Supplementary Table S1. List of primers used for real-time RT-PCR

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Gene transcript | Genebank Accession No. | Forward Primers Sequence  5` - 3` | Reverse Primers Sequence  5` - 3` |
| 1 | human GAPDH | NM\_002046.7 | CAACGACCACTTTGTCAAGCTC | GGTCTACATGGCAACTGTGAGG |
| 2 | human IL6 | NM\_0006000.5 | ACTCCTTCTCCACAAGCGCCTTC | CGTTCTGAAGAGGTGAGTGGCTGTC |
| 3 | human CXCL8 | NM\_000584.4 | TCAGAGACAGCAGAGCACACAAG | GGTCCACTCTCAATCACTCTCAG |
| 4 | human IL15 | NM\_000585.5 | GCCATAGCCAGCTCTTCTTCAATAC | TAGGAAGCCCTGCACTGAAACAGC |
| 5 | human FGF2 | NM\_001361665.2 | CGACCCTCACATCAAGCTACAAC | TCTTCCATCTTCCTTCATAGCCAG |
| 6 | human HBEGF | NM\_001945.3 | GGGACCCATGTCTTCGGAAATAC | CCAGGATGGTTGTGTGGTCATAG |
| 7 | human TGFß1 | NM\_000660.7 | GGGATAACACACTGCAAGTGGAC | AGCTGAAGCAATAGTTGGTGTCC |
| 8 | mouse GAPDH | NM\_008084 | AACGGCACAGTCAAGGCCGA | ACCCTTTTGGCTCCACCCTT |
| 9 | mouse NANOG | NM\_028016.3 | AGGGTCTGCTACTGAGATGCTCTG | CAACCACTGGTTTTTCTGCCACCG |
| 10 | mouse OCT3/4 | NM\_013633.3 | TCTTTCCACCAGGCCCCCGGCTC | TGCGGGCGGACATGGGGAGATCC |
| 11 | mouse SOX2 | NM\_011443.4 | TAGAGCTAGACTCCGGGCGATGA | TTGCCTTAAACAAGACCACGAAA |
| 12 | mouse KLF4 | NM\_010637.3 | GGACTTACAAAATGCCAAGGGGTG | TCGCTTCCTCTTCCTCCGACACA |
| 13 | mouse C-MYC | NM\_001177352.1 | TGACCTAACTCGAGGAGGAGCTGGAATC | AAGTTTGAGGCAGTTAAAATTATGGCTGAAGC |
| 14 | mouse CD44 | NM\_001039150.1 | AGAAAAATGGCCGCTACAGTTATC | TGCATGTTTCAAAACCCTTGC |
| 15 | mouse CD133 | NM\_008935.2 | ACCGGAGAGGGATGGTACTTTG | GCACTGAAGACGAGAGCCATAG |
| 16 | mouse ALDH | NM\_009656.4 | AACACAGGTTGGCAAGTTAATCA | TGCGACACAACATTGGCCTT |
| 17 | mouse ITGB1 | NM\_010578.2 | ATGCCAAATCTTGCGGAGAAT | TTTGCTGCGATTGGTGACATT |
| 18 | mouse ITGA6 | NM\_008397.4 | CCTAACAGAATTGACCTCCGCCAGAAG | ACTGAACTCTCGATGACAACCCTGA |
| 19 | mouse ITGA9 | [NM\_001113514.1](https://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&id=165377055) | TGATTCCAGCAGCTCACAGG | AATCTGAGGACTGTGCTGCC |
| 20 | mouse PTK2 | NM\_007982.2 | CGGACACATGCAGTCTCTGT | ATTGCAACAGCCAAAGCTGG |
| 21 | mouse PIK3CA | NM\_008839.2 | GCCACAGACACTACTGCGTA | CACCGAACAGCAAAACTCCG |
| 22 | mouse PIK3CB | NM\_029094.3 | CTGATTTTACGGCGGCATGG | TGAGGGCCTCGTCAAACTTC |
| 23 | mouse PIK3CG | NM\_001146200.2 | ACCTGTGCCTTCTGCCTTAC | TGCGGCCTGAAACTTTTCTTC |
| 24 | mouse PIK3R1 | NM\_001077495.2 | TGCAGCCAAGGAACCGGG | ACGTGTCCTTCTCAGCAACTT |
| 25 | mouse PIK3R5 | NM\_001142633.3 | AAGTCCTTTGTCAGCAGTCCC | CTGGTAAACCTGCAGCAACAC |
| 26 | mouse CDKN1A | NM\_001111099.2 | GCCCGAGAACGGTGGAACTT | GACAAGGCCACGTGGTCCTC |
| 27 | mouse CCND1 | NM\_001379248.1 | AGAACAAGCAGACCATCCGC | GTCCTTGTTTAGCCAGAGGC |
| 28 | mouse FGF2 | NM\_008006.2 | GGCTGCTGGCTTCTAAGTGT | GTCCCGTTTTGGATCCGAGT |
| 29 | mouse E-cadherin | NM\_009864.3 | CGCACCAGGTATTCAACGCATC | GGCATCTTGTGTTTCCACCACG |
| 30 | mouse CK-19 | [NM\_001313963.1](https://www.ncbi.nlm.nih.gov/nuccore/NM_001313963.1) | ACCAAGTTTGAGACAGAACACG | CTCCTCAGGGCAGTAATTTCC |

