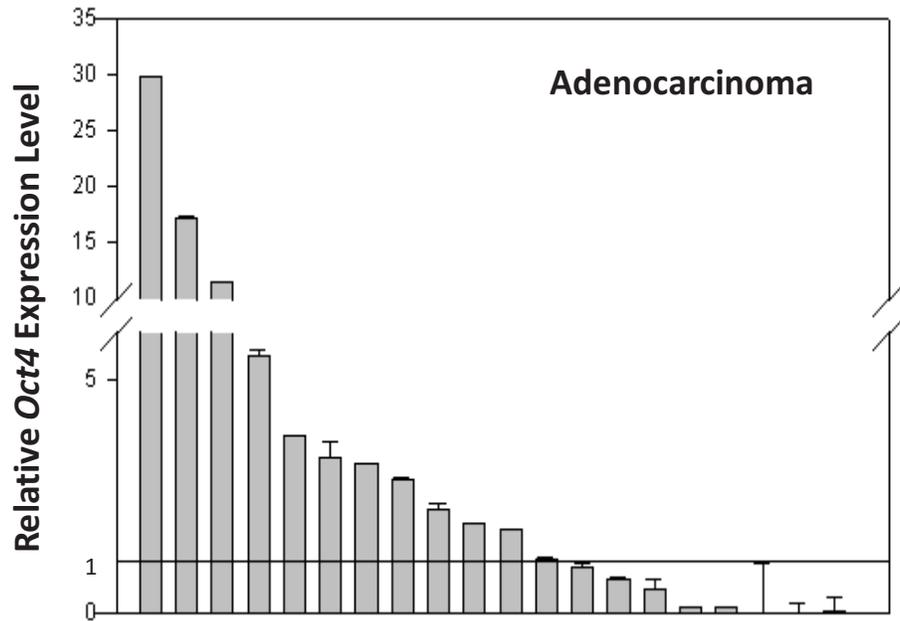
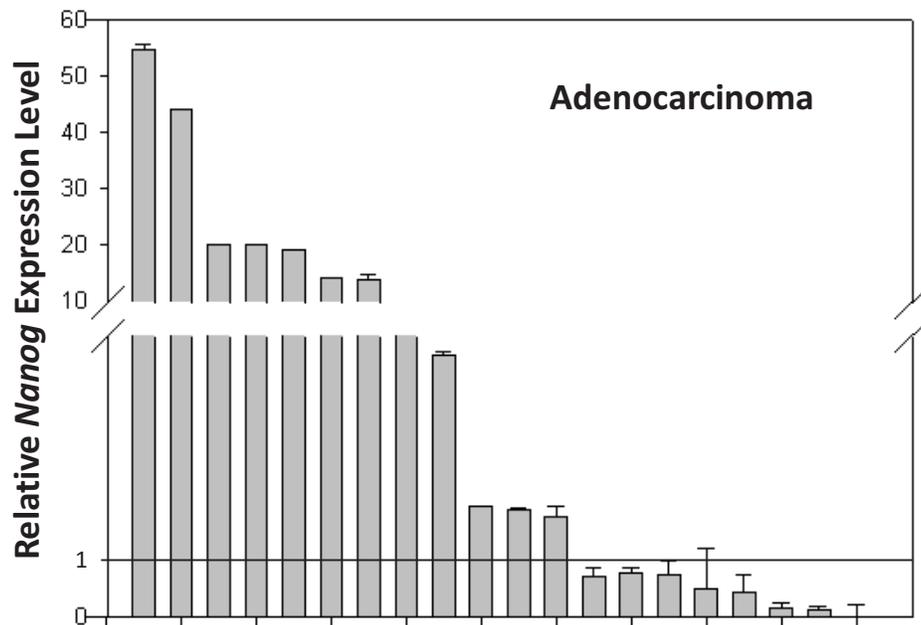


A

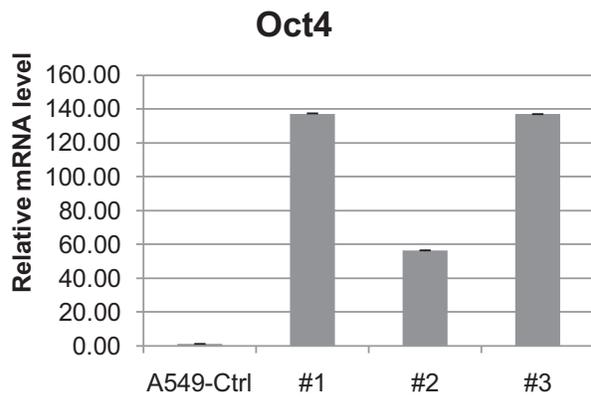
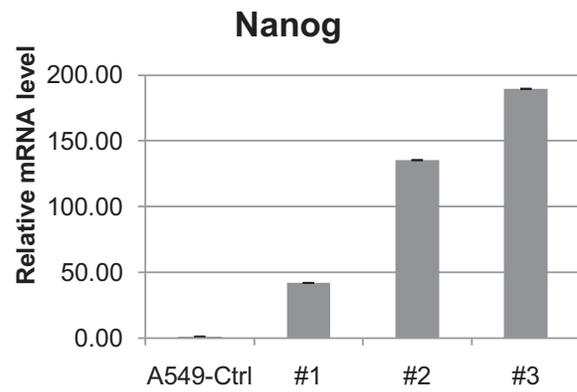
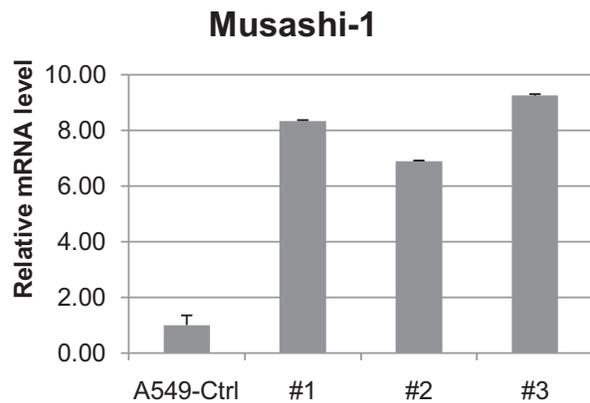
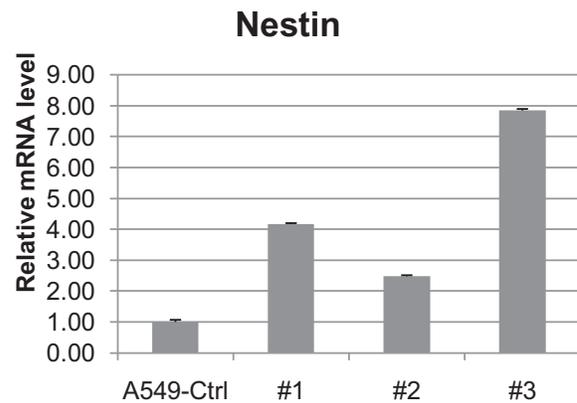


