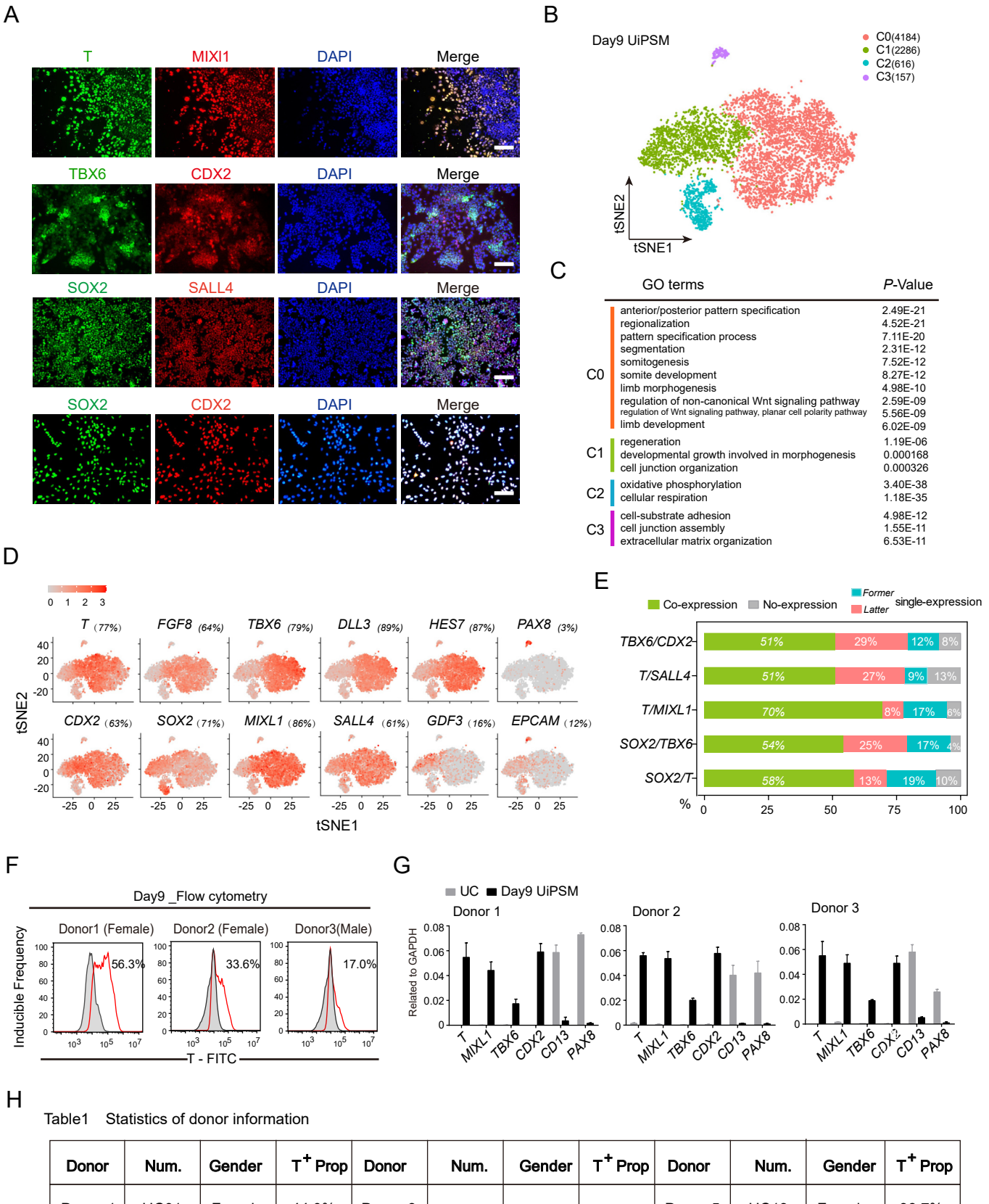
Appendix Table of Content

Appendix Figure S1



H

Table1 Statistics of donor information

Donor	Num.	Gender	T^+ Prop	Donor	Num.	Gender	T^+ Prop	Donor	Num.	Gender	T^+ Prop
Donor 1	UC01	Female	44.0%	Donor 3	UC07	Male	17.0%	Donor 5	UC13	Female	36.7%
Donor 1	UC02	Female	56.3%	Donor 3	UC08	Male	19.3%	Donor 6	UC14	Male	32.2%
Donor 1	UC03	Female	52.8%	Donor 3	UC09	Male	21.6%	Donor 7	UC15	Male	43.3%
Donor 2	UC04	Female	33.6%	Donor 4	UC10	Female	50.5%	Donor 8	UC16	Male	48.8%
Donor 2	UC05	Female	38.5%	Donor 4	UC11	Female	49.4%	Donor 9	UC17	Female	43.9%
Donor 2	UC06	Female	30.3%	Donor 4	UC12	Female	46.7%	Donor 10	UC18	Female	52.7%

Appendix Figure S1. Specification of the Chemical reprogramming from UC to UiPSM.