Supplementary Table S3. Summary of the result of metadata analysis of breast cancer cell line.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Cell line** | **Clinical Subtypes** | | | **Molecular Subtype** | **Tumor** | **Scores of gene expression** | | | | | |
| **ER** | **PR** | **HER2** | **IL-6** | **IL-8** | **IL-15** | **FGF2** | **HBEGF** | **TGFB** |
| **HCC1143** | N | N | N | Basal A | DC | +++ | ++++ | ++ | ++ | ++ | ++ |
| **HCC1937** | N | N | N | Basal A | DC | ++ | + | ++ | +++ | +++ | ++ |
| **MDAMB436** | N | N | N | Basal A | AC | ++++ | ++++ | +++ | +++ | ++ | ++ |
| **MDAMB468** | N | N | N | Basal A | AC | ++ | +++ | ++ | ++ | ++ | ++ |
| **BT549** | N | N | N | Basal B | IDC | +++ | ++++ | ++ | ++++ | +++ | +++ |
| **HCC1395** | N | N | N | Basal B | DC | +++ | +++ | + | ++ | ++ | +++ |
| **Hs578T** | N | N | N | Basal B | IDC | +++ | ++++ | ++ | +++ | +++ | +++ |
| **MCF7** | P | P | N | Luminal A | IDC | + | + | + | + | ++ | +++ |
| **MDAMB134VI** | P | N | N | Luminal A | IDC | + | + | + | + | ++ | ++ |
| **MDAMB175VII** | P | N | N | Luminal A | IDC | + | ++ | + | + | ++ | ++ |
| **MDAMB415** | P | P/N | N | Luminal A | AC | + | + | ++ | + | ++ | + |
| **T47D** | P | P | N | Luminal A | IDC | + | + | + | + | ++ | + |
| **ZR751** | P | P/N | N | Luminal A | IDC | + | + | + | + | ++ | + |
| **BT474** | P | P | P | Luminal B | IDC | + | + | + | + | ++ | + |
| **MDAMD453** | N | N | P | HER2 Positive | AC | + | + | + | + | ++ | + |
| **SKBR3** | N | N | P | HER2 Positive | AC | + | + | ++ | + | ++ | + |

P: positive, N: Negative, P/N: positive or negative

++++: scores more than or equal to10

+++: scores more than or equal to 7.5 and less than 10

++: scores more than or equal to 5 and less than 7.5

+: scores less than 5

Supplementary Table S4. Comparison of the sphere forming ability of the cells.

|  |  |  |
| --- | --- | --- |
| References | Cell type | Estimated sphere frequency |
| Current study | miPS-FGF2 | 1/23 |
| miPS-FGF2-P1-GFP+ | 1/21 |
| miPS-FGF2-P1-GFP- | 1/34 | Cell numbers / injection | | | | | | |
|  | | 100 | 400 | 500 | 1000 | 2000 | 5000 | 10000 |
| Lau et al. 2014 [19] | GC16 CD44+ | 1/140 | 6/8 | - | 3/3 | 3/3 | ND | - | ND |
| GC21 CD44+ | 1/100 | 13/15 | 14/15 | 5/5 | 5/5 | 17/17 | 3/3 | 2/2 |
| GC38 CD44+ | 1/70 | - | - | 4/7 | 4/7 | 12/13 | - | 12/13 |
| Paquet-Fifield et al. 2018 [20] | CRC1-CON | 1/23.4 | - | - | 4/5 | - | - | 3/5 | - |
| CRC1-CLDN2 | 1/7 | - | - | 5/5 | - | - | 5/5 | - |
| Essex et al. 2019 [21] | TOP-GFP Low | 1/39 | 0/4 | - | - | 2/4 | - | 3/4 | ND |
| TOP-GFP High | 1/12 | 0/4 | - | - | 3/4 | - | 2/4 | 2/2 |
| TOP-GFP Low + MFCM | 1/36 | 2/4 | - | - | 2/4 | - | 4/4 | 12/13 |

