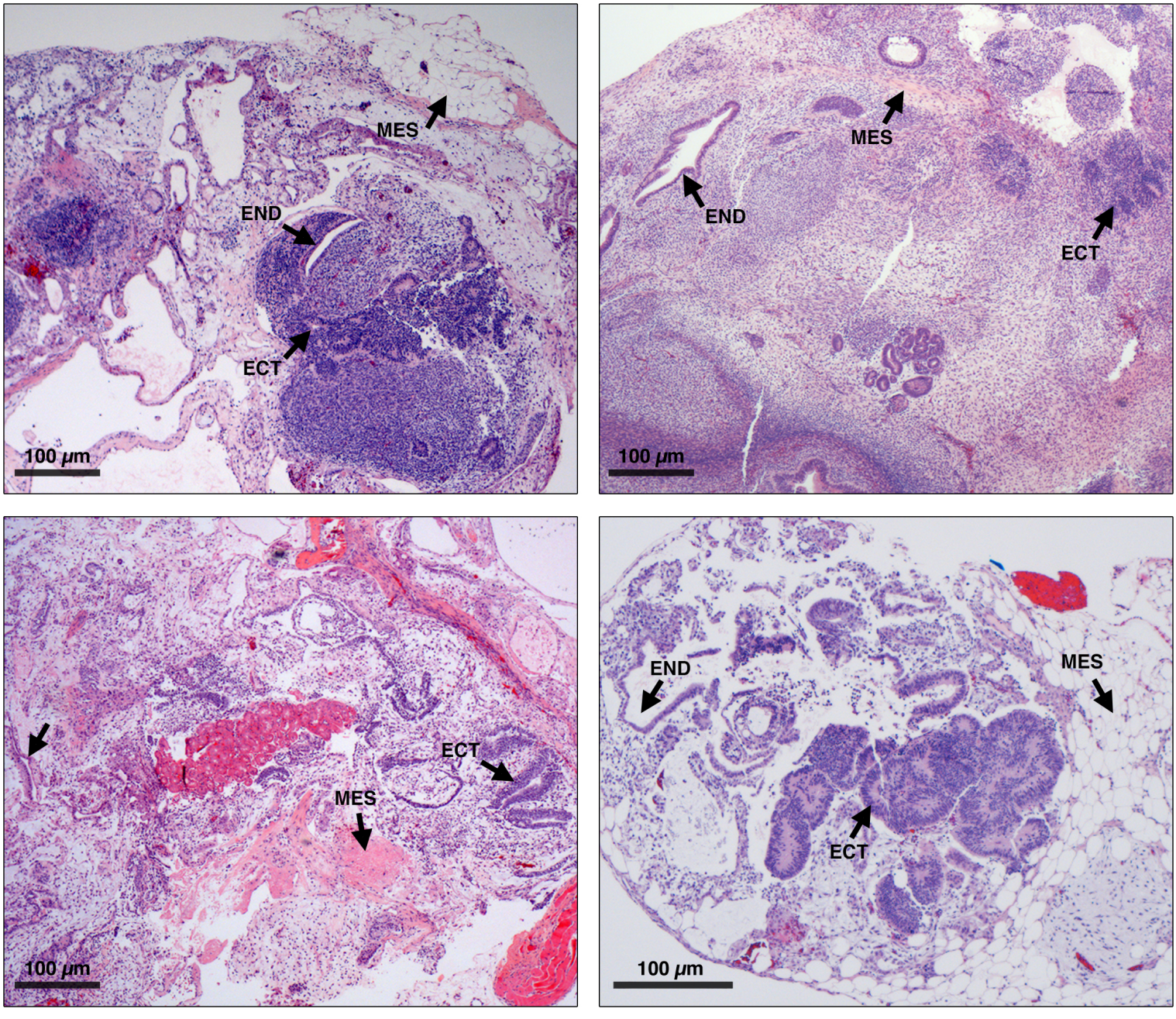
Stem Cell Reports, Volume *4*

Supplemental Information

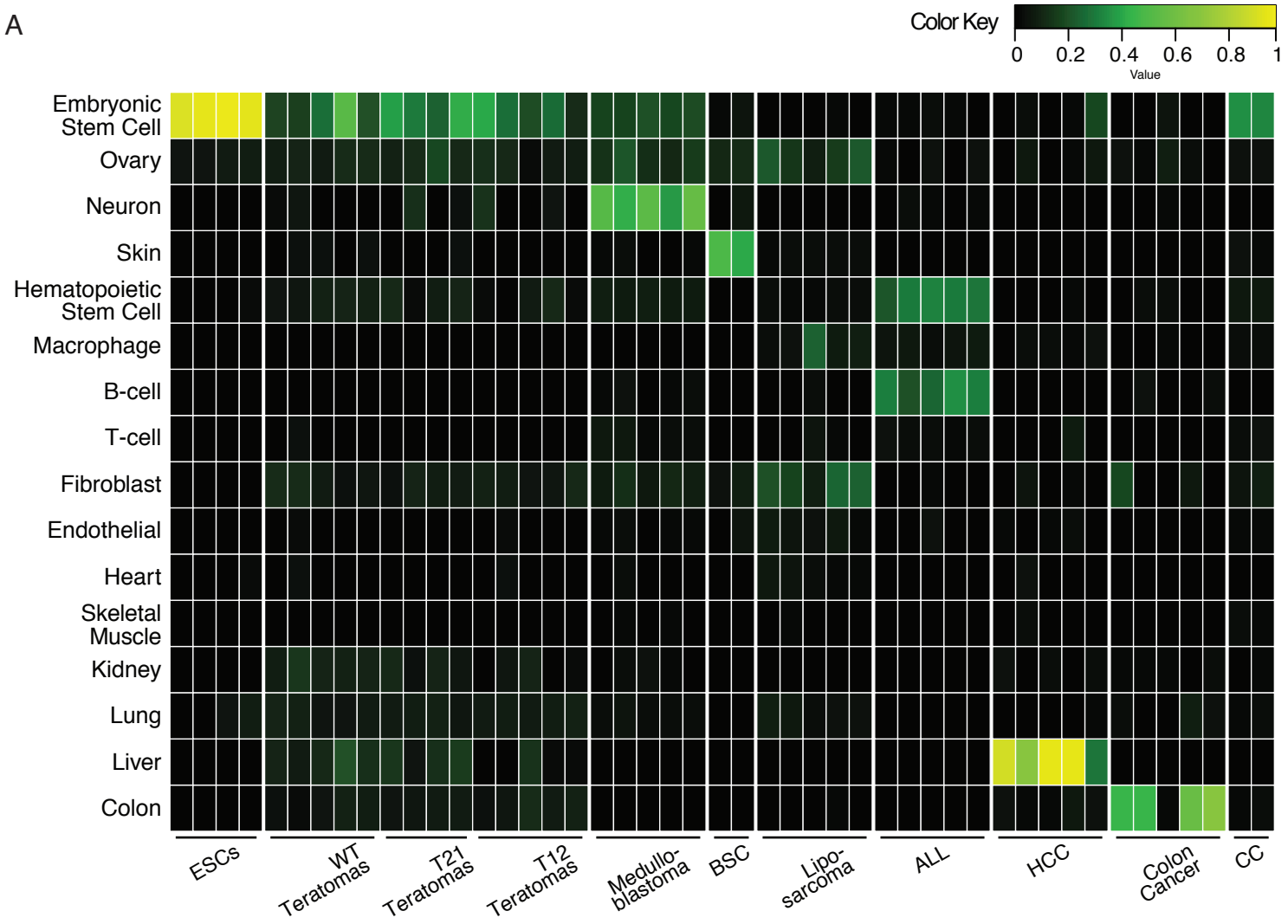
TeratoScore: Assessing the Differentiation Potential of Human Pluripotent Stem Cells by Quantitative Expression Analysis of Teratomas

