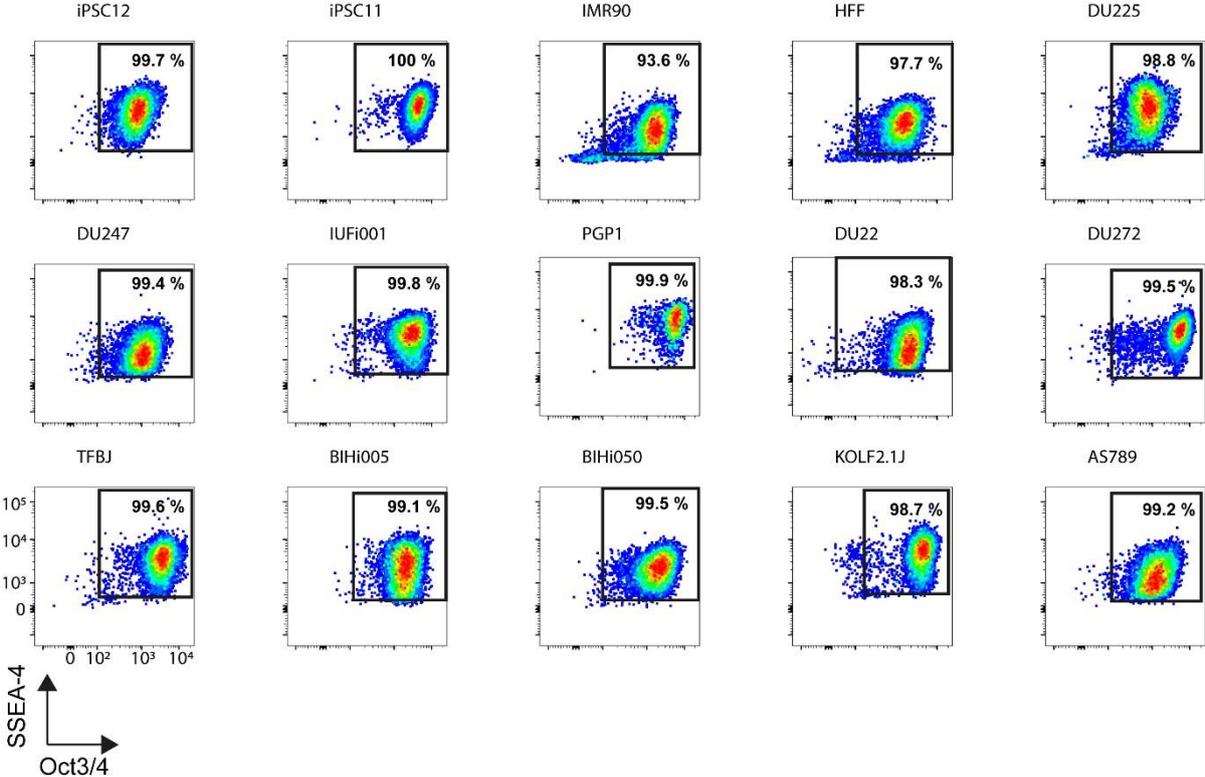
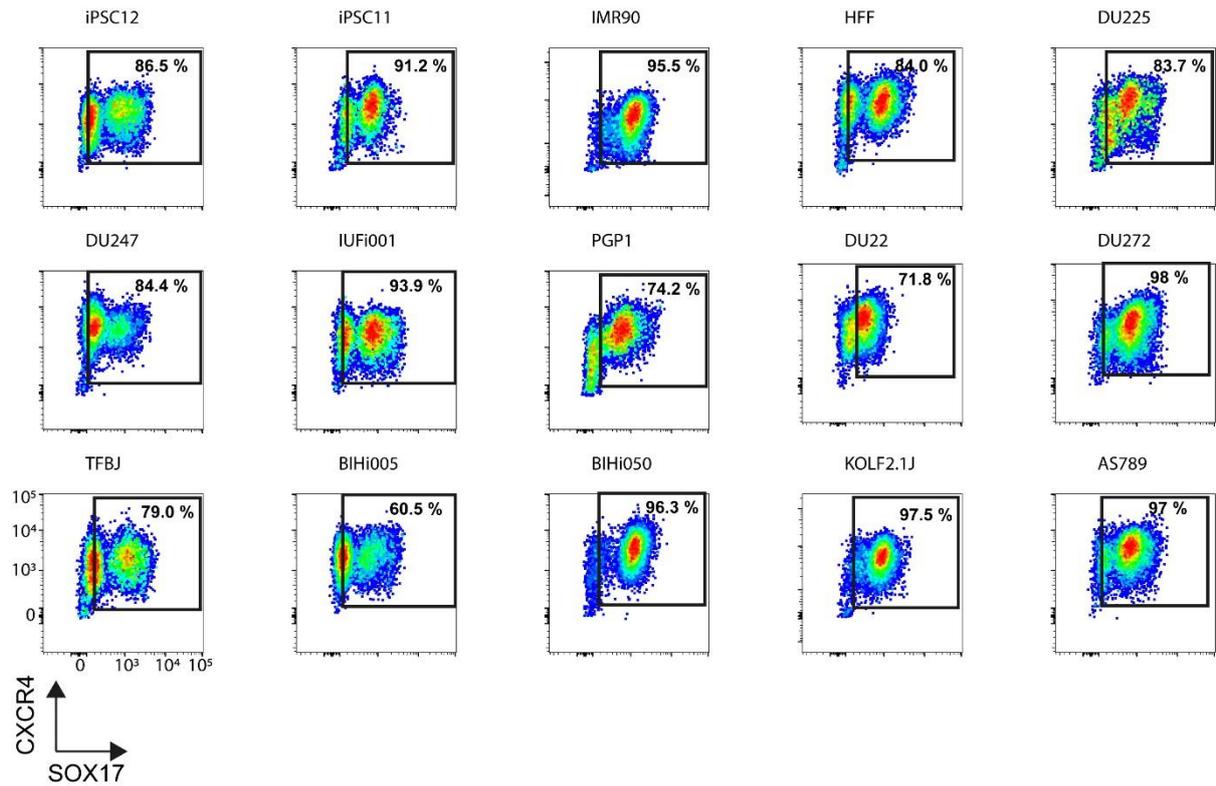


Reassessment of Marker Genes in Human Induced Pluripotent Stem Cells for Enhanced Quality Control