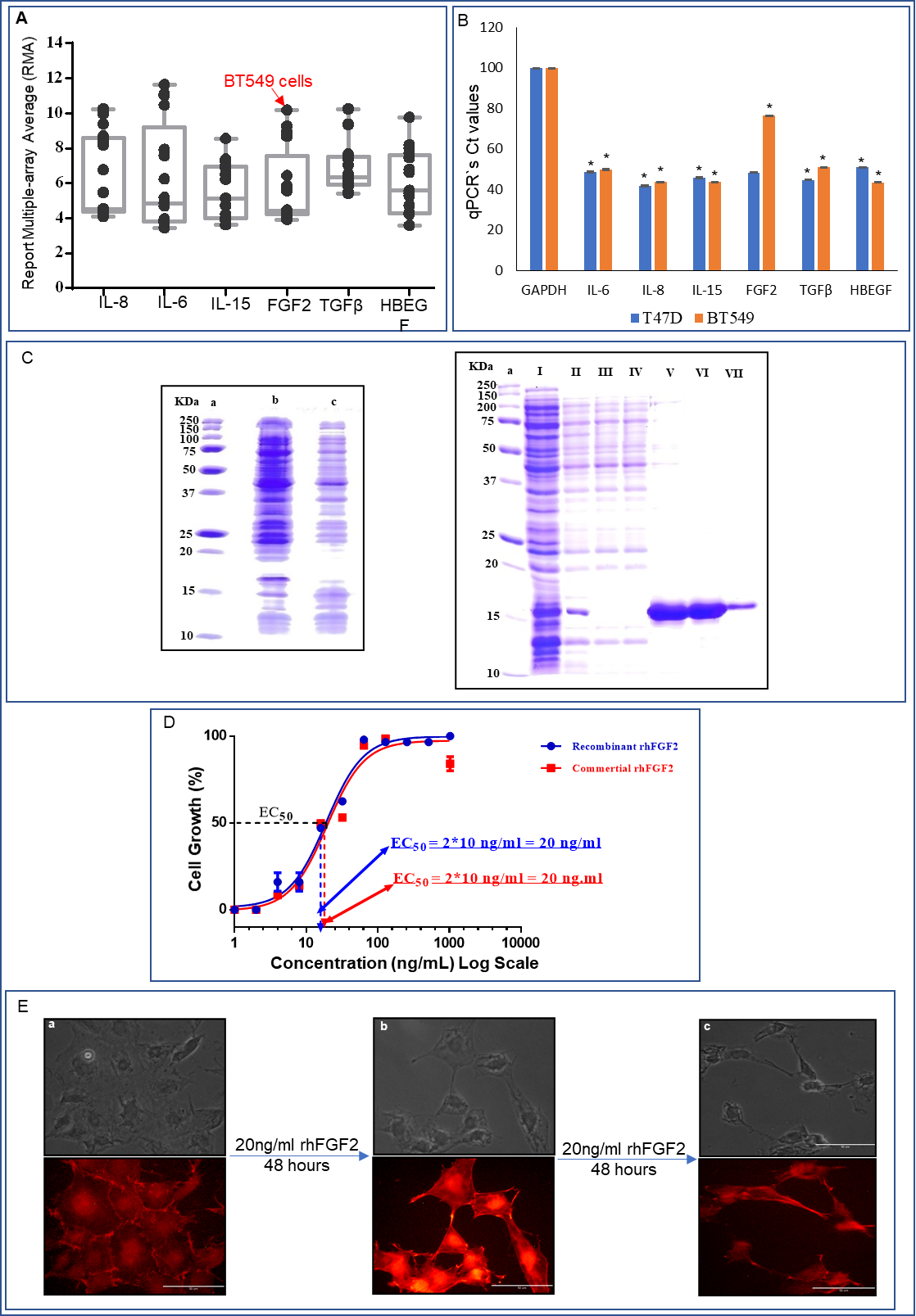
**-**: not described. ND: not detected.

Supplementary Table S6. Expression values of transcripts.