A. Immunofluorescence picture s of the UiPSM at day 9 (n=3 independent experiments). The scale bar represents 200 μm.

B. t-SNE plot of 6824 UiPSM cells at day 9 colored by clusters.

C. GO enrichment analysis showed the potential functions of 4 groups in (B). P value was less than 0.05.

D. t-SNE projection of cells colored by typical presomitic-related gene expressions at day 9.

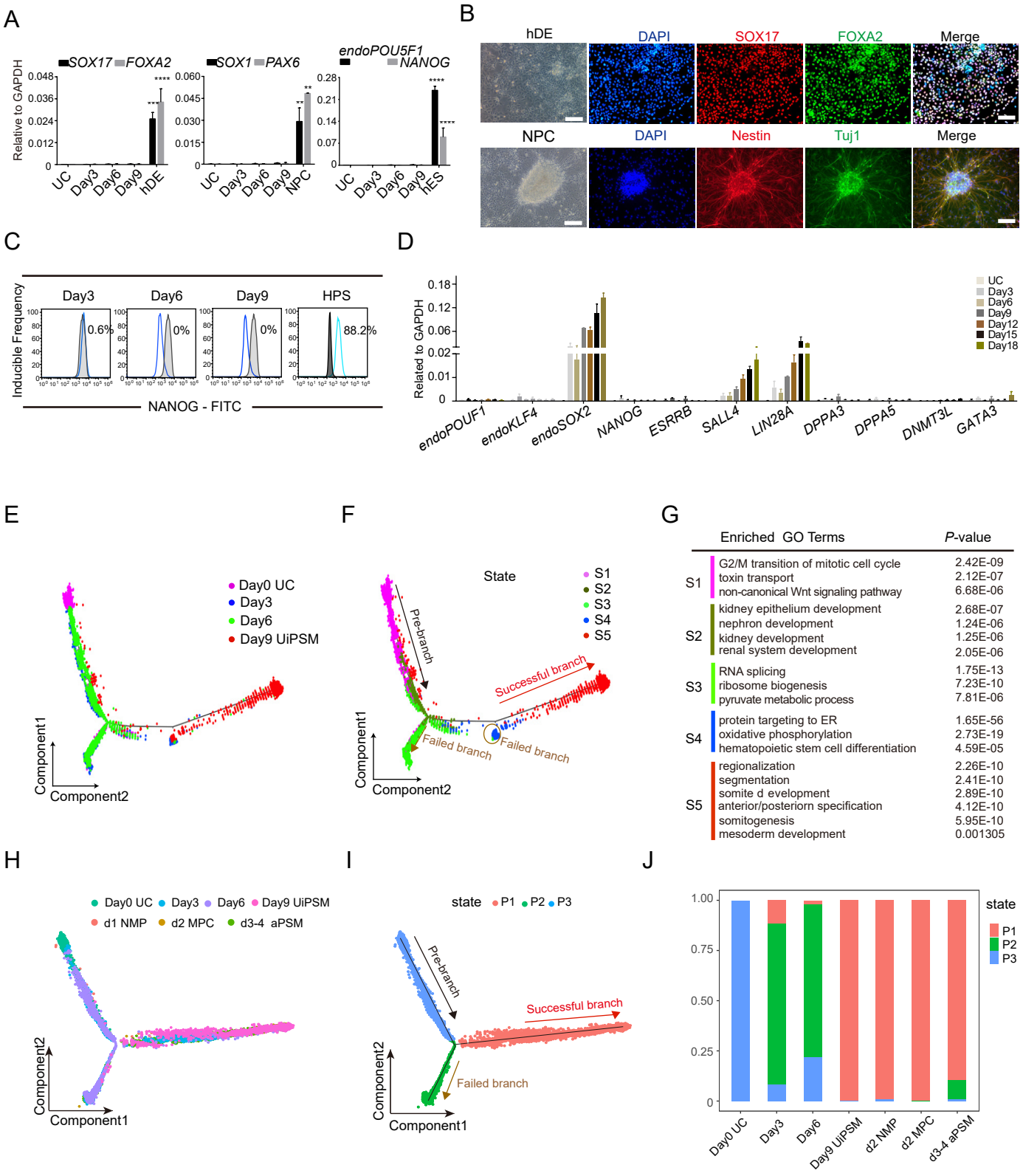
E. The proportion of co-expressed *TBX6* and *CDX2, T* and *SALL4, T* and *MIXL1, SOX2* and *TBX6, SOX2 and T* in UiPSM at day 9. Both expressions greater than 0 SCT value were considered as co-expression (green). Only one expression greater than 0 SCT value in one cell was considered to express (red and blue). Both expressions less than 0.5 SCT value were considered as no-expression (gray).

F. Representative flow cytometric inducible effects on T protein expression of the UiPSM (n=3 independent experiments).

G. qRT-PCR analysis of PSM-specific genes *T, MIXL1, TBX6* and *CDX2* and UC specific genes *CD13* and *PAX8* at day 9 of the UiPSM individually. Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments.

H. The PSM inductive efficiency of 18 urine cell samples from 10 healthy volunteers aged 20-40 years via detecting T positive frequency when reprogramming at day9. T^+Prop, T positive proportion.

Appendix Figure S2



Appendix Figure S2. Characterization of the UiPSM reprogramming process

A. qRT-PCR analysis of layer-associated genes during the UiPSM reprogramming process versus DE (definitive endoderm), NPC (neural progenitor cell) and hESC respectively. The endoderm-associated genes *SOX17* and *FOXA2*; The ectoderm-associated genes *SOX1* and *PAX6*; The pluripotency-associated genes *POU5F1* and *NANOG*; Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments. ***p < 0.001.

B. (Upper) Representative images of hDE, showing the co-expression of SOX17 and FOXA2 scale bar,100μm. (Bottom) Representative images of NPC, showing the co-expression of Nestin and Tuj1, scale bar,100μm. n = 3 independent experiments.

C. Flow cytometric detected the expression of NANOG protein during the UiPSM reprogramming process (n =3 biological experiments).

D. qRT-PCR analysis of pluripotency-associated genes of the UiPSM reprogramming process. Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments.

E. Trajectory reconstructed from all single-cell transcriptomic data throughout the UiPSM reprogramming process by Monocle2. The whole process including all 32692 individual cells at day0, day3, day6, day9, were colored by indicated time points.

F. Cell trajectory plot revealing four branches: pre-branch, two failed branches and a successful branch. The plot was colored by cell states defined by Monocle2.

G. Gene ontology analyses of each cell state cluster in ‘E’. P value was less than 0.05.

H. Trajectory reconstructed from all single-cell transcriptomic data throughout the UiPSM reprogramming process by Monocle2, combined with the published data (d1NMP, d2 MPC and d3-4 aPSM).