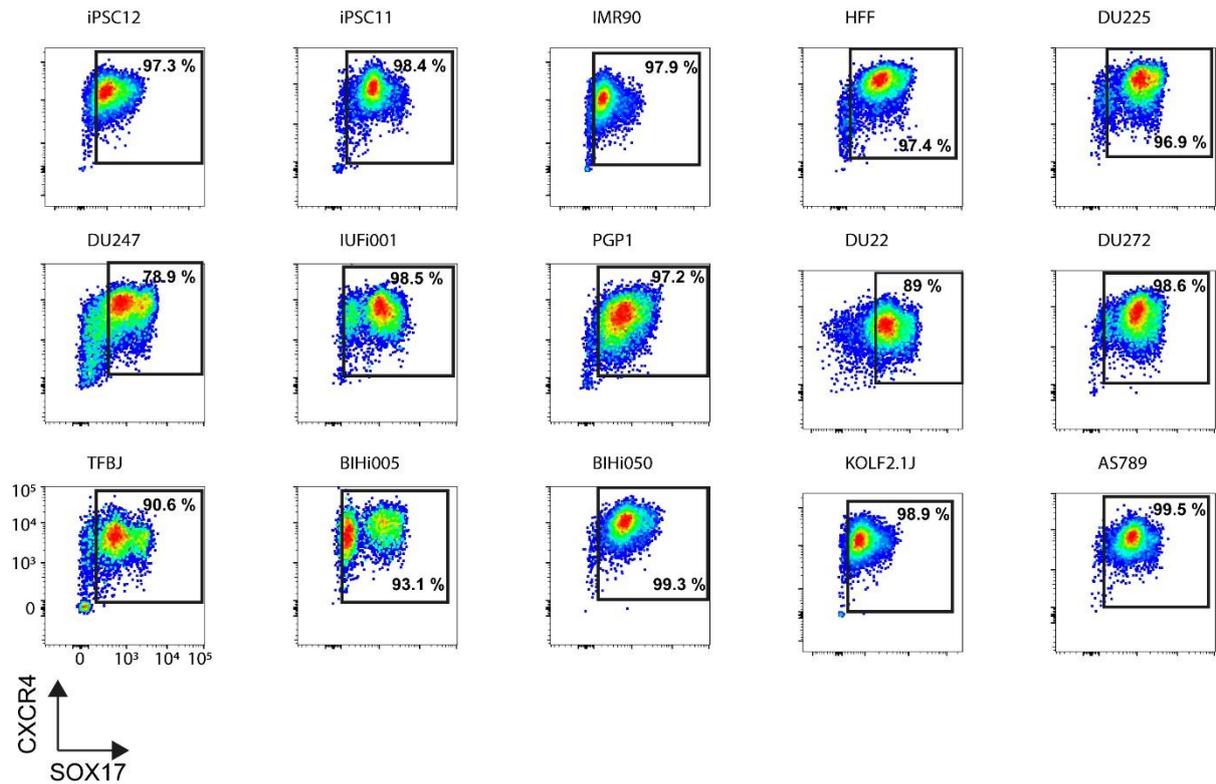
Supplementary Information



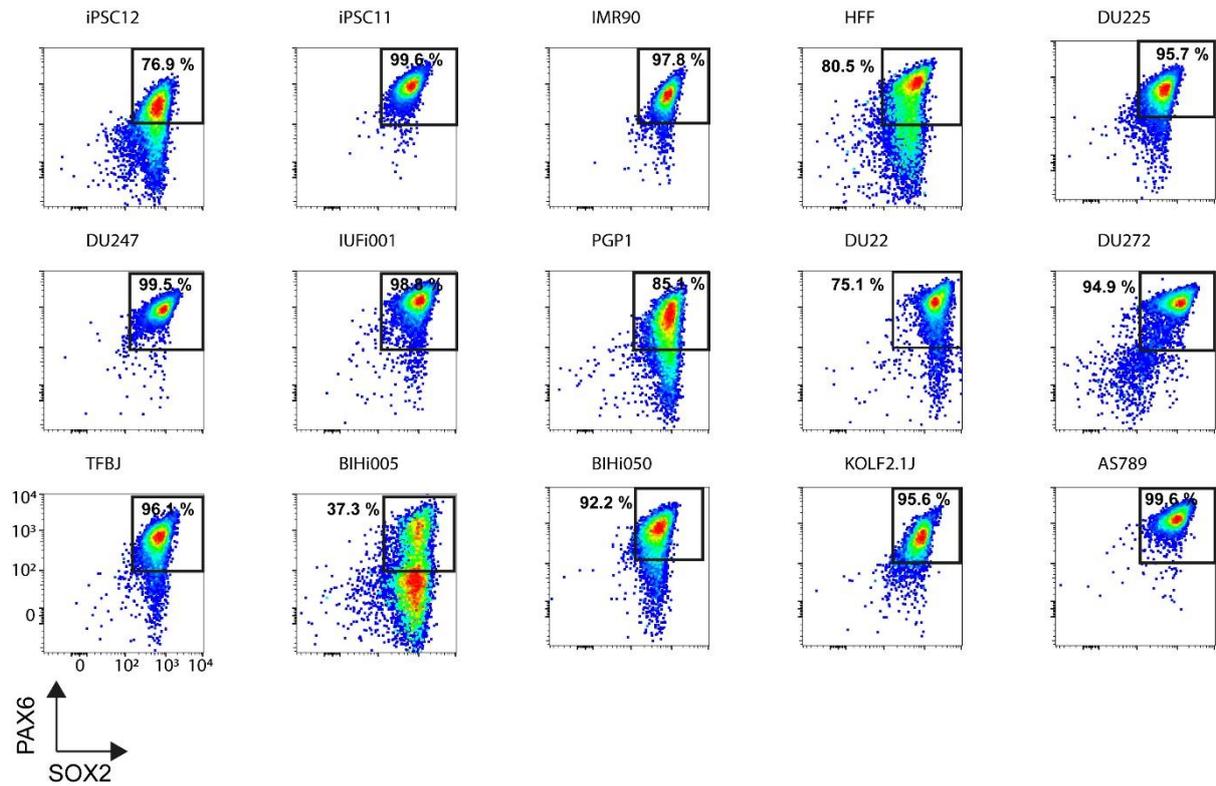
Supplementary Fig. 1. Flow cytometry analysis of undifferentiated human induced pluripotent stem cells used in this study. The 15 different human induced pluripotent stem cell (iPSC) lines used in this study were characterized for protein levels of the two markers of the undifferentiated state SSEA-4 and Oct3/4. The DU22 line expresses a GFP protein. The fluorescent signal was compensated accordingly.



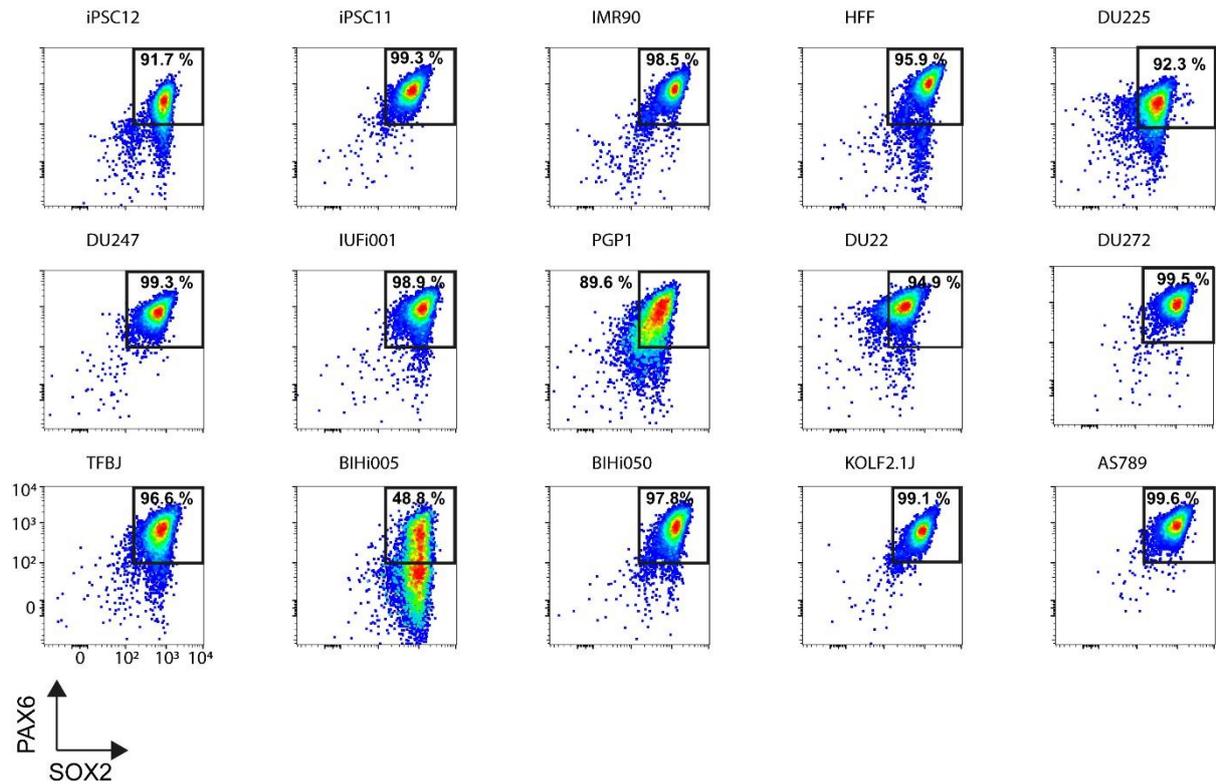
Supplementary Fig. 2. Flow cytometry analysis of endoderm-differentiated human induced pluripotent stem cells used in this study. The 15 different human induced pluripotent stem cell (iPSC) lines used in this study were differentiated into endoderm using the STEMdiff Trilineage differentiation kit and characterized for protein levels of the two markers of the endoderm state CXCR4 and SOX17. The DU22 line expresses a GFP protein. The SOX17 antibody conjugate was adjusted accordingly to PE.



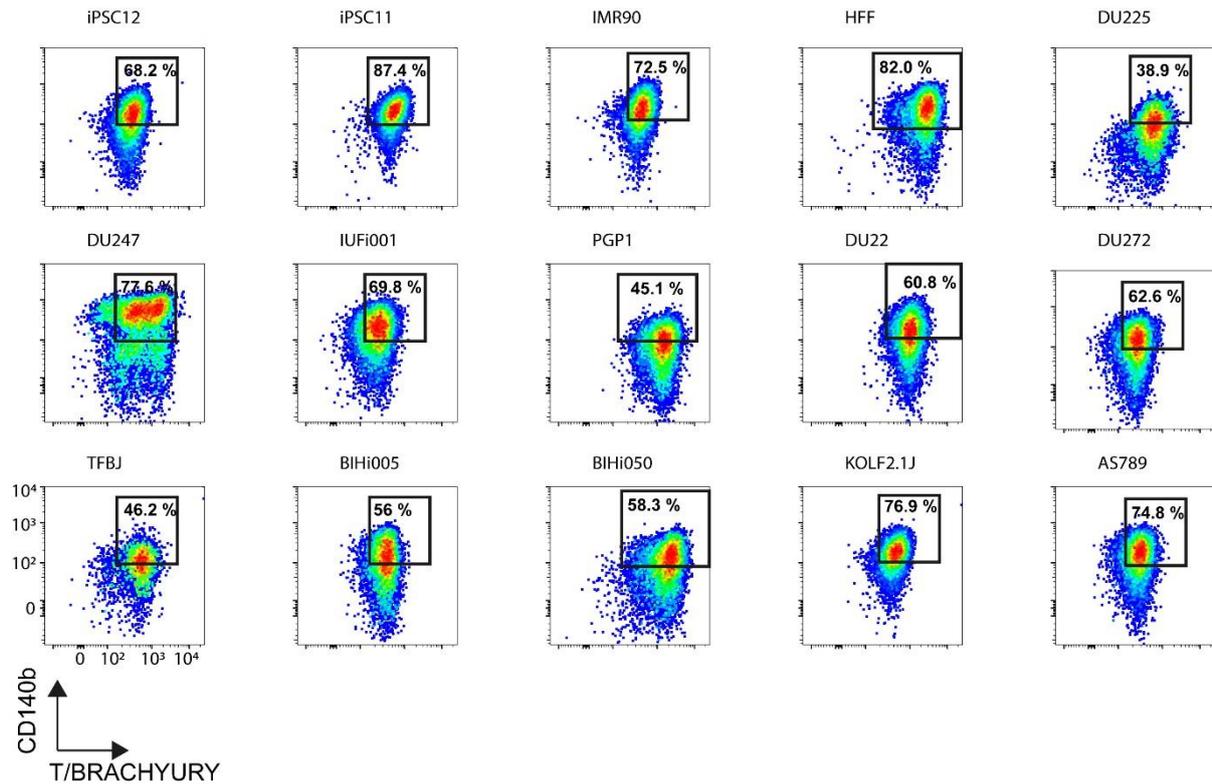
Supplementary Fig. 3. Flow cytometry analysis of endoderm-differentiated human induced pluripotent stem cells used in this study. The 15 different human induced pluripotent stem cell (iPSC) lines used in this study were differentiated into endoderm using the StemMACS Trilineage differentiation kit and characterized for protein levels of the two markers of the endoderm state CXCR4 and SOX17. The DU22 line expresses a GFP protein. The SOX17 antibody conjugate was adjusted accordingly to PE.