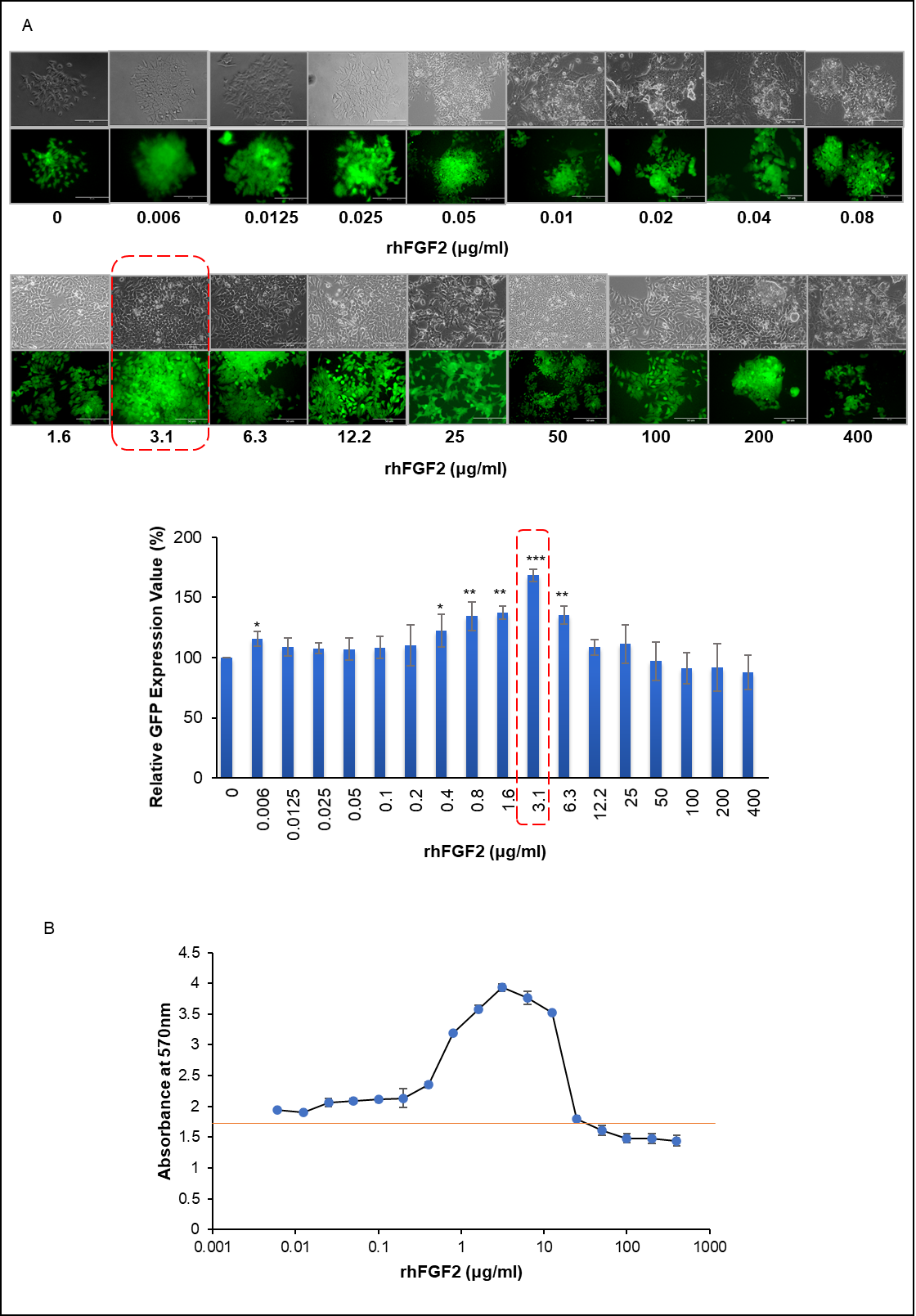
|  |  |  |  |
| --- | --- | --- | --- |
| FPKM\*1 | miPSC | miPS-FGF2-P1-GFP+ | miPS-FGF2-P1-GFP- |
| FGFR1 |  |  |  |
| Variant 1 | 10.3 | 20.5 | 22.7 |
| Variant 2 | 11.5 | 17.7 | 17.4 |
| Variant 3 | 0.0 | 7.8 | 6.5 |
| FGFR2 |  |  |  |
| Variant 1 | 0.6 | 7.3 | 9.1 |
| Variant 2 | 0.0 | 2.7 | 1.9 |
| Variant 3 | NA\*2 | NA | NA |