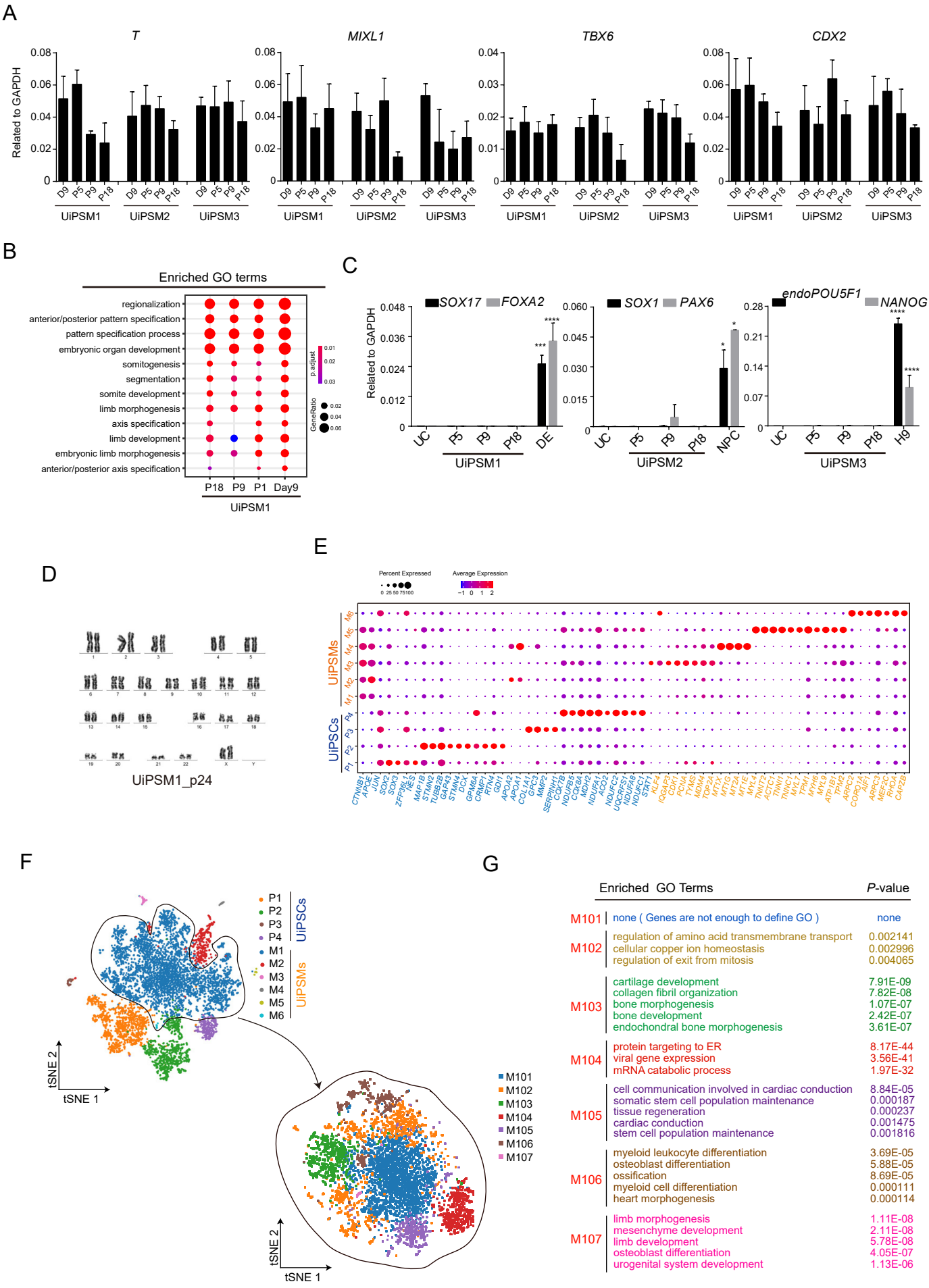
I. Cell trajectory plot revealing three branches in ‘H’: pre-branch, failed branches and a successful branch. The plot was colored by cell states defined by Monocle2.

J. The percentage change of all samples of ‘H’ in each state.

K. Typical presomitic marker genes expression on cell trajectory plot in ‘H’.

L. Gene ontology analyses of each cell state cluster in ‘I’. P value was less than 0.05.

Appendix Figure S3



	Enriched GO Terms	*P*-value
M101	none (Genes are not enough to define GO)	none
M102	regulation of amino acid transmembrane transport	0.002141
	cellular copper ion homeostasis	0.002996
	regulation of exit from mitosis	0.004065
M103	cartilage development	7.91E-09
	collagen fibril organization	7.82E-08
	bone morphogenesis	1.07E-07
	bone development	2.42E-07
	endochondral bone morphogenesis	3.61E-07
M104	protein targeting to ER	8.17E-44
	viral gene expression	3.56E-41
	mRNA catabolic process	1.97E-32
M105	cell communication involved in cardiac conduction	8.84E-05
	somatic stem cell population maintenance	0.000187
	tissue regeneration	0.000237
	cardiac conduction	0.001475
	stem cell population maintenance	0.001816
M106	myeloid leukocyte differentiation	3.69E-05
	osteoblast differentiation	5.88E-05
	ossification	8.69E-05
	myeloid cell differentiation	0.000111
	heart morphogenesis	0.000114
M107	limb morphogenesis	1.11E-08
	mesenchyme development	2.11E-08
	limb development	5.78E-08
	osteoblast differentiation	4.05E-07
	urogenital system development	1.13E-06

Appendix Figure S3. Maintaining expanded UiPSM colonies derived from UCs.

A. Representative presomitic-mesoderm-related gene expressions of the UiPSM colonies derived from three donors. Data were mean ± SD, n = 3 independent experiments.

B. Enriched GO terms and p values of the UiPSM colonies at day 9, P1, P9 and P18.

C. qRT-PCR analysis of layer-associated genes of the UiPSM colonies versus DE, NPC and hESC respectively. Expression levels were normalized to GAPDH. Data were mean ± SD, n = 3 independent experiments. ***p < 0.001.

D. Karyotype analysis of representative UiPSM colonies. n = 3 independent experiments.

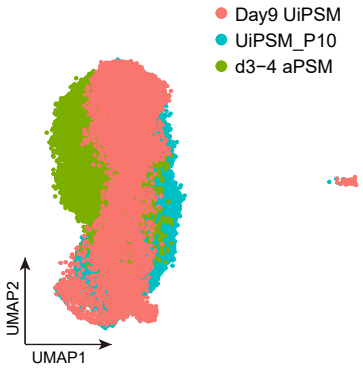
E. Dot plot showed the gene expression and ratio for single-cell RNA-seq data of UiPSM and UiPSC teratoma in each cluster.

F. t-SNE plot of subpopulations of M1 cluster (bottom) in Fig 2H.

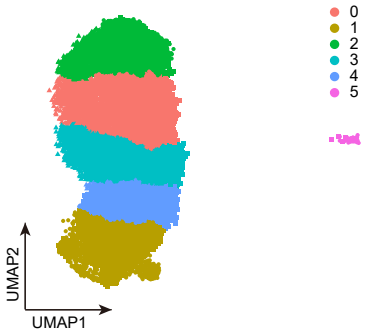
G. Enriched gene ontology analysis of each subcluster. P value was less than 0.05.

