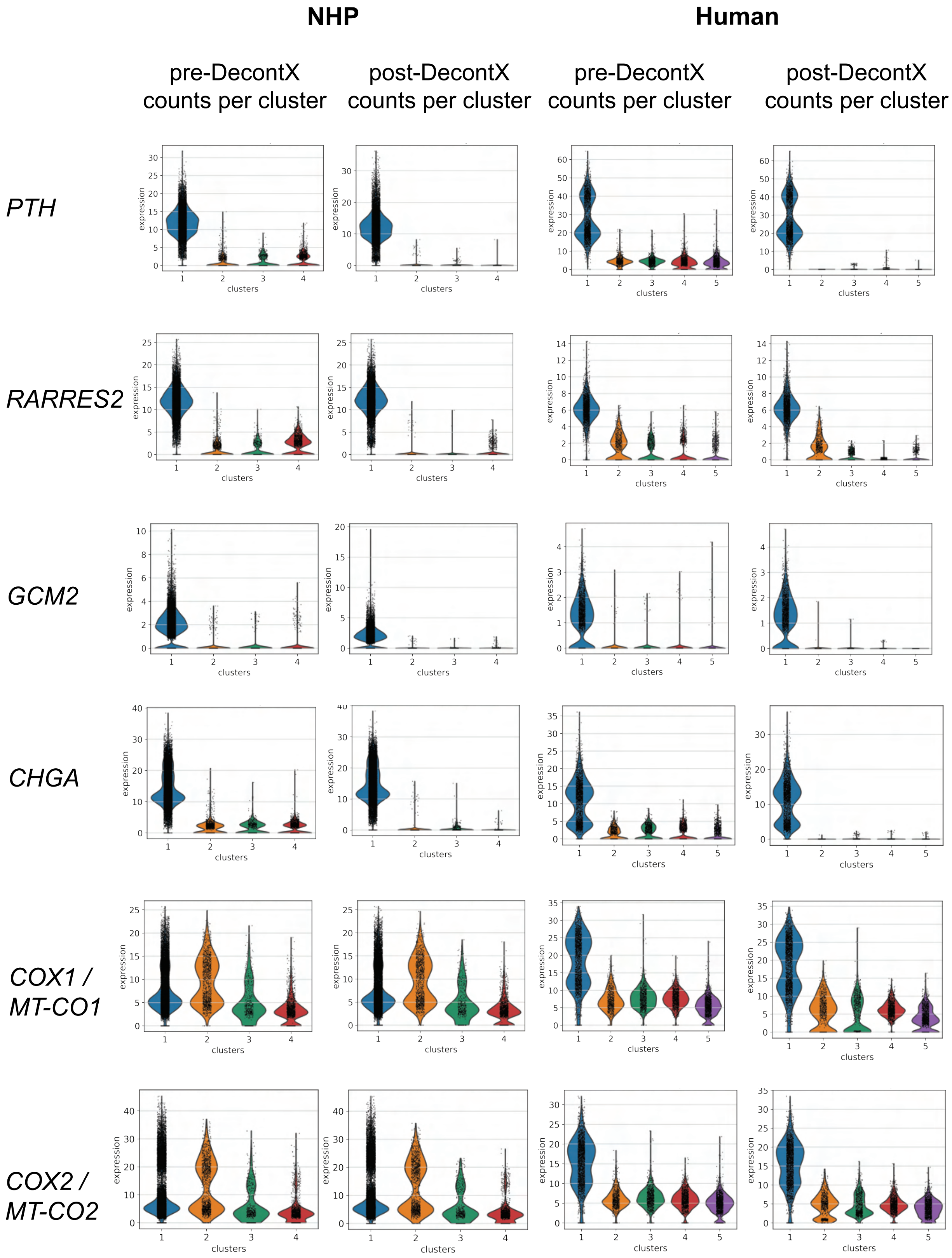
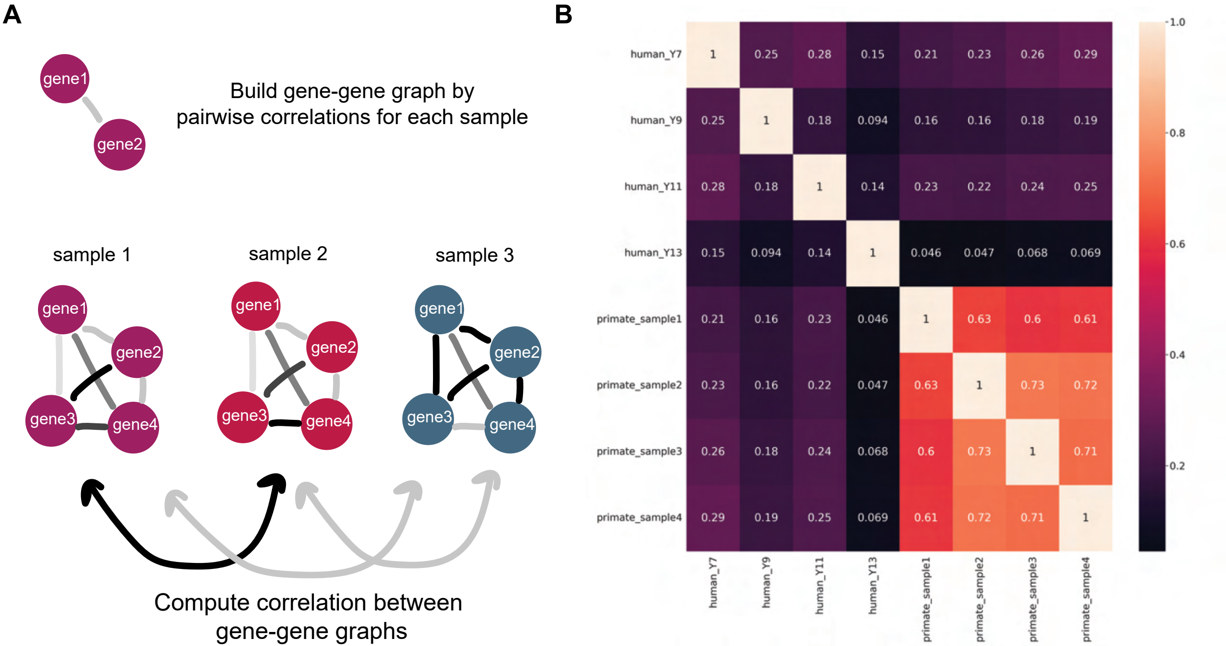