\*1FPKM: Fragments Per Kilobase of exon per Million mapped reads

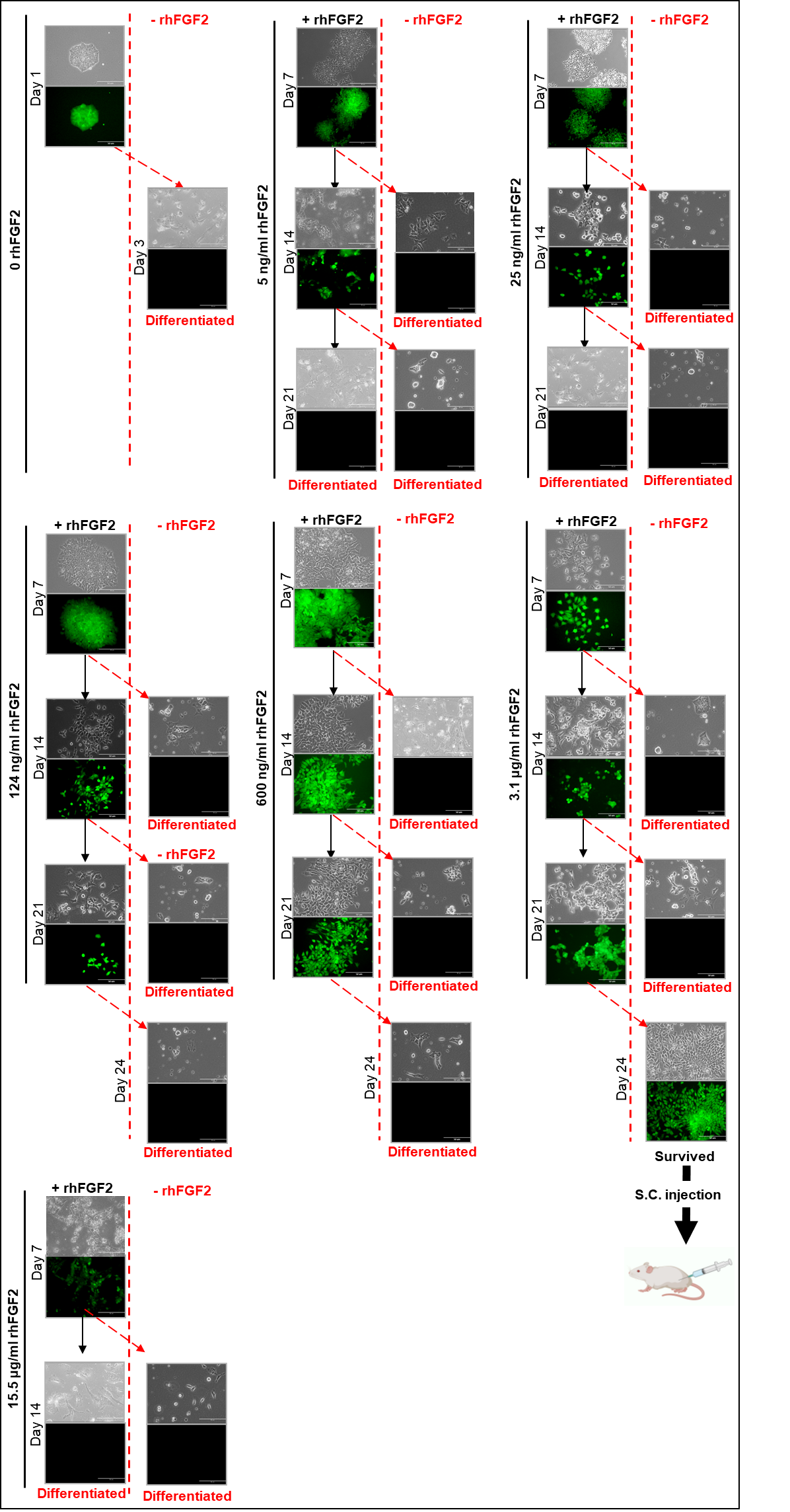
\*2NA: no information available



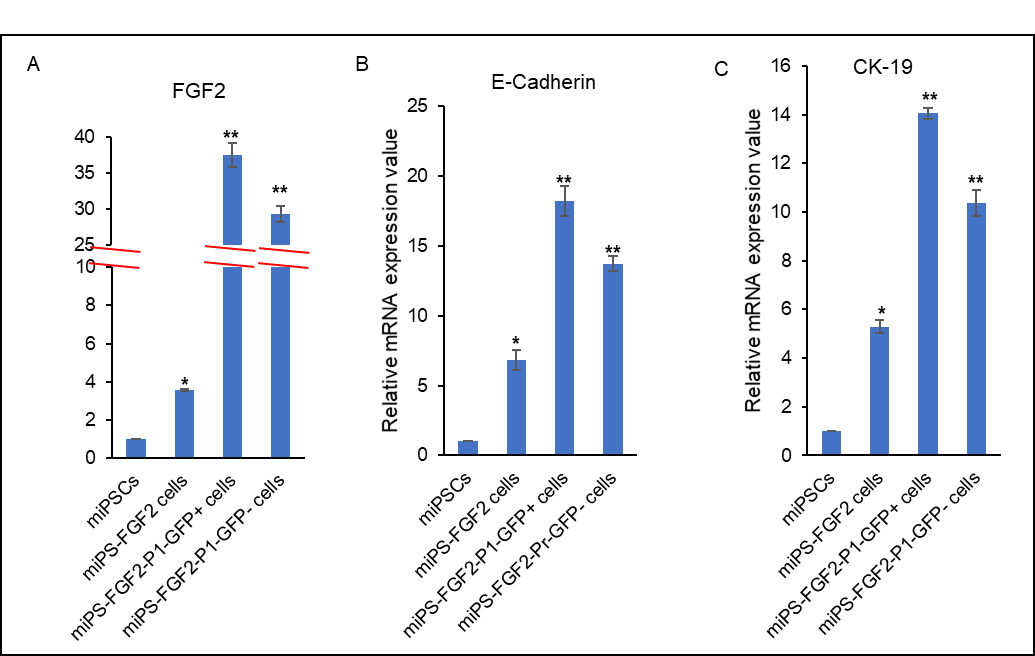
Supplementary Figure S1. (A) The box plot bar of robust multi-array average (RMA) values of cytokines and growth factors. The expression of IL-8, IL-6, IL-15, FGF2, HBEGF, and TGFβ genes in MCF7, MDA-MB-134, MDAMB175VII, MDAMB415, T47D, ZR751, BT474, MDAMD453, SKBR3, HCC1143, HCC1937, MDAMB436, MDAMB468, BT549, HCC1395, Hs578T, and MDAMB231 cells was evaluated from the microarray data in GEO. Boxes denote median values with upper and lower quartiles, and whiskers minimum and maximum outliers. Arrows show the spots from BT549 cells. (B) Comparison of Ct values from RT-qPCR results of BT549 and T47D cells assessing the expression of the genes for IL-8, IL-6, IL-15, FGF2, TFGβ, and HBEGF where the value of GAPDH used as house-keeping gene \**p* < 0.0001. C, SDS-PAGE of samples from the purification of bacterially expressed rhbFGF2. Lane a, standard molecular mass markers (kDa); lanes b and c, fractions from before and after induction with IPTG, respectively. Lane I, supernatant of cell lysate; lane II, fraction passed through DEAE cellulose column; lane III, fraction passed through heparin column; lane IV, fraction eluted from heparin column with 1M NaCl; lanes V, VI, & VII; fractions eluted from heparin column with 2M NaCl. (D) Cell growth ratio evaluated on Balb/c 3T3 cells with MTT in the presence of rhFGF2 (blue curve) in comparison with commercial rhFGF2 (red curve) as a positive control. (E) The effect of rhFGF2 on the morphology of Balb/c 3T3 cells. (a) Untreated cells. (b) Cells after 24-hour treatment with 20 ng/mL rhFGF2. (c) Cells after 48-hour treatment with 20 ng/mL rFGF2. The F-actin filaments were stained with phalloidin. Micrographs of phase contrast (top) and fluorescence (bottom). Scale bar indicates 50 um. Each result is shown as a representative of at least three independent experiments.