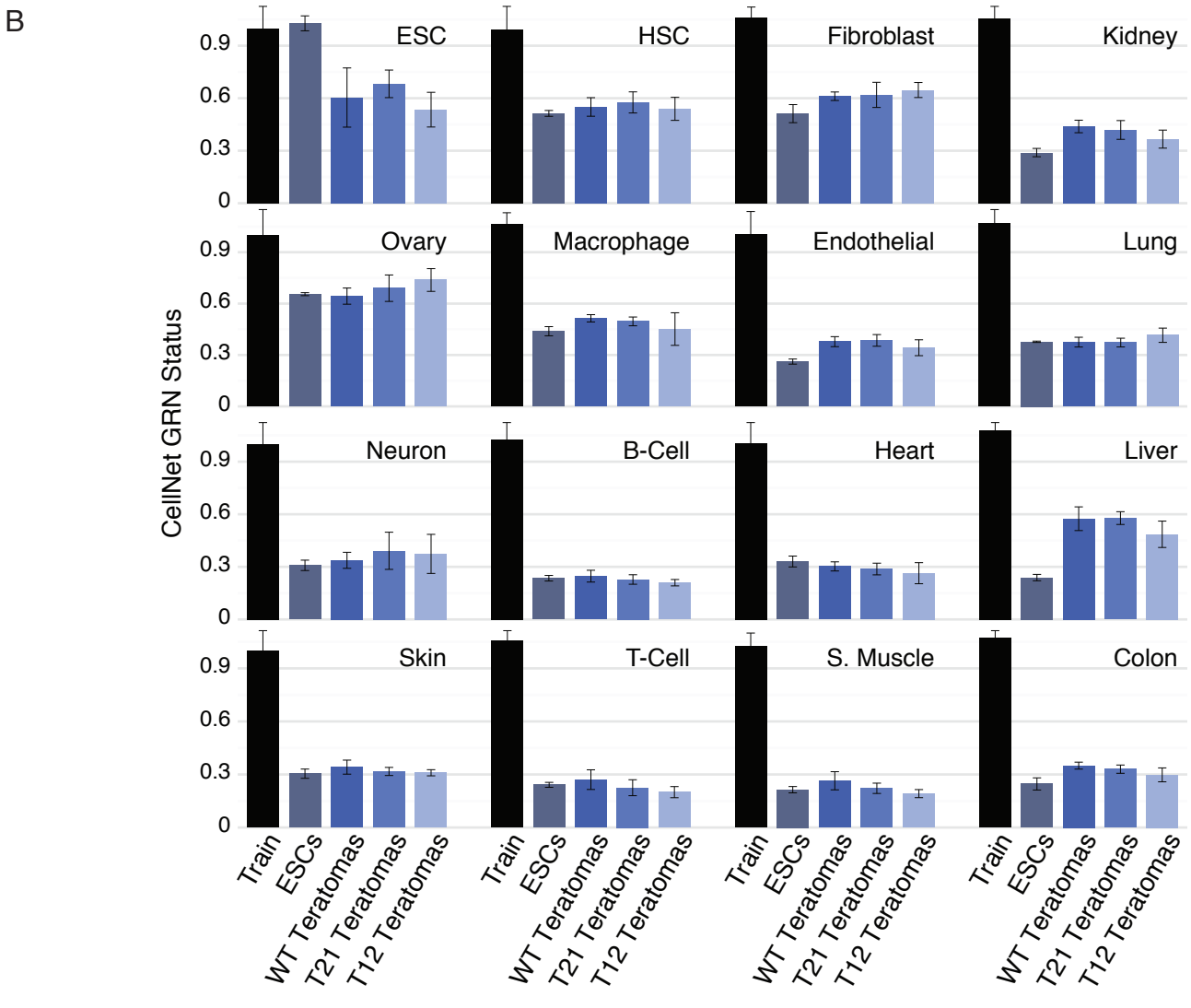
Yishai Avior, Juan Carlos Biancotti, and Nissim Benvenisty

Supplementary Figure 1. Representative Histological Sections of Teratomas Showing Structures from All Three Germ Layers. The teratomas reported in this paper were heterogeneously comprised, with structures originating from endoderm, mesoderm and ectoderm lineages (marked with arrows). 4-μm sections of teratomas stained with hematoxylin and eosin are shown. The teratomas were initiated from CSES32 (top right), CSES21 (top left), CSES26 (bottom left) and CSES25 (bottom right) cells. Abbreviations: END: Endoderm, MES: Mesoderm, ECT: Ectoderm.