NHP
Human
pre-DecontX counts per cluster
post-DecontX counts per cluster
pre-DecontX counts per cluster
post-DecontX counts per cluster
PTH
RARRES2
GCM2
CHGA
COX1 / MT-CO1
COX2 / MT-CO2
expression
clusters
1
2
3
4
5

Figure S1: DecontX removes ambient transcripts from cell suspensions. Counts of UMIs per cluster are shown for parathyroid-relevant genes for primate and human samples, pre- and post-processing with DecontX. Low level *PTH* transcripts are observed in all cell clusters before processing, despite clear delineation of cell types in each cluster. Post-processing counts of parathyroid-relevant genes are preserved post-processing.



	human_Y7	human_Y9	human_Y11	human_Y13	primate_sample1	primate_sample2	primate_sample3	primate_sample4
human_Y7	1	0.25	0.28	0.15	0.21	0.23	0.26	0.29
human_Y9	0.25	1	0.18	0.094	0.16	0.16	0.18	0.19
human_Y11	0.28	0.18	1	0.14	0.23	0.22	0.24	0.25
human_Y13	0.15	0.094	0.14	1	0.046	0.047	0.068	0.069
primate_sample1	0.21	0.16	0.23	0.046	1	0.63	0.6	0.61
primate_sample2	0.23	0.16	0.22	0.047	0.63	1	0.73	0.72
primate_sample3	0.26	0.18	0.24	0.068	0.6	0.73	1	0.71
primate_sample4	0.29	0.19	0.25	0.069	0.61	0.72	0.71	1

Figure S2: Scoring gene correlations between samples. (a) Schematic of how pairwise gene correlations were calculated. (b) Gene-gene correlations computed per replicate based on denoised and pseudobulked measurements, then compared via correlation of gene-gene graphs and visualized.