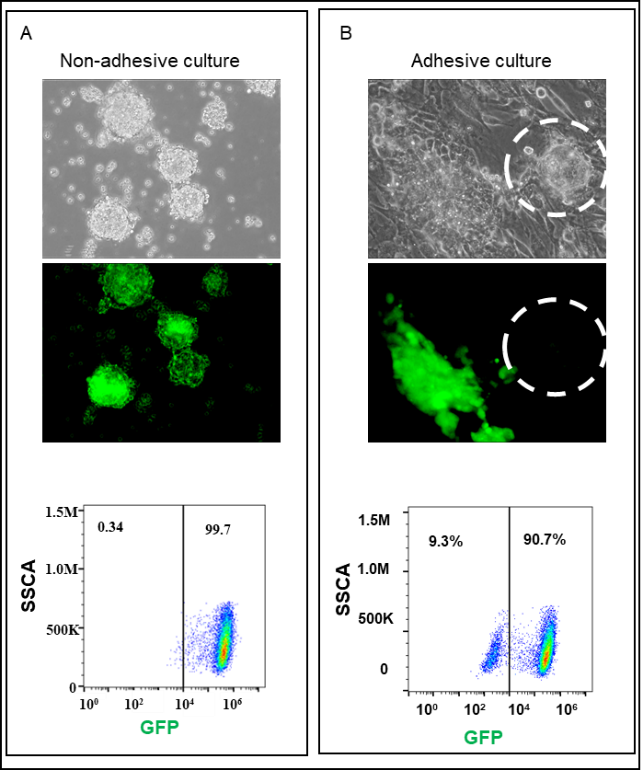
Supplementary Figure S2; The effect of various doses of rhFGF2 on miPSCs. (A) representative phase contrast and GFP fluorescence images of the cells treated with different concentrations of rhFGF2 for one week (top). The intensity of GFP fluorescence from the cells (bottom). (B) Absorbance of MTT at 540 nm was monitored to estimate the live cells by the treatment with different concentrations of rhFGF2 for one week. The cells were seeded in a 96-well plate in triplicates and each plot shows mean + SD. \**p* < 0.01, \*\**p* < 0.001, \*\*\**p* < 0.0001.