B

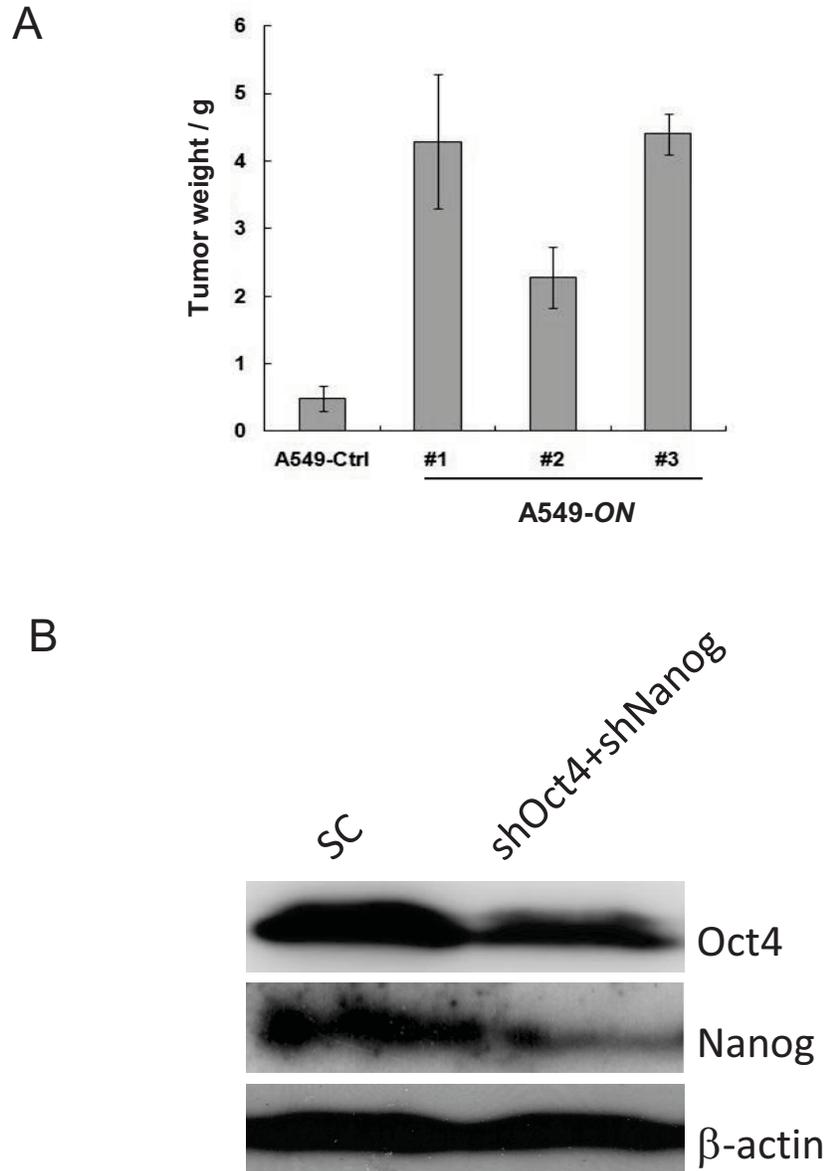


**Supplementary Fig. 1**

Total RNA from 20 pairs of primary LAC and matching adjacent non-tumorous tissues was purified and analyzed for *Oct4* and *Nanog* mRNA expressions by quantitative real-time PCR. The data was presented as relative fold comparing to its adjacent non-tumorous counter parts.

**A****B****C****D****Supplementary Fig. 2**

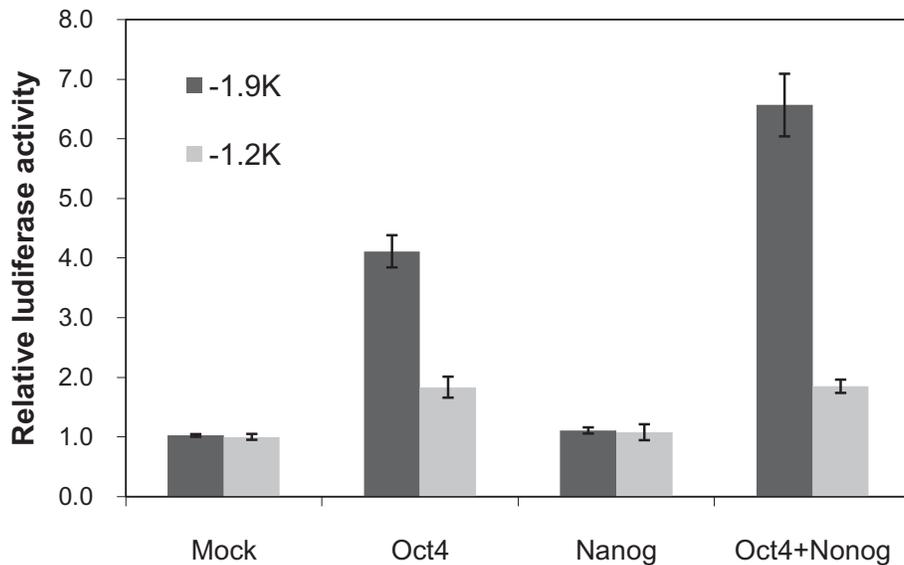
Total RNA was extracted from selected A549-ON clones (#1, #2, #3) and A549-Ctrl cells. The mRNA level of *Oct4* (A), *Nanog* (B), *Musashi-1* (C), and *Nestin* (D) were examined and quantified by quantitative real-time PCR. The graphs were presented as relative folds of gene expression level in A549-ON comparing to that in A549-Ctrl.



**Supplementary Fig. 3**

(A) The A549-Ctrl and the A549-ON clones were subcutaneously injected in 8-week-old male BALB/c nude mice. The tumors were surgically excised and weighted 8 weeks after injection. Each injection was performed in at least three individual mice. The graph demonstrated the mean values of each group with error bars indicating the difference between each mouse within the same experiment. The tumors generated by A549-ON cells were 5 to 8 folds heavier than those generated by A549-Ctrl cells. (B) The knockdown of *Oct4* and *Nanog* in the A549-ON cells transfected with shRNAs against these two stemness factors (shOct4+shNanog) was confirmed by Western blot. Scramble shRNA transfected A549-ON cells (SC) were used as a negative control, whereas  $\beta$ -actin was used as an internal control.