Appendix Figure S4

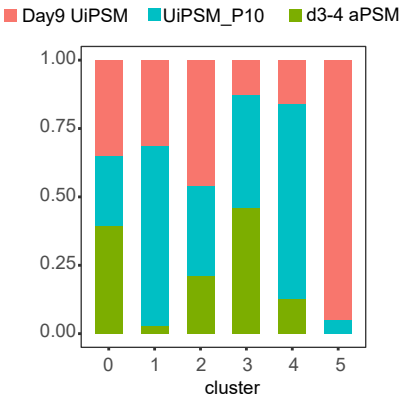
A



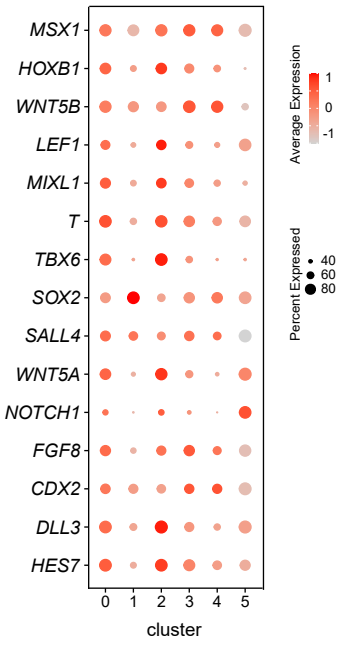
B



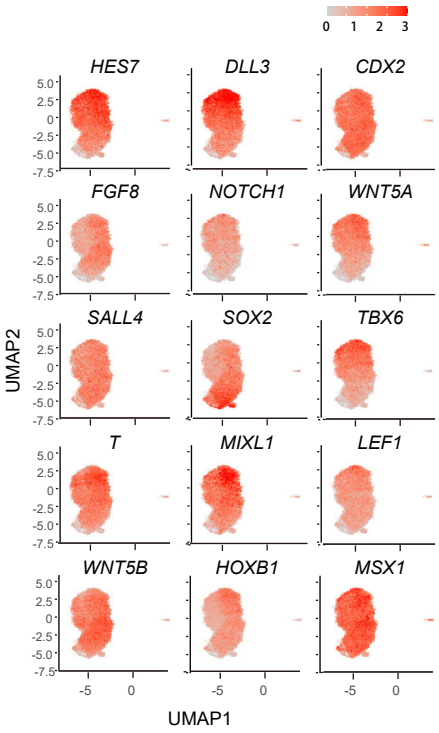
C



D



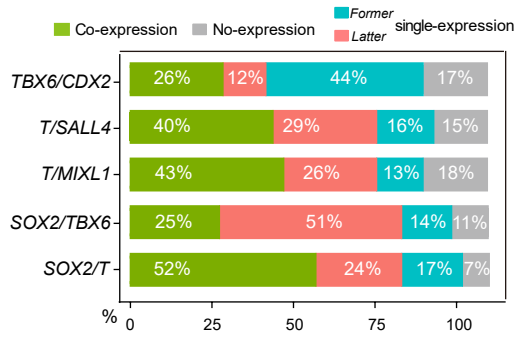
E



F

	GO terms	*P*-Value
C0	regionalization	3.13E-11
	mesoderm development	1.94E-11
	anterior/posterior pattern specification	1.48E-10
	pattern specification process	1.48E-10
	somitogenesis	1.64E-08
	segmentation	3.12E-07
	somite development	1.14E-07
C1	regeneration	3.74E-08
	cellular response to copper ion	1.17E-05
	polyamine biosynthetic process	1.26E-05
C2	regionalization	3.23E-24
	pattern specification process	1.82E-21
	anterior/posterior pattern specification	7.22E-21
	segmentation	1.88E-14
	somitogenesis	2.48E-14
	somite development	4.75E-14
	mesoderm development	3.23E-10
C3	embryonic organ development	5.36E-04
	glucose transmembrane transport	7.14E-04
C4	regulation of chromosome segregation	2.13E-08
	mitotic nuclear division	2.52E-07
	nuclear division	3.74E-07
	chromosome separation	6.13E-07
C5	epithelial cell migration	7.43E-10
	regionalization	1.53E-08
	pattern specification process	1.95E-07
	anterior/posterior pattern specification	2.30E-07
	limb development	9.31E-05
	limb morphogenesis	2.61E-04

