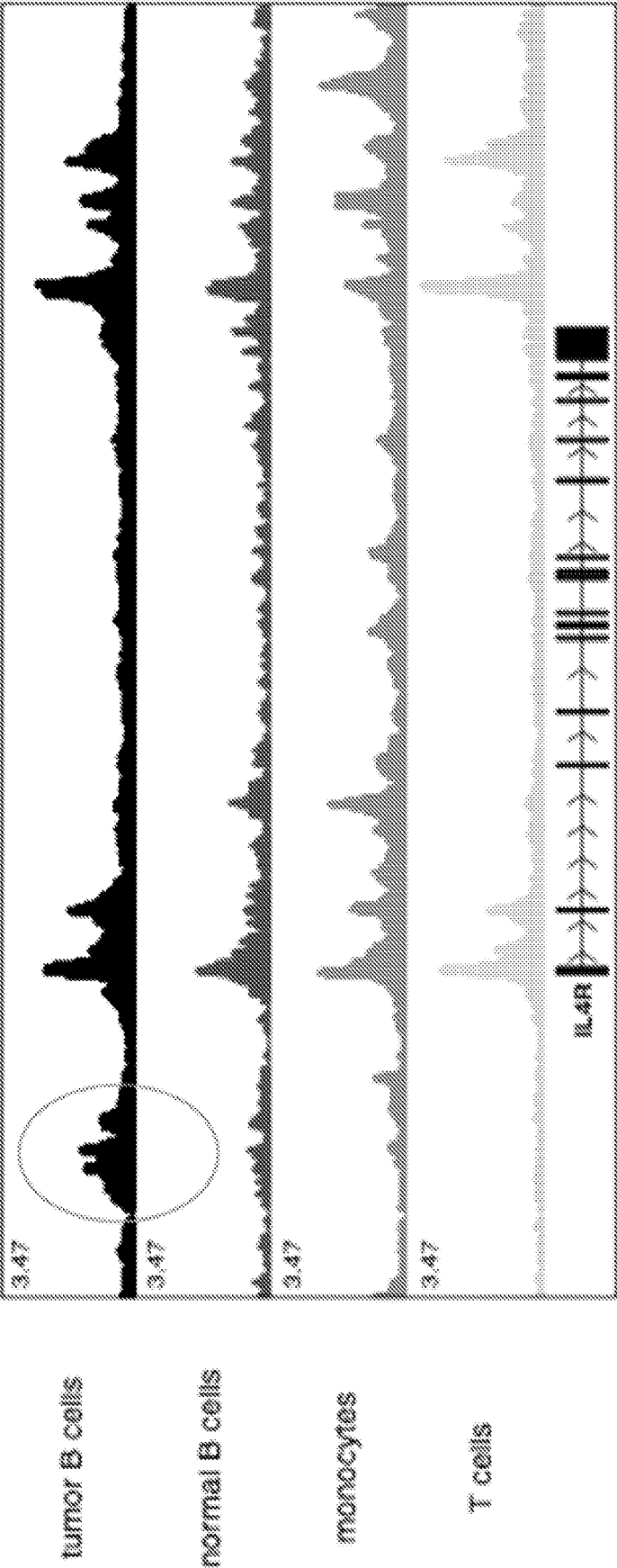
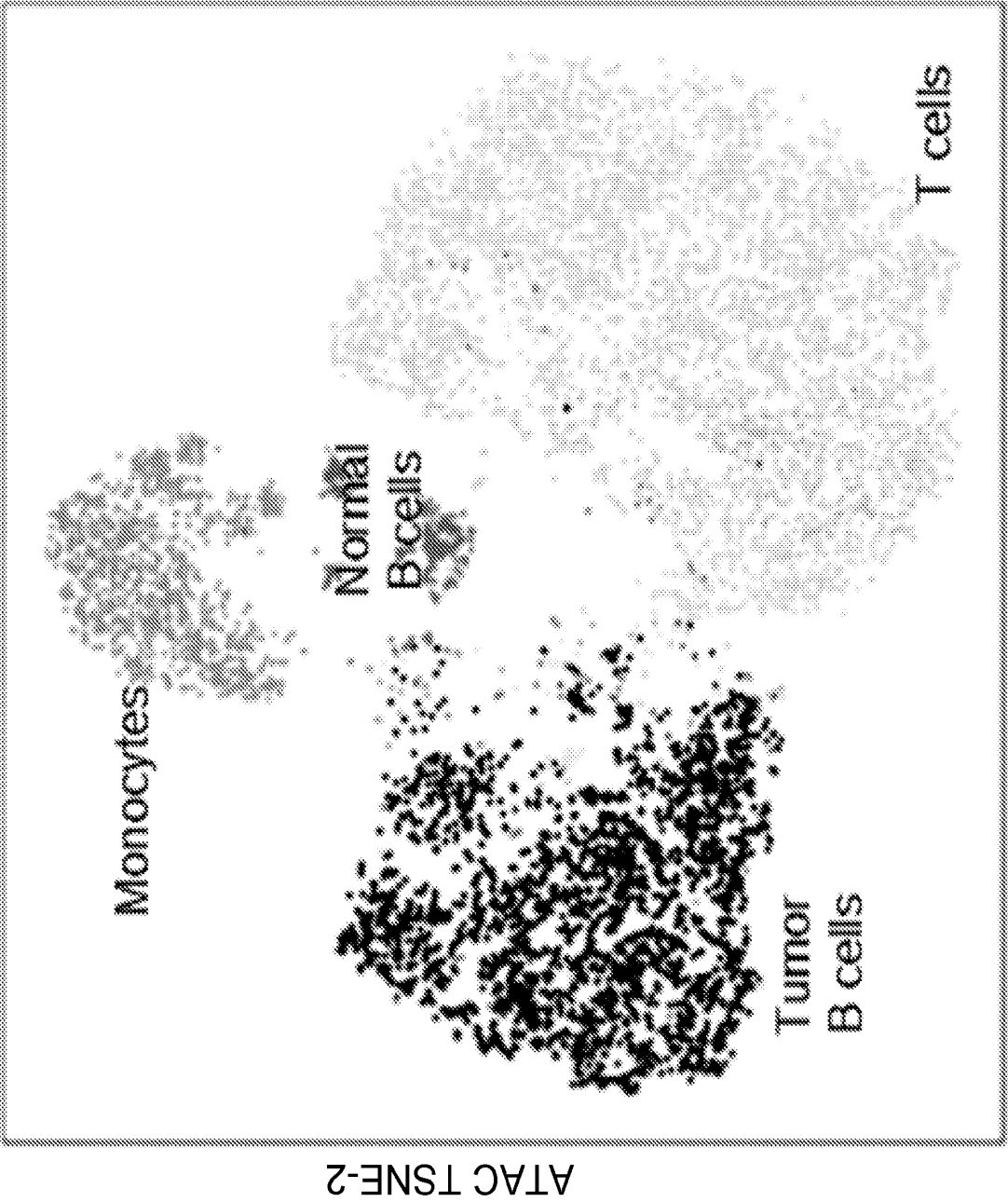