MES
END
ECT
100 μm
MES
END
ECT
100 μm
ECT
MES
100 μm
END
MES
ECT
100 μm

Supplementary Figure 2. CellNet Online Resource Output for hESC, Teratoma and Tumor data. Expression data sets from four human embryonic stem cell lines (ESCs), 14 teratomas and seven different somatic tumors were uploaded to the online CellNet resource (Cahan et al., 2014). While ESCs show a clear resemblance to the ESC categories, all teratomas do not show a distinct similarity to any of the cell types measured, as shown in both a heat map analysis of the individual data sets (A) and histograms of gene regulatory networks found in each teratoma-group representing different cell types (B). Somatic tumors do show a distinct classification according to their initiating tissue (A). Abbreviations: ESC: embryonic stem cells; BSC: basal cell carcinoma, ALL: acute lymphoblastic leukemia, HCC: hepatocellular carcinoma, CC: choriocarcinoma, HSC: Hematopoietic Stem Cells, S. Muscle: Skeletal Muscle.


A
Color Key
0
0.2
0.4
0.6
0.8
1
Value
Embryonic Stem Cell
Ovary
Neuron
Skin
Hematopoietic Stem Cell
Macrophage
B-cell
T-cell
Fibroblast
Endothelial
Heart
Skeletal Muscle
Kidney
Lung
Liver
Colon
ESCs
WT Teratomas
T21 Teratomas
T12 Teratomas
Medullo-blastoma
BSC
Lipo-sarcoma
ALL
HCC
Colon Cancer
CC


B
CellNet GRN Status
ESC
HSC
Fibroblast
Kidney
Ovary
Macrophage
Endothelial
Lung
Neuron
B-Cell
Heart
Liver
Skin
T-Cell
S. Muscle
Colon
0
0.3
0.6
0.9
Train
ESCs
WT Teratomas
T21 Teratomas
T12 Teratomas

Supplementary Figure 3. TeratoScore Grades of Different Normal Tissues. Average tissue gene expression was calculated in respect of its expression values in the original tissue. Lineage expression was calculated as the mean expression of all genes representing tissues in that lineage. Error bars represent standard error. Each tissue shows a correct specific tissue and lineage expression. All tissues produced a low TeratoScore value, reflecting the lack of pluripotent origin.

Central and Peripheral Nervous System

TeratoScore Value: 0.8±0.2

Skin

TeratoScore Value: 6±0.4

Adipose Tissue

TeratoScore Value: 0.9±0.1

Bone Marrow

TeratoScore Value: 0.4±0.0

Liver

TeratoScore Value: 4.1±0.6

Colon

TeratoScore Value: 0.9±0.4

Placenta

TeratoScore Value: 6.5±0.5

Average % Expression of Original Tissue / Lineage

C/PNV
Skin
Adipose Tissue
Bone Marrow
Kidney
Joint
Heart
Skeletal Muscle
Liver
Gut
Lung
Pancreas
Placenta
Ectoderm
Mesoderm
Endoderm
Extraembryonic

Supplementary Table 1. Teratoma Expression Profiles' GEO Sample IDs. Cell lines used to initiate the teratomas used in this study are listed, as well as their karyotype classification and Gene Expression Omnibus (GEO) sample numbers.

Supplementary Table 1

Cell line	Karyotype	Geo Series ID	GEO Sample ID	Reference
CSES7	46, XX	GSE51455	GSM1246012	*
CSES12	46, XX	GSE64669	GSM1576685	*This Study*
CSES15	46, XX	GSE51455	GSM1246013	*
CSES26	46, XX	GSE51455	GSM1246015	*
CSES25	46, XX	GSE51455	GSM1246014	*
CSES22 (p16)	46, XX + 12	GSE64669	GSM1576686	*This Study*
CSES22 (p35)	46, XX + 12	GSE51455	GSM1246017	*
CSES45 (p20)	46, XX + 12	GSE51455	GSM1246018	*
CSES45 (p30)	46, XX + 12	GSE64669	GSM1576687	*This Study*
HiPS18	46, XX + 12	GSE51455	GSM1246019	*
CSES13	46, XX + 21	GSE64669	GSM1576688	*This Study*
CSES20	46, XX + 21	GSE64669	GSM1576689	*This Study*
CSES21	46, XX + 21	GSE64669	GSM1576690	*This Study*
CSES32	46, XX + 21	GSE64669	GSM1576691	*This Study*