## Slug promoter



### Supplementary Fig. 4

Two reporter plasmids containing *Slug* promoter region either from -1.2K (-1.2K) or -1.9K (-1.9K) upstream of its start codon were constructed in the pGL3 plasmid. A549 cells were transiently transfected with either of the *Slug* promoter reporter plasmids, together with plasmids expressing *Oct4* and/or *Nanog* as indicated. Renilla luciferase plasmid (pRL-SV40) was co-transfected as an internal control (mock). Cell lysate was collected 48 hr after transfection, and the dual-luciferase assay was performed according to the manufacturer's protocol (Promega). The graph was presented as relative values of luciferase activity in the *Oct4*- and *Nanog*-transfected cells comparing to that in mock transfected cells. *Oct4* remarkably enhanced the activity of the reporter plasmid containing *Slug* promoter region from -1.9K, but not -1.2K. *Nanog* had little effect on both reporter activities. However, co-transfection of *Nanog* with *Oct4* enhanced the effect of *Oct4* on the -1.9K reporter activity. This data clearly showed that *Oct4* and *Nanog* synergistically enhance *Slug* transcription, and the region between -1.9K and -1.2K upstream of *Slug* start codon is important for this regulation.

**Supplementary Table 1. List of primers used in RT-qPCR**

|           |         |                                  |
|-----------|---------|----------------------------------|
| Oct4      | forward | 5'-TTTAATCCCACATCATGTATCACT-3'   |
|           | reverse | 5'-CTATCTACTGTGTCCCAGG-3'        |
| Nanog     | forward | 5'-CGATCTCCTGACCTTGT-3'          |
|           | reverse | 5'-CACGCCTGTAAATCCCA-3'          |
| CD133     | forward | 5''-CTATTCAGGATATACTCTCAGCATT-3' |
|           | reverse | 5'- TTTCTGTGGATGTAACCTTTCAGTG-3' |
| ABCB1     | forward | 5'-GCTCGTGCCCTTGTTAGAC-3'        |
|           | reverse | 5'-CTCCTTGACTCTGCCATTCTG-3'      |
| ABCG5     | forward | 5'-GCTGAACGCTGTGAATCTG-3'        |
|           | reverse | 5'-TGGCAACAACGCTGAAGG-3'         |
| Musashi-1 | forward | 5'-TCCCTCGGCGAGCACA -3'          |
|           | reverse | 5'-GACAGCCCCCCCACAAA-3'          |
| Nestin    | forward | 5'-ACCCTTCCAGACTCCACTC-3'        |
|           | reverse | 5'-CACTCCTCTTCTCCCTCCTC-3'       |
| Snail     | forward | 5'-CGAGCTGCAGGACTCTAAT-3'        |
|           | reverse | 5'- CCACTGTCCTCATCTGACA -3'      |
| Slug      | forward | 5'-TTCGGAC CCAC ACA TTACCT-3'    |
|           | reverse | 5'- TTGGAGCAGTTTTTGC ACTG-3'     |
| 18S RNA   | forward | 5'-ACAGGTCTGTGATGCC-3'           |
|           | reverse | 5'-ATCGGTAGTAGCGACG-3'           |

**Supplementary Table 2. The predominant processes up- (a) and down-regulated (b) in LC-CD133<sup>+</sup> cells**

**a**

| <b>Term</b>                   | <b>Count</b> | <b>%</b> | <b>P-value</b> |
|-------------------------------|--------------|----------|----------------|
| Mitosis                       | 23           | 9.2      | 6.63E-13       |
| Nuclear division              | 23           | 9.2      | 6.63E-13       |
| M phase of mitotic cell cycle | 23           | 9.2      | 9.57E-13       |
| Organelle fission             | 23           | 9.2      | 1.49E-12       |
| Cell division                 | 25           | 10       | 5.01E-12       |
| M phase                       | 26           | 10.4     | 7.85E-12       |
| Mitotic cell cycle            | 27           | 10.8     | 1.70E-11       |
| Cell cycle phase              | 28           | 11.2     | 3.76E-11       |
| Cell cycle                    | 38           | 15.2     | 6.72E-11       |
| Cell cycle process            | 32           | 12.8     | 9.51E-11       |

**b**

| <b>Term</b>                           | <b>Count</b> | <b>%</b> | <b>P-value</b> |
|---------------------------------------|--------------|----------|----------------|
| Immune system process                 | 29           | 12.2     | 3.08E-05       |
| Immune response                       | 23           | 9.7      | 3.71E-05       |
| Respiratory gaseous exchange          | 6            | 2.5      | 7.23E-05       |
| Cell-cell adhesion                    | 13           | 5.4      | 1.53E-04       |
| Regulation of response to stimulus    | 17           | 7.1      | 1.94E-04       |
| Regulation of immune effector process | 8            | 3.3      | 2.40E-04       |
| Immune effector process               | 9            | 3.7      | 2.43E-04       |
| Regulation of humoral immune response | 4            | 1.6      | 3.67E-04       |
| Cell adhesion                         | 21           | 8.8      | 3.68E-04       |
| Biological adhesion                   | 21           | 8.8      | 3.75E-04       |

**Supplementary Table 3a.**

**The up-expression gene profiling of lung cancer LC-CD133<sup>+</sup> cells as compared to LC-CD133<sup>-</sup> or low-grade LC tissues.**

| Probe SetID | Gene Symbol | Gene Title  | p-value  |
|-------------|-------------|---|----------|
| 205434_s_at | AAK1        | AP2 associated kinase 1   | 1.95E-10 |
| 32836_at    | AGPAT1      | 1-acylglycerol-3-phosphate O-acyltransferase 1 (lysophosphatidic acid acyltransferase, alpha) | 1.99E-09 |
| 202541_at   | AIMP1       | aminoacyl tRNA synthetase complex-interacting multifunctional protein 1                       | 5.03E-10 |
| 223144_s_at | AKIRIN2     | akirin 2  | 1.10E-09 |
| 212224_at   | ALDH1A1     | aldehyde dehydrogenase 1 family, member A1  | 2.75E-10 |
| 223418_x_at | ANKRD13C    | ankyrin repeat domain 13C   | 3.69E-09 |
| 202109_at   | ARFIP2      | ADP-ribosylation factor interacting protein 2   | 8.79E-10 |
| 226513_at   | ASB7        | ankyrin repeat and SOCS box-containing 7  | 4.19E-09 |
| 208079_s_at | AURKA       | aurora kinase A   | 4.20E-09 |
| 211946_s_at | BAT2D1      | BAT2 domain containing 1  | 5.92E-10 |
| 202265_at   | BMI1        | BMI1 polycomb ring finger oncogene  | 8.89E-10 |
| 202102_s_at | BRD4        | bromodomain containing 4  | 1.52E-10 |
| 209974_s_at | BUB3        | budding uninhibited by benzimidazoles 3 homolog (yeast)                                       | 9.28E-10 |
| 234947_s_at | C10orf84    | chromosome 10 open reading frame 84   | 1.83E-09 |
| 225836_s_at | C12orf32    | chromosome 12 open reading frame 32   | 3.35E-09 |
| 221777_at   | C12orf52    | chromosome 12 open reading frame 52   | 1.86E-09 |