Figure 33C



ATAC TSNE-1

Figure 33D

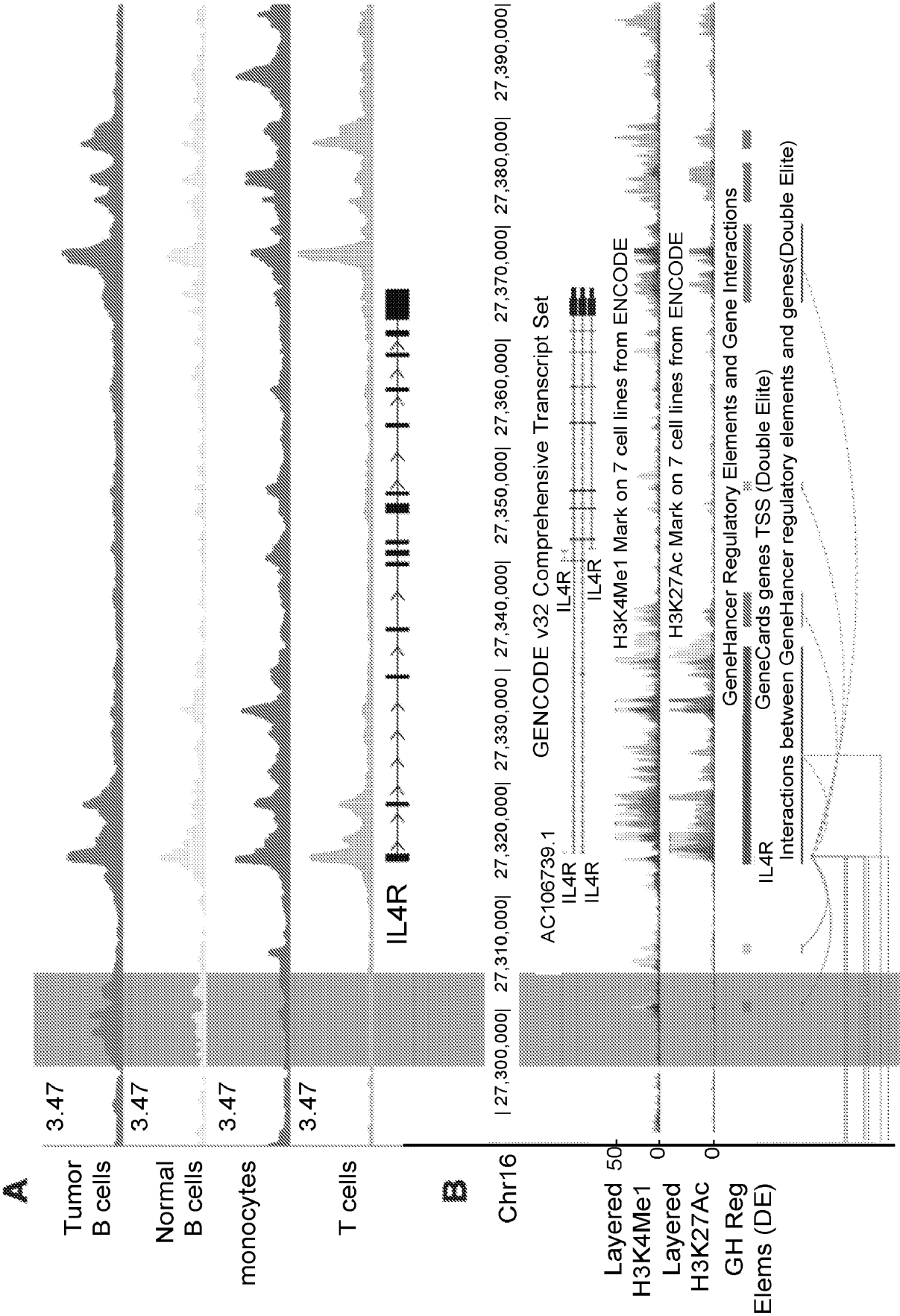


Figure 34