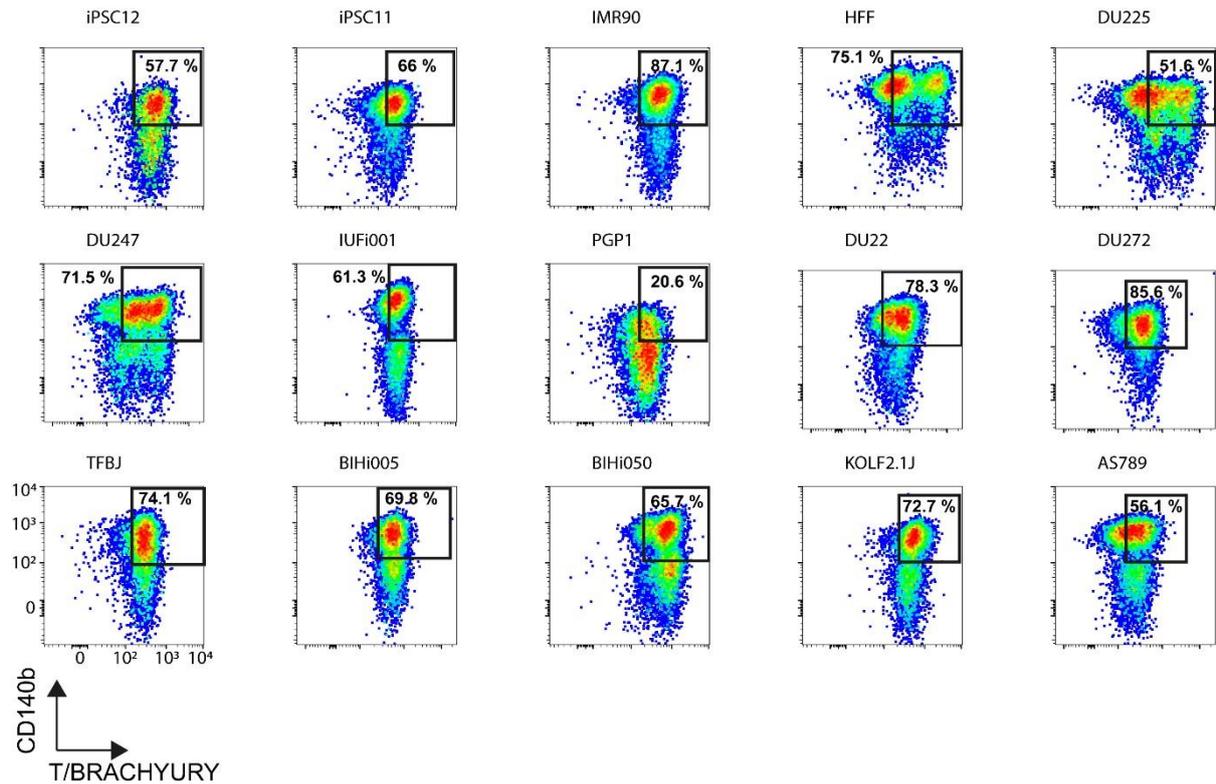
Supplementary Fig. 4. Flow cytometry analysis of ectoderm-differentiated human induced pluripotent stem cells used in this study. The 15 different human induced pluripotent stem cell (iPSC) lines used in this study were differentiated into ectoderm using the STEMdiff Trilineage differentiation kit and characterized for protein levels of the two markers of the ectoderm state PAX6 and SOX2. The DU22 line expresses a GFP protein. The SOX2 antibody conjugate was adjusted accordingly to PE.



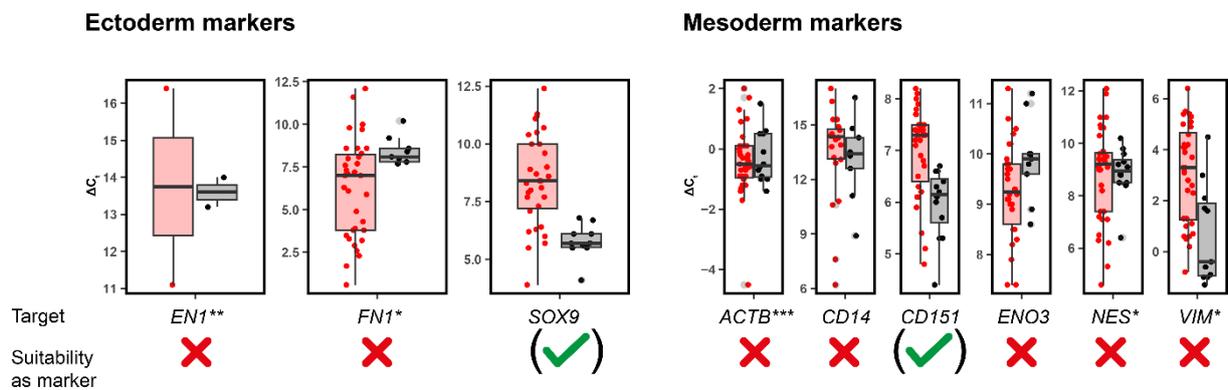
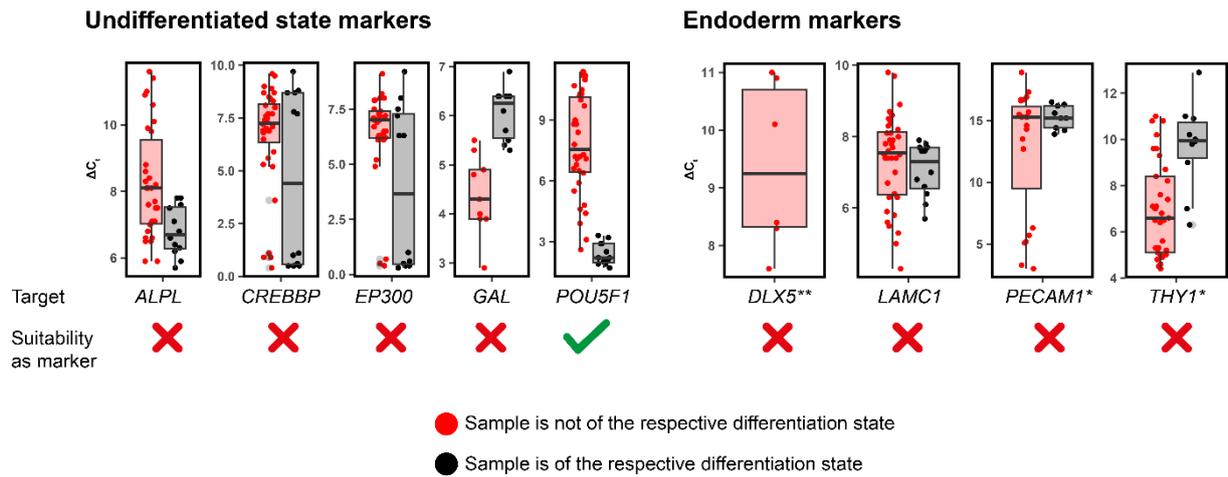
Supplementary Fig. 5. Flow cytometry analysis of ectoderm-differentiated human induced pluripotent stem cells used in this study. The 15 different human induced pluripotent stem cell (iPSC) lines used in this study were differentiated into ectoderm using the StemMACS Trilineage differentiation kit and characterized for protein levels of the two markers of the ectoderm state PAX6 and SOX2. The DU22 line expresses a GFP protein. The SOX2 antibody conjugate was adjusted accordingly to PE.



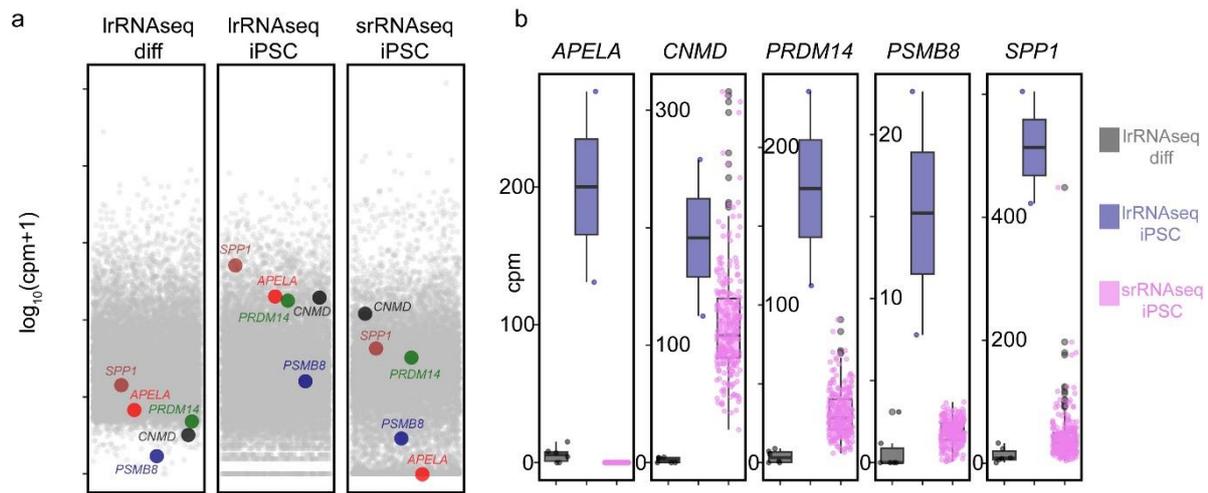
Supplementary Fig. 6. Flow cytometry analysis of mesoderm-differentiated human induced pluripotent stem cells used in this study. The 15 different human induced pluripotent stem cell (iPSC) lines used in this study were differentiated into mesoderm using the STEMdiff Trilineage differentiation kit and characterized for protein levels of the two markers of the mesoderm state CD140b and T/BRACHYURY. The DU22 line expresses a GFP protein. The secondary antibody used for T/BRACHYURY detection was adjusted to PE accordingly.