|             |          |   |          |
|-------------|----------|---|----------|
| 225578_at   | C13orf37 | chromosome 13 open reading frame 37                     | 2.27E-10 |
| 203830_at   | C17orf75 | chromosome 17 open reading frame 75                     | 1.55E-09 |
| 218513_at   | C4orf43  | chromosome 4 open reading frame 43                      | 9.65E-12 |
| 48030_i_at  | C5orf4   | chromosome 5 open reading frame 4                       | 7.55E-10 |
| 228323_at   | CASC5    | cancer susceptibility candidate 5                       | 3.86E-09 |
| 225644_at   | CCDC117  | coiled-coil domain containing 117                       | 1.57E-11 |
| 234995_at   | CCDC52   | coiled-coil domain containing 52                        | 1.08E-09 |
| 213226_at   | CCNA2    | cyclin A2   | 1.50E-09 |
| 229091_s_at | CCNJ     | cyclin J  | 5.67E-10 |
| 212014_x_at | CD44     | CD44 molecule (Indian blood group)                      | 2.29E-10 |
| 203213_at   | CDC2     | cell division cycle 2, G1 to S and G2 to M              | 5.36E-10 |
| 217880_at   | CDC27    | cell division cycle 27 homolog (S. cerevisiae)          | 2.60E-10 |
| 210622_x_at | CDK10    | cyclin-dependent kinase 10                              | 3.61E-09 |
| 218542_at   | CEP55    | centrosomal protein 55kDa                               | 1.01E-10 |
| 201183_s_at | CHD4     | chromodomain helicase DNA binding protein 4             | 4.43E-09 |
| 211136_s_at | CLPTM1   | cleft lip and palate associated transmembrane protein 1 | 1.43E-09 |
| 201117_s_at | CPE      | carboxypeptidase E                                      | 2.79E-09 |
| 204263_s_at | CPT2     | carnitine palmitoyltransferase 2                        | 2.87E-10 |
| 201161_s_at | CSDA     | cold shock domain protein A                             | 3.49E-09 |
| 202574_s_at | CSNK1G2  | casein kinase 1, gamma 2                                | 1.22E-09 |
| 216389_s_at | DCAF11   | DDB1 and CUL4 associated factor 11                      | 1.00E-09 |
| 212919_at   | DCP2     | DCP2 decapping enzyme homolog (S. cerevisiae)           | 2.42E-09 |

|             |                    |  |          |
|-------------|--------------------|--|----------|
| 225228_at   | DRAM2              | DNA-damage regulated autophagy modulator 2   | 1.73E-09 |
| 203693_s_at | E2F3               | E2F transcription factor 3   | 7.96E-11 |
| 201437_s_at | EIF4E              | eukaryotic translation initiation factor 4E  | 7.08E-12 |
| 201936_s_at | EIF4G3             | eukaryotic translation initiation factor 4 gamma, 3  | 4.39E-09 |
| 208706_s_at | EIF5               | eukaryotic translation initiation factor 5   | 2.50E-09 |
| 200879_s_at | EPAS1              | endothelial PAS domain protein 1   | 2.66E-09 |
| 221267_s_at | FAM108A1           | family with sequence similarity 108, member A1   | 7.22E-12 |
| 227864_s_at | FAM125A            | family with sequence similarity 125, member A  | 8.34E-10 |
| 225703_at   | FBRSL1             | fibrosin-like 1  | 2.69E-09 |
| 232693_s_at | FBXO16 /// ZNF395  | F-box protein 16 /// zinc finger protein 395   | 2.14E-09 |
| 225734_at   | FBXO22             | F-box protein 22   | 7.80E-10 |
| 212499_s_at | FCF1 /// MAPK1IP1L | FCF1 small subunit (SSU) processome component homolog (S. cerevisiae) ///<br>mitogen-activated protein kinase 1 interacting protein 1-like | 1.40E-09 |
| 203689_s_at | FMR1               | fragile X mental retardation 1   | 4.68E-10 |
| 212824_at   | FUBP3              | far upstream element (FUSE) binding protein 3  | 2.90E-09 |
| 33579_i_at  | GALR3              | galanin receptor 3   | 3.86E-09 |
| 210005_at   | GART               | phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase,<br>phosphoribosylaminoimidazole synthetase              | 5.52E-11 |
| 200009_at   | GDI2               | GDP dissociation inhibitor 2   | 2.74E-09 |
| 204384_at   | GOLGA2             | golgi autoantigen, golgin subfamily a, 2   | 3.17E-09 |
| 1555913_at  | GON4L              | gon-4-like (C. elegans)  | 4.13E-09 |

|             |   |   |          |
|-------------|---|---|----------|
| 202320_at   | GTF3C1  | general transcription factor IIIC, polypeptide 1, alpha 220kDa  | 5.01E-12 |
| 242685_at   | GTPBP8  | GTP-binding protein 8 (putative)  | 2.03E-09 |
| 211940_x_at | H3F3A /// H3F3B /// LOC440926   | H3 histone, family 3A /// H3 histone, family 3B (H3.3B) /// H3 histone, family 3A pseudogene  | 2.59E-10 |
| 209524_at   | HDGFRP3   | hepatoma-derived growth factor, related protein 3   | 2.75E-09 |
| 200989_at   | HIF1A   | hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)   | 1.49E-10 |
| 216497_at   | HNRNPA1 /// HNRNPA1L2 /// HNRPA1L-2 /// HNRPA1P5 /// LOC100290205 /// LOC120364 /// LOC642817 /// LOC644037 /// LOC728643 | heterogeneous nuclear ribonucleoprotein A1 /// heterogeneous nuclear ribonucleoprotein A1-like 2 /// heterogeneous nuclear ribonucleoprotein A1 pseudogene 5 /// hypothetical protein LOC100290205 //   | 3.19E-09 |
| 217353_at   | HNRNPA1 /// HNRNPA1L2 /// HNRPA1L-2 /// HNRPA1P5 /// LOC120364 /// LOC642817 /// LOC644037 /// LOC728643                  | heterogeneous nuclear ribonucleoprotein A1 /// heterogeneous nuclear ribonucleoprotein A1-like 2 /// heterogeneous nuclear ribonucleoprotein A1 pseudogene 5 /// heterogeneous nuclear ribonucleoprotein A1 pseudogene 5 /// similar to heterogeneous nuclear rib | 4.19E-09 |
| 208765_s_at | HNRNPR  | heterogeneous nuclear ribonucleoprotein R   | 2.46E-09 |
| 208713_at   | HNRNPUL1  | heterogeneous nuclear ribonucleoprotein U-like 1  | 4.19E-09 |
| 210013_at   | HPX   | hemopexin   | 5.95E-10 |
| 202282_at   | HSD17B10  | hydroxysteroid (17-beta) dehydrogenase 10   | 2.10E-09 |
| 218637_at   | IMPACT  | Impact homolog (mouse)  | 7.77E-11 |
| 200995_at   | IPO7  | importin 7  | 1.54E-09 |
| 60528_at    | JMJD7-PLA2G4B /// PLA2G4B   | JMJD7-PLA2G4B readthrough transcript /// phospholipase A2, group IVB (cytosolic)  | 2.01E-10 |

