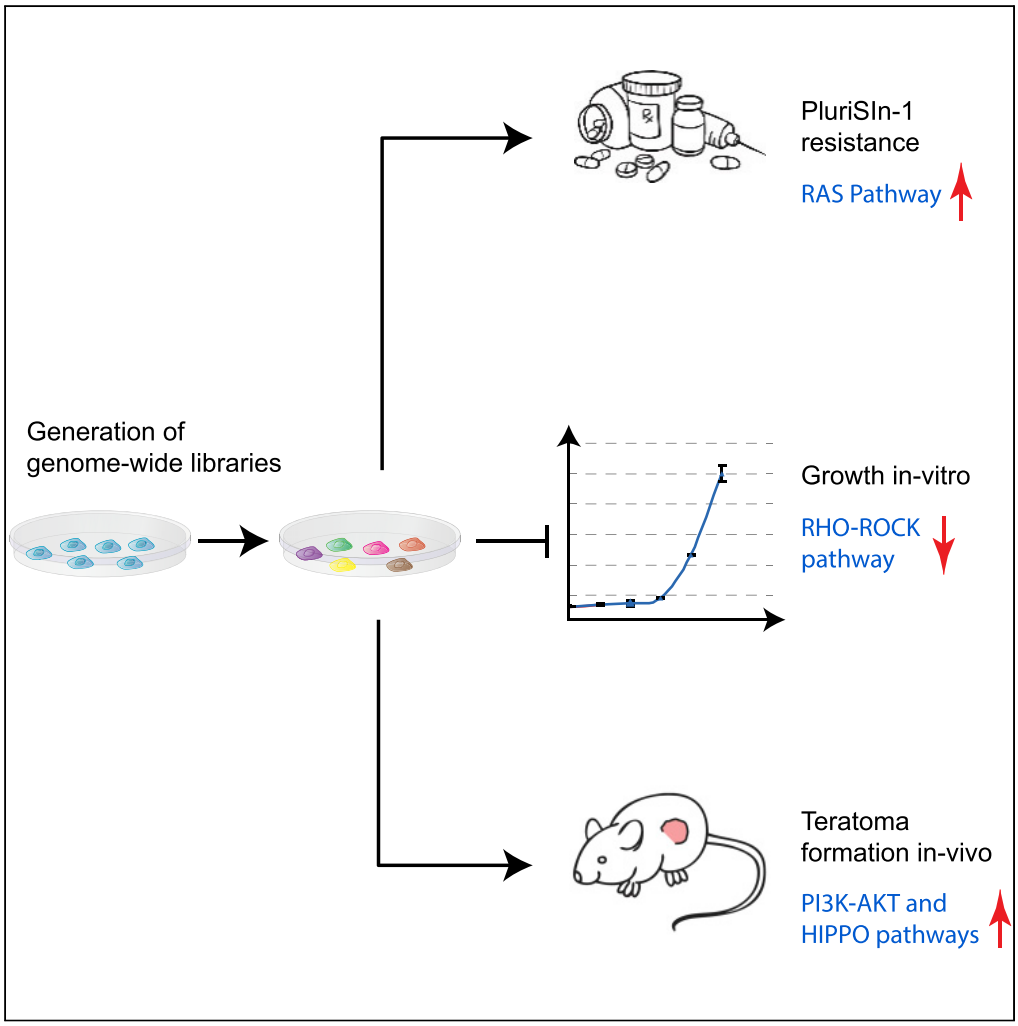
iScience

CellPress

Article

Genome-wide Screen for Culture Adaptation and Tumorigenicity-Related Genes in Human Pluripotent Stem Cells



Uri Weissbein, Mordecai Peretz, Omer Plotnik, Ofra Yanuka, Ido Sagi, Tamar Golan-Lev, Nissim Benvenisty

nissimb@mail.huji.ac.il

HIGHLIGHTS

Large-scale analysis of genes and pathways involved in growth and survival of hPSCs

Activation of the RAS pathways confers enhanced resistance to PluriSIn-1 treatment

Inactivation of the RHO-ROCK pathway gives selective growth advantage to hPSCs

The PI3K-AKT and HIPPO pathways are involved in the process of teratoma formation

Weissbein et al., iScience *11*, 398–408
January 25, 2019
https://doi.org/10.1016/j.isci.2018.12.031

iScience

CellPress

Article

Genome-wide Screen for Culture Adaptation and Tumorigenicity-Related Genes in Human Pluripotent Stem Cells

Uri Weissbein,[1] Mordecai Peretz,[1] Omer Plotnik,[1] Ofra Yanuka,[1] Ido Sagi,[1] Tamar Golan-Lev,[1] and Nissim Benvenisty[1,2,*]

SUMMARY

Human pluripotent stem cells (hPSCs) acquire genetic changes during their propagation in culture that can affect their use in research and future therapies. To identify the key genes involved in selective advantage during culture adaptation and tumorigenicity of hPSCs, we generated a genome-wide screening system for genes and pathways that provide a growth advantage either *in vitro* or *in vivo*. We found that hyperactivation of the RAS pathway confers resistance to selection with the hPSC-specific drug PluriSIn-1. We also identified that inactivation of the RHO-ROCK pathway gives growth advantage during culture adaptation. Last, we demonstrated the importance of the PI3K-AKT and HIPPO pathways for the teratoma formation process. Our screen revealed key genes and pathways relevant to the tumorigenicity and survival of hPSCs and should thus assist in understanding and confronting their tumorigenic potential.

INTRODUCTION

Human embryonic stem cells (hESCs) are isolated from the inner cell mass of the blastocyst. During embryonic development cells of the inner cell mass differentiate to give rise to the entire embryo. *In vitro*, hESCs can be maintained in their pluripotent state using defined culture conditions, or exit pluripotency and differentiate into desired cell types upon receiving external developmental cues. Thus human pluripotent stem cells (hPSCs) have an important role in biomedical research and future regenerative medicine, especially with the ability to reprogram somatic cells into induced pluripotent stem cells (Ben-David et al., 2012).

Several concerns regarding the genomic integrity of hPSCs *in vitro* and their tumorigenicity *in vivo* compromise their proper use. During prolonged culturing and culture adaptation, hPSCs tend to acquire various types of genetic aberrations. These acquired changes might alter the ability of hPSCs to self-renew, to respond to growth factors, and to differentiate, and they can lead to marked changes in their global gene expression (Ben-David and Benvenisty, 2011; Lund et al., 2012; Weissbein et al., 2014). The recurrent genetic abnormalities in hPSCs include gains of chromosomes 1, 12, 17, and X, and duplication of 20q11.21 and 12p13.31 (Baker et al., 2016; Ben-David and Benvenisty, 2011; Lefort et al., 2008; Lund et al., 2012; Mayshar et al., 2010; Närvä et al., 2010; Weissbein et al., 2014). However, the genes driving the positive selection of these alterations and the dramatic changes in the characteristics of the culture-adapted cells are largely unknown.

In vivo transplantation of hPSCs into immunodeficient mice results in tumors called teratomas, which consist of cells from all the three embryonic germ layers (Ben-David and Benvenisty, 2011). Although teratomas are benign tumors, genetic changes such as trisomy of chromosome 12 or duplication of the 20q11.21 region can enhance its aggressiveness (Ben-David et al., 2014; Werbowetski-Ogilvie et al., 2009). Although these tumors are known to be polyclonal, composed of differentiated cells that originate from multiple undifferentiated progenies (Blum and Benvenisty, 2007), the mechanisms underlying tumor formation remain almost completely unknown.

In this study, we apply a genome-wide screen on hPSCs to identify genes that confer selective advantage under various selective pressures. By using the PiggyBac (PB) transposon system, we generated libraries of hESCs with altered gene expression levels on a genomic scale. Using these libraries, we defined the main pathways responsible for selection during chemical treatment, prolonged culturing, and teratoma formation.

[1]The Azrieli Center for Stem Cells and Genetic Research, Department of Genetics, Silberman Institute of Life Sciences, The Hebrew University, Jerusalem, Israel

[2]Lead Contact

*Correspondence: nissimb@mail.huji.ac.il

https://doi.org/10.1016/j.isci.2018.12.031

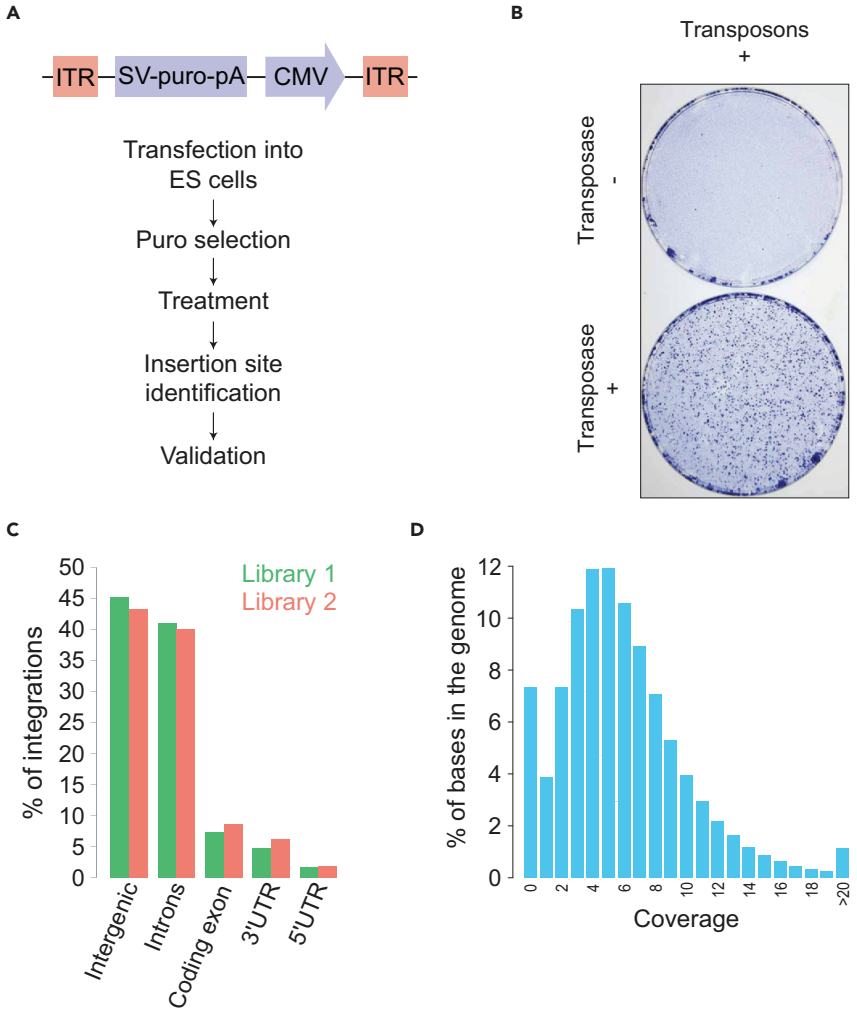


Figure 1. Preparation and Characterization of the PB Libraries

(A) Schematic representation of the constructs used to build the libraries, and the downstream experimental procedure.

(B) MEF culture plates of 10 cm with ESCs electroporated with the transposon construct and with or without the transposase followed by puromycin selection. The plates were stained with methylene blue.

(C) Location distribution of the transposon in different genomic features.

(D) The genomic distribution of integration potential coverage. Each integration was expanded in size 25 kb to each direction, and then the coverage at each position in the genome was calculated.

RESULTS

Construction of PiggyBac Overexpression Libraries

In our screen, we used a PB transposon construct containing a puromycin resistance gene followed by the cytomegalovirus (CMV) enhancer and promoter sequences surrounded by PB inverted terminal repeat sequences (Figure 1A). This system has been shown to have no particular bias toward certain genomic locations and to leave no trace sequence after excision (Chen et al., 2013; Copeland and Jenkins, 2010). Upon co-transfection with PB transposase, this construct may integrate into the genome and activate nearby genes, or alternatively reduce gene expression if integrated intragenically or in regulatory elements. This was previously demonstrated by picking single colonies and analyzing the integration sites parallel to gene expression (Chen et al., 2013). In the presence of transposase, we could achieve high integration efficiency and high number of individual colonies after selection (Figures 1B, S1A, and S1B). To determine integration sites we used splinkerette PCR, a procedure that enables direct amplification of the integration sequences (Uren et al., 2009) (see Methods).

We created two libraries, each containing ~2.5×10^5 individual integrations, named hereafter Library 1 and Library 2, suggesting a transposon integration within every ~10 kb. As the integrated CMV promoter and enhancer are strong inducers of gene expression, able to activate genes at a distance of over 50 kb (Chen et al., 2013), a given gene should be activated by five integrations on average.

To characterize the libraries, we extracted DNA from the total pool of cells in each library and added Illumina flow-cell-binding adaptors to the second splinkerette PCR primers. The PCR products were analyzed using Illumina next-generation sequencing, and the reads were mapped to the reference human genome. In both libraries, the integrations were distributed along the genome with a slight preference to transcribed regions, as was described previously (Ding et al., 2005) (Figure 1C). The mean distance between intergenic integration to the nearest gene is ~50 kb, a distance that allows activation by the CMV promoter (Figure S1C). To gain insight into the comprehensiveness of our libraries, we utilized the fact that the CMV enhancer can activate genes from a distance. We simulated the effective region of each insertion by expanding the integration size by 25 kb to each direction *in silico*. We found that less than 8% of the genome, which consists mainly of heterochromatic regions, is not covered by any integration, whereas most of the genome is covered by two to eight integrations (Figures 1D, S1D, and S1E). Overall, these results indicate that our overexpression libraries in hESCs have high genomic coverage.