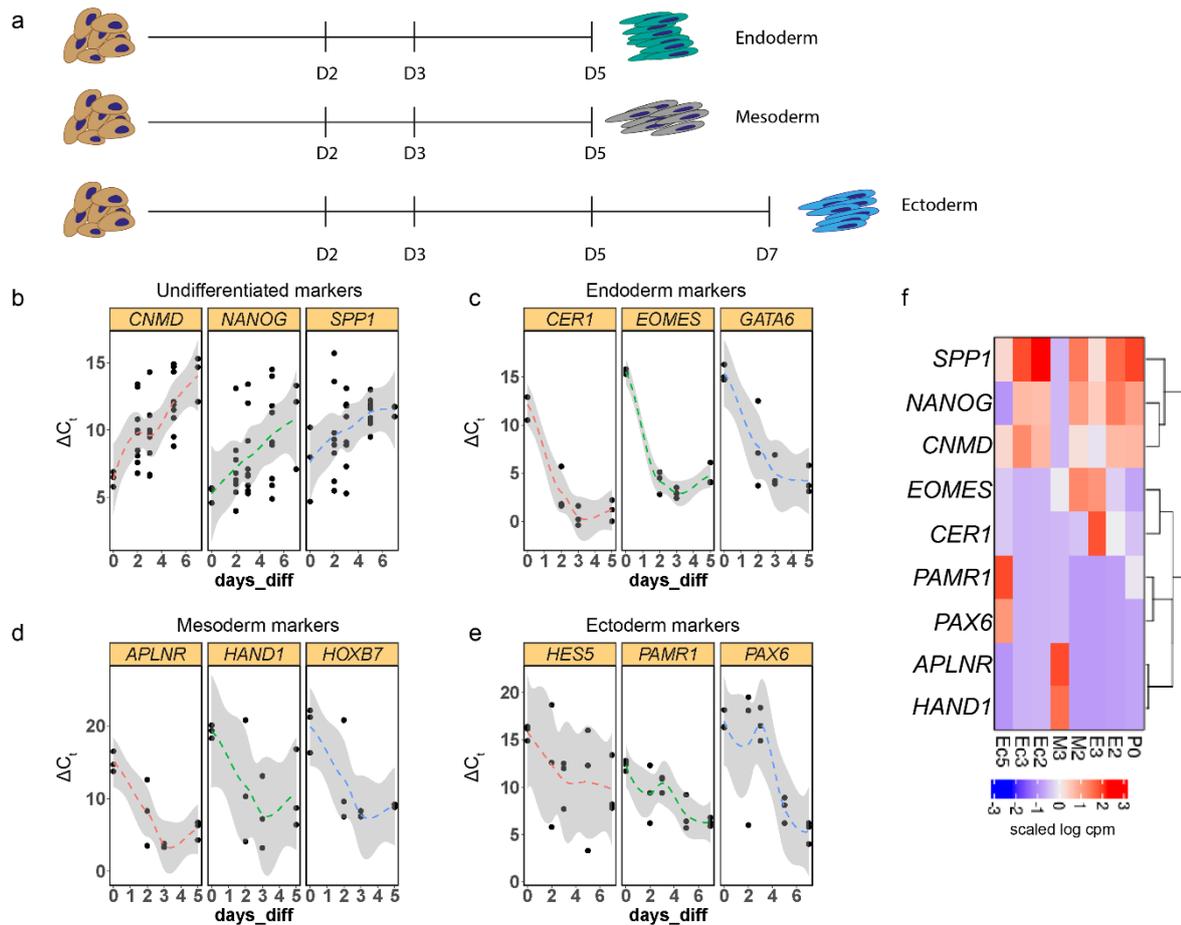
Supplementary Fig. 7. Flow cytometry analysis of mesoderm-differentiated human induced pluripotent stem cells used in this study. The 15 different human induced pluripotent stem cell (iPSC) lines used in this study were differentiated into mesoderm using the StemMACS Trilineage differentiation kit and characterized for protein levels of the two markers of the mesoderm state CD140b and T/BRACHYURY. The DU22 line expresses a GFP protein. The secondary antibody used for T/BRACHYURY detection was adjusted to PE accordingly.



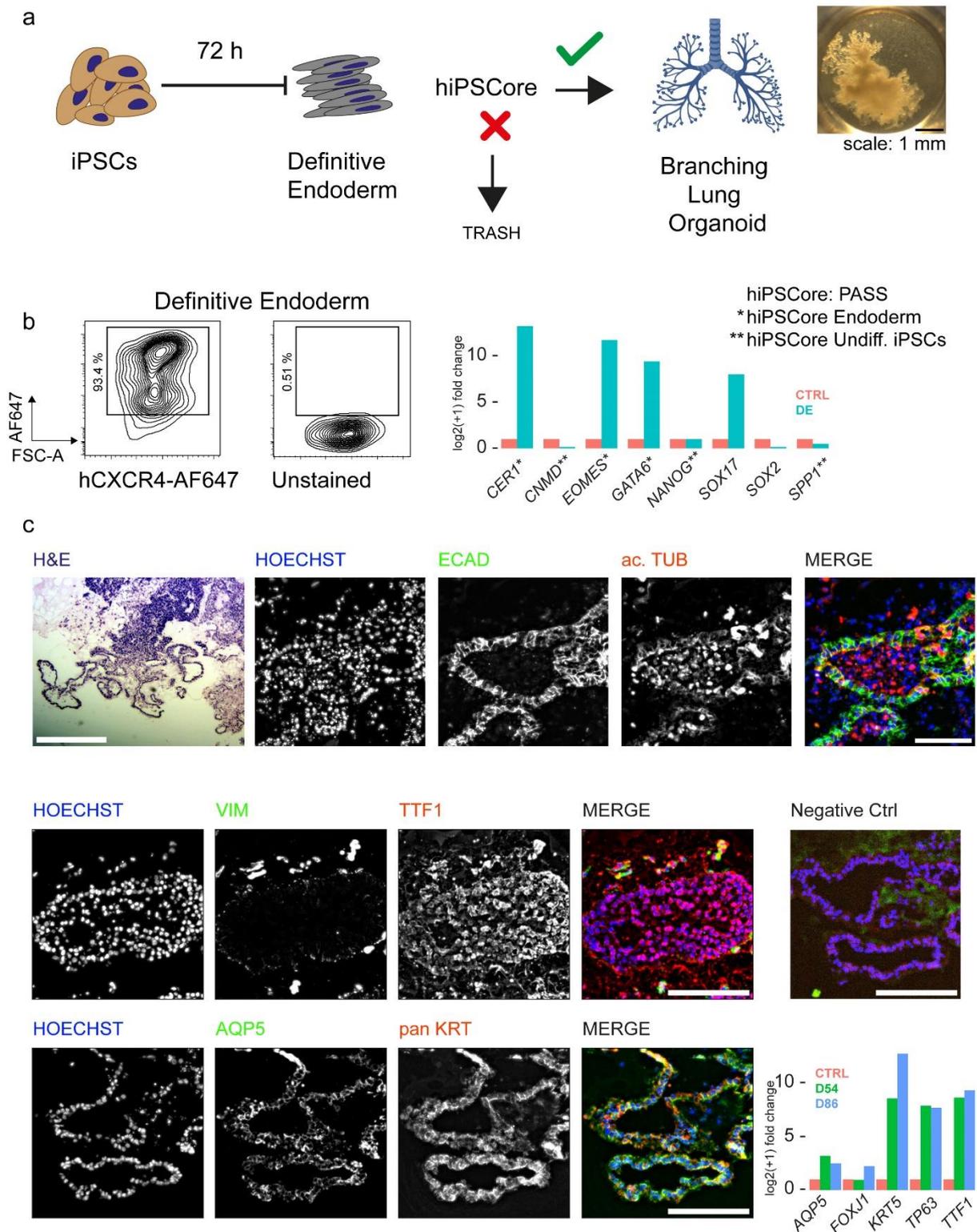
Supplementary Fig. 8. Analysis of randomly selected marker genes by qPCR. Eighteen different marker genes were randomly selected from the marker recommendations to identify human induced pluripotent stem cell (iPSC) differentiation states listed in the 2023 version of the International Society for Stem Cell Research guidelines. These genes were analyzed by qPCR to assess their suitability as markers. Most of the analyzed gene transcripts do not seem to be feasible to discriminate unequivocally between iPSC early differentiation states obtained by directed trilineage differentiation. Only a few (*POU5F1*, *SOX9*, *CD151*) showed potential to unequivocally discriminate between differentiation states. An $n = 6$ independent samples per each germ layer differentiation state and protocol, as well as an $n = 12$ independent undifferentiated samples were analyzed. *: Listed for more than one differentiation state. **: No amplification for the majority of samples. ***: For *ACTB*, *GAPDH* only was used as a reference gene. For the rest of the genes, both *ACTB* and *GAPDH* were used as reference genes.



Supplementary Fig. 9. Comparison between genes of the undifferentiated state identified by short- and long-read sequencing. Data generated by Oxford Nanopore Technologies long-read sequencing ($n = 2$) and publicly available transcriptome sequencing datasets ($n = 247$) based on short-read sequencing of human induced pluripotent stem cells (iPSCs) in the undifferentiated state were compared including differentiated cells summarized as “diff” ($n = 2$ for each germ layer). **a** Plot shows log₁₀ counts per million (cpm) +1-normalized gene counts. Genes with clearly distinct gene expression patterns between short- and long-read sequencing of iPSCs are highlighted. These genes show no overlapping expression with differentiated samples and may thus serve as improved iPSC undifferentiated state markers. **b** “Zoom-in” view of the genes highlighted in **a** to exemplify the clear separation between short- and long-read-based transcriptome sequencing results based on transcript cpm. Some of these genes were further validated. Single dots in **a** represent individual transcripts, whereas single dots in **b** represent individual datasets. diff: differentiated iPSCs (endoderm, ectoderm, mesoderm summarized), IrRNAseq: long-read transcriptome sequencing, srRNAseq: short-read transcriptome sequencing.