|             |              |  |          |
|-------------|--------------|--|----------|
| 228680_at   | KIF3A        | kinesin family member 3A   | 4.35E-10 |
| 220266_s_at | KLF4         | Kruppel-like factor 4 (gut)  | 7.48E-11 |
| 225267_at   | KPNA4        | karyopherin alpha 4 (importin alpha 3)                                 | 5.79E-10 |
| 208975_s_at | KPNB1        | karyopherin (importin) beta 1  | 1.53E-09 |
| 212193_s_at | LARP1        | La ribonucleoprotein domain family, member 1                           | 3.62E-09 |
| 214155_s_at | LARP4        | La ribonucleoprotein domain family, member 4                           | 3.27E-09 |
| 203276_at   | LMNB1        | lamin B1   | 2.93E-09 |
| 213089_at   | LOC100272216 | hypothetical LOC100272216  | 1.03E-09 |
| 202651_at   | LPGAT1       | lysophosphatidylglycerol acyltransferase 1                             | 1.28E-09 |
| 202904_s_at | LSM5         | LSM5 homolog, U6 small nuclear RNA associated ( <i>S. cerevisiae</i> ) | 2.56E-11 |
| 213490_s_at | MAP2K2       | mitogen-activated protein kinase kinase 2                              | 2.39E-09 |
| 226048_at   | MAPK8        | mitogen-activated protein kinase 8                                     | 1.82E-09 |
| 214363_s_at | MATR3        | matrin 3   | 4.00E-09 |
| 229711_s_at | MDM2         | Mdm2 p53 binding protein homolog (mouse)                               | 4.08E-09 |
| 222466_s_at | MRPL42       | mitochondrial ribosomal protein L42                                    | 7.74E-10 |
| 202431_s_at | MYC          | v-myc myelocytomatosis viral oncogene homolog (avian)                  | 3.22E-10 |
| 220184_at   | NANOG        | Nanog homeobox   | 6.73E-10 |
| 208754_s_at | NAPIL1       | nucleosome assembly protein 1-like 1                                   | 7.09E-10 |
| 225065_x_at | NCRNA00188   | non-protein coding RNA 188   | 1.69E-10 |
| 225304_s_at | NDUFA11      | NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 11, 14.7kDa        | 3.27E-09 |
| 227062_at   | NEAT1        | nuclear paraspeckle assembly transcript 1 (non-protein coding)         | 6.92E-10 |

|             |  |   |          |
|-------------|--|---|----------|
| 216263_s_at | NGDN   | neuroguidin, EIF4E binding protein  | 4.22E-09 |
| 225921_at   | NIN  | ninein (GSK3B interacting protein)  | 3.82E-09 |
| 222497_x_at | NMD3   | NMD3 homolog (S. cerevisiae)  | 3.34E-09 |
| 225047_at   | NUPL1  | nucleoporin like 1  | 6.45E-10 |
| 201282_at   | OGDH   | oxoglutarate (alpha-ketoglutarate) dehydrogenase (lipoamide)  | 4.66E-10 |
| 44617_at    | OGFOD2   | 2-oxoglutarate and iron-dependent oxygenase domain containing 2   | 5.66E-10 |
| 201544_x_at | PABPN1   | poly(A) binding protein, nuclear 1  | 2.50E-09 |
| 1555823_at  | PACS2  | phosphofurin acidic cluster sorting protein 2   | 1.84E-09 |
| 229043_at   | PAPD5  | PAP associated domain containing 5  | 1.58E-09 |
| 213372_at   | PAQR3  | progesterin and adipoQ receptor family member III   | 4.76E-10 |
| 205060_at   | PARG   | poly (ADP-ribose) glycohydrolase  | 1.09E-09 |
| 204992_s_at | PFN2   | profilin 2  | 1.04E-09 |
| 1554316_at  | PGLS   | 6-phosphogluconolactonase   | 4.40E-09 |
| 213074_at   | PHIP   | pleckstrin homology domain interacting protein  | 1.23E-09 |
| 206138_s_at | PI4KB  | phosphatidylinositol 4-kinase, catalytic, beta  | 2.28E-10 |
| 208502_s_at | PITX1  | paired-like homeodomain 1   | 9.41E-11 |
| 204886_at   | PLK4   | polo-like kinase 4 (Drosophila)   | 6.15E-10 |
| 179_at      | PMS2L11  | postmeiotic segregation increased 2-like 11 pseudogene  | 2.46E-10 |
| 208286_x_at | POU5F1 /// POU5F1B /// POU5F1P3 ///<br>POU5F1P4///OCT4 | POU class 5 homeobox 1 /// POU class 5 homeobox 1B /// POU class 5 homeobox 1<br>pseudogene 3 /// POU class 5 homeobox 1 pseudogene 4 | 2.86E-09 |
| 218010_x_at | PPDPF  | pancreatic progenitor cell differentiation and proliferation factor homolog (zebrafish)   | 3.39E-09 |

|             |              |  |          |
|-------------|--------------|--|----------|
| 208932_at   | PPP4C        | protein phosphatase 4 (formerly X), catalytic subunit        | 1.21E-09 |
| 205628_at   | PRIM2        | primase, DNA, polypeptide 2 (58kDa)                          | 1.20E-09 |
| 213951_s_at | PSMC3IP      | PSMC3 interacting protein                                    | 5.30E-11 |
| 200733_s_at | PTP4A1       | protein tyrosine phosphatase type IVA, member 1              | 2.28E-10 |
| 222077_s_at | RACGAP1      | Rac GTPase activating protein 1                              | 1.01E-09 |
| 213205_s_at | RAD54L2      | RAD54-like 2 ( <i>S. cerevisiae</i> )                        | 3.19E-10 |
| 209050_s_at | RALGDS       | ral guanine nucleotide dissociation stimulator               | 4.14E-09 |
| 200749_at   | RAN          | RAN, member RAS oncogene family                              | 9.96E-10 |
| 210120_s_at | RANBP3       | RAN binding protein 3  | 1.45E-09 |
| 222810_s_at | RASAL2       | RAS protein activator like 2                                 | 9.02E-10 |
| 225751_at   | RBM17        | RNA binding motif protein 17                                 | 1.90E-09 |
| 212104_s_at | RBM9         | RNA binding motif protein 9                                  | 1.20E-09 |
| 225813_at   | RC3H2        | ring finger and CCCH-type zinc finger domains 2              | 1.73E-09 |
| 242079_at   | RGS12        | regulator of G-protein signaling 12                          | 2.28E-09 |
| 214700_x_at | RIF1         | RAP1 interacting factor homolog (yeast)                      | 8.55E-11 |
| 218499_at   | RP6-213H19.1 | serine/threonine protein kinase MST4                         | 3.48E-09 |
| 211720_x_at | RPLP0        | ribosomal protein, large, P0                                 | 3.01E-10 |
| 200024_at   | RPS5         | ribosomal protein S5   | 3.00E-09 |
| 219037_at   | RRP15        | ribosomal RNA processing 15 homolog ( <i>S. cerevisiae</i> ) | 3.43E-09 |
| 222465_at   | RSL24D1      | ribosomal L24 domain containing 1                            | 1.32E-10 |
| 201542_at   | SAR1A        | SAR1 homolog A ( <i>S. cerevisiae</i> )                      | 1.25E-09 |