Genes Involved in Resistance to PluriSIn-1 Drug Treatment

We initially utilized the mutant libraries for the analysis of resistance of hESCs to drug treatment. We thus used PluriSIn-1, which is a small molecule inhibitor of stearoyl-CoA desaturase-1 (SCD1), a key enzyme in fatty acid metabolism, triggering apoptosis specifically in pluripotent cells (Ben-David et al., 2013). PluriSIn-1 blocks the production of oleic acid (OA) and induces the endoplasmic reticulum (ER) unfolded protein response, leading to cell death by apoptosis (Ben-David et al., 2013). This drug was proved to be highly efficient in eliminating pluripotent cells from cell cultures and preventing teratoma formation (Ben-David et al., 2013).

To understand which mechanisms can confer resistance to PluriSIn-1, we treated our libraries with two rounds of 15 μM or 20 μM of PluriSIn-1. At the end of the experiment, all the samples were sequenced and analyzed by targeted DNA sequencing (DNA-seq) (Table S1). Global analysis of the samples revealed strong selection to certain integration sites following the treatment (Figures S2A and S2B). Unsupervised hierarchical clustering uncovered that in each library the treated samples clustered together and apart from the untreated control (Figure S2C). Principle component analysis (PCA) showed that the abundances of integrations before the treatments are relatively similar between the libraries, and different from those found after the treatments (Figure 2A). Next, we assembled a list of enriched integrations that could mediate the selection and defined the genes that might have been affected. We could assign genes for the majority of these integrations, although some integrations were located far from any known transcript and are probably passenger mutations (Figure S2D). Out of the hits, we found multiple integrations inside MED13A, a component of the mediator complex. MED13A was shown to be recruited to promoters by ATF6α, the master regulator of the ER unfolded protein response, to allow upregulation of stress response genes (Sela et al., 2012). In addition, we detected integration upstream of GNPNAT1, a target of the transducer of the unfolded protein response signal XBP1 (Wang et al., 2014). Intriguingly, 7 of the 27 significantly enriched integrations were related to the RAS signaling pathway including *MRAS* itself, guanine exchange factors (GEFs), and downstream targets (Figures 2B, 2C, and S2E). Overall, these integrations seemed to have an activating effect on the RAS signaling pathway (Table S2).

The initial screening with the libraries was performed on mouse embryonic fibroblast (MEF) feeder cells with ESC medium, which contains high levels of OA. However, elimination of hESCs grown on Matrigel-coated plates without MEFs and with the OA-poor mTeSR medium requires much lower concentration of PluriSIn-1 (Ben-David et al., 2013). This high sensitivity of the cells can be reduced by supplementing OA to the medium (Ben-David et al., 2013). To test whether RAS activation could confer resistance to PluriSIn-1, we cloned *MRAS* and transiently expressed it in hPSCs for 24 h (Figure 2D). Then, the cells were supplemented with different concentrations of OA and PluriSIn-1 for an additional 48 h followed by a viability test. Comparison of cells overexpressing *MRAS* to controls revealed greater resistance of *MRAS*-overexpressing cells to PluriSIn-1, when grown on mTeSR supplemented with 50 μM OA. This effect was dependent on the addition of OA to the medium, because it was not observed when supplementing the medium with 10 μM OA (Figure 2E). This significant resistance was observed even with transient

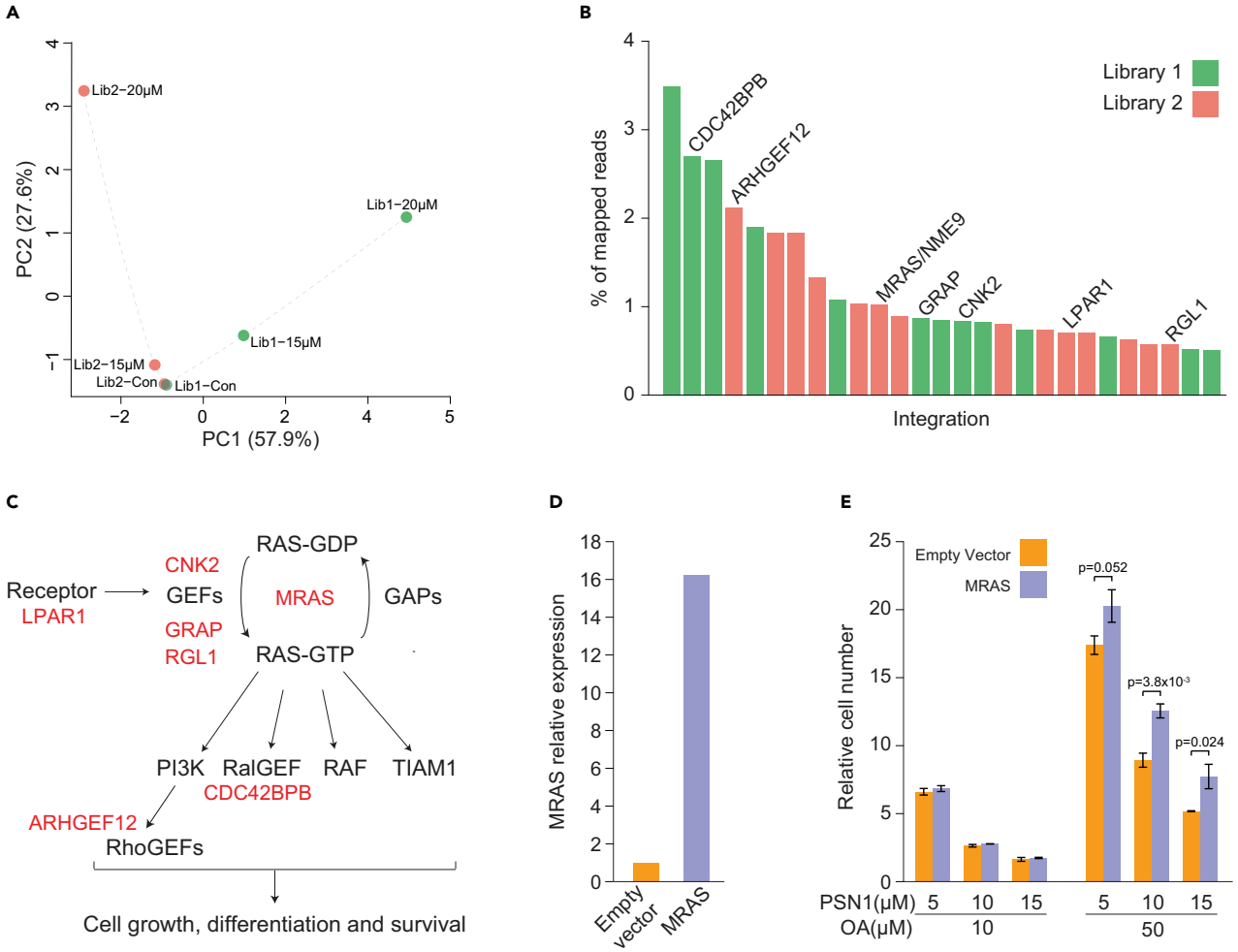


Figure 2. Activation of RAS Pathways Confers Resistance to PluriSIn-1

(A) PCA analysis of the integration patterns of the different PluriSIn-1 treatments.

(B) All the integrations that occupied more than 0.5% of the mapped reads at least at high concentrations of PluriSIn-1 and are more enriched in the 20-μM treatment than in the 15-μM treatment.

(C) Schematic illustration of the RAS signaling pathway. Genes with detected integrations are marked in red.

(D) Quantification of MRAS levels after transient overexpression.

(E) Relative cell number after treatment with PluriSIn-1 in medium supplemented with different concentrations of OA. Data presented are an average of three biological repeats. p values were calculated with a one-tailed Student's t test, standard error is shown as error bars. The experiment was performed in feeder-free Matrigel-coated plates.

transfection of the *MRAS* plasmid (~30% of cells are transfected), which suggests a much greater effect than measured. This result suggests that RAS pathway activation leads to enhanced resistance to SCD1 inhibition. As the protective effects of *MRAS* occurred only in high concentration of OA in the medium, it is likely mediated by OA import into the cells. These conclusions are in line with the knowledge obtained from RAS-driven cancers (Kamphorst et al., 2013; Salloum et al., 2014; White, 2013).

Genes Involved in Selection during Prolonged Culturing

To identify which genes are involved in the selection of hESCs during their culture adaptation, we grew both mutant libraries on plates with MEF feeder cells for 20 consecutive passages. Cells were collected every five passages and analyzed with high-throughput DNA-seq to determine the abundance of each of the mutant cells during the process of growth selection. The selection process affecting the cultured libraries was already evident from the change in the growth rate of cells (Figures S3A–S3C). PCA analysis and unsupervised hierarchical clustering of the integration patterns in each passage showed that both libraries

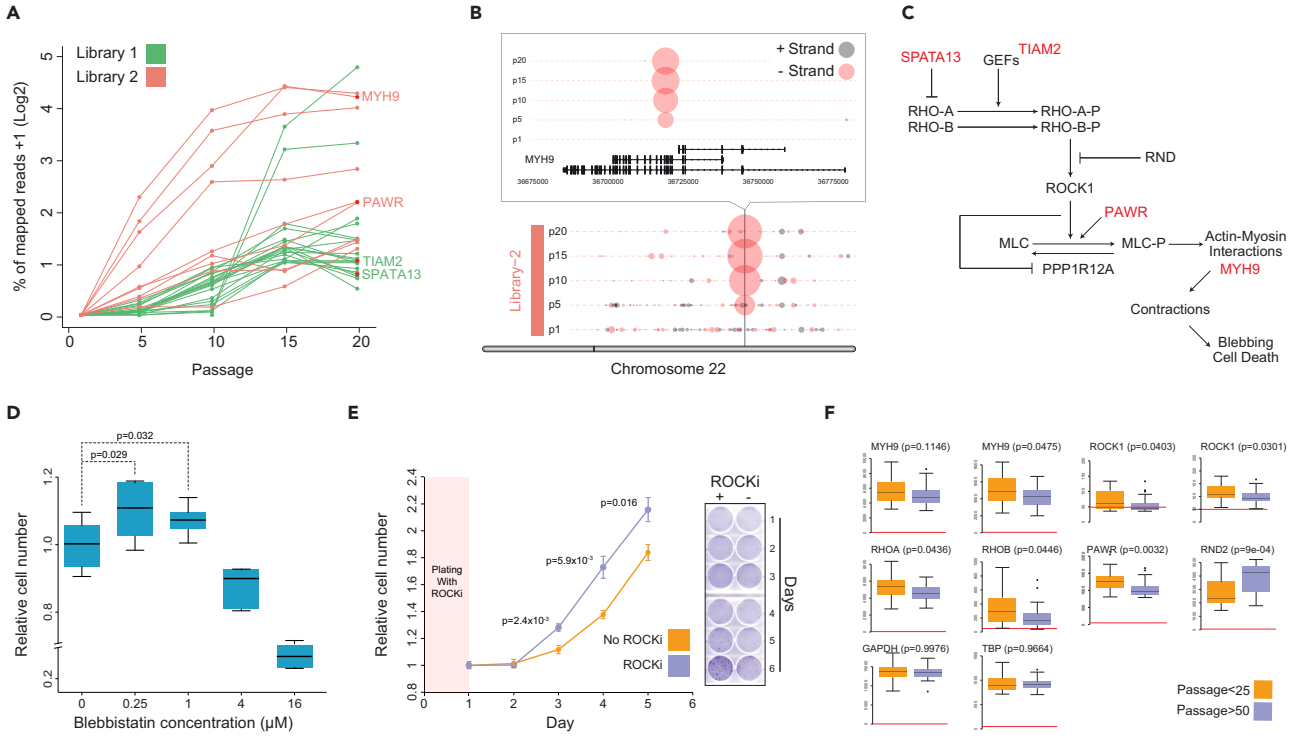


Figure 3. Changes in the PB Libraries during Long Culturing

(A) The dynamics of the integration patterns during prolonged culturing of both libraries is shown as the percentage of each integration out of total mapped reads. Integrations that constitute at least 1% of the mapped reads at passage 5 to 20 are shown.

(B) Dot plot of the entire map of integrations in chromosome 22 at different time points during culturing, with enlargement of the MYH9 locus. Each integration is plotted as a dot where the area represents the integration frequency out of the total mapped reads. Each horizontal line of dots corresponds to different passage. The color of each dot indicates the integration orientation.

(C) Schematic representation of the RHO-ROCK pathway. Genes with detected integrations are marked in red.

(D) Box plot depicting relative cell number after 2 days of treatment with increasing concentration of blebbistatin. A total of 10,000 cells were plated in Matrigel-coated 96-well plate with 6 repeats per concentration. Cell numbers were estimated using CellTiter-Glo. p values were calculated using a one-tailed Student's t test.

(E) Quantification of the effect of ROCKi on proliferation. On the left graph 1,000 cells were plated in a 96-well plate with 6 repeats per treatment. The right panel shows a growth assay demonstrating the effect of ROCKi on cell proliferation with 50,000 cells that were plated in a 6-well dish. In all experiments, ROCKi was added for the first 24 h. Staining and quantification were performed using methylene blue. p values were calculated using a two-tailed Student's t test, standard error is shown as error bars.

(F) Gene expression analysis of microarray of low- (p<=25) and high-passage (p ≥ 50) hPSCs revealing changes in gene expression in RHO/ROCK pathway genes. p values were calculated using a two-tailed Student's t test.

changed in a similar fashion during culturing (Figures S3D and S3E). In addition, we observed a gradual increase in the frequency of certain integrations throughout passaging (Figures 3A and S3F).