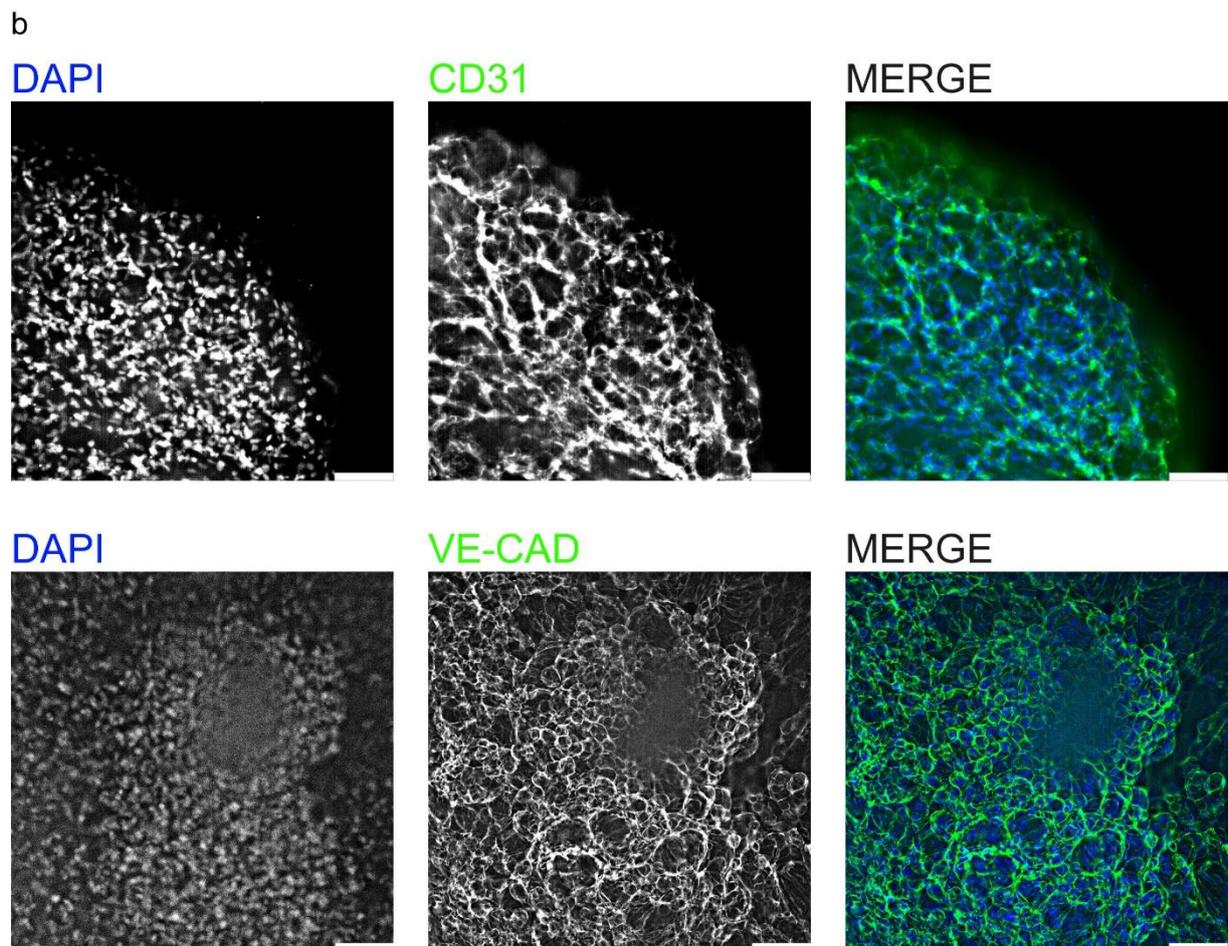
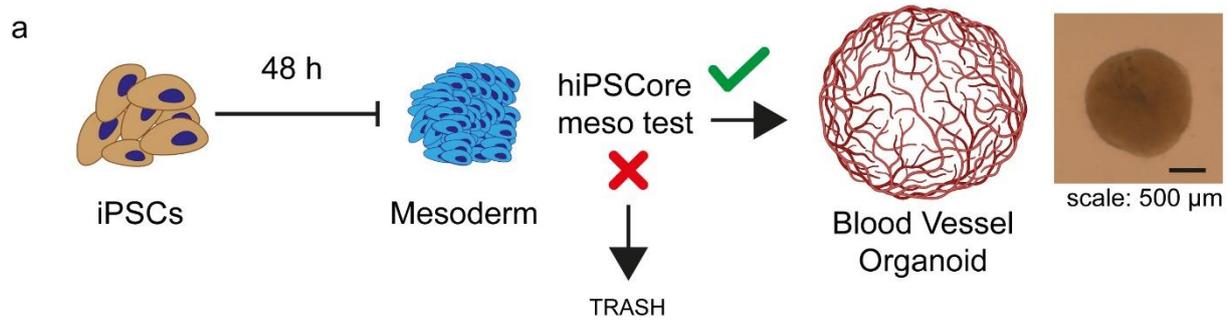


Supplementary Fig. 10. Gene expression dynamics of validated marker genes for pluripotency testing of human induced pluripotent stem cells. **a** $n = 3$ different human induced pluripotent stem cell (iPSC) lines were differentiated into endoderm, ectoderm, and mesoderm using the STEMdiff Trilineage differentiation kit. Cells were harvested at day 2, 3, 5 (endoderm, mesoderm, ectoderm), and 7 (ectoderm) and the relative gene expression patterns were analyzed by qPCR. **b** Undifferentiated markers expression levels decline after two days of differentiation until the final day. **c** Endoderm markers expression levels increase after two days of differentiation. **d** Mesoderm markers expression levels increase after two days of differentiation. **e** Ectoderm markers expression levels increase after five days of differentiation until the final day. **b - e** Transcript levels were normalized to the two reference genes *ACTB*, and *GAPDH*. High ΔC_t values indicate low expression and vice versa. **f** Long-read transcriptome sequencing experiment of $n = 1$ trilineage-differentiated iPSCs (iPSC12) demonstrates the same pattern as shown in **b - e**, indicating upregulation of endo- and mesodermal genes after two days, and of ectodermal genes after five to seven days. P0: undifferentiated iPSCs. E2/3: endoderm-differentiated at day 2/3. M2/3: mesoderm-differentiated at day 2/3. Ec2/3/5: ectoderm-differentiated at day 2/3/5.

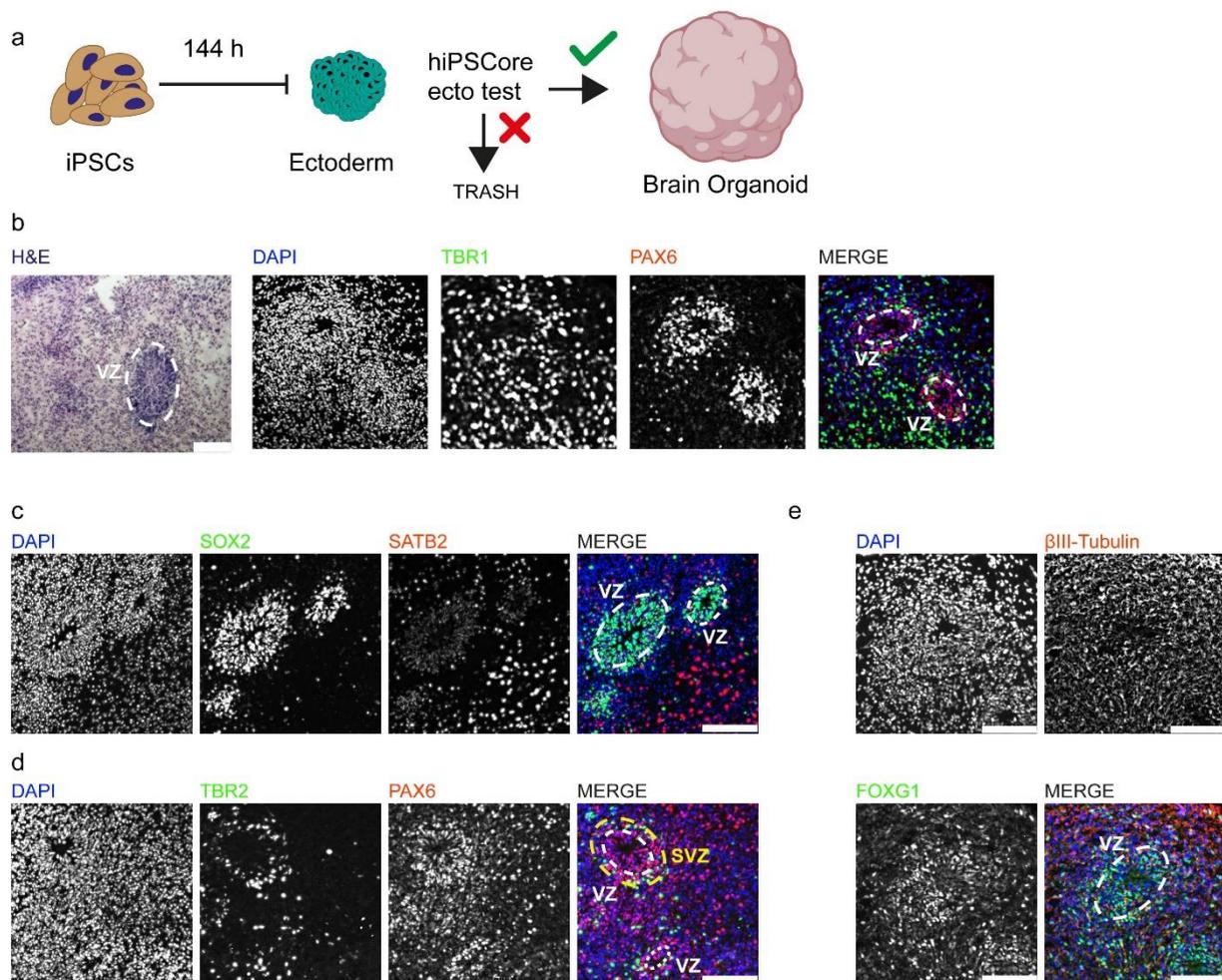


Supplementary Fig. 11. Human induced pluripotent stem cells differentiate into branching lung organoids. **a** Human induced pluripotent stem cells (iPSCs) need to reach the definitive endoderm stage before they can differentiate further into branching lung organoids. The hiPSCore test can be used to assess success of endoderm specification and to decide whether iPSC-derived cells are fit to continue differentiation. **b** The success of definitive endoderm differentiation was assessed by flow cytometry analysis of the endodermal