|             |          |   |          |
|-------------|----------|---|----------|
| 212329_at   | SCAP     | SREBF chaperone   | 3.71E-09 |
| 227369_at   | SERBP1   | SERPINE1 mRNA binding protein 1   | 7.55E-10 |
| 224625_x_at | SERF2    | small EDRK-rich factor 2  | 2.55E-09 |
| 1569105_at  | SETD5    | SET domain containing 5   | 1.00E-09 |
| 211784_s_at | SFRS1    | splicing factor, arginine/serine-rich 1   | 1.58E-10 |
| 40149_at    | SH2B1    | SH2B adaptor protein 1  | 4.12E-09 |
| 209900_s_at | SLC16A1  | solute carrier family 16, member 1 (monocarboxylic acid transporter 1)                            | 7.21E-11 |
| 223649_s_at | SLC25A39 | solute carrier family 25, member 39   | 2.43E-09 |
| 225223_at   | SMAD5    | SMAD family member 5  | 8.36E-10 |
| 212520_s_at | SMARCA4  | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4 | 1.79E-11 |
| 224842_at   | SMG1     | SMG1 homolog, phosphatidylinositol 3-kinase-related kinase (C. elegans)                           | 6.86E-10 |
| 200071_at   | SMNDC1   | survival motor neuron domain containing 1   | 2.34E-09 |
| 219480_at   | SNAI1    | snail homolog 1 (Drosophila)  | 3.77E-09 |
| 213139_at   | SNAI2    | snail homolog 2 (Drosophila)  | 2.82E-09 |
| 201770_at   | SNRPA    | small nuclear ribonucleoprotein polypeptide A   | 3.92E-09 |
| 213721_at   | SOX2     | SRY (sex determining region Y)-box 2  | 1.27E-09 |
| 219055_at   | SRBD1    | S1 RNA binding domain 1   | 1.32E-09 |
| 201224_s_at | SRRM1    | serine/arginine repetitive matrix 1   | 7.83E-10 |
| 208610_s_at | SRRM2    | serine/arginine repetitive matrix 2   | 1.24E-11 |
| 1554686_at  | STAU2    | staufen, RNA binding protein, homolog 2 (Drosophila)  | 1.16E-09 |
| 41657_at    | STK11    | serine/threonine kinase 11  | 3.45E-09 |

|             |        |  |          |
|-------------|--------|--|----------|
| 209358_at   | TAF11  | TAF11 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 28kDa | 5.45E-10 |
| 213811_x_at | TCF3   | transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)     | 4.81E-10 |
| 222439_s_at | THRAP3 | thyroid hormone receptor associated protein 3                                    | 2.35E-09 |
| 228619_x_at | TIPRL  | TIP41, TOR signaling pathway regulator-like ( <i>S. cerevisiae</i> )             | 2.81E-10 |
| 213883_s_at | TM2D1  | TM2 domain containing 1  | 2.88E-09 |
| 212279_at   | TMEM97 | transmembrane protein 97   | 8.04E-10 |
| 240044_x_at | TNRC6B | Trinucleotide repeat containing 6B   | 4.80E-10 |
| 200990_at   | TRIM28 | tripartite motif-containing 28   | 4.20E-09 |
| 203983_at   | TSNAX  | translin-associated factor X   | 2.71E-11 |
| 202096_s_at | TSPO   | translocator protein (18kDa)   | 1.39E-09 |
| 213943_at   | TWIST1 | twist homolog 1 ( <i>Drosophila</i> )  | 3.33E-09 |
| 209096_at   | UBE2V2 | ubiquitin-conjugating enzyme E2 variant 2  | 1.59E-09 |
| 209088_s_at | UBN1   | ubiquitin 1  | 3.45E-09 |
| 214881_s_at | UBTF   | upstream binding transcription factor, RNA polymerase I                          | 8.76E-10 |
| 225655_at   | UHRF1  | ubiquitin-like with PHD and ring finger domains 1                                | 4.43E-11 |
| 203519_s_at | UPF2   | UPF2 regulator of nonsense transcripts homolog (yeast)                           | 1.02E-09 |
| 201903_at   | UQCRC1 | ubiquinol-cytochrome c reductase core protein I                                  | 1.43E-10 |
| 201831_s_at | USO1   | USO1 homolog, vesicle docking protein (yeast)                                    | 3.30E-09 |
| 222387_s_at | VPS35  | vacuolar protein sorting 35 homolog ( <i>S. cerevisiae</i> )                     | 2.47E-09 |
| 213390_at   | ZC3H4  | zinc finger CCCH-type containing 4   | 1.38E-10 |

|             |        |   |          |
|-------------|--------|---|----------|
| 226496_at   | ZCCHC7 | zinc finger, CCHC domain containing 7   | 4.08E-09 |
| 212764_at   | ZEB1   | zinc finger E-box binding homeobox 1    | 1.70E-10 |
| 213065_at   | ZFC3H1 | zinc finger, C3H1-type containing       | 4.36E-10 |
| 222407_s_at | ZFP106 | zinc finger protein 106 homolog (mouse) | 1.71E-09 |
| 208019_at   | ZNF157 | zinc finger protein 157                 | 2.87E-09 |
| 228785_at   | ZNF281 | Zinc finger protein 281                 | 1.59E-09 |
| 212366_at   | ZNF292 | zinc finger protein 292                 | 1.89E-09 |
| 1553703_at  | ZNF791 | zinc finger protein 791                 | 2.04E-10 |