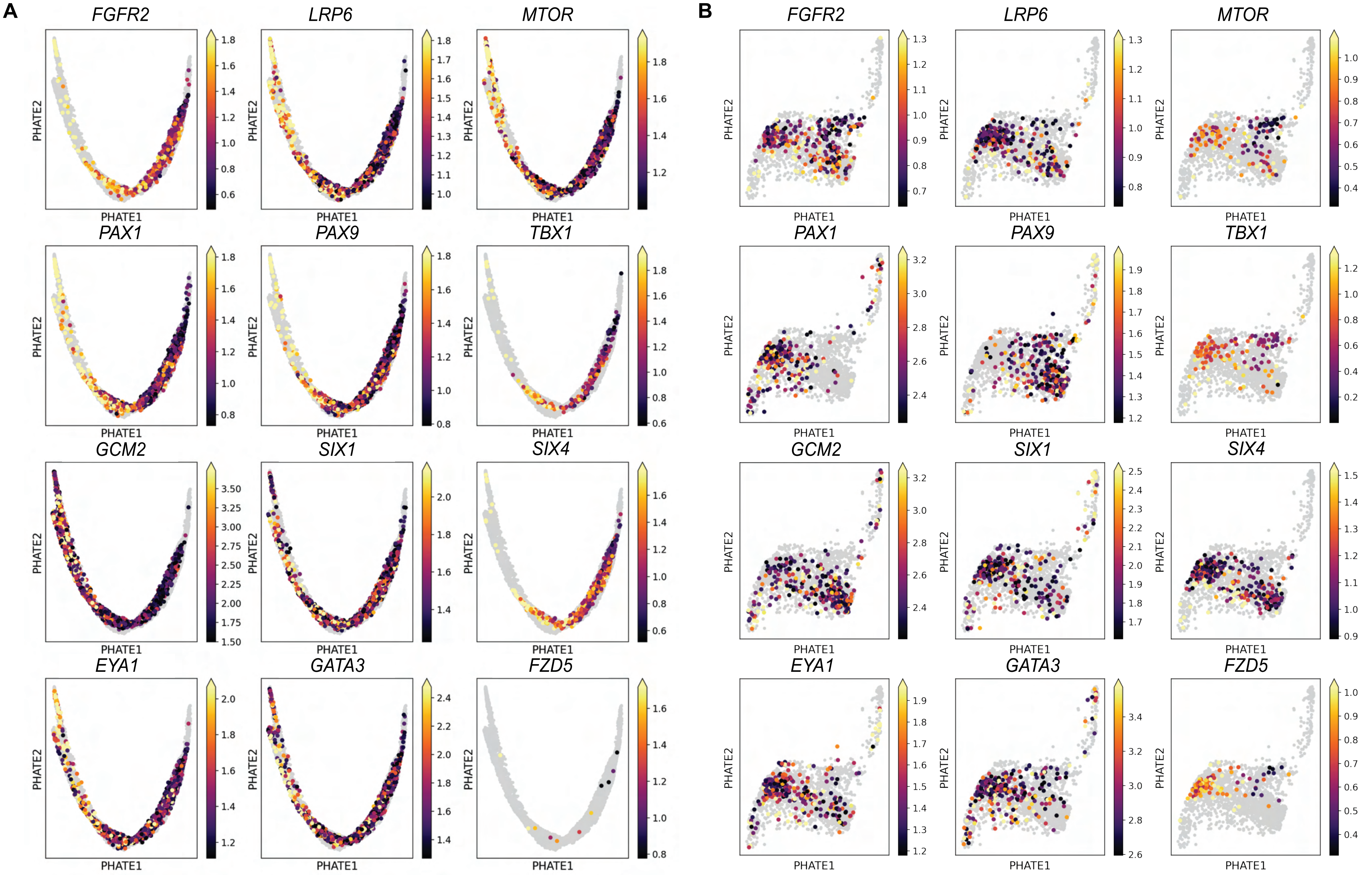

A
FGFR2
1.8 1.6 1.4 1.2 1.0 0.8 0.6
LRP6
1.8 1.7 1.6 1.5 1.4 1.3 1.2 1.1 1.0
MTOR
1.8 1.6 1.4 1.2
PAX1
1.8 1.6 1.4 1.2 1.0 0.8
PAX9
1.8 1.6 1.4 1.2 1.0 0.8
TBX1
1.8 1.6 1.4 1.2 1.0 0.8 0.6
GCM2
3.50 3.25 3.00 2.75 2.50 2.25 2.00 1.75 1.50
SIX1
2.0 1.8 1.6 1.4
SIX4
1.6 1.4 1.2 1.0 0.8 0.6
EYA1
2.0 1.8 1.6 1.4 1.2
GATA3
2.4 2.2 2.0 1.8 1.6 1.4
FZD5
1.6 1.4 1.2 1.0 0.8
B
FGFR2
1.3 1.2 1.1 1.0 0.9 0.8 0.7
LRP6
1.3 1.2 1.1 1.0 0.9 0.8
MTOR
1.0 0.9 0.8 0.7 0.6 0.5 0.4
PAX1
3.2 3.0 2.8 2.6 2.4
PAX9
1.9 1.8 1.7 1.6 1.5 1.4 1.3 1.2
TBX1
1.2 1.0 0.8 0.6 0.4 0.2
GCM2
3.2 3.0 2.8 2.6 2.4
SIX1
2.5 2.4 2.3 2.2 2.1 2.0 1.9 1.8 1.7
SIX4
1.5 1.4 1.3 1.2 1.1 1.0 0.9
EYA1
1.9 1.8 1.7 1.6 1.5 1.4 1.3 1.2
GATA3
3.4 3.2 3.0 2.8 2.6
FZD5
1.0 0.9 0.8 0.7 0.6 0.5 0.4
PHATE1
PHATE2

Figure S3: Genes implicated in parathyroid development enriched in $GCM2^{hi}PTH^{lo}$ in NHP samples only. The top 10% of cells expressing developmentally relevant genes in (a) NHP cells and (b) human cells.