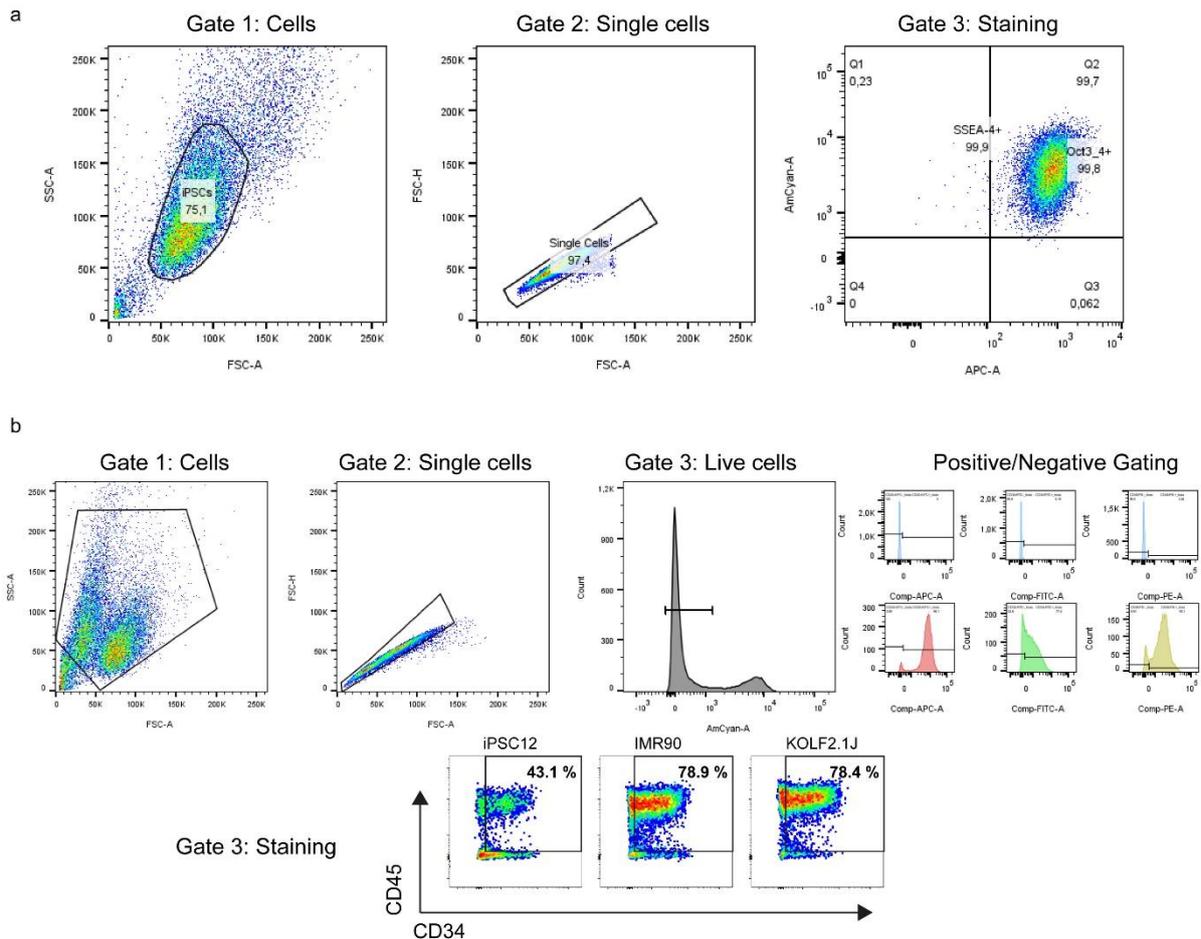
surface marker CXCR4 (left), as well as the qPCR-based hiPSCore endoderm subtest and further genes important in the context of iPSC differentiation (right). **c** A 54 day old branching lung organoid was fixed, frozen in OCT, and characterized by hematoxylin and eosin staining (H&E), as well as immunofluorescent (IF) imaging of cell type-specific markers indicative of successful branching lung organoid differentiation. Gene expression of cell type-specific genes was analyzed by qPCR. In summary, cells passing the qPCR-based hiPSCore endoderm subtest can be subjected to branching lung organoid differentiation and long-term maturation. H&E, IF imaging, and flow cytometry data are representative of $n = 3$ independent differentiations using two different iPSC lines. Scale bar H&E: 500 μm , scale bar IF images: 100 μm . The branching lung organoid icon (in Panel **a**) was created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.



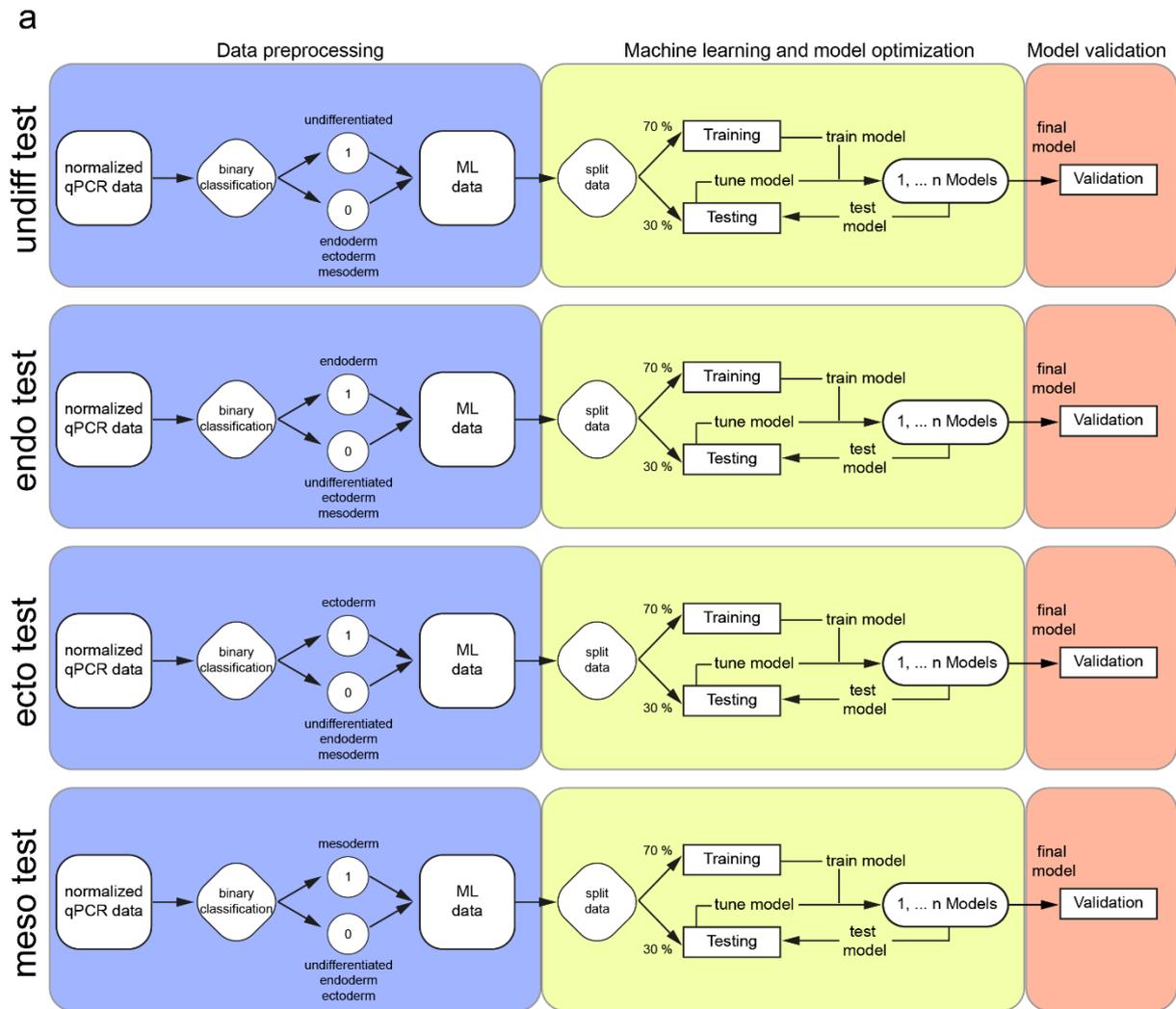
Supplementary Fig. 12. Human induced pluripotent stem cells differentiate into blood vessel organoids. **a** Human induced pluripotent stem cells (iPSCs) need to reach the mesoderm stage before they can differentiate further into blood vessel organoids. The hiPSCore test can be used to assess success of mesoderm specification and to decide whether iPSC-derived cells are fit to continue differentiation. **b** Success of blood vessel organoid formation was assessed at day 18 by whole mount staining of blood vessel organoids with endothelial-specific markers CD31/PECAM-1 and VE-CAD/CD144 by immunofluorescent (IF) imaging. IF images are representative of $n = 3$ independent differentiations using three different iPSC lines. Scale bar: 100 μ m. The blood vessel organoid icon (in Panel **a**) was created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.



Supplementary Fig. 13. Human induced pluripotent stem cells differentiate into brain organoids. **a** Human induced pluripotent stem cells (iPSCs) need to reach the ectoderm stage before they can differentiate further into brain organoids. The hiPSCore test can be used to assess success of ectoderm specification and to decide whether iPSC-derived cells are fit to continue differentiation. **b** Success of brain organoid formation was assessed at day 80 by staining of cryosections with hematoxylin & eosin (H&E), and assessing staining patterns of neuroectodermal/dorsal pallium progenitors (PAX6), and deep layer neurons (TBR1) by immunofluorescent (IF) imaging. **c** IF staining patterns of neural progenitor/astrocyte (SOX2), and superficial layer neurons (SATB2) in day 80 brain organoids. **d** IF staining patterns of neuroectodermal/dorsal pallium progenitors (PAX6), and intermediate progenitors (TBR2) in day 80 brain organoids. **e** IF staining patterns of neuron-specific markers β III-Tubulin and FOXG1 in day 80 brain organoids. IF images are representative of $n = 3$ independent differentiations using three different iPSC lines. Scale bar: 100 μ m. The brain organoid icon (in Panel **a**) was created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license.