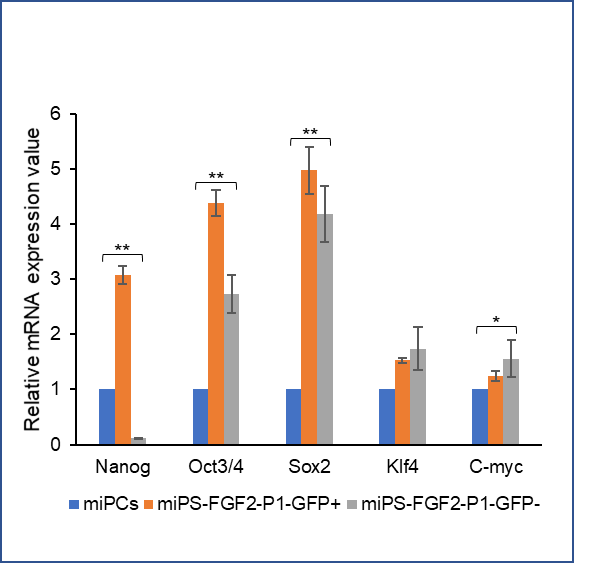
Supplementary Figure S3; Representative phase-contrast (top) and fluorescence (bottom) images for cells treated with different concentrations of rFGF2 for one, two, and three weeks. The cells were kept cultured without rhFGF2 at the end of each week and images were taken to assess GFP expression and cell survival. Cells treated with 5 and 25 ng/ml, failed to survive beyond week two. The treatment with 124 and 600 ng/ml maintained GFP expression and miPSCs survived beyond week three but they failed to survive without rhFGF2. rhFGF2 at 3.1 μg/ml maintained GFP expression for more than three weeks and miPSCs kept proliferated. After 21 days of treatment, miPSCs maintained GFP expression without rhFGF2. The survived cells were cultured for more 3 days without rhFGF2 and injected subcutaneously (s.c.) into BALBc nu/nu mice. Cells treated with 15.5 μg/ml failed to survive beyond one week.



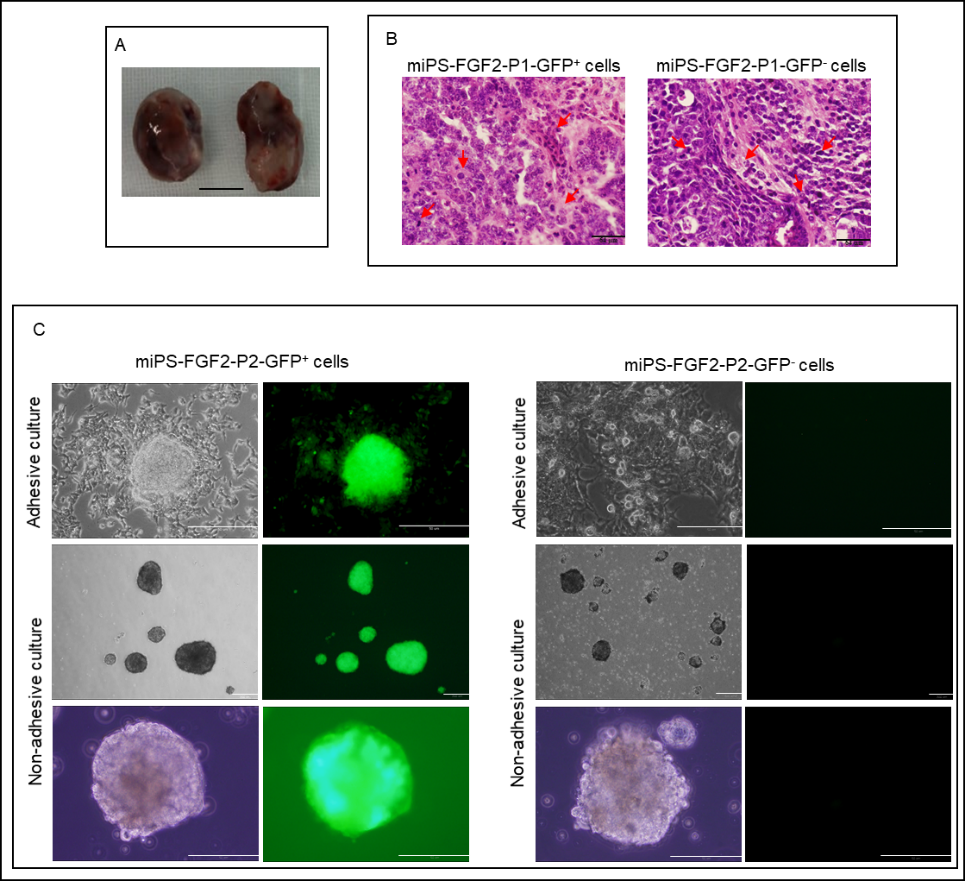
Supplementary Figure S4; Expression levels of FGF2, E-Cadherin and CK-19 compared among miPS, miPS-FGF2 miPS-FGF2-P1-GFP+, and miPS-FGF2-P1-GFP- cells by RT-qPCR. The expression of GAPDH gene used to normalize the expression level of each gene. \*p < 0.01, \*\*p < 0.001.