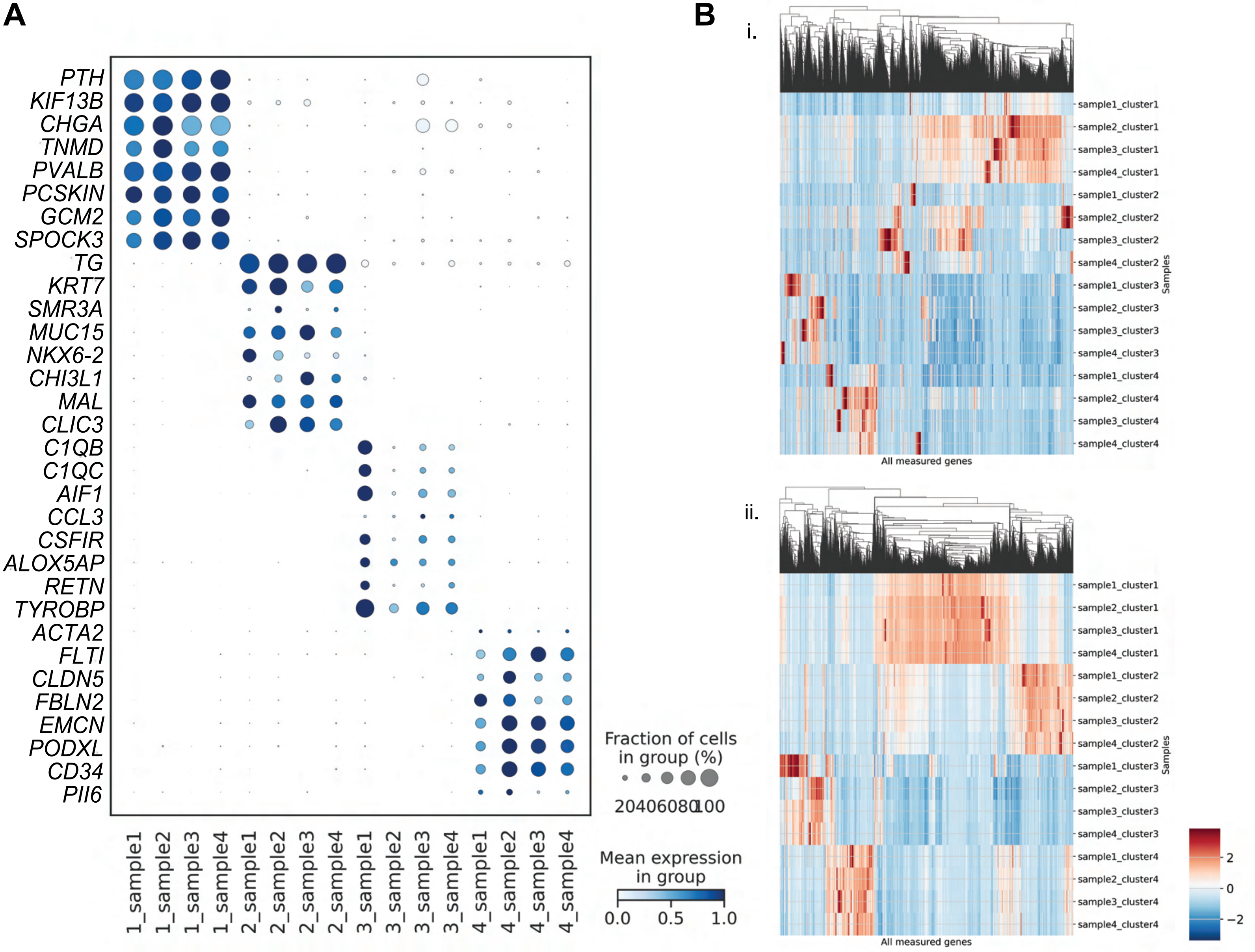

A
PTH
KIF13B
CHGA
TNMD
PVALB
PCSKIN
GCM2
SPOCK3
TG
KRT7
SMR3A
MUC15
NKX6-2
CHI3L1
MAL
CLIC3
C1QB
C1QC
AIF1
CCL3
CSFIR
ALOX5AP
RETN
TYROBP
ACTA2
FLTI
CLDN5
FBLN2
EMCN
PODXL
CD34
PII6
1_sample1
1_sample2
1_sample3
1_sample4
2_sample1
2_sample2
2_sample3
2_sample4
3_sample1
3_sample2
3_sample3
3_sample4
4_sample1
4_sample2
4_sample3
4_sample4
Fraction of cells in group (%)
20 40 60 80 100
Mean expression in group
0.0 0.5 1.0
B
i.
sample1_cluster1
sample2_cluster1
sample3_cluster1
sample4_cluster1
sample1_cluster2
sample2_cluster2
sample3_cluster2
sample4_cluster2
sample1_cluster3
sample2_cluster3
sample3_cluster3
sample4_cluster3
sample1_cluster4
sample2_cluster4
sample3_cluster4
sample4_cluster4
Samples
All measured genes
ii.
sample1_cluster1
sample2_cluster1
sample3_cluster1
sample4_cluster1
sample1_cluster2
sample2_cluster2
sample3_cluster2
sample4_cluster2
sample1_cluster3
sample2_cluster3
sample3_cluster3
sample4_cluster3
sample1_cluster4
sample2_cluster4
sample3_cluster4
sample4_cluster4
Samples
All measured genes
2
0
−2

Figure S4: Comparison of cell type reproducibility across NHP samples. (a) Visualization of top differentially expressed genes per cluster, resolved per replicate for NHP samples. (b) Genome-scale heatmaps (i) before batch correction and (ii) after batch correction.

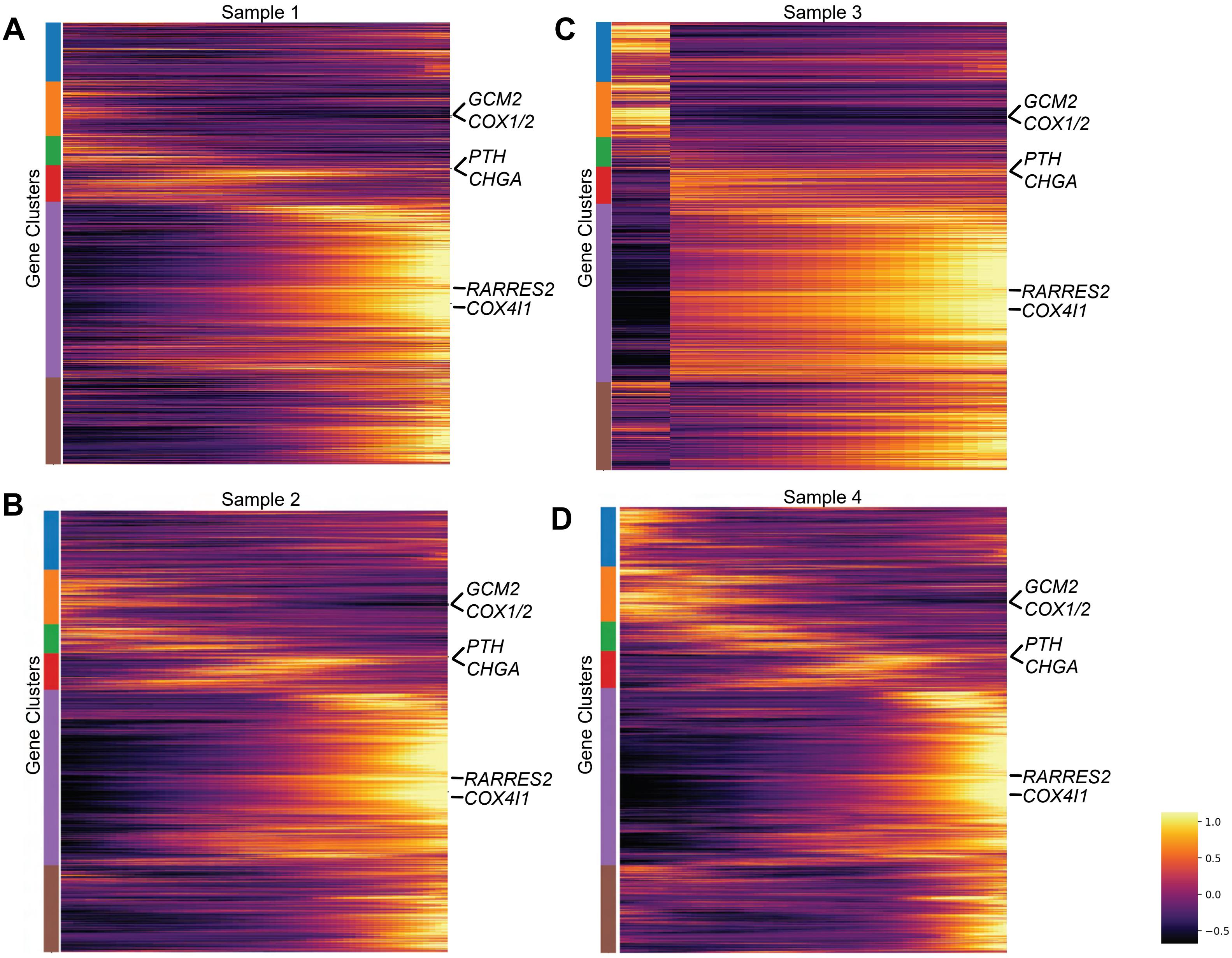

A
Sample 1
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
B
Sample 2
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
C
Sample 3
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
D
Sample 4
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
1.0
0.5
0.0
−0.5

Figure S5: Comparison of gene trend reproducibility across NHP samples. Visualization of gene expression dynamics across time per sample (a) Sample 1 (b) Sample 2 (c) Sample 3 (d) Sample 4, ordered by gene heatmap in Figure 3.