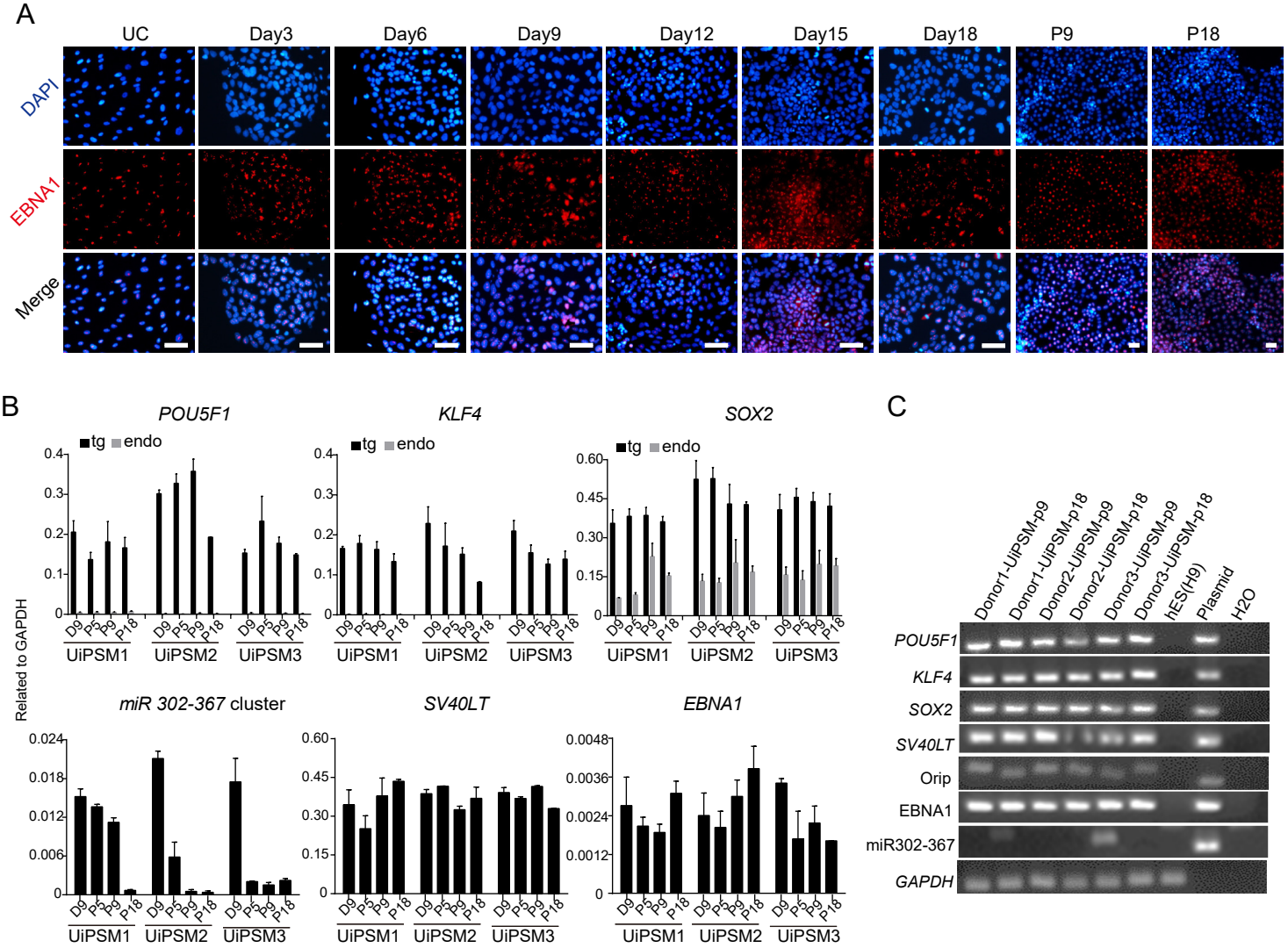
G



Appendix Figure S4. Transcriptomic analysis of UiPSM clone

A. UMAP projection of 22957 cells from UiPSM at day 9 (6824 cells) and UiPSM_P10 (10289 cells) and the published data (d3-4 aPSM, 5844cells).
B. All cells in (A) grouped into 6 clusters by Louvian clustering.
C. The percentage change of UiPSM_P10 and d3-4 aPSM in each cluster
D. Dot plot showed the expression and ratio of PSM specific genes in 6 clusters
E. Expression levels of PSM marker genes in UMAP projection.
F. Gene ontology analysis of 6 clusters based on differentially expressed genes indicated significant enriched terms ranked by P value. P value less than 0.05.
G. The proportion change of cells co-expressed *TBX6* and *CDX2, T* and *SALL4, T* and *MIXL1, SOX2* and *TBX6, SOX2 and T* in UiPSM_P10.

Appendix Figure S5



Appendix Figure S5. The episomal vectors work on UiPSM clone

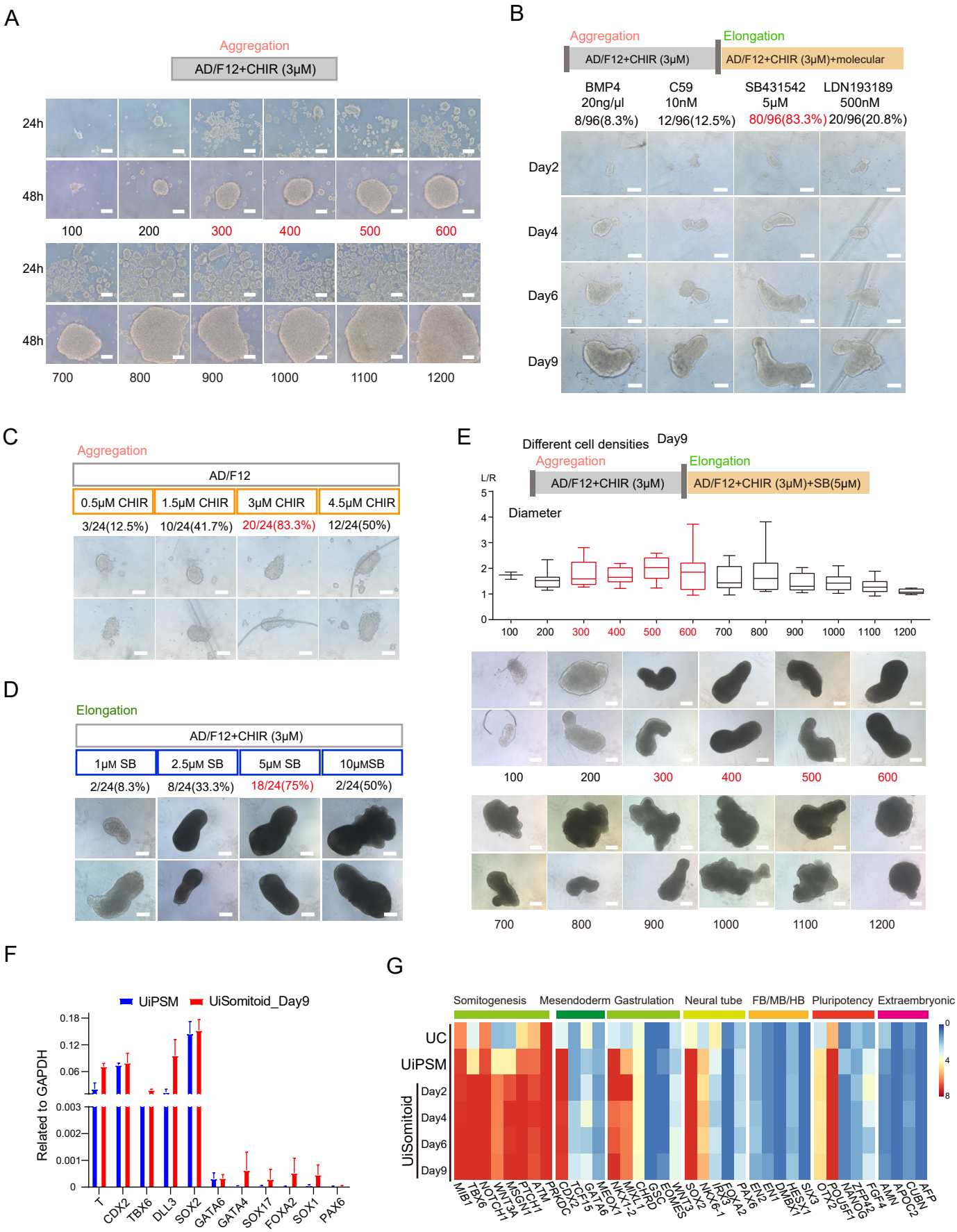
A. Immunofluorescence of EBNA1 (red) of the UiPSM reprogramming process and UiPSM colonies at P9 and P18. the scale bar represents 200 μm. n = 3 independent experiments.