***** *Ben-David et. al, 2014*

Supplementary Table 2. Expression Profiles of Tissues and Tumors Used in this Study. Expression profiles of tissues representing the three germ layers, the placenta and embryonic stem cells used to develop the TeratoScore are listed with their Gene Expression Omnibus (GEO) sample and series IDs. Tumors expression profiles used to compare TeratoScores are also listed with their Gene Expression Omnibus (GEO) sample and series IDs.

Supplementary Table 3. Genes Enriched for Tissues from Three Germ Layers and Placenta Used to Calculate TeratoScore. 100 genes were selected to represent tissues from all three germ layers and extraembryonic tissue in order to calculate tumors' TeratoScore. Genes enriched for each tissue were selected computationally and verified as being tissue-specific using Amazonia! online database (Le Carrour et al., 2010). Tissues with similar gene-expression were merged into a single, larger category. An additional set of ten genes representing pluripotent markers was selected. These genes are not taken into account when calculating TeratoScore. Listed are gene symbols, approved names and chromosomal locations.

Supplementary Table 4. Tissue-Specific Genes Up- and Downregulated in Teratomas Initiated from Aneuploid Pluripotent Cells. Tissue-specific genes differentially expressed in chromosomally aberrant teratomas compared to teratomas initiated from diploid cells. Out of the top 200 expressed genes in each tissue, a gene was considered differentially expressed when changed over 2-fold (either upregulated or downregulated) compared to control.

Supplemental Experimental Procedures

Cell Culture. Human embryonic stem cell lines CSES7, CSES 12, CSES13, CSES15, CSES20, CSES21, CSES22, CSES25, CSES26, CSES32 and CSES45 (Biancotti et al., 2010; Narwani et al., 2010) and iPSC line hiPS18 (Chin et al., 2009) were cultured using standard conditions as previously described (Mayshar et al., 2010). In brief, cells were grown on mouse embryonic fibroblast feeder (MEF) layer in KnockOut Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 15% KnockOut SR (Invitrogen), 1 mM glutamine, 0.1 mM b-mercaptoethanol (Sigma-Aldrich), 1% nonessential amino acids stock (Biological Industries), penicillin (50 U/ml), streptomycin (50 µg/ml) and FGF2 (4 ng/ml, Invitrogen). Cells were passaged using 0.25% trypsin solution (Biological Industries).

Karyotyping. A confluent, 10-cm plate of hESCs was used to analyze each cell line. Cells culture media were supplemented with KaryoMAX colcemid (100 ng/ml, Biological Industries) for 30 min at 37°C (with 5% CO2). Cultures were then trypsinized, treated with hypotonic solution and fixed using methanol:acetic acid solution (3:1). Metaphases were spread on glass slides and chromosomes were stained using standard G-banding technique. Chromosomal calssification was made using GenASIs software (Applied Spectral Imaging) according to the ISCN (International System for Human Cytogenetic Nomenclature).

Tumor Formation and Analysis. Undifferentiated hPSC colonies cultured on MEFs were trypsinized into single cells. ~3x106 cells were re-suspended in 50–100 µl of phosphate buffered saline (PBS) and injected under the kidney capsule of 6- to 8-week-old NOD-SCID mice. Three to four weeks post-injection, mice were sacrificed and resulting teratomas were extracted and dissected. Tumors were dissected into smaller pieces from different areas. Half of the tumor pieces were mixed together and taken for RNA extraction while the rest were saved for H&E staining, carried out as previously described (Kopper et al., 2010). The RNA portion was immediately homogenized in RLT buffer (RNeasy Mini kit, Qiagen) using a Polytron PT 1200 E system (Kinematica).