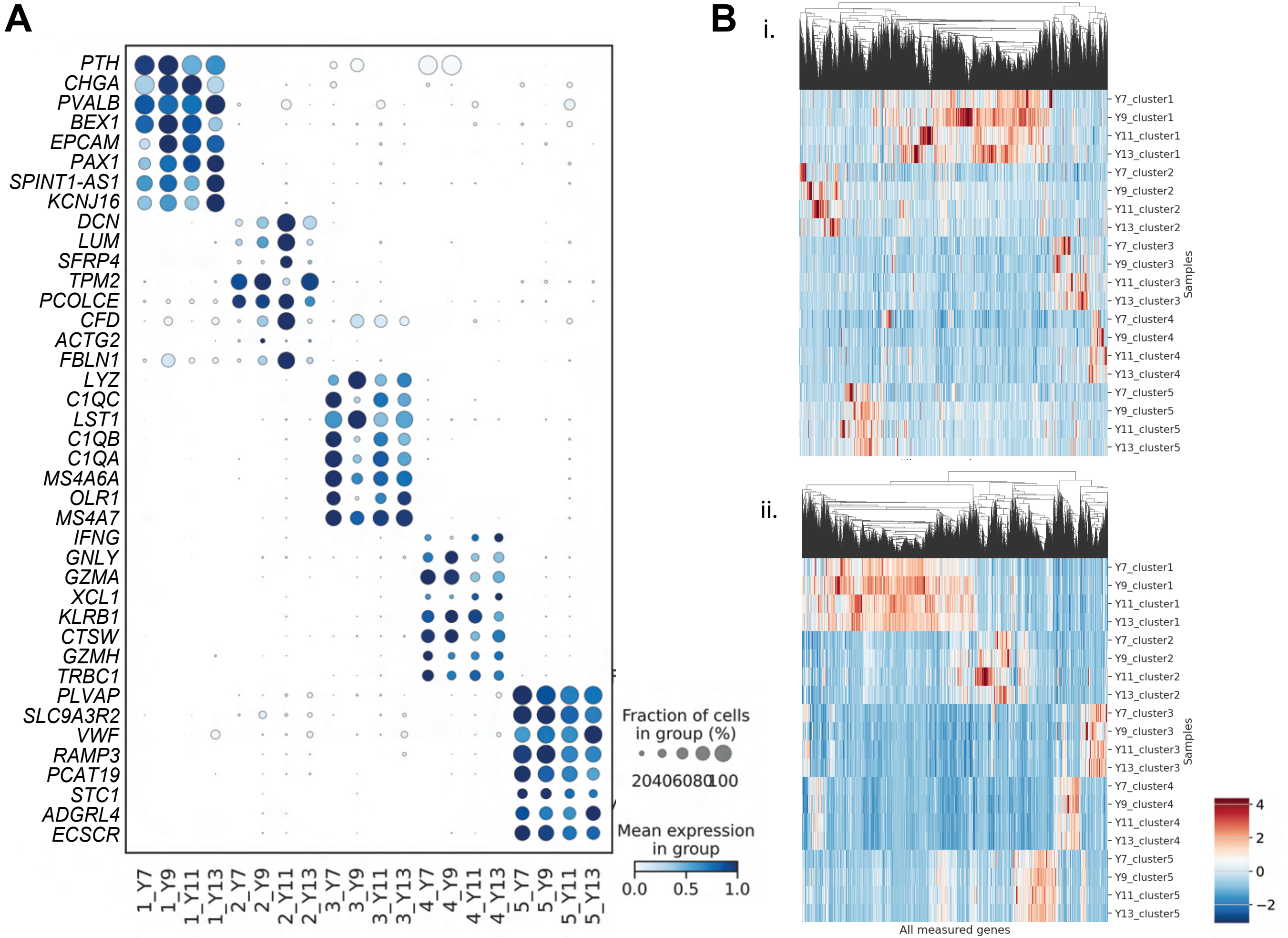

A
PTH
CHGA
PVALB
BEX1
EPCAM
PAX1
SPINT1-AS1
KCNJ16
DCN
LUM
SFRP4
TPM2
PCOLCE
CFD
ACTG2
FBLN1
LYZ
C1QC
LST1
C1QB
C1QA
MS4A6A
OLR1
MS4A7
IFNG
GNLY
GZMA
XCL1
KLRB1
CTSW
GZMH
TRBC1
PLVAP
SLC9A3R2
VWF
RAMP3
PCAT19
STC1
ADGRL4
ECSCR
1_Y7
1_Y9
1_Y11
1_Y13
2_Y7
2_Y9
2_Y11
2_Y13
3_Y7
3_Y9
3_Y11
3_Y13
4_Y7
4_Y9
4_Y11
4_Y13
5_Y7
5_Y9
5_Y11
5_Y13
Fraction of cells in group (%)
20406080100
Mean expression in group
0.0
0.5
1.0
B
i.
Y7_cluster1
Y9_cluster1
Y11_cluster1
Y13_cluster1
Y7_cluster2
Y9_cluster2
Y11_cluster2
Y13_cluster2
Y7_cluster3
Y9_cluster3
Y11_cluster3
Y13_cluster3
Y7_cluster4
Y9_cluster4
Y11_cluster4
Y13_cluster4
Y7_cluster5
Y9_cluster5
Y11_cluster5
Y13_cluster5
Samples
ii.
Y7_cluster1
Y9_cluster1
Y11_cluster1
Y13_cluster1
Y7_cluster2
Y9_cluster2
Y11_cluster2
Y13_cluster2
Y7_cluster3
Y9_cluster3
Y11_cluster3
Y13_cluster3
Y7_cluster4
Y9_cluster4
Y11_cluster4
Y13_cluster4
Y7_cluster5
Y9_cluster5
Y11_cluster5
Y13_cluster5
Samples
All measured genes
4
2
0
−2

Figure S6: Comparison of cell type reproducibility across human samples. (a) Visualization of top differentially expressed genes per cluster, resolved per sample for human samples. (b) Genome-scale heatmap for human samples.