The increased growth rate of the cells can be attributed to several mechanisms that may occur simultaneously in these heterogeneous cultures. Some mutations may enhance proliferation, some may reduce apoptosis, and some may suppress spontaneous differentiation. Many mutations can be harmful and thus are eliminated from the population. To understand the effect of each mutation that increases cell growth (abundance higher than 1% of the mapped reads at passage five or later), we identified the potentially affected genes (Figure S3G). In our list, we found an integration upstream of the transcription factor SP3, which was shown to regulate expression levels of pluripotency master regulators NANOG and OCT4 in the mouse (Pesce et al., 1999; Wu and Yao, 2006); an activating integration near the insulin-like growth factor 1 receptor, which may confer enhanced growth and survival; and an integration in the first part of CDK13, which plays a role in regulating the cell cycle. In this gene list four genes were part of the RHO-Rho-associated kinase (ROCK) pathway (Table S3). One of the most highly enriched integrations was within the non-muscle myosin II heavy chain II-a (*MYH9*) gene, which is an important component of actin-myosin motors (Figure 3B). This gene, which encodes an important part of the cytoskeleton, is involved in multiple

processes, including cell motility and maintenance of cell shape (Quintin et al., 2008). Contractions are achieved by the binding of MYH9 to phosphorylated myosin light chain, which is activated by kinases such as the ROCK. ROCK is activated by phosphorylated RHO, which belongs to a family of GTPases (Amano et al., 2010). These proteins are active when phosphorylated by GEF proteins (Figure 3C). In addition to *MYH9*, two RHO-GEF proteins had enriched integrations: *TIAM2* had integrations that were mapped to the beginning of the gene body and *SPATA13* had an activating integration upstream of the promoter. SPATA13 acts in multiple pathways and was shown to inhibit RHOA activity upon overexpression (Bristow et al., 2009). Additional integration mapped inside PAWR, a gene involved in actin-myosin contractility (Vetterkind and Morgan, 2009; Vetterkind et al., 2010) (Figures 3C and S3H). hPSCs require a balance between the forces of the actin-myosin contractions and the opposing anchorage of the actin filaments through integrins and E-cadherins (Okeyo et al., 2009). Altered balance leads to cell blebbing and apoptosis (Figure 3C).

Inhibition of MYH9 by high concentrations of blebbistatin (a nonmuscle myosin II heavy chain inhibitor) or small interfering RNA (siRNAs), was previously shown to increase cell survival and cloning efficiency of hESCs after colony dissociation (Chen et al., 2010). Prolonged knockdown of MYH9 by siRNA was shown to suppress cell growth (Chen et al., 2010). As the integration inside MYH9 is heterozygous, we hypothesized that partial MYH9 inhibition may be beneficial for cell growth. To test this hypothesis, we treated hESCs with increasing concentrations of blebbistatin for 2 days. Indeed, high concentrations of blebbistatin efficiently inhibit cell viability, However, low concentrations of blebbistatin significantly improve cell growth (Figure 3D).

Inhibition of ROCK using a specific inhibitor (Y27632, ROCKi) was shown to prevent apoptosis after dissociation into single cells and to increase colony formation capabilities (Chen et al., 2010; Watanabe et al., 2007). To show that ROCK inhibition can increase cell growth, we plated cells with ROCKi for 24 h and then tested the effect of the ROCKi by maintaining or withdrawing the ROCKi supplement. We found that in addition to the enhancement of the cloning efficiency, inhibition of ROCK also provided a significant growth advantage (Figure 3E). These results were also seen after initial seeding of 48 h, ensuring that the effect is not on colony formation but on cell proliferation (Figure S3I).

Our results demonstrate that genetic modifications in the RHO-ROCK pathway can affect cell survival and growth, mimicking the effect of inhibition of ROCK by small molecules. To further demonstrate this effect, we compared microarray gene expression levels of low- and high-passage hPSCs (Nazor et al., 2012). The euploidy of the cell lines was validated by e-karyotyping (Weissbein et al., 2017). We discovered that activators of this pathway, such as *MYH9*, *RHO-A*, *RHO-B,* and *ROCK1*, are downregulated during culturing, whereas *RND2*, which is an inhibitor of ROCK1, becomes upregulated (Figure 3F). These results are not due to methylation differences between the groups (Figure S4A). To further validate these results, we performed an additional analysis with more than 700 RNA sequencing datasets (Figure S4B). As the vast majority of the studies does not report the passage number of the samples, we assessed the time in culture based on the expression of TSPYL5 and CAT, genes that are expressed in low-passage hPSCs and recurrently silenced by methylation during prolonged culturing (Konki et al., 2016; Weissbein et al., 2017) (Figures S4B and S4C). Again, we detected reduction in the RHO-ROCK pathways signaling (Figures S4D and S4E). Overall, we show that the RHO-ROCK pathway can be altered during culturing, and this alteration may lead to enhanced cell survival.

Genes Involved in Selection during Teratoma Formation *In Vivo*

In addition to selection *in vitro*, we used our libraries to examine cell selection *in vivo*. In this assay, as a complementary approach, we also utilized a gene-trap mutant library constructed after enrichment for haploid hESCs (Sagi et al., 2016). The gene-trap construct inserts a splicing acceptor followed by a puromycin resistance gene and a poly(A) signal. This library targets expressed genes and creates a complete loss of function upon integration. To further characterize this library, we amplified the integration sites and analyzed them by high-throughput DNA-seq. As expected, integrations were enriched in expressed genes, and most of the affected genes contained more than one independent integration (Figure 4A). To study the mechanisms governing teratoma formation, we injected the overexpression libraries, as well as the haploid knockout library, under the skin of severe combined immunodeficient mice and obtained tumors composed of cells from the three germ layers after 8 weeks (Figure S5A). From each teratoma, DNA was extracted from one to three distinct areas.

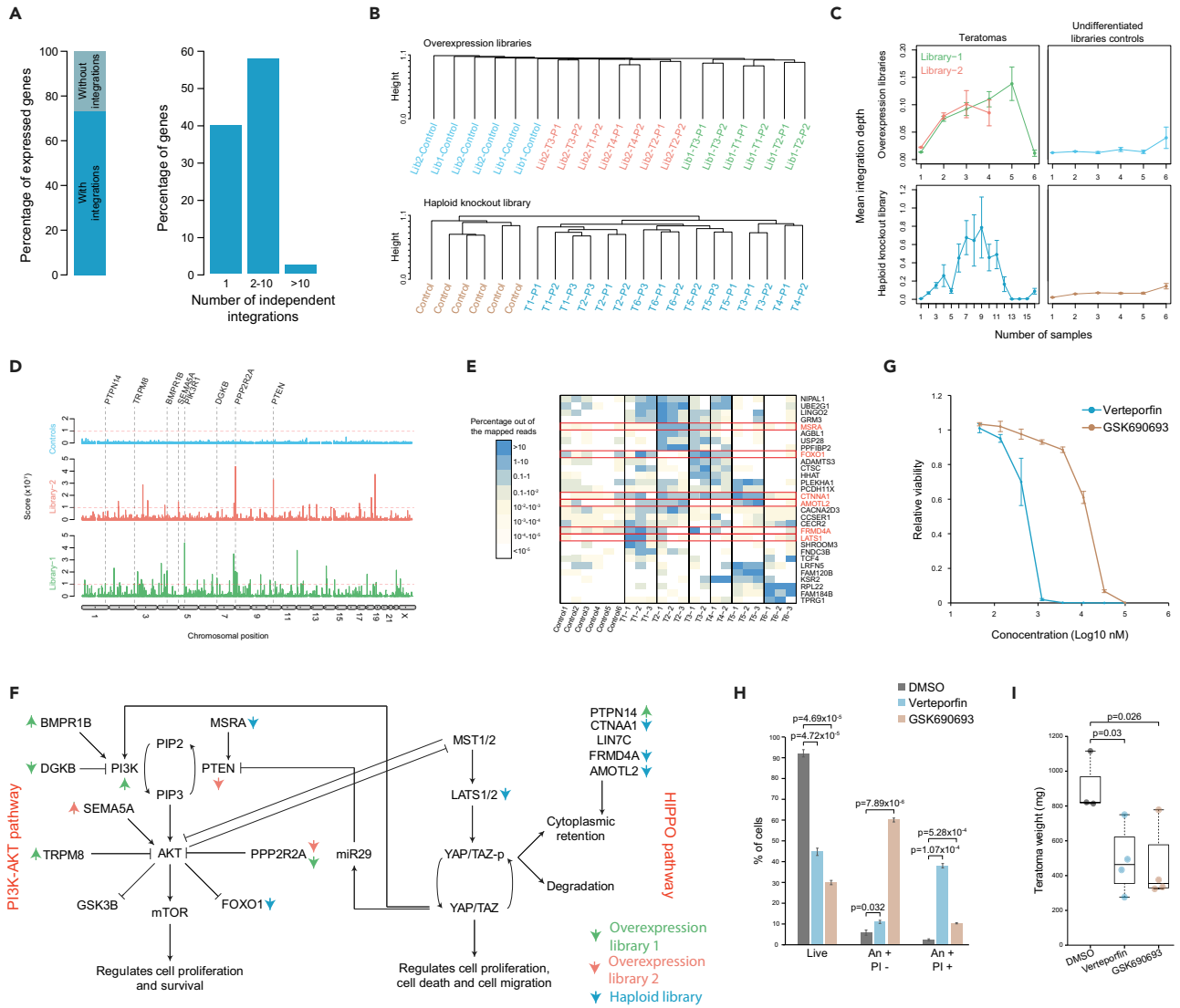


Figure 4. Selection of Specific Integrations during Teratoma Formation

(A) Analysis of the gene-trap mutant library in haploid hESCs. The left bar plot shows the percentage of expressed genes fragments Per Kilobase of transcript per Million mapped reads (FPKM)>0.5 with integrations. The right bar plot shows number of independent integrations in all the integration containing expressed genes.

(B) Unsupervised hierarchical clustering of the different pieces (P) from the different teratomas (T). Also included are undifferentiated libraries as controls.

(C) Recurrent integrations display higher coverage in the PB libraries, but not in the undifferentiated controls. In each group of samples, the integrations are divided according to the number of teratomas that each integration appears in (x axis). The y axis shows the mean percentage out of the total mapped reads of the integrations included in this group. Error bars represent standard error.

(D and E) (D) Specific genomic regions are important for teratoma formation. Shown are scores of 30,000-bp windows along the genome. The red dashed line marks the score that is 2-fold higher than the maximal score of the control group. The positions of the PI3K-AKT related genes are indicated in the graph. (E) Heatmap of the coverage of the genes shows that their scores are significantly higher in the teratoma samples than in the control. Highlighted are genes related to the PI3K-AKT and HIPPO pathways.

(F) Schematic illustration of the PI3K-AKT and HIPPO pathways. Genes affected in the screen are marked with arrows. The direction of the arrow marks up- or down-regulation.

(G) Cell viability curve after treating teratoma-derived cell line with the indicated drug for 96 h. The results are relative to control treated only with DMSO. Each time point is an average of three biological repeats. Error bars represent standard error.

(H) Quantification of cell death by annexin V (An) and propidium iodide (PI) staining. Cells were treated with 5 μM VP for 24 h, 20 μM GSK690693 for 72 h, or with DMSO for 72 h. Data presented are an average of three biological repeats. p values were calculated with a two-tailed Student's t test. Error bars represent standard error. (I) Tumor weights after the indicated treatments.

Analysis of the integration patterns by unsupervised hierarchical clustering showed that the teratomas clustered distinctly from control undifferentiated libraries, and samples from the same library, as well as different pieces from the same teratoma, clustered together (Figure 4B). These results show that although each teratoma developed separately, some integrations were shared among the different tumors and even between the different libraries. The different pieces of each tumor clustered together, but with clear distance, implying both shared and separated selection mechanisms. Our results support previous reports showing that hPSC-induced teratomas are polyclonal tumors derived from multiple precursor cells (Blum and Benvenisty, 2007). However, our results demonstrate that genetic changes may affect the extent to which each cell may contribute to the teratoma mass.

To determine whether integrations shared between different teratoma samples may be relevant to the tumor development, we analyzed all the teratoma samples together. We also included six DNA-seq results of the undifferentiated libraries before injection as controls. We found that integrations that appeared in multiple control samples did not show an increase in the mean sequencing depth, implying that they were not positively selected. However, integrations that appeared in multiple teratoma samples had higher mean sequencing depth, pointing to their positive selection in the tumors (Figure 4C).

To find genomic loci that could be relevant to teratoma formation in the overexpression libraries, we scored windows of 30 kb along the genome (see Methods). As opposed to the control samples, wherein no significantly enriched regions were detected, we could detect multiple regions with high score in both libraries (Figure 4D). Interestingly, a few of these regions corresponded to regulators of the phosphatidylinositol 3-kinase (PI3K)-AKT pathway, including key genes such as PTEN and PI3K (Figures 4D, S5B, and S5C and Table S4).

To identify genes whose knockout contributes to teratoma formation in the haploid library, we scored each gene both in the teratomas and in the control samples (see Methods). Among the genes with significantly higher scores in the teratomas, two genes were part of the PI3K-AKT pathway and four were in the HIPPO pathway (Figure 4E). The PI3K-AKT pathway and the HIPPO pathway cooperate to regulate cell proliferation, survival, and organ size, via a coordinated cross talk (Csibi and Blenis, 2012; Kelleher and O'Sullivan, 2014) (Figure 4F). These results suggest a functional role of the PI3K-AKT and HIPPO signaling pathways during teratoma formation and growth.