C

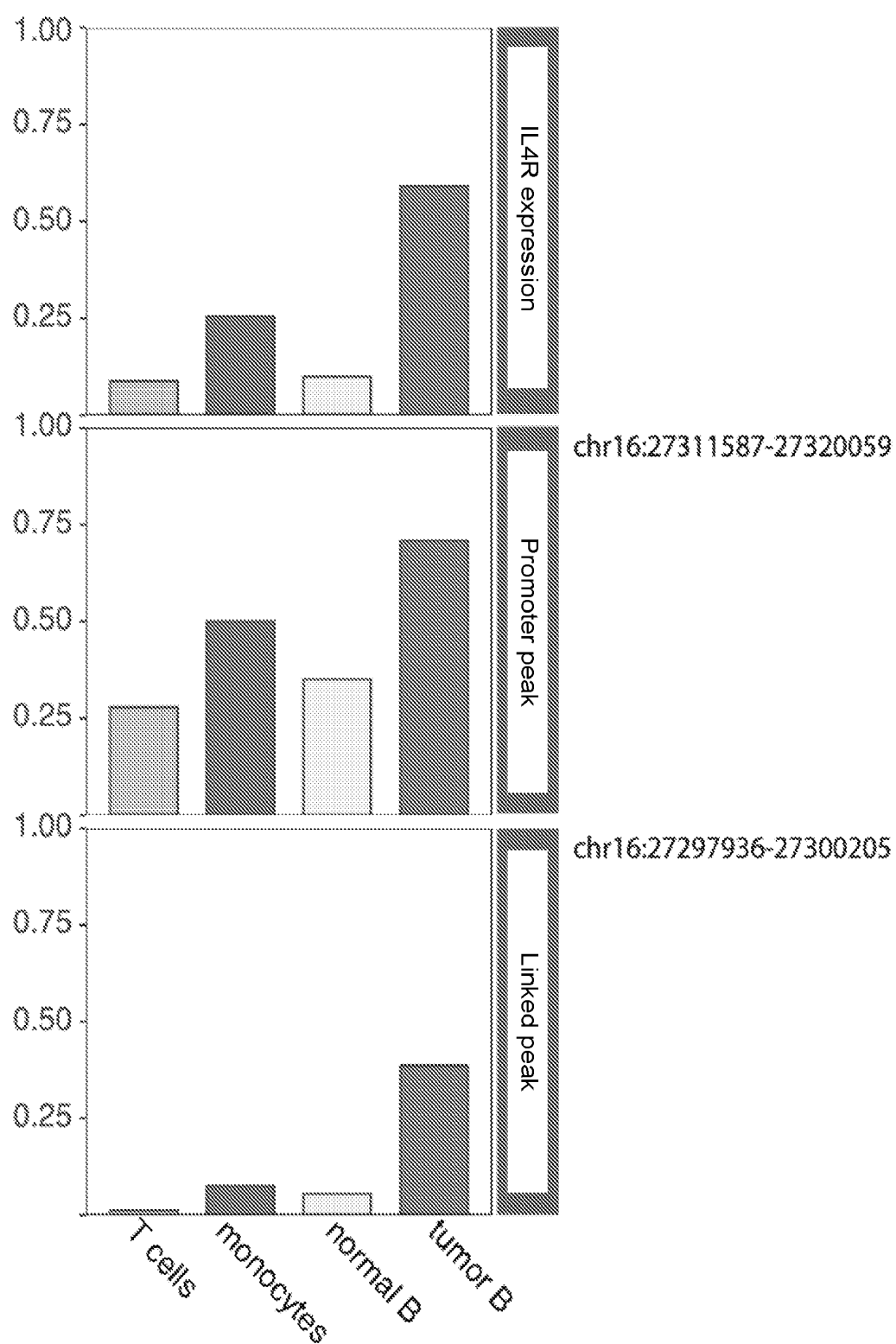


Figure 34C
SUBSTITUTE SHEET (RULE 26)