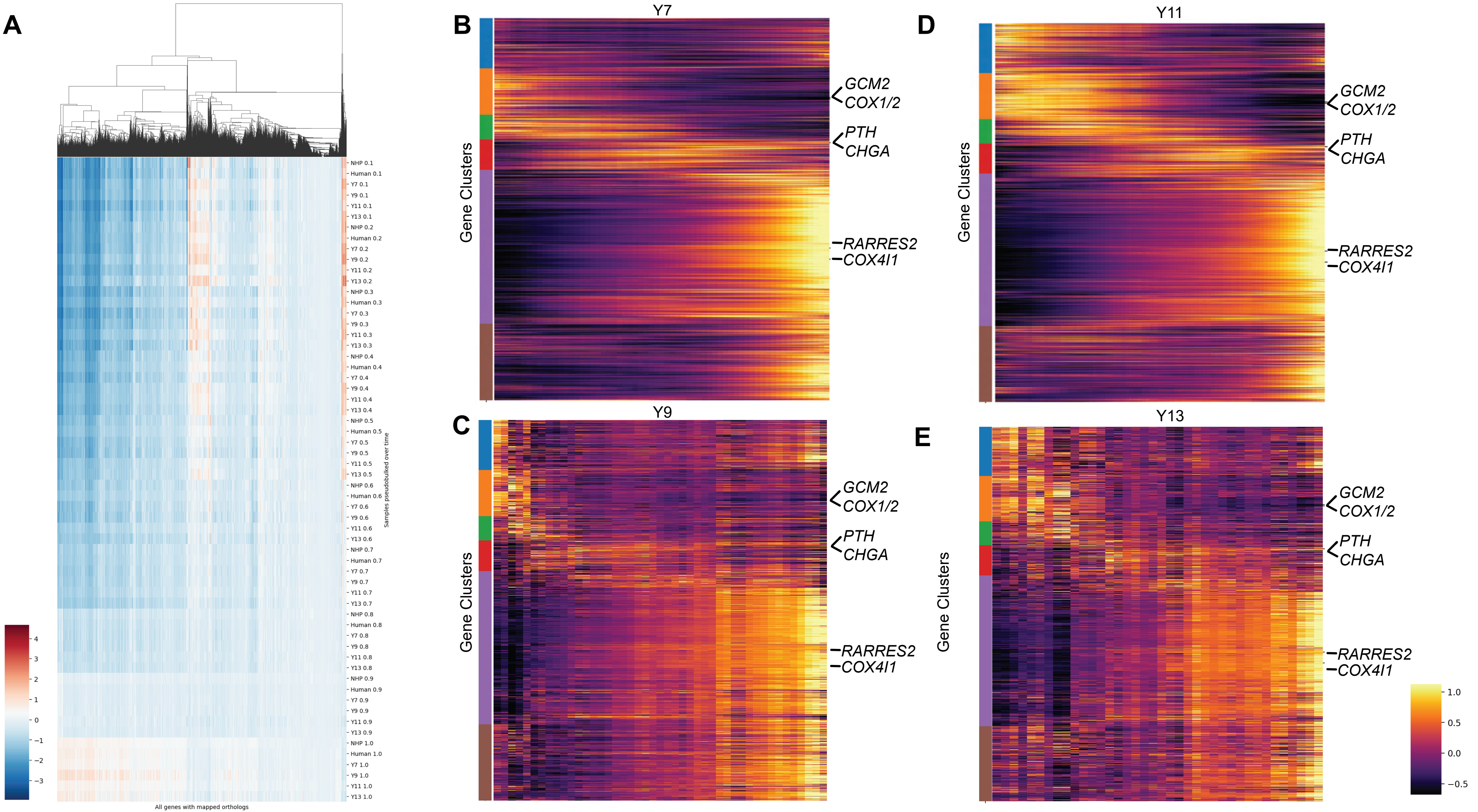

A
NHP 0.1
Human 0.1
Y7 0.1
Y9 0.1
Y11 0.1
Y13 0.1
NHP 0.2
Human 0.2
Y7 0.2
Y9 0.2
Y11 0.2
Y13 0.2
NHP 0.3
Human 0.3
Y7 0.3
Y9 0.3
Y11 0.3
Y13 0.3
NHP 0.4
Human 0.4
Y7 0.4
Y9 0.4
Y11 0.4
Y13 0.4
NHP 0.5
Human 0.5
Y7 0.5
Y9 0.5
Y11 0.5
Y13 0.5
NHP 0.6
Human 0.6
Y7 0.6
Y9 0.6
Y11 0.6
Y13 0.6
NHP 0.7
Human 0.7
Y7 0.7
Y9 0.7
Y11 0.7
Y13 0.7
NHP 0.8
Human 0.8
Y7 0.8
Y9 0.8
Y11 0.8
Y13 0.8
NHP 0.9
Human 0.9
Y7 0.9
Y9 0.9
Y11 0.9
Y13 0.9
NHP 1.0
Human 1.0
Y7 1.0
Y9 1.0
Y11 1.0
Y13 1.0
Samples pseudobulked over time
All genes with mapped orthologs
4
3
2
1
0
−1
−2
−3
B
Y7
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
C
Y9
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
D
Y11
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
E
Y13
Gene Clusters
GCM2
COX1/2
PTH
CHGA
RARRES2
COX4I1
1.0
0.5
0.0
−0.5

Figure S7: Comparison of NHP samples and mapped human samples. (a) Genome-scale heatmap of NHP samples and mapped human samples, pseudobulked over time. Visualization of gene expression dynamics per sample (b) Y7 (c) Y9 (d) Y11 (e) Y13.