To validate the involvement of HIPPO and AKT pathways in the formation of the tumor, we treated a teratoma-derived cell line with either a HIPPO-YAP inhibitor, verteporfin (VP) (Feng et al., 2016; Liu-Chittenden et al., 2012; Wang et al., 2016; Zhang et al., 2015), or with an AKT1-3 inhibitor, GSK690693 (Altomare et al., 2010; Carol et al., 2010; Levy et al., 2009; Rhodes et al., 2008). Indeed, inhibition of these pathways led to massive cell death (Figures 4G and 4H). The EC50 values of VP and GSK690693 are 0.58μM and 13.8μM, respectively, concentrations that are not lethal to all cell types (Cani et al., 2015; Donohue et al., 2013), suggesting that functional HIPPO and AKT signaling pathways are required for the survival of teratoma-derived cells. We then injected hESCs into immunodeficient mice, and 5 days after the injection, we started treating the mice three times a week with either HIPPO or AKT inhibitor. Analysis of the tumors after 6 weeks revealed a significant decrease in the weight of teratomas treated with either inhibitor (Figure 4I). The treatments did not affect the differentiation potential of the cells as we detected structures from the three embryonic germ layers in the treated teratomas (Figure S5D). Overall, we demonstrated that these pathways are involved in the teratoma formation and growth.

DISCUSSION

hPSCs continuously acquire genomic changes during their growth in culture. These changes are the consequence of selection of specific clones that take over the culture. This may cause changes in the behavior of the cells, leading to false interpretation of experimental results and compromise future therapies. Genome-wide genetic screens enable the identification of genes that confer drug resistance and cellular selection *in vitro* and *in vivo*. PluriSIn-1 was suggested as a chemical compound that can eliminate pluripotent cells from differentiated cultures and reduce tumorigenic risks. Its specificity is achieved by the dependency of hPSCs on OA. Although it is highly specific to pluripotent cells, we show here that alteration in the RAS pathway can affect the sensitivity of the cells to this compound. These results strengthened the data obtained in RAS-driven cancers, which indicate reduced sensitivity of these tumors to SCD1 inhibitory treatments, due to increased import of lipid into the cells (Kamphorst et al., 2013; Salloum et al., 2014;

White, 2013). In addition, we suggest using growth medium completely lacking OA in future treatments of hPSCs with PluriSIn-1, which will render RAS-dependent resistance less effective.

By using the PB system, we sought to understand some of the molecular processes that take place during hPSC culture adaptation. We demonstrated the importance of the RHO-ROCK1 pathway in hPSCs survival. Although it is well established that ROCKi can be used during cell passaging to increase colony formation efficiency, we could show that inhibition of ROCK1 also increases hPSC growth. As this is the only pathway with multiple hits, we suggest that this is the main pathway involved in selection during culturing. We also showed that this pathway is altered in cultured cells at high passage, in the absence of any external genetic manipulation, supporting the importance of this pathway.

Our analysis emphasizes the strength of selection forces during growth in culture. In the long culturing experiment, already in a rather small number of passages (5 to 15 splits), few cells with advantageous integrations populated a high percentage of the culture. These rapid changes in the diversity of the population are a consequence of two factors: (1) the selective advantage of certain integrations and (2) the splitting process that drastically reduces the population size and eliminates cells with low representation in the cell culture (population bottleneck effect). Inhibition of the ROCK pathway (genetically or chemically) can alter the population bottleneck effect in both directions: increasing the genetic diversity because more cells survive the passage and reducing the genetic diversity because it allows harsher splitting. In addition, as we demonstrated, inhibition of the ROCK pathway affects the cell growth rate. Altogether, this demonstrates how quickly a mutation or chromosomal aberration can spread in the cell culture, emphasizing the importance of routine assays for genomic integrity.

Although we could detect significant hits, none of them mapped to the critical regions that drive the selection of the recurrently seen large chromosomal aberrations. This result may indicate that the selection advantage of our most significant hits is far greater than that of the genes that drive the selection of the common chromosomal aberrations. Alternatively, the positive selection of the recurrent large duplications may be generated by more than one gene, and thus cannot be detected in our assay.

Examination of the mechanisms leading to the formation of teratoma *in vivo* is extremely challenging. By following the positive selection in the tumor, we revealed the pathways that participate in tumor formation. We have shown the involvement of known oncogenes and pathways, especially the PI3K-AKT and HIPPO pathways in the formation and development of hPSC-induced teratomas. Previous experiments in mouse models have demonstrated that alterations in the PI3K-AKT pathway in the male primordial germ cells or in the female granulosa cells promote testicular or ovarian teratomas, respectively (Balakrishnan and Chaillet, 2013; Kimura et al., 2003). In line with these findings, our results show that the PI3K-AKT pathway is also involved in the development of human teratomas. Our results suggest that to reduce the risk of teratoma formation in future cell injection therapies, inhibition of PI3K-AKT and HIPPO pathway may be beneficial. Although the specific means of inhibition of these pathways should be calibrated, and the specific effects on the desired differentiated cell should be analyzed, these results suggest applicable approaches to reduce the risk of teratoma formation. Last, our results may have beneficial contribution to treating naturally occurring teratomas in humans, as they share induction mechanisms.

Limitations of the Study

Genome-wide genetic libraries are a powerful tool to perform unbiased screens for genetic factors that regulate specific phenotypes. Transposable-elements-mediated genetic screen, as opposed to CRISPR-Cas9-based screens, does not rely on prior selection of the targets; however, the effects of the integrations are less predictable. In addition, in the system used in this study, each cell received more than one transposon on average, which complicates the process of identifying the driving genes. Last, in our system, integrations can cause both overexpression and downregulation, depending on the relative location and orientation of the integration from the gene. We chose not to validate the effect of each enriched integration, but rather to bioinformatically predict the effect on gene expression from the integration location and orientation relative to the nearby genes. This approach might introduce some noise to the experimental system. To overcome these difficulties, in each part of the study, we focused on pathways in which we can find integrations in multiple genes and directly validated the results by chemical or genetic approaches. Additional experiments using genome-edited cell lines or single-cell analysis would provide further support for the observed results.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods, five figures, and five tables and can be found with this article online at https://doi.org/10.1016/j.isci.2018.12.031.

ACKNOWLEDGMENTS

We thank all members of The Azrieli Center for Stem Cells and Genetic Research for their input and critical reading of the manuscript. This work was partially supported by the US-Israel Binational Science Foundation (grant no. 2015089), by the Rosetrees Trust, and by the Azrieli Foundation. U.W. was supported by the Clore Fellowship., I.S. was supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, and N.B. is the Herbert Cohn Chair in Cancer Research.

AUTHOR CONTRIBUTIONS

U.W. and N.B. designed the experiments and wrote the manuscript. U.W. performed the bioinformatic analysis. U.W. and O.P. performed the experiments. M.P., I.S. and O.Y. developed the haploid knockout library and performed the experiments with it. T.G.-L. performed the sections and stainings of the tumors.

DECLARATION OF INTERESTS

The authors declare no competing interests. N.B. is the CSO of NewStem Ltd..

Received: September 25, 2018
Revised: November 19, 2018
Accepted: December 27, 2018
Published: January 25, 2019

REFERENCES

Altomare, D.A., Zhang, L., Deng, J., Di Cristofano, A., Klein-Szanto, A.J., Kumar, R., and Testa, J.R. (2010). GSK690693 delays tumor onset and progression in genetically defined mouse models expressing activated Akt. Clin. Cancer Res. *16*, 486–496.

Amano, M., Nakayama, M., and Kaibuchi, K. (2010). Rho-kinase/ROCK: a key regulator of the cytoskeleton and cell polarity. Cytoskeleton (Hoboken) *67*, 545–554.

Baker, D., Hirst, A.J., Gokhale, P.J., Juarez, M.A., Williams, S., Wheeler, M., Bean, K., Allison, T.F., Moore, H.D., Andrews, P.W., et al. (2016). Detecting genetic mosaicism in cultures of human pluripotent stem cells. Stem Cell Rep. *7*, 998–1012.

Balakrishnan, A., and Chaillet, J.R. (2013). Role of the inositol polyphosphate-4-phosphatase type II Inpp4b in the generation of ovarian teratomas. Dev. Biol. *373*, 118–129.

Ben-David, U., and Benvenisty, N. (2011). The tumorigenicity of human embryonic and induced pluripotent stem cells. Nat. Rev. Cancer *11*, 268–277.

Ben-David, U., Kopper, O., and Benvenisty, N. (2012). Expanding the boundaries of embryonic stem cells. Cell Stem Cell *10*, 666–677.

Ben-David, U., Gan, Q.-F., Golan-Lev, T., Arora, P., Yanuka, O., Oren, Y.S., Leikin-Frenkel, A., Graf, M., Garippa, R., Boehringer, M., et al. (2013). Selective elimination of human pluripotent stem cells by an oleate synthesis inhibitor discovered in a high-throughput screen. Cell Stem Cell *12*, 167–179.

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.

Blum, B., and Benvenisty, N. (2007). Clonal analysis of human embryonic stem cell differentiation into teratomas. Stem Cells *25*, 1924–1930.

Bristow, J.M., Sellers, M.H., Majumdar, D., Anderson, B., Hu, L., and Webb, D.J. (2009). The Rho-family GEF Asef2 activates Rac to modulate adhesion and actin dynamics and thereby regulate cell migration. J. Cell Sci. *122*, 4535–4546.

Cani, A., Simioni, C., Martelli, A.M., Zauli, G., Tabellini, G., Ultimo, S., McCubrey, J.A., Capitani, S., and Neri, L.M. (2015). Triple Akt inhibition as a new therapeutic strategy in T-cell acute lymphoblastic leukemia. Oncotarget *6*, 6597–6610.

Carol, H., Morton, C.L., Gorlick, R., Kolb, E.A., Keir, S.T., Reynolds, C.P., Kang, M.H., Maris, J.M., Billups, C., Smith, M.A., et al. (2010). Initial testing (stage 1) of the Akt inhibitor GSK690693 by the pediatric preclinical testing program. Pediatr. Blood Cancer *55*, 1329–1337.

Chen, G., Hou, Z., Gulbranson, D.R., and Thomson, J.A. (2010). Actin-myosin contractility is responsible for the reduced viability of dissociated human embryonic stem cells. Cell Stem Cell *7*, 240–248.

Chen, L., Stuart, L., Ohsumi, T.K., Burgess, S., Varshney, G.K., Dastur, A., Borowsky, M., Benes, C., Lacy-Hulbert, A., and Schmidt, E.V. (2013). Transposon activation mutagenesis as a screening tool for identifying resistance to cancer therapeutics. BMC Cancer *13*, 93.

Copeland, N.G., and Jenkins, N.A. (2010). Harnessing transposons for cancer gene discovery. Nat. Rev. Cancer *10*, 696–706.

Csibi, A., and Blenis, J. (2012). Hippo-YAP and mTOR pathways collaborate to regulate organ size. Nat. Cell Biol. *14*, 1244–1245.

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., and Xu, T. (2005). Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell *122*, 473–483.

Donohue, E., Thomas, A., Maurer, N., Manisali, I., Zeisser-Labouebe, M., Zisman, N., Anderson, H.J., Ng, S.S.W., Webb, M., Bally, M., et al. (2013). The autophagy inhibitor verteporfin moderately enhances the antitumor activity of gemcitabine in a pancreatic ductal adenocarcinoma model. J. Cancer *4*, 585–596.

Feng, J., Gou, J., Jia, J., Yi, T., Cui, T., and Li, Z. (2016). Verteporfin, a suppressor of YAP-TEAD complex, presents promising antitumor properties on ovarian cancer. Onco. Targets Ther. *9*, 5371–5381.

Kamphorst, J.J., Cross, J.R., Fan, J., de Stanchina, E., Mathew, R., White, E.P., Thompson, C.B., and Rabinowitz, J.D. (2013). Hypoxic and Ras-transformed cells support growth by scavenging unsaturated fatty acids from lysophospholipids. Proc. Natl. Acad. Sci. U S A *110*, 8882–8887.

Kelleher, F.C., and O'Sullivan, H. (2014). Oxford and the Savannah: can the hippo provide an explanation for Peto's paradox? Clin. Cancer Res. *20*, 557–564.

Kimura, T., Suzuki, A., Fujita, Y., Yomogida, K., Lomeli, H., Asada, N., Ikeuchi, M., Nagy, A., Mak, T.W., and Nakano, T. (2003). Conditional loss of PTEN leads to testicular teratoma and enhances embryonic germ cell production. Development *130*, 1691–1700.

Konki, M., Pasumarthy, K., Malonzo, M., Sainio, A., Valensisi, C., Söderström, M., Emani, M.R., Stubb, A., Närvä, E., Ghimire, B., et al. (2016). Epigenetic silencing of the key antioxidant enzyme catalase in karyotypically abnormal human pluripotent stem cells. Sci. Rep. *6*, 22190.

Lefort, N., Feyeux, M., Bas, C., Féraud, O., Bennaceur-Griscelli, A., Tachdjian, G., Peschanski, M., and Perrier, A.L. (2008). Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. Nat. Biotechnol. *26*, 1364–1366.

Levy, D.S., Kahana, J.A., and Kumar, R. (2009). AKT inhibitor, GSK690693, induces growth inhibition and apoptosis in acute lymphoblastic leukemia cell lines. Blood *113*, 1723–1729.

Liu-Chittenden, Y., Huang, B., Shim, J.S., Chen, Q., Lee, S., Anders, R.A., Liu, J.O., and Pan, D. (2012). Genetic and pharmacological disruption of the TEAD-YAP complex suppresses the oncogenic activity of YAP. Genes Dev. *26*, 1300–1305.

Lund, R.J., Närvä, E., and Lahesmaa, R. (2012). Genetic and epigenetic stability of human pluripotent stem cells. Nat. Rev. Genet. *13*, 732–744.

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.

Närvä, E., Autio, R., Rahkonen, N., Kong, L., Harrison, N., Kitsberg, D., Borghese, L., Itskovitz-Eldor, J., Rasool, O., Dvorak, P., et al. (2010). High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. Nat. Biotechnol. *28*, 371–377.

Nazor, K.L., Altun, G., Lynch, C., Tran, H., Harness, J.V., Slavin, I., Garitaonandia, I., Müller, F.-J., Wang, Y.-C., Boscolo, F.S., et al. (2012). Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. Cell Stem Cell *10*, 620–634.