Supplementary Fig. 14. Gating strategy for flow cytometry experiments. a. For staining of undifferentiated human induced pluripotent stem cells (iPSCs) and trilineage-differentiated iPSCs, cells were counted before fixation and viability determined by trypan blue staining ($\geq 90\%$ viability). Gate 1 (population gate) was set using SSC-A and FSC-A gating. Gate 2 (single cell gate) was set using FSC-H and FSC-A gating. Gate 3 (staining gate) was set according to fluorescence minus one (FMO) stainings. **b.** For staining of iPSC-derived hematopoietic progenitor cells, gating for population and single cells was performed according to **a**. An additional gate (Gate 3) was set to discriminate between live and dead cells using DAPI staining immediately added before measurement (only dead cells are stained). Positive/negative gating was set according to FMO staining.



b

$$hiPSCore = \sum \begin{matrix} (class_{raw} * class_p)_{undiff\ test} \\ (class_{raw} * class_p)_{endo\ test} \\ (class_{raw} * class_p)_{ecto\ test} \\ (class_{raw} * class_p)_{meso\ test} \end{matrix}$$

Supplementary Fig. 15. Workflow of the hiPSCore pluripotency test machine learning modelling. **a.** The machine learning models underlying hiPSCore consist of four individual models for undifferentiated (undiff test) human induced pluripotent stem cells (iPSCs), iPSC-derived endoderm (endo test), iPSC-derived ectoderm (ecto test), and iPSC-derived mesoderm (meso test) samples. For each of these tests, validated ΔC_T -normalized qPCR gene expression data was classified using dummy-coded binary classification: Samples which belonged to a certain differentiation state, e.g. undifferentiated iPSCs for the undiff test, endoderm cells for the endo test, etc., were classified as “1”, whereas all other samples were classified as “0”. Afterwards, preprocessed data was split into training and testing data (70:30) and machine learning models trained. After fine tuning and selection of the best model for a

given subtest, each model was validated on an independent flow cytometry-validated dataset.

b. The final hiPSCore is calculated as the sum of each of the individual subtests, ensuring optimization of reagent use.

Supplementary Table 1: Cell lines used in this study.

ID	Cell line	Type	Sex origin	Source	hPSCreg	Reference
1	iPSC11	Commercially available	Male	Cell App.	-	¹
2	iPSC12	Commercially available	Female	Cell App.	IUFi004-A	²
3	IMR90	Commercially available	Female	WiStar	WISCi004-B	-
4	HFF	Reprogrammed from commercially available (Merck) fibroblasts	Male	in-house	-	Unpublished
5	DU225	Genome-edited (parental: 2), <i>ERCC6</i> ^{R683X/R683X}	Female	in-house	-	³
6	DU247	Genome-edited (parental: 2), <i>KCNK2</i> KI	Female	in-house	-	Unpublished
7	IUFi001	Reprogrammed from patient fibroblasts	Male	in-house	IUFi001	⁴
8	PGP1	Commercially available	Male	Synthego	-	-
9	DU22	Genome-edited (parental: 3), Life-Act-GFP	Female	in-house	-	⁵
10	D1	Genome-edited (parental: 3), <i>AHR</i> KO	Female	in-house	-	Unpublished
11	TFBJ	Reprogrammed from commercially available (ATCC) fibroblasts	Male	in-house	HHUUKDi009-A	⁶
12	BIHi005-A24	Genome-edited (parental: BIHi005-A), doxycyclin-inducible <i>NGN2</i>	Male	in-house	BIHi005-A-24	⁷
13	BIHi050	Reprogrammed from patient fibroblasts	Male	In-house	BIHi050-A	-
14	KOLF2.1J	Commercially available	Male	JAX	WTSli018-B-12	⁸
15	AS789	Reprogrammed from patient fibroblasts	Male	In-house	-	⁹

		(disease: Cockayne syndrome B)				
--	--	--------------------------------	--	--	--	--

Supplementary Table 2: Gene set enrichment analysis of directed trilineage-differentiated human induced pluripotent stem cells.

Germlayer (vs. Undifferentiated)	Ontology	ID	Description	NES (mean \pm sd)
Endoderm	BP	GO:0007492	endoderm development	2.23 \pm 0.02
Endoderm	BP	GO:0001706	endoderm formation	2.17 \pm 0.02
Endoderm	BP	GO:0035987	endodermal cell differentiation	1.96 \pm 0.02
Ectoderm	BP	GO:0061351	neural precursor cell proliferation	2.44 \pm 0.02
Ectoderm	BP	GO:0097485	neuron projection guidance	2.18 \pm 0.03
Ectoderm	BP	GO:0030182	neuron differentiation	2.05 \pm 0.02
Mesoderm	BP	GO:0001568	blood vessel development	2.74 \pm 0.01
Mesoderm	BP	GO:0043534	blood vessel endothelial cell migration	2.13 \pm 0.01
Mesoderm	BP	GO:0048534	hematopoietic or lymphoid organ development	1.79 \pm 0.01

NES: normal enrichment score. BP: biological process. GO: gene ontology.

Supplementary Table 3: Gene list of the validated hiPSCore gene panel.

Target	Germ layer	Source	ISSCR Guidelines
<i>CNMD</i>	Undifferentiated	This study	-
<i>NANOG</i>	Undifferentiated	This study/Literature ¹⁰	Listed as undifferentiated state marker
<i>SPP1</i>	Undifferentiated	This study	-
<i>CER1</i>	Endoderm	This study	-
<i>EOMES</i>	Endoderm	This study/Literature ¹¹	Listed as Trophoblast marker
<i>GATA6</i>	Endoderm	Literature ¹²	Listed as Undifferentiated state and endoderm marker
<i>HES5</i>	Ectoderm	Literature ¹³	-
<i>PAMR1</i>	Ectoderm	This study	-
<i>PAX6</i>	Ectoderm	Literature ¹⁴	Listed as ectoderm and endoderm marker
<i>APLN</i>	Mesoderm	This study	-
<i>HAND1</i>	Mesoderm	This study	Listed as ectoderm and mesoderm marker
<i>HOXB7</i>	Mesoderm	This study	-

Supplementary Table 4: hiPSCore substest performance on previously unknown flow cytometry-characterized samples (validation test set).

iPSC line	Endoderm		Ectoderm		Mesoderm	
hPSCreg	CXCR4/SOX17 double pos. (%)	hiPSCore endo *	PAX6/SOX2 double pos. (%)	hipSCore ecto *	CD140b/CD144 single pos. (%)	hiPSCore meso *
BIHi242-C	97	1.00	93	1.00	62.5/34.6	1.00
BIHi245-B	93.2	1.00	98.4	0.99	89.9/9.3	0.96
BIHi258-B	89.8	1.00	97.5	1.00	89.8/2.3	0.96
BIHi005-A24	92.9	1.00	97.5	1.00	23.8/68.9	1.00
UCSFi001-A	89.5	1.00	82.7	1.00	92.6/3.1	1.00
BIHi043-A	93.7	1.00	88.6	1.00	87.2/8.31	0.99
BIHi256-A	95.9	1.00	98.9	1.00	45.7/48.2	0.99
BIHi050-A	88	1.00	81.2	0.99	49.8/45.2	1.00
BIHi005-A3	97.5	1.00	99.3	1.00	21/75	0.99
BIHi049-A	80.6	1.00	85.1	0.90	61.1/26.4	1.00

*: Likelihood of the sample to be endoderm/ectoderm/mesoderm.

Supplementary Table 5: Antibodies used in this study.

Antibody	Company	Isotype	Conjugate	Catalog Number/Dilution
Acetylated Tubulin	Sigma	Mouse IgG _{2B}	Unconjugated	T7451/1:500
Anti-Goat	Jackson	Donkey IgG	DyLight 405	705-475-003/1:400
Anti-Goat	ThermoFisher	Donkey IgG	AlexaFluor 488	A-11055/1:500
Anti-mouse	Dianova	Mouse IgG	AlexaFluor 647	715-605-150/1:250
Anti-rabbit	Thermo Fisher	Goat IgG	AlexaFluor 488	A-11034/1:250
AQP5	Alomone	Rabbit IgG	Unconjugated	AQP-005/1:500
CD140b	Miltenyi Biotec	Human IgG1	APC	130-121-128/1:50

CD31	Santa Cruz	Mouse IgG1	Unconjugated	sc-376764/1:50
CD34	STEMCELL Technologies	Mouse IgG1	FITC	60013Fl.1/1:20
CD43	STEMCELL Technologies	Mouse IgG1	APC	60085AZ.1/1:20
CD45	STEMCELL Technologies	Mouse IgG1	PE	60018PE.1/1:20
CXCR4	Miltenyi Biotec	Human IgG1	APC	130-120-778/1:50
CXCR4	R&D	Mouse IgG _{2B}	Unconjugated	MAB172/1:200
E-CAD	Cell Signaling Technologies	Rabbit IgG	Unconjugated	3195/1:200
OCT3/4	Miltenyi Biotec	Human IgG1	APC	130-117-821/1:50
Pan-cytoKRT	Thermo Fisher	Mouse IgG1	Unconjugated	MA1-82041/1:100
PAX6	Miltenyi Biotec	Human IgG1	APC	130-123-328/1:50
PAX6	Santa Cruz	Mouse IgG1	Unconjugated	sc-81649/1:200
SATB2	Santa Cruz	Mouse IgG _{2A}	Unconjugated	sc-518006/1:200
SOX17	Miltenyi Biotec	Human IgG1	Vio B515	130-111-147/1:50
SOX17	Miltenyi Biotec	Human IgG1	PE	130-111-032/1:50
SOX2	Miltenyi Biotec	Human IgG1	FITC	130-120-790/1:50
SOX2	Miltenyi Biotec	Human IgG1	PE	130-121-053/1:50

SOX2	Cell Signaling Technologies	Rabbit IgG	Unconjugated	3579/1:200
SSEA-4	Miltenyi Biotec	Human IgG1	VioGreen	130-124-073/1:50
T/BRACHYURY	R&D	Goat IgG	Unconjugated	AF2095/1:800
TBR1	Abcam	Rabbit IgG	Unconjugated	ab183032/1:200
TBR2	Cell Signaling Technologies	Rabbit IgG	Unconjugated	81493/1:200
TTF1	Thermo Fisher	Mouse IgG1	Unconjugated	MA5-31938/1:500
Tyrosine Hydroxylase	Merck Millipore	Rabbit IgG	Unconjugated	AB152/1:500
VE-CAD	Thermo Fisher	Mouse IgG1	Unconjugated	14-1449-82/1:100
VIM	Cell Signaling Technologies	Rabbit IgG	Unconjugated	5741/1:200
β 3 Tubulin	Sigma	Mouse IgG _{2A}	Unconjugated	T8578/1:2000

Supplementary Table 6: Primer sequences used in this study.