B. The expression of episomal-vector-carried genes in the UiPSM colonies of various passages from three donors. 'tg' stands for the exogenous expression levels of episomal-vector-carried factors. Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments.

C. The gel identification of insertions in UiPSM colonies. (Negative control: hES(H9); Positive control: plasmid, pEP4EO2SET2K and pCEP4-miR302-367 cluster; Blank: H2O). n ⩾ 10 independent experiments.

D-E. qRT-PCR analyzed the expression levels of PSM-specific genes in POU5F1-knocked out (D) and SALL4-knocked out (E) UiPSM colonies. Expression levels were normalized to GAPDH. Data were mean ± SD, n = 3 independent experiments. ****p < 0.0001.

Appendix Figure S6



Appendix Figure S6. The elongation of UiSomitoids rely on Nodal signaling pathway regulation.

A. The aggregation of UiPSM cultured in DM medium for 2 days under different initial seeding density. Scale bars, 200 μm. n=3 independent experiments.

B. UiPSM cells (400 cells in a well) cultured in DM medium for 2 days formed aggregates, and elongated with different molecules for 7 days. Scale bars, 200 μm. 24 UiSomitoids were collected in one experiment and the elongated efficiency was calculated in all 4 independent experiments.

C. The aggregation efficiency under the gradient treatment of CHIR99021 via calculating 24 aggregates in one experiment. Scale bars, 200 μm. n=3 independent experiments.

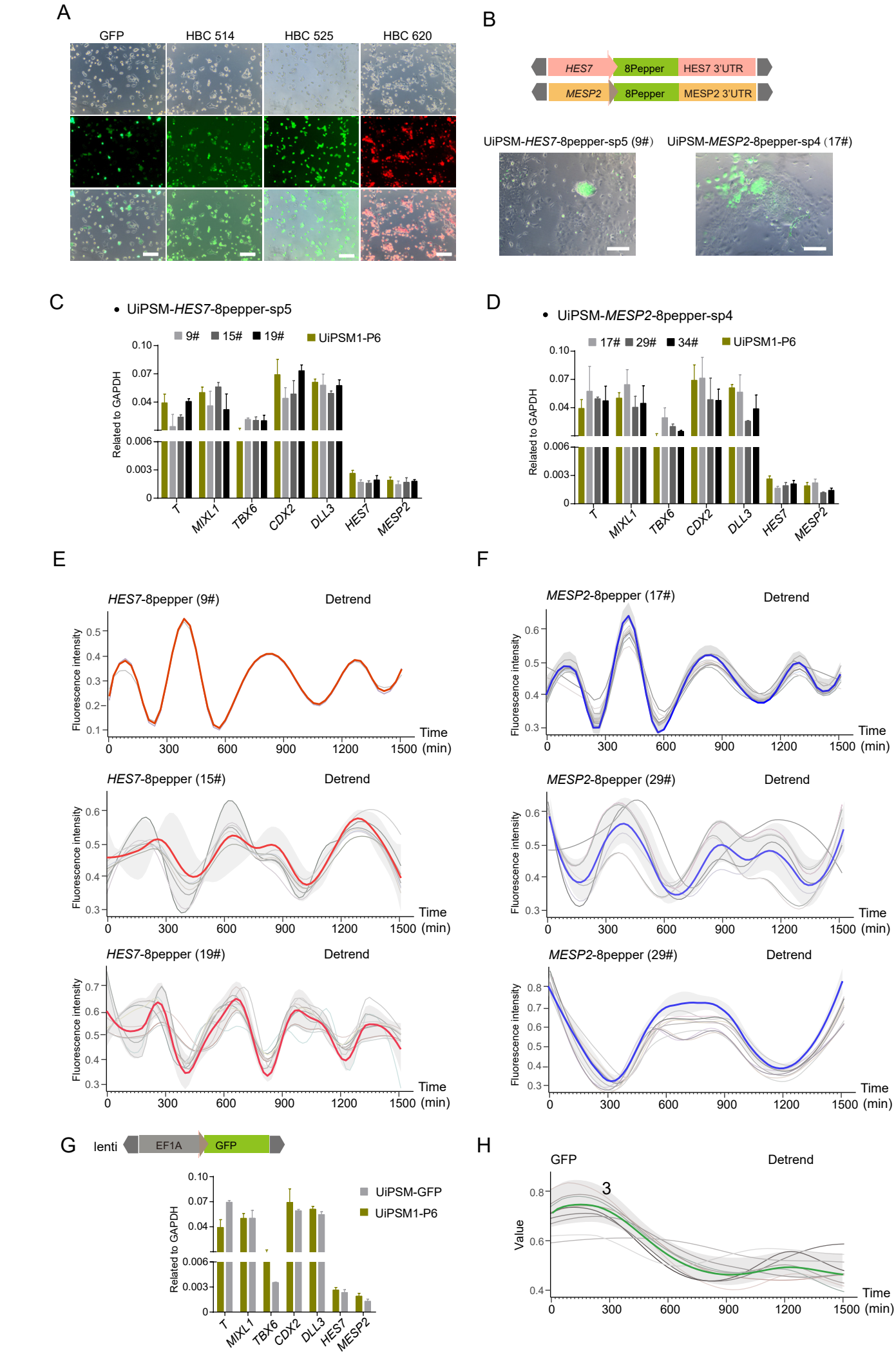
D. The elongated efficiency changes with increasing concentration of SB431542 calculated with 24 aggregates in one experiment. Scale bars, 200 μm. n=3 independent experiments.

E. The elongation rate of the aggregates treated with 3μM CHIR99021+ 5μM SB431542 under different initial seeding density. Scale bars, 200 μm. 24 aggregates in one experiment. n=3 independent experiments.

F. qRT-PCR analysis of PSM-associated genes (*T, CDX2, TBX6, SOX2*) and mesendoderm-related genes (*GATA4* and *GATA6*), endodermal makers (*SOX17* and *FOXA2*), ectodermal markers (*SOX1* and *PAX6*) of the UiPSM cells and elongated UiSomitoids at day9. Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments.