Okeyo, K.O., Adachi, T., Sunaga, J., and Hojo, M. (2009). Actomyosin contractility spatiotemporally regulates actin network dynamics in migrating cells. J. Biomech. *42*, 2540–2548.

Pesce, M., Marin Gomez, M., Philipsen, S., and Schöler, H.R. (1999). Binding of Sp1 and Sp3 transcription factors to the Oct-4 gene promoter. Cell. Mol. Biol. (Noisy-le-grand) *45*, 709–716.

Quintin, S., Gally, C., and Labouesse, M. (2008). Epithelial morphogenesis in embryos: asymmetries, motors and brakes. Trends Genet. *24*, 221–230.

Rhodes, N., Heerding, D.A., Duckett, D.R., Eberwein, D.J., Knick, V.B., Lansing, T.J., McConnell, R.T., Gilmer, T.M., Zhang, S.-Y., Robell, K., et al. (2008). Characterization of an Akt kinase inhibitor with potent pharmacodynamic and antitumor activity. Cancer Res. *68*, 2366–2374.

Sagi, I., Chia, G., Golan-lev, T., Peretz, M., Weissbein, U., Sui, L., Sauer, M.V., Yanuka, O., Egli, D., and Benvenisty, N. (2016). Derivation and differentiation of haploid human embryonic stem cells. Nature *532*, 107–111.

Salloum, D., Mukhopadhyay, S., Tung, K., Polonetskaya, A., and Foster, D.A. (2014). Mutant Ras elevates dependence on serum lipids and creates a synthetic lethality for rapamycin. Mol. Cancer Ther. *13*, 733–741.

Sela, D., Chen, L., Martin-Brown, S., Washburn, M.P., Florens, L., Conaway, J.W., and Conaway, R.C. (2012). Endoplasmic reticulum stress-responsive transcription factor ATF6α directs recruitment of the mediator of RNA polymerase II transcription and multiple histone acetyltransferase complexes. J. Biol. Chem. *287*, 23035–23045.

Uren, A.G., Mikkers, H., Kool, J., van der Weyden, L., Lund, A.H., Wilson, C.H., Rance, R., Jonkers, J., van Lohuizen, M., Berns, A., et al. (2009). A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. Nat. Protoc. *4*, 789–798.

Vetterkind, S., and Morgan, K.G. (2009). The pro-apoptotic protein Par-4 facilitates vascular contractility by cytoskeletal targeting of ZIPK. J. Cell. Mol. Med. *13*, 887–895.

Vetterkind, S., Lee, E., Sundberg, E., Poythress, R.H., Tao, T.C., Preuss, U., and Morgan, K.G. (2010). Par-4: a new activator of myosin phosphatase. Mol. Biol. Cell *21*, 1214–1224.

Wang, C., Zhu, X., Feng, W., Yu, Y., Jeong, K., Guo, W., Lu, Y., and Mills, G.B. (2016). Verteporfin inhibits YAP function through up-regulating 14-3-3σ sequestering YAP in the cytoplasm. Am. J. Cancer Res. *6*, 27–37.

Wang, Z.V., Deng, Y., Gao, N., Pedrozo, Z., Li, D.L., Morales, C.R., Criollo, A., Luo, X., Tan, W., Jiang, N., et al. (2014). Spliced X-box binding protein 1 couples the unfolded protein response to hexosamine biosynthetic pathway. Cell *156*, 1179–1192.

Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J.B., Nishikawa, S., Nishikawa, S., Muguruma, K., et al. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. Nat. Biotechnol. *25*, 681–686.

Weissbein, U., Benvenisty, N., and Ben-David, U. (2014). Genome maintenance in pluripotent stem cells. J. Cell Biol. *204*, 153–163.

Weissbein, U., Plotnik, O., Vershkov, D., and Benvenisty, N. (2017). Culture-induced recurrent epigenetic aberrations in human pluripotent stem cells. PLoS Genet. *13*, e1006979.

Werbowetski-Ogilvie, T.E., Bossé, M., Stewart, M., Schnerch, A., Ramos-Mejia, V., Rouleau, A., Wynder, T., Smith, M.-J., Dingwall, S., Carter, T., et al. (2009). Characterization of human embryonic stem cells with features of neoplastic progression. Nat. Biotechnol. *27*, 91–97.

White, E. (2013). Exploiting the bad eating habits of Ras-driven cancers. Genes Dev. *27*, 2065–2071.

Wu, D.Y., and Yao, Z. (2006). Functional analysis of two Sp1/Sp3 binding sites in murine Nanog gene promoter. Cell Res. *16*, 319–322.

Zhang, H., Ramakrishnan, S.K., Triner, D., Centofanti, B., Maitra, D., Győrffy, B., Sebolt-Leopold, J.S., Dame, M.K., Varani, J., Brenner, D.E., et al. (2015). Tumor-selective proteotoxicity of verteporfin inhibits colon cancer progression independently of YAP1. Sci. Signal. *8*, ra98.

ISCI, Volume *11*

Supplemental Information

Genome-wide Screen for Culture Adaptation and Tumorigenicity-Related Genes in Human Pluripotent Stem Cells

Uri Weissbein, Mordecai Peretz, Omer Plotnik, Ofra Yanuka, Ido Sagi, Tamar Golan-Lev, and Nissim Benvenisty

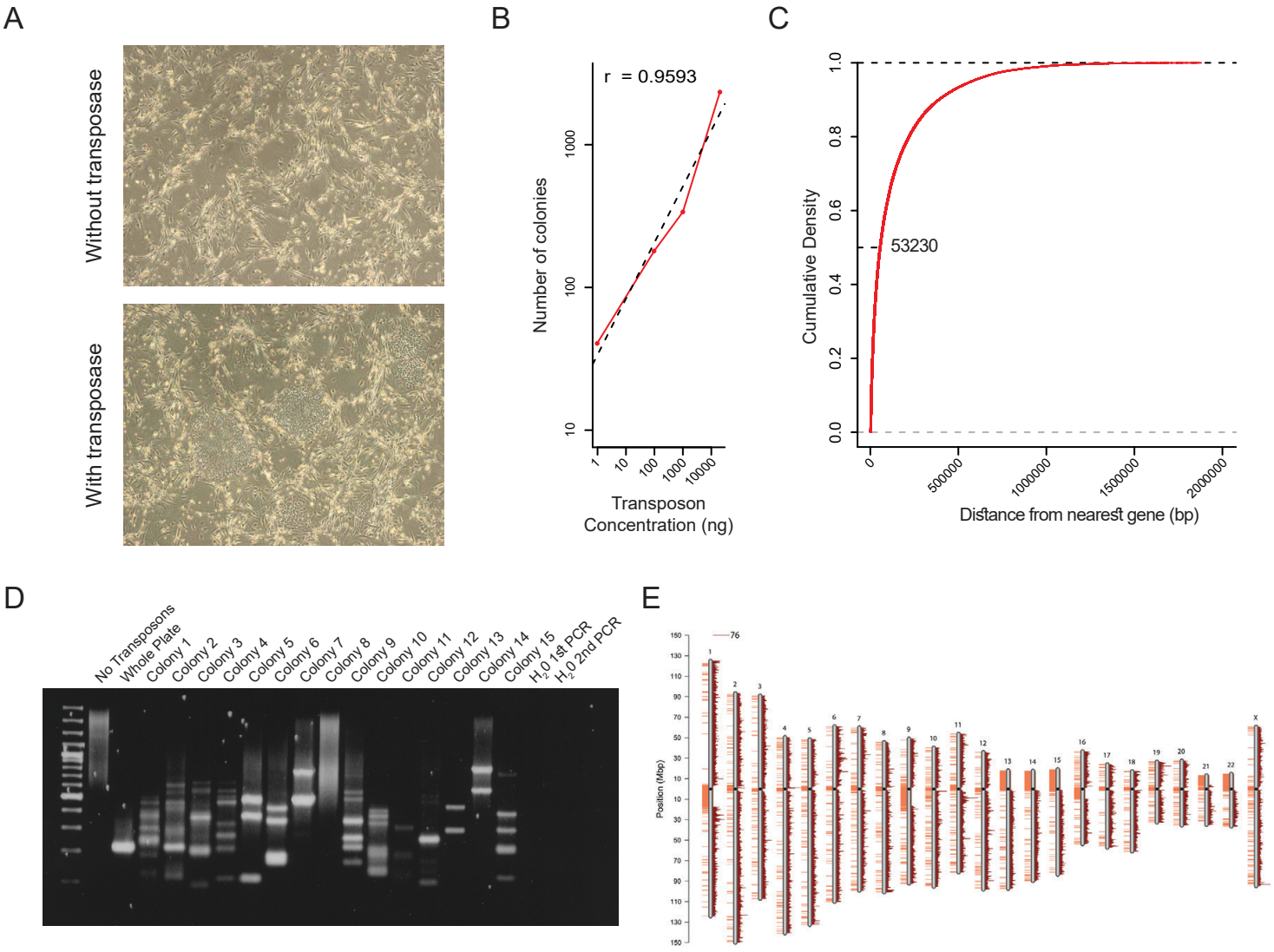


Figure S1 - Characterization of the transposon libraries, Related to Figure 1. (**A**) hESCs colonies morphology after electroporation and antibiotic selection. (**B**) Number of resistant colonies relative to the transposon plasmid concentration. A fixed amount of 20µg of transposase plasmid was used. (**C**) Cumulative distribution of the distance of each integration from the nearest gene. Integrations inside genes were omitted from the analysis. The average value is indicated. (**D**) Splinkerette-PCR and gel electroporation of a whole plate of resistant colonies, cells without transposons and 15 individual puromycin-resistant colonies. This analysis demonstrates the different integrations in each colony. (**E**) Map of the potential genomic coverage of the libraries. Each integration coordinates were expanded by 25kb upstream and downstream. The amount of expended integration covering each base in the genome is presented as a red histogram on the right side of each chromosome. Regions with no coverage are also highlighted as an orange line in the left side of each chromosome.

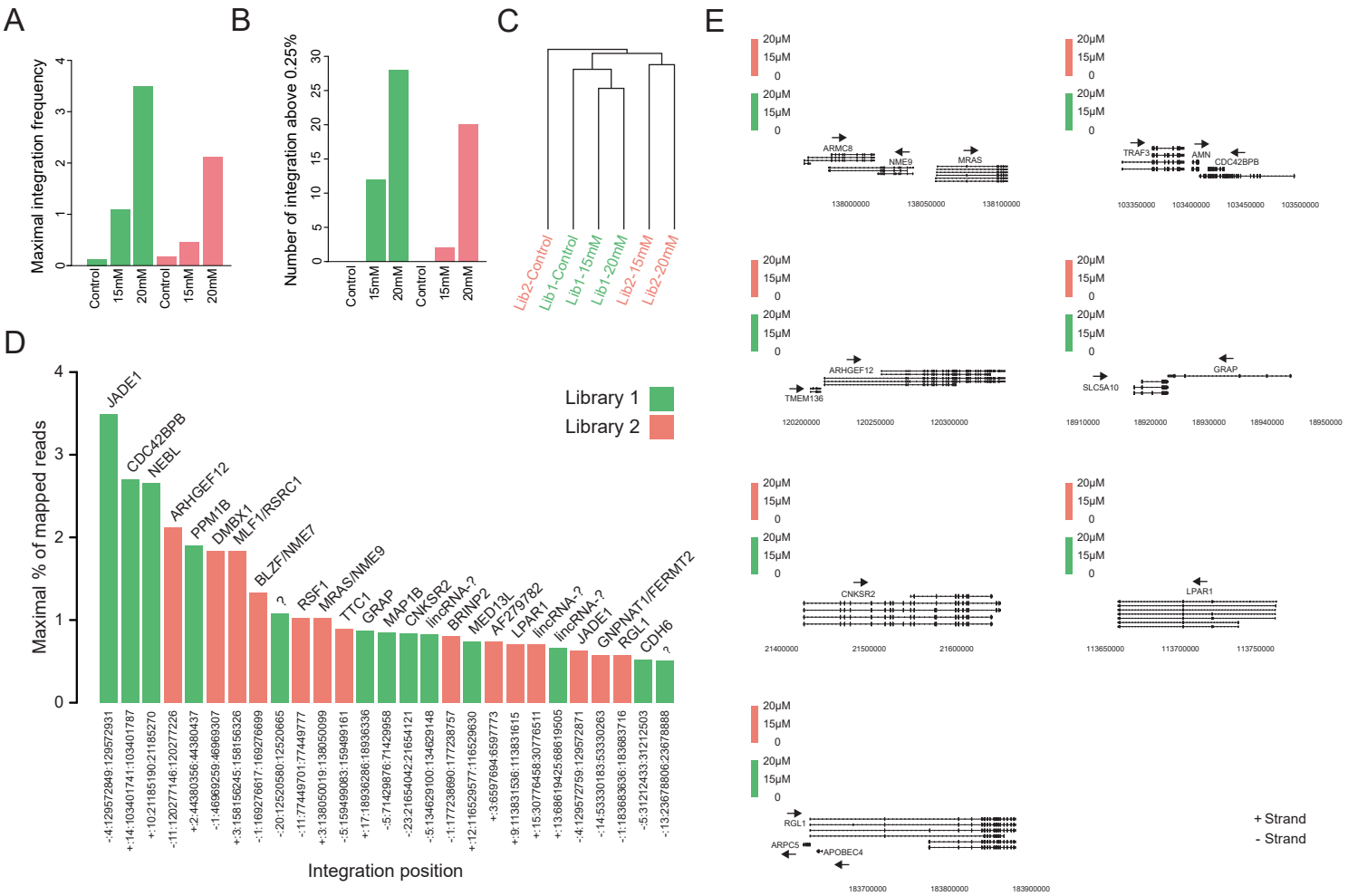


Figure S2 - Integration pattern changes after PluriSIn-1 selection, Related to Figure 2. (**A**) Shown are the frequencies of the most enriched integrations in each treatment. (**B**) Number of Integrations that correspond to more than 0.25% of the mapped reads. (**C**) Unsupervised hierarchical clustering of the integration pattern after PluriSIn-1 treatments. (**D**) All the integrations that occupy more than 0.5% of the mapped reads at least at 20μM PluriSIn-1 concentration, and are more enriched in 20μM than in 15μM treatment. (**E**) Dot plot of the genomic region next to genes related to RAS pathway that are selected during PluriSIn-1 treatments. Each integration is plotted as a dot where the area represents the integration frequency out of the total mapped reads. Each line of dots corresponds to a different sample. The color of each dot indicates the integration strand.