Target	Forward primer (5'->3')	Reverse primer (5'->3')	Amplicon size (bp)
<i>ACTB</i>	TGAGGCACTCTTCCAGCCTTC	CGGCAATGCCAGGGTACATG	158
<i>ALB</i>	GATGAGATGCCTGCTGACTTGC	CACGACAGAGTAATCAGGATGCC	147
<i>ALPL</i>	GCTGTAAGGACATCGCCTACCA	CCTGGCTTTCTCGTCACTCTCA	131
<i>APLNR</i>	TTGCAGAGTGGGTGACAGAG	CTGGTTGTCTGCCCCATAGT	122

<i>AQP5</i>	TCCATTGGCCTGTCTGTCAC	AACCGATTCATGACCACCGC	107
<i>CD14</i>	CTGGAACAGGTGCCTAAAGGAC	GTCCAGTGTGAGGTTATCCACC	120
<i>CD151</i>	GGAGAACCTGAAGGACACCATG	CAGTCCTGTGAGTTGTTGCTGC	123
<i>CER1</i>	CCCATCAAAAGCCATGAAGT	AATGAACAGACCCGCATTTTC	133
<i>CNMD</i>	CCGTGACCAAACAGAGCATCTC	CTGTTGTCCTTCACAGGCTGATC	112
<i>CREBBP</i>	AGTAACGGCACAGCCTCTCAGT	CCTGTCGATACAGTGCTTCTAGG	115
<i>DLX5</i>	TACCCAGCCAAAGCTTATGCCG	GCCATTCACCATTCTCACCTCG	138
<i>EN1</i>	GTGGTCAAACTGACTCGCAGC	CCGCTTGTCTCCTTCTCGTTC	135
<i>ENO3</i>	TGGGAAGGATGCCACCAATGTG	GCGATAGAACTCAGATGCTGCC	160
<i>EOMES</i>	GGCTGTCTCCTAGCAACTCC	GCATAATACCCTCCCATGCCT	112
<i>EP300</i>	GATGACCCTTCCCAGCCTCAAA	GCCAGATGATCTCATGGTGAAGG	151
<i>FN1</i>	ACAACACCGAGGTGACTGAGAC	GGACACAACGATGCTTCCTGAG	143
<i>FOXJ1</i>	GTTTGGGTTTGGTGGTTTGG	CCAGTTAAGCCTCAGCTACAG	99
<i>GAL</i>	TGGCAACCACAGGTCATTCAGC	TCAGGTATGGACCTGTCAAAGCT	105
<i>GAPDH</i>	GTCTCCTCTGACTTCAACAGCG	ACCACCCTGTTGCTGTAGCCAA	131
<i>GATA3</i>	GCGATAGAACTCAGATGCTGCC	TCGGTTTCTGGTCTGGATGCCT	132

<i>GATA6</i>	TGTGCGTTCATGGAGAAGATCA	TTTGATAAGAGACCTCATGAACCGACT	83
<i>GDF3</i>	TCTCCCGAGACTTATGCTACG	AGTAGAGGAGCTTCTGCAGG	135
<i>HAND1</i>	ACATCGCCTACCTGATGGAC	CGGCTCACTGGTTAACTCC	225
<i>HES5</i>	CTGCTCAGCCCCAAAGAG	GCTCGATGCTGCTGTTGAT	85
<i>HOXB7</i>	ATCTACCCCTGGATGCGAAGCT	GCGTCAGGTAGCGATTGTAGTG	115
<i>KRT5</i>	CATGGAGATTGCCTCTTCTAGG	TCTCCTGGGAACCAAAGAATG	109
<i>LAMC1</i>	CTGTGAGGTCAACCACTTTGGG	AGCCTTCTCTGCATTCACAGCG	119
<i>NANOG</i>	TCCAACATCCTGAACCTCAG	ACCATTGCTATTCTTCGGCC	108
<i>NCAM1</i>	CCAGCTGACCATCAAAAAGG	ATTCCATGGCAGTCTGGTTC	152
<i>NES</i>	TTAATGGCCAGGGTCCCAAC	GCTCCAGCCCGTTCATCACT	72
<i>OTX2</i>	GCCAATCCTTGTTGAATCTTAGG	CAATCAGTCACACAATTCACACAGC	120
<i>PAMR1</i>	TTGCCAGCAGAATGGAGAGTGG	CTTGACTGAACCTGCATCGGAAG	117
<i>PAX6</i>	GAGGTCAGGCTTCGCTAATG	TTGCTTGAAGACCACAATGG	183/211
<i>PECAM1</i>	AAGTGGAGTCCAGCCGCATATC	ATGGAGCAGGACAGGTTTCAGTC	133
<i>POU5F1</i>	AAAGCGAACCAGTATCGAGAAC	GCCGGTTACAGAACCACACT	146/149
<i>SOX17</i>	TCATGGTGTGGGCTAAGGA	TTGTAGTTGGGGTGGTCCTG	184

SOX9	TGACCTATCCAAGCGCATTAC	GCTTGCTCTGAAGAGGGTTTA	117
SPP1	CGAGGTGATAGTGTGGTTTATGG	GCACCATTCAACTCCTCGCTTTC	128
TF	TCAGCAGAGACCACCGAAGACT	GACCACACTTGCCCGCTATGTA	103
THY1	GAAGGTCCTCTACTTATCCGCC	TGATGCCCTCACACTTGACCAG	143
TP63	AACAGCCATGCCCAGTATGTA	CGGTTTCATCCCTCCAACACA	146
TTF1	GCAAAGAGGACTCGCTTGTA	AGTGACAAGTGGGTTATGTTGA	103
TTR	CGTGTCATGTGTTTCAGAAAGGCTG	CTCCTCAGTTGTGAGCCCATGC	100
VIM	AATCCAAGTTTGCTGACCTCTCTG	TCATTGGTTCCTTTAAGGGCATCC	139

Supplementary Table 7: Sequencing parameters of long-read sequencing experiments.

Sample	Reads	Mbs	Mean length	Detected transcripts
Ectoderm n1	873,070	1,096	1255.4	10,424
Ectoderm n2	1,232,615	1,560	1265.9	10,640
Endoderm n1	970,694	1,178	1213.1	10,627
Endoderm n2	793,478	936.8	1180.7	10,463
Mesoderm n1	1,026,458	1,312	1278.5	10,778
Mesoderm n2	1,381,219	1,348	976.2	10,820
Undifferentiated n1	807,913	1,006	1245.8	10,458
Undifferentiated n2	1,115,741	1,470	1317.4	10,633
E2	179,059	189.2	1056.4	7,420
E3	221,753	238	1073.5	7,521
M2	137,288	89.8	653.9	7,270
M3	131,570	141	1071.5	7,396
Ec2	150,464	167.1	1110.8	7,463
Ec3	167,761	191.8	1143	6,760

Ec5	182,564	200.3	1097.3	6,792
P0	137,537	142	1032.2	6,792

Supplementary References

1. Ma, X. *et al.* Deterministically patterned biomimetic human iPSC-derived hepatic model via rapid 3D bioprinting. *Proceedings of the National Academy of Sciences* **113**, 2206–2211 (2016).
2. Binder, S., Ramachandran, H., Hildebrandt, B., Dobner, J. & Rossi, A. Prime-Editing of human ACTB in induced pluripotent stem cells to model human ACTB Loss-of-Function diseases and compensatory mechanisms. *Stem Cell Res* **75**, 103304 (2024).
3. Dobner, J. *et al.* Mitochondrial DNA integrity and metabolome profile are preserved in the human induced pluripotent stem cell reference line KOLF2.1J. *Stem Cell Reports* **19**, 343–350 (2024).
4. Martins, S. *et al.* Generation of an induced pluripotent stem cell line (IUFi001) from a Cockayne syndrome patient carrying a mutation in the ERCC6 gene. *Stem Cell Res* **55**, (2021).
5. Tigges, J. Academic application of Good Cell Culture Practice for induced pluripotent stem cells. *ALTEX* **38**, 595-614 (2021).
6. Lorenz, C. *et al.* Generation of four iPSC lines from four patients with Leigh syndrome carrying homoplasmic mutations m.8993T > G or m.8993T > C in the mitochondrial gene MT-ATP6. *Stem Cell Res* **61**, 102742 (2022).
7. Lickfett, S. *et al.* High-content analysis of neuronal morphology in human iPSC-derived neurons. *STAR Protoc* **3**, 101567 (2022).
8. Pantazis, C. B. *et al.* A reference human induced pluripotent stem cell line for large-scale collaborative studies. *Cell Stem Cell* **29**, 1685-1702.e22 (2022).
9. Szepanowski, L.-P. *et al.* Cockayne Syndrome Patient iPSC-Derived Brain Organoids and Neurospheres Show Early Transcriptional Dysregulation of Biological Processes Associated with Brain Development and Metabolism. *Cells* **13**, 591 (2024).
10. Takahashi, K. *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**, 861–872 (2007).
11. Teo, A. K. K. *et al.* Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev* **25**, 238–250 (2011).
12. Fisher, J. B., Pulakanti, K., Rao, S. & Duncan, S. A. GATA6 is essential for endoderm formation from human pluripotent stem cells. *Biol Open* **6**, 1084-1095 (2017) doi:10.1242/bio.026120.
13. Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* **27**, 275–280 (2009).
14. Kuang, Y.-L. *et al.* Evaluation of commonly used ectoderm markers in iPSC trilineage differentiation. *Stem Cell Res* **37**, 101434 (2019).