CycleGAN CV	Total Corr.	Y7 Corr.	Y9 Corr.	Y11 Corr.	Y13 Corr.
0	-0.066	-0.071	-0.211	-0.042	-0.159
1	0.000	0.029	0.135	-0.086	-0.012
2	-0.003	0.028	-0.013	-0.006	-0.150
3	0.031	0.049	0.003	0.039	0.046
4	-0.161	-0.157	-0.334	-0.167	-0.226
5	0.068	0.131	0.092	0.057	0.057
6	-0.121	-0.156	-0.208	-0.105	-0.131
7	-0.029	-0.008	-0.158	0.004	-0.197
8	0.067	0.082	0.164	0.024	-0.054
9	0.035	0.034	0.045	0.066	0.007
mean ± std	-0.018 ± 0.077	-0.004 ± 0.096	-0.049 ± 0.169	-0.022 ± 0.077	-0.082 ± 0.103

scMMGAN CV	Total Corr.	Y7 Corr.	Y9 Corr.	Y11 Corr.	Y13 Corr.
0	0.173	0.210	0.328	0.167	0.264
1	0.212	0.235	0.385	0.176	0.229
2	0.194	0.203	0.294	0.170	0.275
3	0.173	0.232	0.333	0.144	0.179
4	0.185	0.202	0.406	0.160	0.165
5	0.155	0.205	0.287	0.155	0.064
6	0.169	0.175	0.263	0.145	0.148
7	0.173	0.217	0.382	0.153	0.238
8	0.175	0.192	0.300	0.156	0.171
9	0.170	0.182	0.285	0.166	0.122
mean ± std	**0.178** ± 0.016	**0.205** ± 0.019	**0.326** ± 0.049	**0.159** ± 0.011	**0.185** ± 0.066

Figure S8: Ten-fold cross-validation (CV) for CycleGAN (top) versus scMMGAN (bottom). CV computed for gene correlation between input and mapped human samples across all cells (Total Corr.) and resolved per human sample.