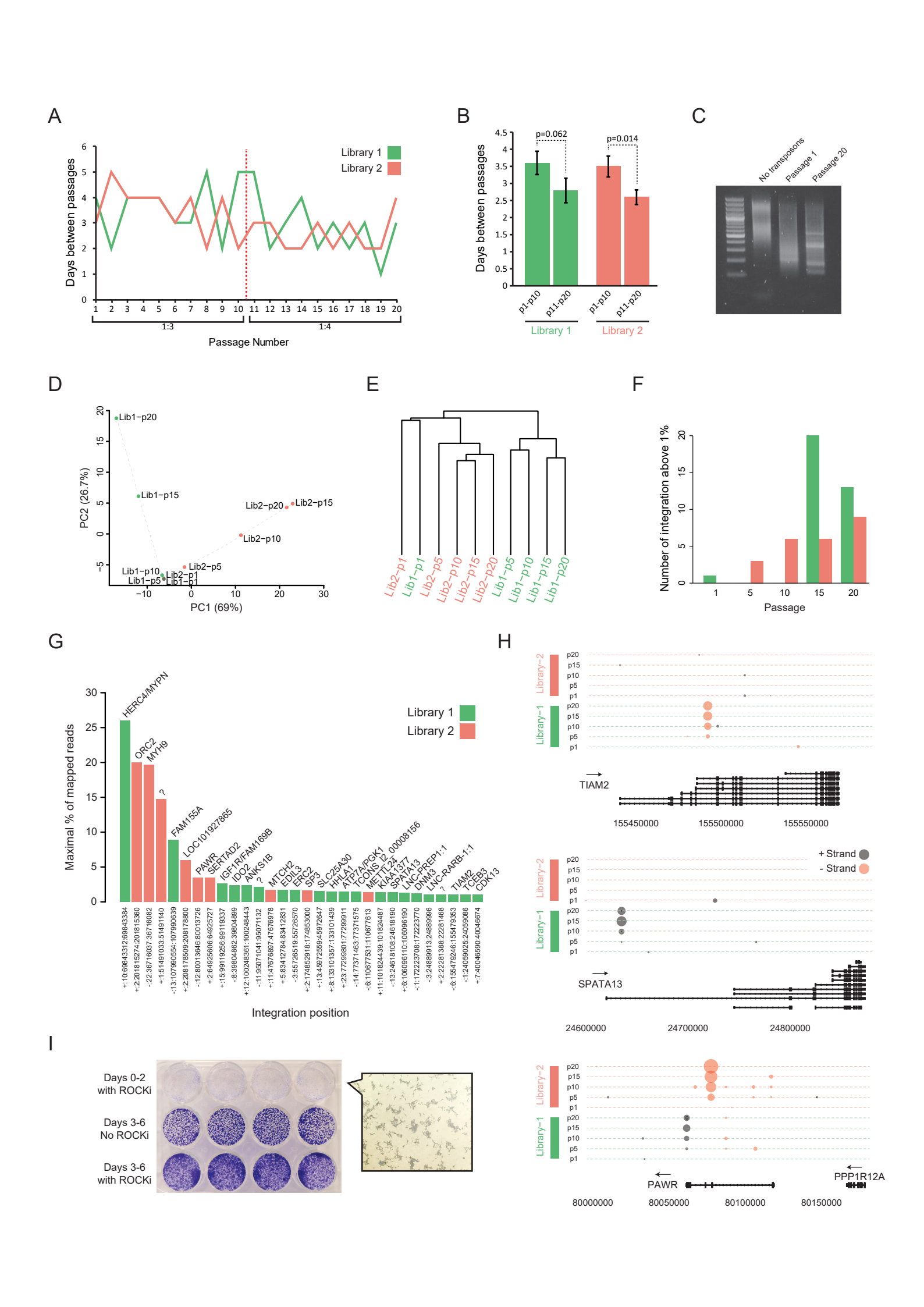

A
Days between passages
Library 1
Library 2
1:3
1:4
Passage Number
B
Days between passages
p=0.062
p=0.014
p1-p10
p11-p20
p1-p10
p11-p20
Library 1
Library 2
C
No transposons
Passage 1
Passage 20
D
PC2 (26.7%)
PC1 (69%)
Lib1−p20
Lib1−p15
Lib2−p20
Lib2−p15
Lib2−p10
Lib2−p5
Lib1−p10
Lib2−p1
Lib1−p5
Lib1−p1
E
Lib2−p1
Lib1−p1
Lib2−p5
Lib2−p10
Lib2−p15
Lib2−p20
Lib1−p5
Lib1−p10
Lib1−p15
Lib1−p20
F
Number of integration above 1%
Passage
G
Maximal % of mapped reads
Library 1
Library 2
HERC4/MYPN
ORC2
MYH9
FAM155A
LOC101927865
PAWR
SERTAD2
IGF1R/FAM169B
IDO2
ANKS1B
MTCH2
EDIL3
ERC2
SP3
SLC25A30
HHLA1
ATP7A/PGK1
TCONS_l2_00008156
METTL24
KIAA1377
SPATA13
LNC-PREP1:1
DNM3
LNC-RARB-1:1
TIAM2
TCEB3
CDK13
Integration position
H
Library−2
Library−1
p20
p15
p10
p5
p1
TIAM2
155450000
155500000
155550000
+ Strand
- Strand
SPATA13
24600000
24700000
24800000
PAWR
PPP1R12A
80000000
80050000
80100000
80150000
I
Days 0-2
with ROCKi
Days 3-6
No ROCKi
Days 3-6
with ROCKi

Figure S3 - Integration pattern changes during prolong culturing, Related to Figure 3. (**A**) The number of days between subsequent passages during the prolong culturing of the PB libraries. The ratio of the cell split is indicated below the passage number. (**B**) The difference in culturing duration of each passage between first and second half of the experiment. P-values were calculated with one tailed students T-test. (**C**) Splinkerette-PCR gel of Library-2 at first passage and after 20 passages in culture. (**D**) PCA analysis of the integration patterns of both libraries during culturing. (**E**) Unsupervised hierarchical clustering of the integration patterns during culturing. (**F**) Number of integrations that correspond to more than 1% of the mapped reads in both libraries at different time point during culturing. (**G**) All the integrations that occupies more the 1% of the mapped reads at any time point during culturing and the genes that can be affected by them. (**H**) Dot plot of the genomic region next to TIAM2, SPATA13 or PAWR during the prolong culturing. Each integration is plotted as a dot where the area represents the integration frequency out of the total mapped reads. Each line of dots corresponds to a different sample. The color of each dot indicates the integration strand. (**I**) Growth assay of cells plated for 48h with ROCKi and then treated with or without ROCKi for further 96h. High magnification of the cells after 48h is displayed on the right panel, showing the cells already in colonies.

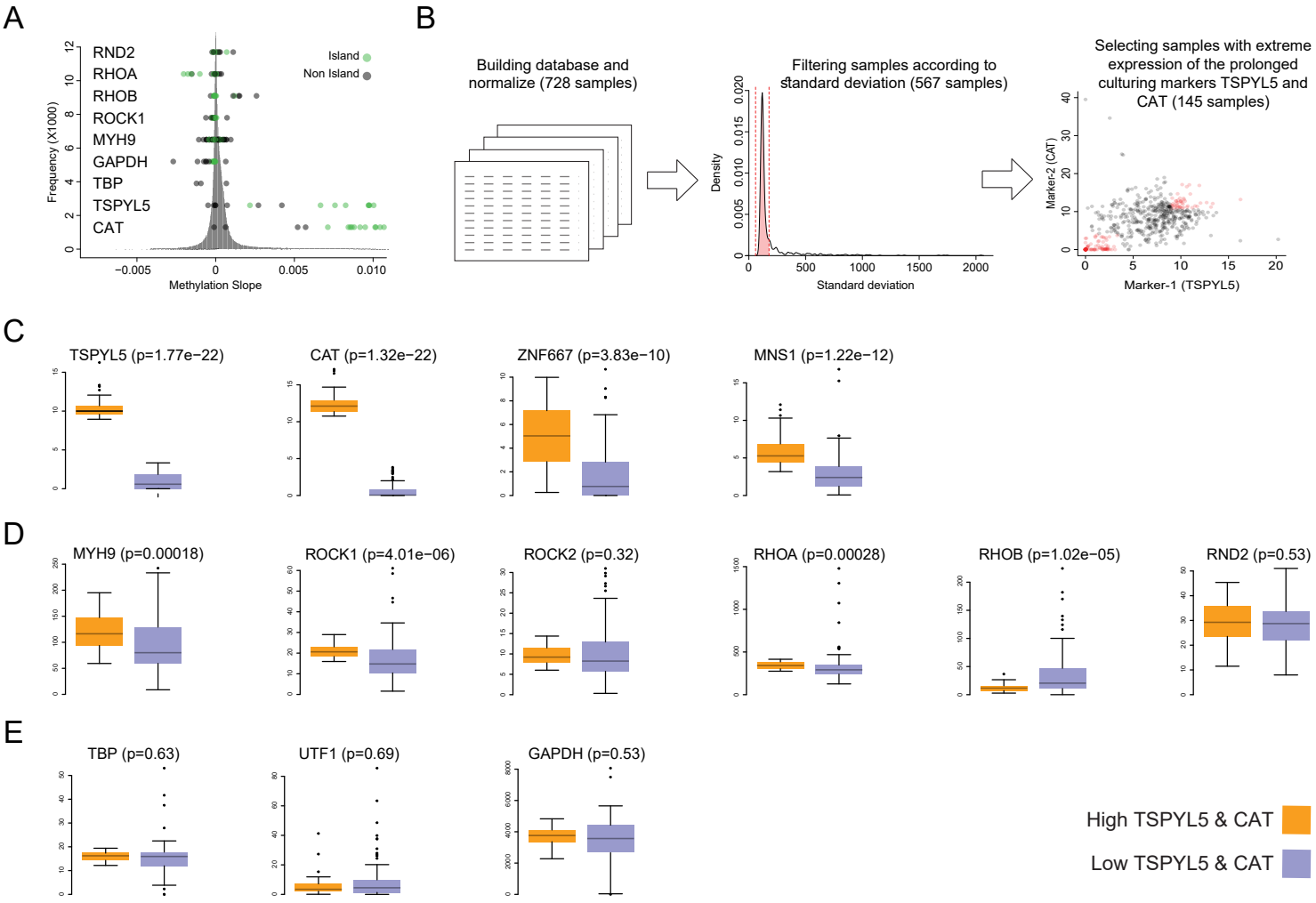


Figure S4 - Gene expression changes in the RHO/ROCK pathway genes, Related to Figure 3. (**A**) Methylation changes during culturing of genes from the RHO/ROCK pathway. For each probe of the 450K Methylation beadChips (Illumina), a linear regression between the methylation values and the passage numbers was performed. The slopes of the regression lines describe the overall change in methylation levels during the culturing. The grey histogram describes the distribution of the methylation slopes of all methylation probes. Methylation slopes of probes associated with the RHO/ROCK pathway genes are shown in the histogram. Two additional genes, TSPYL5 and CAT, which gain methylation during prolonged culturing, are also presented. (**B**) The steps in the preparation of RNA-Seq data for analysis. In red are the samples that passed each filtration step (**C-E**) Comparison of the gene expression levels between the TSPYL5 and CAT high and low hPSCs samples. P-values were calculated using Wilcoxon signed-rank test. Shown are genes that gain methylation and silenced expression during culturing (**C**), genes related to the RHO/ROCK pathway (**D**) and housekeeping or pluripotency genes as controls (**E**).