**Supplementary Table 3b.**

**The down-expression gene profiling of lung cancer LC-CD133<sup>+</sup> cells as compared to LC-CD133<sup>-</sup> or low-grade LC tissues.**

| Probe SetID | Gene Symbol | Gene Title  | p-value  |
|-------------|-------------|---|----------|
| 1552800_at  | ABHD11      | abhydrolase domain containing 11  | 4.20E-09 |
| 208637_x_at | ACTN1       | actinin, alpha 1  | 3.22E-10 |
| 209612_s_at | ADH1B       | alcohol dehydrogenase 1B (class I), beta polypeptide                                | 3.99E-09 |
| 38447_at    | ADRBK1      | adrenergic, beta, receptor kinase 1   | 2.79E-10 |
| 238642_at   | ANKRD13D    | ankyrin repeat domain 13 family, member D   | 4.30E-10 |
| 202652_at   | APBB1       | amyloid beta (A4) precursor protein-binding, family B, member 1 (Fe65)              | 1.95E-09 |
| 215579_at   | APOBEC3G    | apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G                 | 3.20E-09 |
| 200602_at   | APP         | amyloid beta (A4) precursor protein   | 4.43E-09 |
| 34206_at    | ARAP1       | ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 1                           | 3.20E-09 |
| 202956_at   | ARFGEF1     | ADP-ribosylation factor guanine nucleotide-exchange factor 1(brefeldin A-inhibited) | 1.74E-09 |
| 231090_s_at | ARID2       | AT rich interactive domain 2 (ARID, RFX-like)                                       | 7.91E-11 |
| 221230_s_at | ARID4B      | AT rich interactive domain 4B (RBP1-like)   | 1.77E-10 |
| 202208_s_at | ARL4C       | ADP-ribosylation factor-like 4C   | 2.07E-09 |
| 1555279_at  | ARMC8       | armadillo repeat containing 8   | 1.79E-10 |
| 232197_x_at | ARSB        | arylsulfatase B   | 5.22E-10 |
| 224791_at   | ASAP1       | ArfGAP with SH3 domain, ankyrin repeat and PH domain 1                              | 1.74E-09 |
| 212297_at   | ATP13A3     | ATPase type 13A3  | 4.42E-10 |
| 212255_s_at | ATP2C1      | ATPase, Ca <sup>++</sup> transporting, type 2C, member 1                            | 1.65E-09 |

|              |           |  |          |
|--------------|-----------|--|----------|
| 212041_at    | ATP6V0D1  | ATPase, H+ transporting, lysosomal 38kDa, V0 subunit d1      | 1.93E-09 |
| 214149_s_at  | ATP6V0E1  | ATPase, H+ transporting, lysosomal 9kDa, V0 subunit e1       | 3.99E-11 |
| 205198_s_at  | ATP7A     | ATPase, Cu++ transporting, alpha polypeptide                 | 2.13E-09 |
| 205416_s_at  | ATXN3     | ataxin 3   | 2.86E-09 |
| 91920_at     | BCAN      | brevican   | 2.19E-09 |
| 202331_at    | BCKDHA    | branched chain keto acid dehydrogenase E1, alpha polypeptide | 2.89E-09 |
| 203140_at    | BCL6      | B-cell CLL/lymphoma 6  | 8.70E-10 |
| 233093_s_at  | BIRC6     | baculoviral IAP repeat-containing 6                          | 1.70E-09 |
| 209308_s_at  | BNIP2     | BCL2/adenovirus E1B 19kDa interacting protein 2              | 8.16E-10 |
| 242515_x_at  | C11orf17  | chromosome 11 open reading frame 17                          | 1.97E-10 |
| 231970_at    | C14orf118 | Chromosome 14 open reading frame 118                         | 4.00E-10 |
| 223474_at    | C14orf4   | chromosome 14 open reading frame 4                           | 8.07E-10 |
| 1570552_at   | C18orf50  | chromosome 18 open reading frame 50                          | 1.97E-09 |
| 55705_at     | C19orf22  | chromosome 19 open reading frame 22                          | 3.81E-10 |
| 204700_x_at  | C1orf107  | chromosome 1 open reading frame 107                          | 3.88E-09 |
| 1558692_at   | C1orf85   | chromosome 1 open reading frame 85                           | 7.53E-10 |
| 214123_s_at  | C4orf10   | chromosome 4 open reading frame 10                           | 2.26E-09 |
| 62987_r_at   | CACNG4    | calcium channel, voltage-dependent, gamma subunit 4          | 3.19E-09 |
| 1552863_a_at | CACNG6    | calcium channel, voltage-dependent, gamma subunit 6          | 2.52E-09 |
| 216535_at    | CADM3     | cell adhesion molecule 3                                     | 3.03E-09 |
| 37012_at     | CAPZB     | capping protein (actin filament) muscle Z-line, beta         | 3.14E-09 |
| 226036_x_at  | CASP2     | caspase 2, apoptosis-related cysteine peptidase              | 1.22E-09 |

|             |          |  |          |
|-------------|----------|--|----------|
| 208728_s_at | CDC42    | cell division cycle 42 (GTP binding protein, 25kDa)                                      | 4.93E-10 |
| 229120_s_at | CDC42SE1 | CDC42 small effector 1   | 6.70E-10 |
| 223249_at   | CLDN12   | claudin 12   | 3.40E-10 |
| 219529_at   | CLIC3    | chloride intracellular channel 3   | 1.03E-09 |
| 228751_at   | CLK4     | CDC-like kinase 4  | 3.84E-10 |
| 225572_at   | CREB1    | cAMP responsive element binding protein 1  | 3.54E-10 |
| 208978_at   | CRIP2    | cysteine-rich protein 2  | 7.87E-10 |
| 223270_at   | CTDSPL2  | CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase like 2 | 4.25E-10 |
| 202901_x_at | CTSS     | cathepsin S  | 6.86E-10 |
| 217028_at   | CXCR4    | chemokine (C-X-C motif) receptor 4   | 1.60E-11 |
| 48580_at    | CXXC1    | CXXC finger 1 (PHD domain)   | 2.69E-10 |
| 209164_s_at | CYB561   | cytochrome b-561   | 7.47E-10 |
| 206539_s_at | CYP4F12  | cytochrome P450, family 4, subfamily F, polypeptide 12                                   | 3.74E-09 |
| 214334_x_at | DAZAP2   | DAZ associated protein 2   | 3.33E-09 |
| 234295_at   | DBR1     | debranching enzyme homolog 1 (S. cerevisiae)   | 1.07E-10 |
| 214384_s_at | DCTN2    | Dynactin 2 (p50)   | 1.38E-11 |
| 33768_at    | DMWD     | dystrophia myotonica, WD repeat containing   | 3.01E-10 |
| 238336_s_at | DNAJC21  | DnaJ (Hsp40) homolog, subfamily C, member 21   | 1.59E-09 |
| 227018_at   | DPP8     | dipeptidyl-peptidase 8   | 3.44E-09 |
| 215732_s_at | DTX2     | deltex homolog 2 (Drosophila)  | 1.01E-09 |
| 224616_at   | DYNC1LI2 | dynein, cytoplasmic 1, light intermediate chain 2  | 7.23E-10 |
| 220926_s_at | EDEM3    | ER degradation enhancer, mannosidase alpha-like 3  | 7.98E-10 |