G. RNA-seq data analysis about the expression levels of somitogenesis-, mesendoderm-, gastrulation-, neural tube-, forebrain/middle-brain/hind-brain-(FB/MB/HB), pluripotency- and extraembryonic-related genes during the UiSomitoid elongation. P value is less than 0.05.

Appendix Figure S7



Appendix Figure S7. The construction of *HES7* and *MESP2* oscillation reporter cell lines.

A. Pepper aptamers (fluorescent RNAs) can bind and activate fluorescent dyes HBC 514,525,620 in UiPSM (n =3 independent experiments). Scale bars, 200 μm.

B. Reporter cell lines for tracking gene *HES7* and *MESP2* oscillations. Negative control: UiPSM cell line overexpressing lenti-GFP. Representative images showed UiPSM-*HES7*-8pepper colonies and UiPSM-*MESP2*-8pepper colonies. Scale bars, 200 μm.

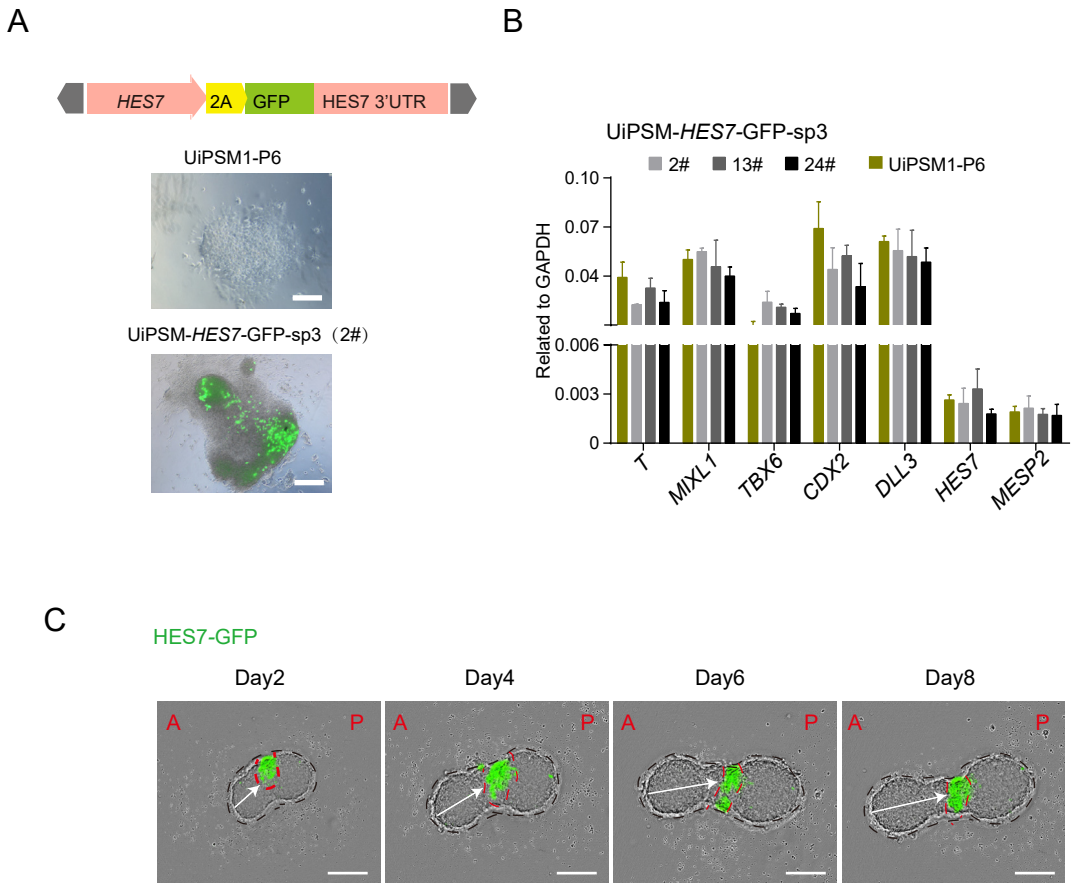
C-D. qRT-PCR analysis of PSM-associated genes of the UiPSM-*HES7*-8pepper colonies and UiPSMMESP2-8pepper colonies. Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments.

E-F. Oscillatory *HES7* (E)and *MESP2* (F) reporter activity measured with Incucyte® S3 Live-Cell Analysis System. Mean intensity profiles for three UiSomitoids were normalized and detrended. Red (or blue) line stands for trend line and each grey line was measured in one experiment. Shaded regions indicate means ± SD, n = 7,12,10 independent experiments for *HES7* and n = 11,9,9 independent experiments for *MESP2*.

G. qRT-PCR analysis of PSM-associated genes of the UiPSM-*GFP* colonies. Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments.

H. The GFP intensity measured with Incucyte® S3 Live-Cell Analysis System. Green line stands for the trend line, each grey line was measured in one experiment. Shaded regions indicate means ± SD, n = 10 independent experiments.