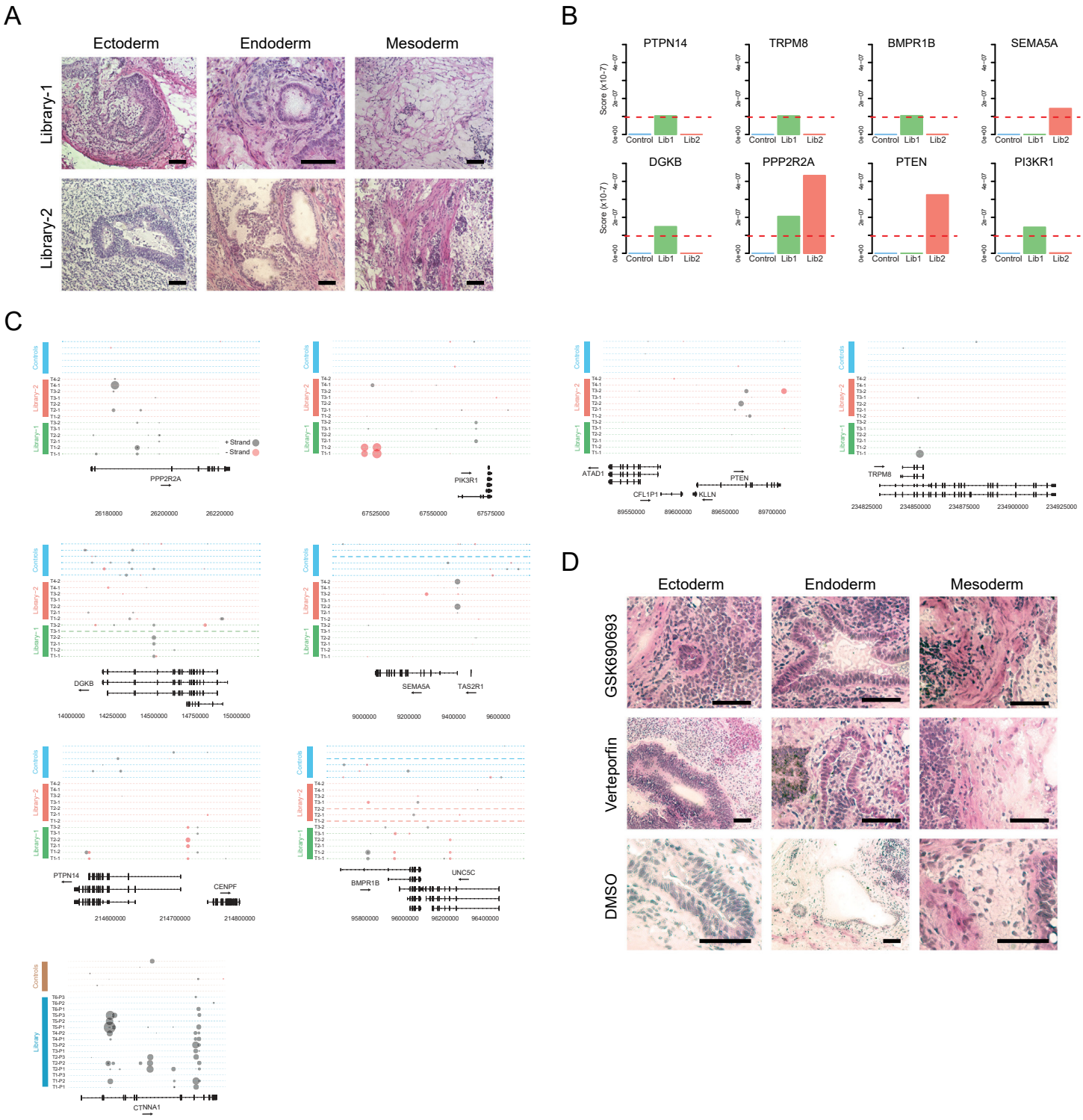


Figure S5 – Selection during teratoma formation, Related to Figure 4 (**A**) Histological analysis of sections from teratomas derived from the libraries with hematoxylin and eosin staining. Scale bar represents 100μm. (**B**) The scores that each gene from the PI3K-AKT pathway received. The red dashed line marks the score which is 2-fold higher than the maximal score of the control group. (**C**) Dot plot of the genomic region of the significant genes in the different teratoma samples and undifferentiated libraries as controls. Each integration is plotted as a dot where the area represents the integration frequency out of the total mapped reads. Each line of dots represents a different sample. The color of each dot indicates the integration strand. (**D**) Hematoxylin and eosin staining of teratomas that developed in mice treated with GSK690693, VP or DMSO. Scale bar represents 100μm.

Table S1 - Adaptors used in the study, Related to Figure 1

Sample	Long ssDNA
Lib1 - Long Culture , p18	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ACATTG**C TAG**CAATGT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Long Culture , p23	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ACTATG**C TAG**CATAGT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Long Culture , p28	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ACTGAG**C TAG**CTCAGT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Long Culture , p33	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**AGAGCT**C TAG**AGCTCT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Long Culture , p38	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**AGGCGT**C TAG**ACGCCT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Long Culture , p18	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**AGGCTG**C TAG**CAGCCT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Long Culture , p23	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**AGTACT**C TAG**AGTACT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Long Culture , p28	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**AGTGAT**C TAG**ATCACT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Long Culture , p33	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**AGTGTA**C TAG**TACACT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Long Culture , p38	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATAGTG**C TAG**CACTAT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Teratoma, T1-1	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATATAT**C TAG**ATATAT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Teratoma, T1-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATATGC**C TAG**GCATAT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Teratoma, T2-1	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATCAGT**C TAG**ACTGAT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Teratoma, T2-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATCATG**C TAG**CATGAT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Teratoma, T3-1	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATGATC**C TAG**GATCAT**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - Teratoma, T3-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATGGAT**C TAG**ATCCAT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Teratoma, T1-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATGTTG**C TAG**CAACAT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Teratoma, T2-1	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATTAAG**C TAG**CTTAAT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Teratoma, T2-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATTGAC**C TAG**GTCAAT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Teratoma, T3-1	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**ATTTAA**C TAG**TTAAAT**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Teratoma, T3-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CATAGT**C TAG**ACTATG**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Teratoma, T4-1	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CATGAT**C TAG**ATCATG**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - Teratoma, T4-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CATGTA**C TAG**TACATG**ATGCGCAGTTTTTTTGCAAAAA
Lib1 – Control-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CATTAG**C TAG**CTAATG**ATGCGCAGTTTTTTTGCAAAAA
Lib2 – Control-2	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CCTGGC**C TAG**GCCAGG**ATGCGCAGTTTTTTTGCAAAAA
Lib1 – Control	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CTATGA**C TAG**TCATAG**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - psn15	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CTCACG**C TAG**CGTGAG**ATGCGCAGTTTTTTTGCAAAAA
Lib1 - psn20	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CTCTGT**C TAG**ACAGAG**ATGCGCAGTTTTTTTGCAAAAA
Lib2 – Control	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CTGTCT**C TAG**AGACAG**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - psn15	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CTTCAG**C TAG**CTGAAG**ATGCGCAGTTTTTTTGCAAAAA
Lib2 - psn20	GTAATACGACTCACTATAGGGCACGCGTGGTCGACTGCGCAT**CTTTGC**C TAG**GCAAAG**ATGCGCAGTTTTTTTGCAAAAA

Foreign tax identifying

Table S2 - Integrations in genes from the RAS pathway, Related to Figure 2

Gene Name	Integration location	Effect on gene expression	Effect on RAS Signaling
MRAS	Upstream	Overexpression	Increased signalling
RGL1	5'UTR	Overexpression?	Increased signalling
LPAR1	Upstream	Overexpression	Increased signalling
GRAP	Intragenic-middle of the gene	Downregulation	Reduced signalling
CNKSR2	Intragenic-end of the gene	?	?
ARHGEF12	5'UTR/Intragenic-beginning of the gene	Overexpression?	Increased signalling
CDC42BPB	Intragenic-end of the gene	?	?

Table S3 - Integrations in genes from the RHO-ROCK pathway, Related to Figure 3

Gene Name	Integration location	Effect on gene expression	Effect on ROCK Signaling
MYH9	Intragenic-middle of the gene	Downregulation	Reduced signalling
SPATA13	Upstream	Overexpression	Reduced signalling
TIAM2	5'UTR/Intragenic-beginning/middle of the gene	Downregulation?	Reduced signalling
PAWR	Intragenic-middle of the gene	Downregulation	Reduce contractility of the cells

Table S4 - Enriched integrations in the teratomas, Related to Figure 4

Library	Chr	Window start	# Integrations	# Teratoma	# Coverage	Score	Gene
1	23	28845001	2	5	1.809499796	2.13E-07	IL1RAPL1
1	23	28860001	2	5	1.809499796	2.13E-07	IL1RAPL1
1	23	65130001	3	4	1.004536882	1.42E-07	
1	20	3810001	1	2	4.798393287	1.13E-07	AP5S1,MAVS, CDC25B
1	20	3825001	1	2	4.798393287	1.13E-07	AP5S1,MAVS
1	19	36885001	2	3	1.397599812	9.85E-08	ZFP82,ZFP14, LOC644189
1	19	36900001	2	3	1.397599812	9.85E-08	ZFP82,ZFP14, LOC644189
1	18	57315001	2	3	2.546799121	1.80E-07	CCBE1
1	18	57330001	2	3	2.546799121	1.80E-07	CCBE1
1	15	29940001	3	2	1.905173004	1.34E-07	
1	15	29955001	3	2	1.905173004	1.34E-07	
1	14	60540001	3	3	2.331915807	2.47E-07	LRRC9,PCNXL4
1	14	60555001	3	3	2.331915807	2.47E-07	LRRC9,PCNXL4
1	14	64995001	2	4	1.505073801	1.41E-07	PPP1R36,ZBTB1, HSPA2,ZBTB25
1	14	65010001	2	4	1.505073801	1.41E-07	PPP1R36,ZBTB1, HSPA2
1	12	39300001	4	4	1.988914033	3.74E-07	CPNE8
1	12	39315001	3	3	1.98626692	2.10E-07	CPNE8
1	12	97170001	3	4	0.877176213	1.24E-07	
1	12	76335001	3	3	0.97589706	1.03E-07	
1	12	90705001	2	4	1.043517364	9.81E-08	
1	12	90720001	2	4	1.043517364	9.81E-08	
1	11	20025001	3	3	0.917402682	9.70E-08	NAV2
1	10	59025001	1	3	3.736022532	1.32E-07	
1	10	59040001	1	3	3.736022532	1.32E-07	MIR3924
1	10	106755001	2	2	2.558034761	1.20E-07	SORCS3
1	10	106770001	2	2	2.558034761	1.20E-07	SORCS3
1	9	113325001	2	5	1.497055029	1.76E-07	SVEP1
1	9	113340001	2	5	1.497055029	1.76E-07	SVEP1
1	9	18465001	3	4	0.958447837	1.35E-07	ADAMTSL1
1	9	8640001	3	4	0.725493785	1.02E-07	PTPRD
1	9	8655001	3	4	0.725493785	1.02E-07	PTPRD
1	8	10890001	4	3	2.452424113	3.46E-07	XKR6,MIR598
1	8	26160001	6	5	0.580798023	2.05E-07	PPP2R2A
1	8	26175001	6	5	0.580798023	2.05E-07	PPP2R2A
1	8	41265001	3	3	1.822261152	1.93E-07	
1	8	10875001	3	2	2.252059147	1.59E-07	XKR6,MIR598
1	7	14490001	2	5	1.259480805	1.48E-07	DGKB
1	7	130635001	3	3	1.185936136	1.25E-07	LINC-PINT
1	6	142335001	2	3	2.521969786	1.78E-07	
1	6	66165001	2	4	1.422110366	1.34E-07	EYS

1	6	66180001	2	4	1.422110366	1.34E-07	EYS
1	6	4695001	2	2	2.800890608	1.32E-07	CDYL
1	6	4710001	2	2	2.800890608	1.32E-07	CDYL
1	6	139455001	1	5	2.056310552	1.21E-07	HECA
1	6	139470001	1	5	2.056310552	1.21E-07	HECA
1	5	67275001	2	3	6.176404331	4.35E-07	
1	5	67260001	2	3	3.151731161	2.22E-07	
1	5	67290001	1	2	6.163249352	1.45E-07	
1	4	41400001	2	3	3.806773301	2.68E-07	LIMCH1
1	4	95805001	4	3	1.474853697	2.08E-07	BMPR1B
1	4	71520001	8	5	0.383968132	1.80E-07	ENAM,JCHAIN
1	4	71505001	7	4	0.383275765	1.26E-07	ENAM,JCHAIN
1	4	21165001	3	2	1.510022585	1.06E-07	KCNIP4
1	4	21180001	3	2	1.510022585	1.06E-07	KCNIP4
1	4	95790001	3	2	1.47416133	1.04E-07	BMPR1B
1	3	21045001	2	3	2.842459382	2.00E-07	
1	3	160680001	2	2	2.929189152	1.38E-07	PPM1L
1	3	55215001	4	3	0.972026551	1.37E-07	
1	3	60585001	3	4	0.920335172	1.30E-07	FHIT
1	3	55200001	3	3	0.954088094	1.01E-07	
1	3	70635001	2	4	1.07139803	1.01E-07	
1	3	70650001	2	4	1.07139803	1.01E-07	
1	3	39525001	2	3	1.412675228	9.96E-08	MOBP
1	2	46485001	2	3	3.249047418	2.29E-07	
1	2	46500001	2	3	3.249047418	2.29E-07	EPAS1
1	2	98250001	1	3	3.132466959	1.10E-07	ACTR1B,COX5B
1	2	98265001	1	3	3.132466959	1.10E-07	ACTR1B,COX5B, LINC01125
1	2	234825001	2	2	2.189762703	1.03E-07	TRPM8
1	2	234840001	2	2	2.189762703	1.03E-07	TRPM8
1	1	24030001	3	3	1.247970827	1.32E-07	RPL11
1	1	214710001	2	5	0.876995305	1.03E-07	PTPN14
1	1	214725001	2	5	0.876995305	1.03E-07	PTPN14
2	2	85245001	4	2	2.191130486	1.48E-07	KCMF1
2	2	85260001	4	2	2.191130486	1.48E-07	KCMF1
2	2	85770001	4	3	1.067400691	1.08E-07	PARTICL,GGCX, MAT2A
2	3	69135001	5	4	1.671716112	2.83E-07	ARL6IP5,LMOD3, UBA3
2	3	114105001	3	3	2.033268914	1.55E-07	ZBTB20-AS1, ZBTB20
2	4	74370001	2	3	2.298049913	1.17E-07	AFM,LOC728040
2	4	74385001	2	3	2.298049913	1.17E-07	AFM,LOC728040
2	5	9525001	2	5	1.702885558	1.44E-07	SNORD123,SNHG18, SEMA5A
2	5	9540001	2	5	1.702885558	1.44E-07	SNORD123,SNHG18, SEMA5A
2	8	13020001	2	3	2.617152442	1.33E-07	DLC1
2	8	13035001	2	3	2.617152442	1.33E-07	DLC1