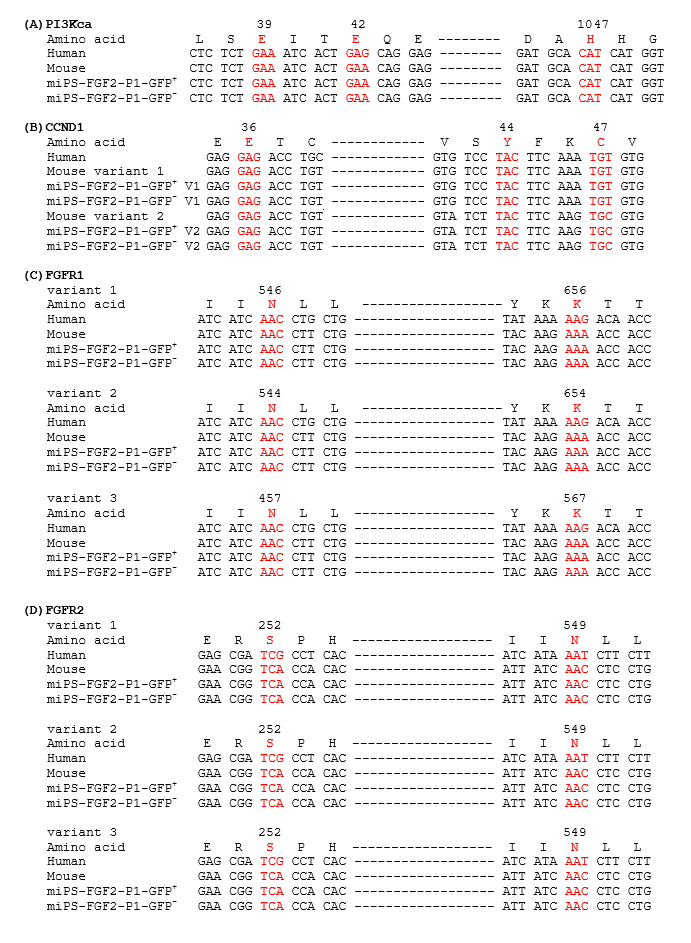
Supplementary Figure S5; Repsentative image of in miPS-FGF2-P1-GFP+. (A) Non- adhesive culture of miPS-FGF2-P1-GFP+ cells and (B) Adhesive culture miPS-FGF2-P1-GFP+ cells after puromycin treatment. White dashed circle, represent GFP- colony.



Supplementary Figure S6; Stem gene expression levels of Nanog, Oct3/4, Sox2, Klf4, and C-myc compared among miPS, miPS-FGF2-P1-GFP+, and miPS-rFGF2-P1-GFP- cells by RT-qPCR. \*p < 0.05, \*\* p<0.0001.



Supplementary Figure S7. Tumors and the primary cultures of miPS-FGF2-P1-GFP+/- cells. (A) Tumors of miPS-FGF2-P1-GFP+ cells (left) and miPS-FGF2-P1-GFP- cells (right). 1×106 of cells were injected subcutaneously into BALB/c-nu mice. Representative macroscopic features were photographed 30 days after injection. (B) H&E staining of the sections of the tumors in A. Red arrows indicate mitotic figures and severe nuclear atypia. Scale bars = 64 μm. (C) Representative picture of the adhesive culture and non-adhesive culture of miPS-FGF2-P2-GFP+/- cells. (20× and 4× magnification).



Supplementary Figure S8. Sequence comparison of the hot spots of frequent oncogenic mutations. The counter parts of human mRNA sequences of the hot spots in the COSMIC database (see text in detail) are compared with mouse sequences. (A) PI3Kca mRNA in human [NM\_006218.4], mouse [NM\_008839.3], miPS-FGF2-P1-GFP+ cells, and miPS-FGF2-P1-GFP- cells (B) CCND1 mRNA in human [NM\_053056.3], mouse variant 1[NM\_001379248.1], mouse variant 2 [NM\_007631.3], miPS-FGF2-P1-GFP+, miPS-FGF2-P1-GFP- cells. (C) FGFR1 mRNA in human variant 1 [NM\_023110.3], human variant 2 [NM\_015850.4], human variant 3 [NM\_023105.3], mouse variant 1 [NM\_010206.3], mouse variant 2 [NM\_001079908.2], mouse variant 3 [NM\_001079909.2], miPS-FGF2-P1-GFP+, miPS-FGF2-P1-GFP- cells. (D) FGFR2 mRNA in human variant 1 [NM\_000141.5], human variant 2 [NM\_022970.3], human variant 3 [NM\_001144913.1], mouse variant 1 [NM\_010207.2], mouse variant 2 [NM\_201601.2], mouse variant 3 [NM\_001347638.1], miPS-FGF2-P1-GFP+, miPS-FGF2-P1-GFP- cells. The translated amino acids from human sequences are depicted above of each codon with amino acid number. The frequent oncogenic spots of mutations are depicted in red. Each representative mRNA sequence mouse from database was identical with that obtained by NGS from miPSCs used in this study.