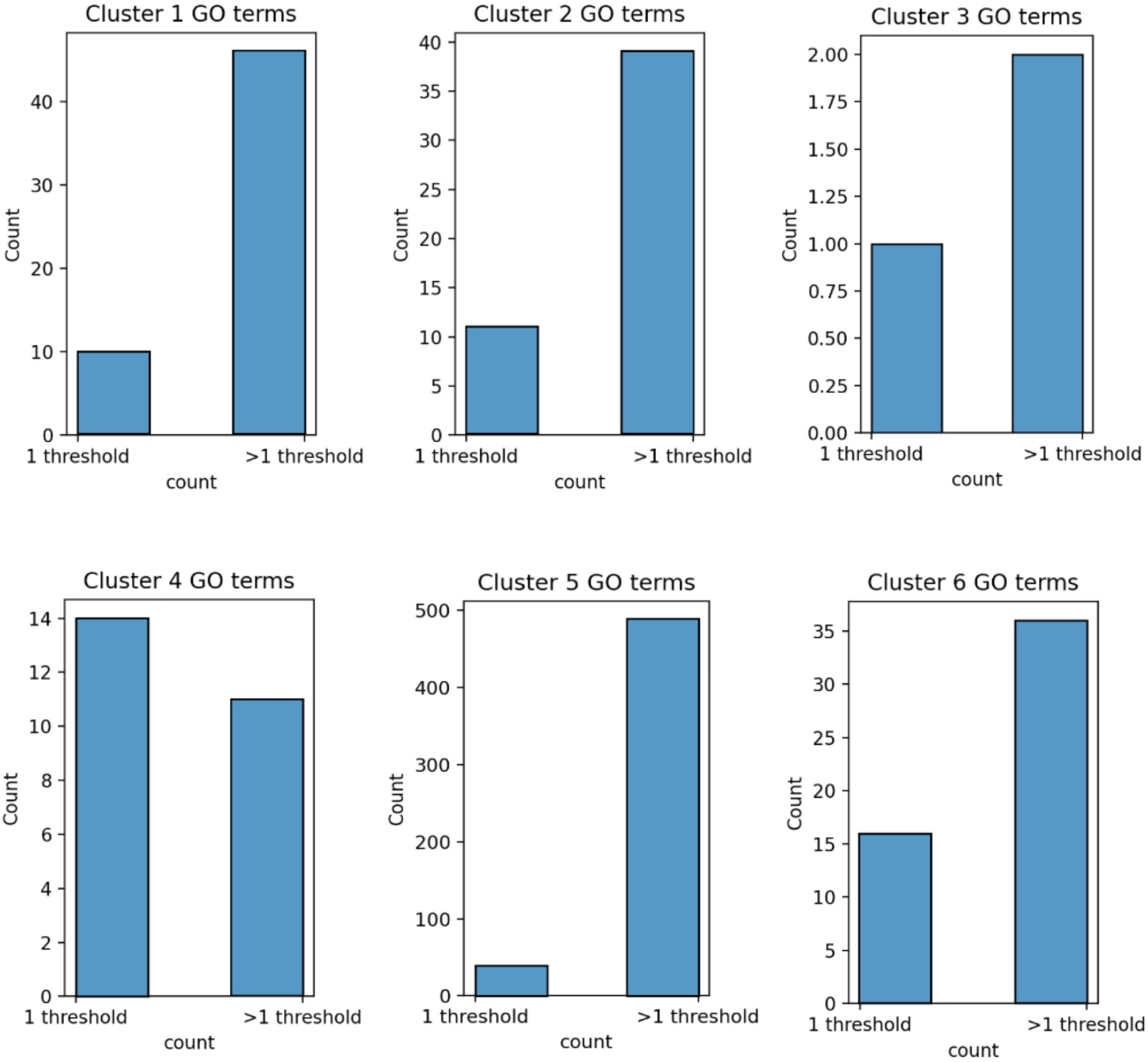

Cluster 1 GO terms
Count
0
10
20
30
40
1 threshold
>1 threshold
count
Cluster 2 GO terms
Count
0
5
10
15
20
25
30
35
40
1 threshold
>1 threshold
count
Cluster 3 GO terms
Count
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
1 threshold
>1 threshold
count
Cluster 4 GO terms
Count
0
2
4
6
8
10
12
14
1 threshold
>1 threshold
count
Cluster 5 GO terms
Count
0
100
200
300
400
500
1 threshold
>1 threshold
count
Cluster 6 GO terms
Count
0
5
10
15
20
25
30
35
1 threshold
>1 threshold
count

Figure S9: Number of GO terms enriched when using one threshold cutoff for variable genes ($p < 0.05$) compare to more than one highly variable gene cutoff value (in this case 0.01, 0.025, 0.075, and 0.1 were used).

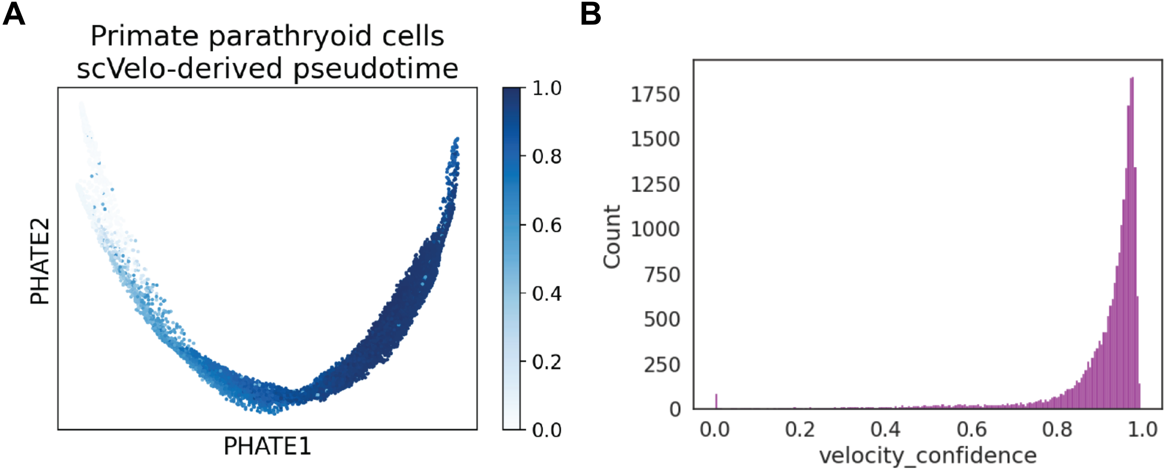

A
Primate parathryoid cells
scVelo-derived pseudotime
PHATE2
PHATE1
1.0
0.8
0.6
0.4
0.2
0.0
B
Count
1750
1500
1250
1000
750
500
250
0
0.0
0.2
0.4
0.6
0.8
1.0
velocity_confidence

Figure S10: scVelo analysis of time axis. (a) Pseudotime axis computed by scVelo (b) distribution of velocity estimation confidence across all cells.

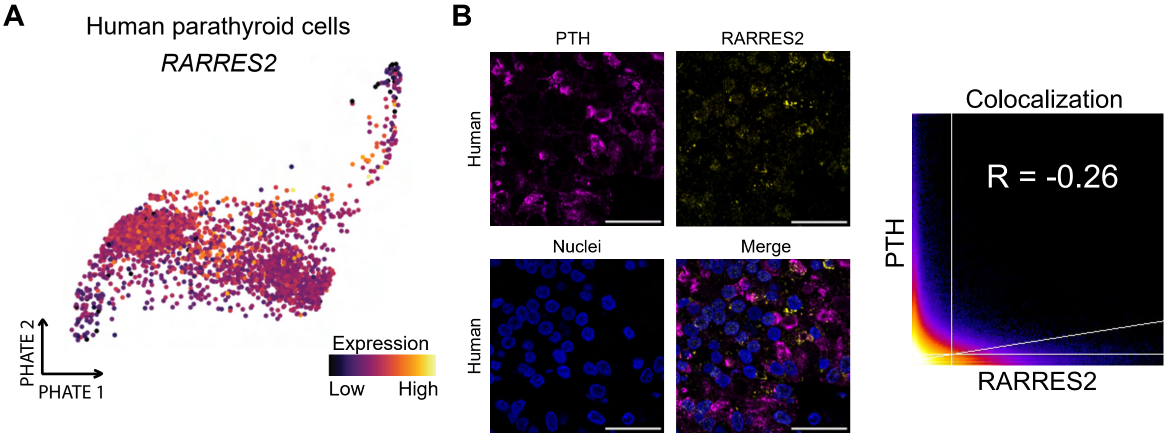

A
Human parathyroid cells
RARRES2
PHATE 2
PHATE 1
Expression
Low
High
B
PTH
RARRES2
Human
Nuclei
Merge
Human
Colocalization
R = -0.26
PTH
RARRES2

Figure S11: (A) Relative RNA transcript distribution of *RARRES2* in human hyperfunctioning parathyroid cells across scRNA-seq datasets. (B) Confocal fluorescence image of immunolabelled human hyperfunctioning parathyroid tissue and pixel-wise colocalization heatmap. Vertical and horizontal lines indicate thresholds for colocalization, and the diagonal line is the linear best-fit (R = -0.26) of pixels above the thresholds (Scale bars = 30 μm).