2	8	26145001	2	4	1.997500544	1.35E-07	PPP2R2A
2	8	26160001	5	5	2.043734072	4.32E-07	PPP2R2A
2	10	89670001	6	5	1.2800062	3.25E-07	PTEN
2	12	94110001	3	2	2.18170888	1.11E-07	CRADD
2	12	94125001	3	2	2.18170888	1.11E-07	CRADD
2	13	20640001	4	4	0.910411191	1.23E-07	ZMYM2
2	13	20655001	4	4	0.910411191	1.23E-07	ZMYM2
2	13	84690001	3	3	1.615119034	1.23E-07	LINC00333
2	13	84705001	3	3	1.615119034	1.23E-07	LINC00333
2	14	102420001	3	4	0.984976832	1.00E-07	DYNC1H1,PPP2R5C
2	14	102435001	3	4	0.984976832	1.00E-07	DYNC1H1
2	17	66525001	3	3	1.733230975	1.32E-07	FAM20A,PRKAR1A
2	19	17805001	1	3	7.027084823	1.78E-07	UNC13A,MAP1S
2	19	17820001	1	3	7.027084823	1.78E-07	UNC13A,MAP1S
2	19	52380001	4	5	2.179779523	3.69E-07	ZNF649,ZNF577
2	19	52395001	2	3	2.177047198	1.11E-07	ZNF649,ZNF577

Table S5 - Primers used in the study, Related to Figure 1

1st PCR primer F	GTAATACGACTCACTATAGGGCACG
1st PCR primer R	CGCTATTTAGAAAGAGAGAGCAATATTTCA
2nd PCR primer F	AATGATACGGCGACCACCGACACTCTACTATAGGGCACGCGTGGT
2nd PCR primer R	CAAGCAGAAGACGGCATACGAGCTCTTCAGAATGCATGCGTCAATTTTACGCAGAC
Sequencing primer	ACTATAGGGCACGCGTGGTCGACTGCGCAT

Transparent Methods

Cell culture

Human ESCs (CSES cell lines) (Lavon et al., 2008) were cultured on mouse embryonic fibroblast treated with mitomycin-C with medium containing KnockOut Dulbecco's modified Eagle's medium (Gibco-Invitrogen, CA) supplemented with 15% KnockOut-SR (Gibco-Invitrogen, CA), 1mM glutamine, 0.1mM ß-mercaptoethanol (Sigma-Aldrich, MO), 1% non-essential amino-acid stock (Gibco-Invitrogen, CA), 50U/ml penicillin (Biological Industries, Beit Haemek, Israel), 50 mg/ml streptomycin (Biological Industries, Beit Haemek, Israel), and 8 ng/ml fibroblast growth factor 2 (Gibco-Invitrogen, CA). In some experiments the cells were plated on Matrigel-coated plates cultured with mTeSR medium (STEMCELL Technologies). Cells were passaged using trypsin-EDTA (Biological Industries, Beit Haemek, Israel). Teratoma derived cell line was cultured in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal bovine serum (Gibco-Invitrogen, CA), 50U/ml penicillin and 50mg/ml streptomycin, (Biological Industries, Beit Haemek, Israel).

Overexpression Libraries construction

Transposon plasmid pPB-SB-CMV-puro-SD was obtain from Schmidt lab (Chen et al., 2013). pCMV-PBase were obtained from the Wellcome Trust Sanger Institute. CSES7 passage 16 from 25 10cm plates were trypsinized, electroporated with 20µg of each plasmid (320V, 250µF and Resistance ∞Ω) and plated on 50 10cm plates with ROCKi (10µM) for the first 24h. After 48h puromycin selection (0.34µM) was initiated in order to recover resistant colonies. Next, each 25 plates where trypsinized, mixed and freeze separately in multiple vials, consisting the two libraries. For the analysis of the effect of the transposon plasmid concentration on the number of puromycin resistant colonies, different amounts of transposon plasmid were electroporated together with 20µg transposase plasmid. Colonies were stained with methylene blue and counted using OpenCFU softwere (Geissmann, 2013). From the constructed libraries, 15 individual puromycin-resistant colonies were picked and amplified. By Splinkerrette-PCR we could estimate that each colony have 4.4 integrations on average. The estimation that each library contains $2.6x10^5$ individual integrations was achieved by multiplication of 25 plates, ~2337 colonies per plate, 4.4 integration per colony.

Libraries selection treatments

For each of the experiments, cells from libraries were thawed and plated on two 10cm plates. For PluriSIn-1 treatments, ESc medium was supplemented with PluriSIn-1 (Roche and BioGems) at 15 µM or 20µM until the massive cell death stage stopped.

Then we let the surviving cells to recover and form colonies without PluriSIn-1. Next, we split the cells to new plates with medium containing PluriSIn-1 and let resistant cells form new colonies. Then the cells were trypsinized, and genomic DNA (gDNA) was extracted. For long culturing assay, the cells were maintained for 20 passages. Each 5 passages cell were collected for gDNA extraction and freezing. In order to produce teratomas, $4x10^6$ cells were resuspended in 100µl ESc medium and 100µl Matrigel (Corning) and injected under the skin of NOD-SCID Il2rg−/− immunodeficient mice (Jackson Laboratory). Eight weeks after injection, mice were sacrificed and the teratomas were dissected. Each teratoma was divided into 2-3 pieces and gDNA was extracted separately from each piece.

GDNA extraction, Splinkerrette-PCR and DNA seq

gDNA was extracted using NucleaSpin Tissue kit (Macherey-Nagel). 1µg DNA was digested with 1µl CSP6I (Thermo Fisher Scientific) in 20µl reaction at 37° for 3h with addition of 1µl enzyme after 1.5h, followed by inactivation at 65° for 20min. The adaptors used for the Splinkerrette-PCR contained variable 6nt barcoded region which allows analysis of multiple samples simultaneously. The two single-strandedDNA parts of the adaptors were diluted to 50µM in 5xNEB buffer 2 (NEB). Prior to adaptor ligation, adaptor mix was heated to 95°C for 5min, cooled to 25°C at -0.1°C/sec and stored on ice, and the gDNA was heated to 65°C for 20min. The ligation was performed with 300ng of digested DNA, 1µl of the adaptors, 2µl T4 DNA ligase (Thermo Fisher Scientific) in 25µl reaction at 16°C overnight. The ligation reaction was purified using MEGAquick-spin Total Fragment DNA Purification Kit (iNtRON Biotechnology), and served as a template for the first PCR. Both PCR was performed KAPA HiFi PCR kit (Kapa Biosystems) in 25µl. The first PCR was performed using 5µl DNA with the following conditions: 3min 95°C, 12 cycles of 20sec 98°C, 15sec 72°C (-0.5°C each cycle) and 15sec 72°C, 15 cycles of 20sec 98°C, 15sec 65°C and 15sec 72°C, 1min 72°C. PCR product was diluted to 100µl and 1µl served as a template for the second nested PCR reaction, which adds the adaptors for the Illumina flowcell, with the following conditions: 3min 95°C, 12 cycles of 20sec 98°C, 15sec 72°C (-0.5°C each cycle) and 15sec 72°C, 19 cycles of 20sec 98°C, 15sec 65°C and 15sec 72°C, 1min 72°C. Sequences are listed in **Table S5**. PCR products from all experiments were pulled together and size selected from agarose gel. High-throughput sequencing was performed using Illumina Next-Seq500 with a custom primer (**Table S5**) generating single-end 85bp reads.

Analysis of the DNA-sequencing data

The sequenced reads were de-multiplexed according to the barcodes using FASTX barcode splitter. Then for each sample, reads were trimmed to remove barcode sequence using FASTX trimmer, transposon sequence was removed using Cutadapt

leaving reads of 7bp or more. Reads were align to the HG19 reference human genome with BWA, output SAM file was converted to BAM file, indexed, and converted to BED file. Then each integration was counted. Integrations which overlap in their genomic locations were merged. For each integration the frequency out of the total mapped reads was calculated. All analyses were performed and visualized in R statistical software (http://www.r-project.org/).

Analysis of the Teratomas data

In addition to the teratoma samples, 6 samples of untreated libraries were included in the analysis for comparison. For scoring genomic regions which may be relevant to teratoma formation in the overexpression libraries, rolling window of 30000bp was used with overlap of 15000bp between each window. In the haploid library we scored each gene. Three parameters were taken into account: the number of samples that had integration (#Samples), the number of unique integrations (#Integrations) and the sum of their percentage out of the mapped reads (#Coverage). The score was calculated as follows:

$$Score = \frac{\#Samples\ x\ \#Integrations\ x\ \#Coverage}{Total\ number\ of\ samples^2\ x\ Total\ number\ of\ integrations\ x\ 100}$$

In the overexpression libraries, we looked for windows which scored in at least one of the teratoma libraries >2 fold more than the maximal score of the controls. In the haploid libraries, we looked for genes with score in the teratoma at least 10 fold higher than in the controls, and required that the scores in the teratoma group is higher than the highest percentile of the controls.

MRAS Cloning

For *MRAS* Overexpression experiments, *MRAS* was PCR amplified from hESCs cDNA and cloned into pcDNA3 construct (Invitrogen) using standard cloning technique. Primer used to amplify MRAS were: GGTCTGACCTACGAGAAACA and TCAGGCCTGTCACAAGA.

Cell viability and proliferation assays

For proliferation assays with or without ROCKi (Y27632), cells were treated with ROCKi 1h prior to the trypsinization and then seeded at density of 50000 cells/well (6 well plate) or 1000 cells/well (96 well plate) with ROCKi for the first 24h. To estimate the relative cell number, cells were fixated with 0.5% glutardialdehyde (Sigma-Aldrich) and staining with methylene blue (Sigma-Aldrich) dissolved in 0.1 M boric acid, pH

8.5. Color was extracted with 0.1M hydrochloric acid, and quantitated by measuring absorbance at 650nM. For MRAS experiments, 30000 cells/well were transiently transfected with MRAS construct or empty pcDNA3 plasmid using X-tremeGENE 9 DNA Transfection Reagent (Roche) and plated on Matrigel coated 96 well plate for 24h. Then, cells were treated with different concentrations of PluriSIn-1 and OA (Sigma-Aldrich) for 48h with medium replacement after 24h. For Blebbistatin experiment, cells were plated in matrigel-coated 96 well plate at a concentration of 10000 cells/well and cultured with mTeSR for 24h. Then, the cells were treated with different concentration of blebbistatin (Cayman Chemical) for two days maintaining constant concentration of DMSO in all wells. Cell number was estimated using CellTiter-Glo Luminescent Cell Viability Assay (Promega). To assess the growth of teratoma-derived cell line after treatment with HIPPO or AKT inhibitors, pES10-teratoma derived cells were treated with different doses of either VP or GSK690693 (AdooQ BioScience) for 96h in triplicates, and then viability was measured with CellTiter-Glo (Promega).

Apoptosis analysis

Cells were treated with either 5μM VP for 24h, 20μM GSK690693 for 72h or with only DMSO for 72h. The cells were then dissociated using TrypLE-Select (Thermo Fisher Scientific), washed with PBS supplemented with 10% FBS and filtered through a 70μm cell strainer. Apoptosis levels were quantified using MEBCYTO Apoptosis Kit (MBL) according to the manufacturer's instructions. Analysis was performed with BD FACSAria III.

Teratoma formation assay

All experimental procedures in animals were approved by the ethics committee of the Hebrew University. $1.8x10^6$ human ESCs were resuspended in 100 μl medium and 100 μl Matrigel and subcutaneously injected into NOD-SCID Il2rg−/− immunodeficient mice. Mice were injected intraperitoneally with either 100mg/kg VP, 30mg/kg GSK690693 or with equal volume of medium with 3.34% DMSO, three times a week for six weeks.

Microarray expression profiles analysis

Published microarray profile GSE30654 was downloaded from the Gene Expression Omnibus database and RMA normalized using Affymetrix Expression Console. Expression values below 20 were raised to 20. Samples with large chromosomal aberration as detected by e-Karyotyping analysis were removed from the analysis (Ben-David et al., 2013), as well as samples of iPSCs. The samples were divided

according to passage number (below or equal to 25 and qual or above 50). The significance of the expression difference between the two groups was obtained by t-test.

RNA-Sequencing gene-expression analysis

FASTQ files were obtained from the SRA (https://www.ncbi.nlm.nih.gov/sra) database (Wheeler et al., 2008), aligned to the genome using TopHat2 (Kim et al., 2013) and FPKM valued were obtained using Cufflink (Trapnell et al., 2010). Each sample was normalized to the mean FPKM value. Samples with outlying standard division values were omitted. Then, the samples were filtered according to TSPYL5 and CAT expression levels, genes that gain methylation and silenced during culturing, hence they serve as markers for prolonged culturing. Samples with high (first quantile) or low (forth quantile) expression of both TSPYL5 and CAT were chosen for the analysis. Expression difference between the group of samples that express TSPYL5 and CAT highly to those who express these genes lowly were calculated and p-value was obtain with Wilcoxon signed-rank test.

Supplemental Reference

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.

Chen, L., Stuart, L., Ohsumi, T.K., Burgess, S., Varshney, G.K., Dastur, A., Borowsky, M., Benes, C., Lacy-Hulbert, A., and Schmidt, E. V (2013). Transposon activation mutagenesis as a screening tool for identifying resistance to cancer therapeutics. BMC Cancer *13*, 93.

Geissmann, Q. (2013). OpenCFU, a New Free and Open-Source Software to Count Cell Colonies and Other Circular Objects. PLoS One *8*, 1–10.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Lavon, N., Narwani, K., Golan-Lev, T., Buehler, N., Hill, D., and Benvenisty, N. (2008). Derivation of Euploid Human Embryonic Stem Cells from Aneuploid Embryos. Stem Cells *26*, 1874–1882.

Trapnell, C., Williams, B. a, Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., et al. (2008). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. *36*, D13–D21.