Appendix Figure S8



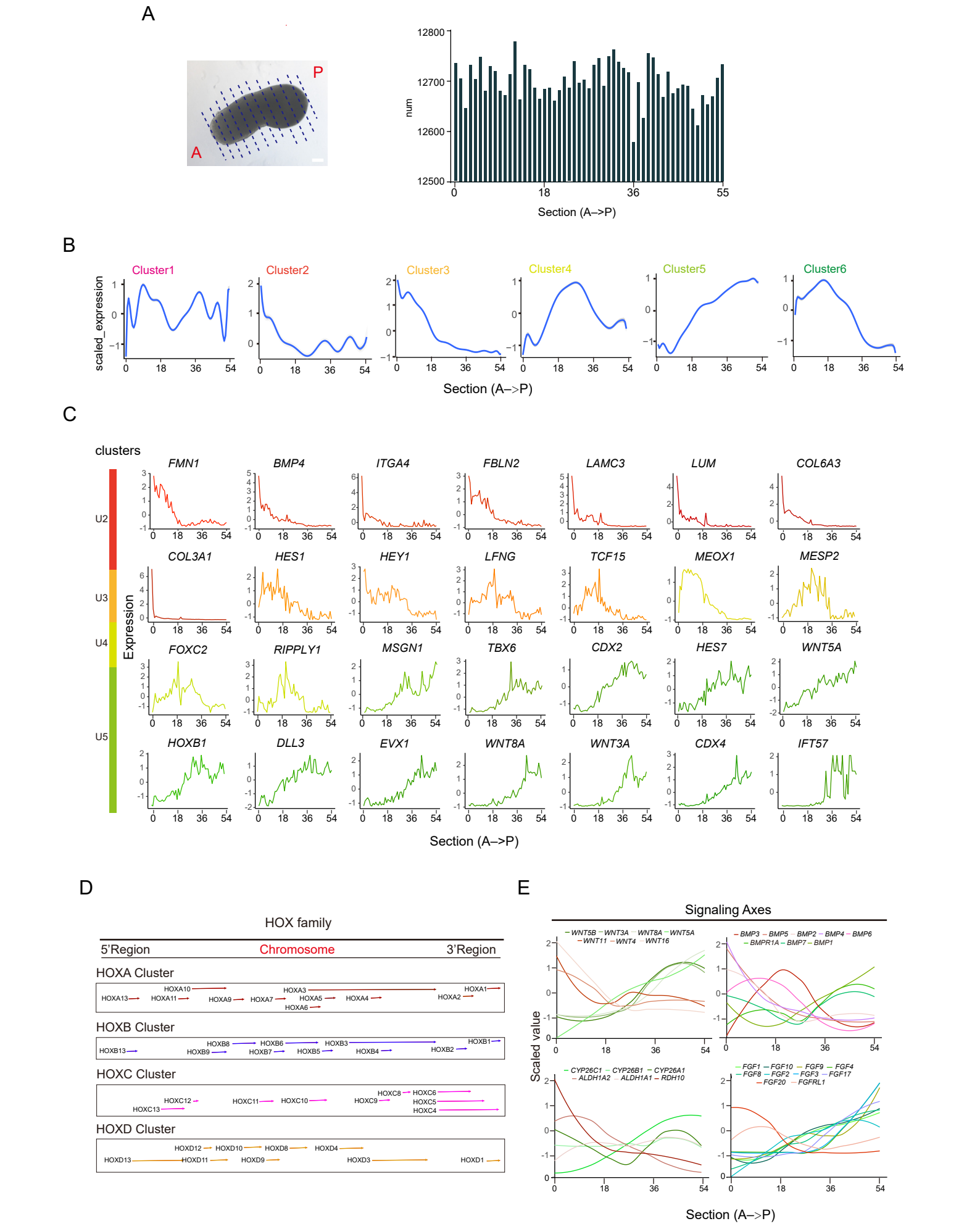
Appendix Figure S8. The elongation of UiSomitoids along the A-P axis.

A. Representative images showed UiPSM colonies and UiPSM-*HES7*-GFP colonies. Scale bars, 100 μm.

B. qRT-PCR analysis of PSM-associated genes of the UiPSM-*HES7*-GFP colonies. Expression levels were normalized to *GAPDH*. Data were mean ± SD, n = 3 independent experiments.

C. Fluorescence images showing the expression change of HES7 along the A-P axis of the UiSomitoid (green, HES7- GFP). n =3 independent experiments. Scale bars, 400 μm.

Appendix Figure S9



Appendix Figure S9. Transcriptomic anteroposterior organization of UiSomitoid

A. Representative image illustrating the sections of the sequencing sample. Scale bars, 400 μm. Quantification of the number of detectable genes in each section along the anteroposterior axis of UiSomitoid.
B. The gene expression patterns of 6 clusters along the anterior-posterior axis of UiSomitoid.
C. Line charts illustrating the trends of typical genes in selected gene clusters.
D. Schematic overview of the genomic distribution of the HOX family.
E. The expressional distribution of Wnt, BMP, RA and FGF signaling pathways along the anteroposterior axis.

Appendix Table S1. qPCR primers information

All qPCR primers were designed by Beacon Designer software.

qPCR primers		
Gene	Forward	Reverse
GAPDH	GTGGACCTGACCTGCCGTCT	GGAGGAGTGGGTGTCGCTGT
T	TGCTTCCCTGAGACCCAGTT	GATCACTTCTTTCCTTTGCATCAAG
MIXL1	AGGCAGGAGAATCACTTG	GGCTACCACAGAACCATAG
TBX6	GTGACAGCCTACCAGAAC	GTGACAGCCTACCAGAAC
CDX2	TTCACTACAGTCGCTACATC	TTTCCTCTCCTTTGCTCTG
HES7	GAATATAAGGCTCCAGGC	CTATTCTCAGCTCGATCC
DLL3	TCGTCCGTAGATTGGAAT	GTAGATGGAAGGAGCAGATA
MESP2	GTCTCTCTGTGTCTCCAG	CAGTCTCTGGCATGATGG
CD13	TTCAGCACATCAGGCAAT	TTCTCTTCGTCGTAGTTCAC
PAX8	CTGTGACAATGACACTGT	ATAGGGAGGTTGAATGGT
Endo*POU5F1*	CCTCACTTCACTGCACTGTA	GGAGGAGTGGGTGTCGCTGT
Endo*KLF4*	GCCTTGCTGATTGTCTAT	AAGTCAACGAAGAGAAGAA
Endo*SOX2*	TTCAAGGAGAGGCTTCTT	AACAAGACCACAGAGATG
NANOG	AAGGTCCCGGTCAAGAAACAG	CTTCTGCGTCACACCATTGC
ESRRB	CTCAGAGAGCAGCCCATA	CCAGTAGGTATGAGACAATCTTG
SALL4	TCACCGAACCAACACATC	CAGCATCACGGCATTAGT
LIN28A	TTAAGAAGTCAGCCAAGG	CTCTCACTCCCAATACAG
DPPA3	TGAAAGAAGACCAACAAACA	TCCATTAGACACGCAGAA
DPPA5	ATCTCGAATCCCTTACAT	TACAAATAGGAGCCGTAA
DNMT3L	TTCGGAAGAAGAATTGTC	GAAGTGAGTTCTGTTGAA
GATA3	AAGGCATCCAGACCAGAA	AAGTCCTCCAGTGAGTCAT
SOX1	TTTCCCCTCGCTTTCTCA	TTTCCCCTCGCTTTCTCA
PAX6	TTGCTTGGGAAATCCGAG	TGCCCGTTCAACATCCTT
SOX17	CTCCGGTGTGAATCTCCCC	CTCCGGTGTGAATCTCCCC
FOXA2	ACAGCAGTCTTCTTCACC	AGCAGGAGTCTACACAGTA