Microarray and Data Analysis. Total RNA was extracted from hPSC-derived teratomas using RNeasy mini kit (Qiagen), according to the manufacture instruction. RNA was subjected to Human Genome U133 Plus 2.0 microarray platform (Affymetrix); washing and scanning were performed according to the manufacturer's protocol. Original microarray data are accessible at the NCBI Gene Expression Omnibus database under the accession numbers GSM1576685 and GSE51455 (Table S1). Microarray data of the various body tissues, hPSC and previously published teratomas and tumors were downloaded from the Gene Expression Omnibus database (Table S2). Arrays were normalized using RMA algorithm in the Affymetrix Expression Console. Probe sets with expression values lower than 20 were raised to this level. The karyotype of the analyzed in-house cell lines was validated by virtual karyotyping (Ben-David et al., 2013).

Following averaging of samples from the same tissue, an expression ratio was calculated between each tissue and the average of all other body tissues, indicating tissue-specificity. Scorecard genes for each tissue were chosen by having a high tissue specificity (over 3 folds) and by being expressed in teratoma but not in hPSCs. Tissue-specific gene expression was further validated using Amazonia!, with a requirement for distinct tissue expression (an order of magnitude over all or most other tissues) (Le Carrour et al., 2010). Gene lists for central and peripheral nervous system, small intestine and colon, fetal brain and brain and fetal liver and liver were each merged together, as their specific-gene expression was considerably similar. Average expression for each lineage was calculated as the mean expression of all genes representing that lineage. TeratoScore grades were calculated as the multiplication of these means and dividing this product by 103. One-tail t-tests between tissue expressions were done using alpha 0.05. Top-200 tissue enriched genes were chosen as the top unique genes in the tissue/body ratio mentioned above. Genes were considered differentially expressed when changed over 2-fold compared to WT. To determine significance a chi-square independence test was performed followed by Bonferroni correction for multiple testing. Statistical calculations and figures were made in R.

Ben-David, U., Mayshar, Y., and Benvenisty, N. (2013). Virtual karyotyping of pluripotent stem cells on the basis of their global gene expression profiles. Nat Protoc 8, 989-997.

Ben-David, U., Arad, G., Weissbein, U., Mandefro, B., Maimon, A., Golan-Lev, T., Narwani, K., Clark, A.T., Andrews, P.W., Benvenisty, N., et al. (2014). Aneuploidy induces profound changes in gene expression, proliferation and tumorigenicity of human pluripotent stem cells. Nat Commun 5, 4825.

Biancotti, J.C., Narwani, K., Buehler, N., Mandefro, B., Golan-Lev, T., Yanuka, O., Clark, A., Hill, D., Benvenisty, N., and Lavon, N. (2010). Human embryonic stem cells as models for aneuploid chromosomal syndromes. Stem Cells 28, 1530-1540.

Cahan, P., Li, H., Morris, S.A., Lummertz da Rocha, E., Daley, G.Q., and Collins, J.J. (2014). CellNet: network biology applied to stem cell engineering. Cell 158, 903-915.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell 5, 111-123.

Kopper, O., Giladi, O., Golan-Lev, T., and Benvenisty, N. (2010). Characterization of gastrulation-stage progenitor cells and their inhibitory crosstalk in human embryoid bodies. Stem Cells 28, 75-83.

Le Carrour, T., Assou, S., Tondeur, S., Lhermitte, L., Lamb, N., Rème, T., Pantesco, V., Hamamah, S., Klein, B., and De Vos, J. (2010). Amazonia!: An Online Resource to Google and Visualize Public Human whole Genome Expression Data. Open Bioinformatics Journal 4.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., and Benvenisty, N. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell Stem Cell 7, 521-531.

Narwani, K., Biancotti, J.C., Golan-Lev, T., Buehler, N., Hill, D., Shifman, S., Benvenisty, N., and Lavon, N. (2010). Human embryonic stem cells from aneuploid blastocysts identified by pre-implantation genetic screening. In Vitro Cell Dev Biol Anim 46, 309-316.

Schuldiner, M., and Benvenisty, N. (2003). Factors Controlling Human Embryonic Stem Cell Differentiation. Methods Enzymol 365, 446-461.