|              |  |   |          |
|--------------|--|---|----------|
| 396_f_at     | EPOR   | erythropoietin receptor   | 8.40E-10 |
| 222765_x_at  | ESF1   | ESF1, nucleolar pre-rRNA processing protein, homolog (S. cerevisiae)  | 1.31E-09 |
| 230102_at    | ETV5   | Ets variant 5   | 4.93E-10 |
| 1552889_a_at | EXOC3L2  | exocyst complex component 3-like 2  | 3.29E-09 |
| 212981_s_at  | FAM115A  | family with sequence similarity 115, member A   | 2.86E-09 |
| 223019_at    | FAM129B  | family with sequence similarity 129, member B   | 2.12E-09 |
| 212929_s_at  | FAM21A ///<br>FAM21B ///<br>FAM21C ///<br>FAM21D | family with sequence similarity 21, member A /// family with sequence similarity 21, member B /// family with sequence similarity 21, member C /// family with sequence similarity 21, member D | 3.04E-10 |
| 217916_s_at  | FAM49B   | family with sequence similarity 49, member B  | 3.71E-09 |
| 226294_x_at  | FAM91A1  | family with sequence similarity 91, member A1   | 1.78E-09 |
| 235169_at    | FBXO27   | F-box protein 27  | 3.08E-09 |
| 212987_at    | FBXO9  | F-box protein 9   | 3.06E-09 |
| 223050_s_at  | FBXW5  | F-box and WD repeat domain containing 5   | 2.37E-09 |
| 1563639_a_at | FHAD1  | forkhead-associated (FHA) phosphopeptide binding domain 1   | 2.36E-09 |
| 1552664_at   | FLCN   | folliculin  | 4.79E-10 |
| 222693_at    | FNDC3B   | fibronectin type III domain containing 3B   | 4.71E-10 |
| 218843_at    | FNDC4  | fibronectin type III domain containing 4  | 1.44E-09 |
| 211074_at    | FOLR1  | folate receptor 1 (adult)   | 4.08E-09 |
| 228188_at    | FOSL2  | FOS-like antigen 2  | 7.90E-11 |
| 226613_at    | GATSL3 ///                                       | GATS protein-like 3 /// TBC1 domain family, member 10A  | 1.50E-09 |

| TBC1D10A     |          |   |          |
|--------------|----------|---|----------|
| 218114_at    | GGA1     | golgi associated, gamma adaptin ear containing, ARF binding protein 1       | 3.90E-09 |
| 200852_x_at  | GNB2     | guanine nucleotide binding protein (G protein), beta polypeptide 2          | 3.93E-09 |
| 221966_at    | GPR137   | G protein-coupled receptor 137  | 3.88E-09 |
| 204137_at    | GPR137B  | G protein-coupled receptor 137B   | 1.11E-09 |
| 201348_at    | GPX3     | glutathione peroxidase 3 (plasma)   | 3.52E-10 |
| 217509_x_at  | GRIK5    | glutamate receptor, ionotropic, kainate 5                                   | 8.50E-10 |
| 201501_s_at  | GRSF1    | G-rich RNA sequence binding factor 1  | 8.69E-10 |
| 202210_x_at  | GSK3A    | glycogen synthase kinase 3 alpha  | 2.84E-09 |
| 200696_s_at  | GSN      | gelsolin (amyloidosis, Finnish type)  | 1.09E-10 |
| 205752_s_at  | GSTM5    | glutathione S-transferase mu 5  | 4.04E-09 |
| 222604_at    | GTF3C3   | general transcription factor IIIC, polypeptide 3, 102kDa                    | 2.50E-10 |
| 1552980_at   | HAS3     | hyaluronan synthase 3   | 2.77E-09 |
| 200643_at    | HDLBP    | high density lipoprotein binding protein                                    | 2.18E-09 |
| 210428_s_at  | HGS      | hepatocyte growth factor-regulated tyrosine kinase substrate                | 2.26E-11 |
| 202772_at    | HMGCL    | 3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase                           | 2.81E-09 |
| 211932_at    | HNRNPA3  | heterogeneous nuclear ribonucleoprotein A3                                  | 1.40E-09 |
| 202146_at    | IFRD1    | interferon-related developmental regulator 1                                | 4.09E-09 |
| 36004_at     | IKBKG    | inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma | 3.21E-09 |
| 1568513_x_at | IL23A    | Interleukin 23, alpha subunit p19   | 1.97E-09 |
| 214021_x_at  | ITGB5    | Integrin, beta 5  | 2.13E-09 |
| 201362_at    | IVNS1ABP | influenza virus NS1A binding protein  | 2.58E-09 |

|              |              |  |          |
|--------------|--------------|--|----------|
| 231183_s_at  | JAG1         | Jagged 1 (Alagille syndrome)   | 1.32E-09 |
| 212447_at    | KBTBD2       | kelch repeat and BTB (POZ) domain containing 2   | 1.75E-09 |
| 208987_s_at  | KDM2A        | lysine (K)-specific demethylase 2A   | 3.64E-11 |
| 213242_x_at  | KIAA0284     | KIAA0284   | 3.51E-09 |
| 231807_at    | KIAA1217     | KIAA1217   | 1.13E-09 |
| 243927_x_at  | KIAA1429     | KIAA1429   | 1.57E-09 |
| 226254_s_at  | KIAA1430     | KIAA1430   | 2.18E-09 |
| 209212_s_at  | KLF5         | Kruppel-like factor 5 (intestinal)   | 5.52E-11 |
| 1553212_at   | KRT78        | keratin 78   | 2.61E-09 |
| 216641_s_at  | LAD1         | ladinin 1  | 5.11E-10 |
| 1554600_s_at | LMNA         | lamin A/C  | 7.58E-10 |
| 230292_at    | LOC100131993 | similar to hCG2020760  | 1.39E-09 |
| 227715_at    | LOC100132288 | hypothetical protein LOC100132288 /// MAFF interacting protein<br>/// MAFIP  | 2.41E-09 |
| 39854_r_at   | LOC100293124 | similar to patatin-like phospholipase domain containing 2 /// patatin-like phospholipase domain containing 2<br>/// PNPLA2 | 1.63E-10 |
| 1558641_at   | LOC202051    | hypothetical protein LOC202051   | 2.72E-10 |
| 1557207_s_at | LOC283177    | hypothetical protein LOC283177   | 3.78E-09 |
| 1558688_at   | LOC441461    | hypothetical LOC441461   | 3.64E-11 |
| 1557987_at   | LOC641298    | SMG1 homolog, phosphatidylinositol 3-kinase-related kinase pseudogene  | 3.09E-09 |
| 242331_x_at  | LOC642236    | similar to FRG1 protein (FSHD region gene 1 protein)   | 2.53E-09 |
| 1557094_at   | LOC728449    | hypothetical protein LOC728449   | 6.29E-10 |

|             |        |  |          |
|-------------|--------|--|----------|
| 221956_at   | LRCH4  | leucine-rich repeats and calponin homology (CH) domain containing 4          | 1.41E-10 |
| 202625_at   | LYN    | v-src-1 Yamaguchi sarcoma viral related oncogene homolog                     | 1.09E-09 |
| 202292_x_at | LYPLA2 | lysophospholipase II   | 1.21E-09 |
| 203514_at   | MAP3K3 | mitogen-activated protein kinase kinase kinase 3                             | 1.42E-10 |
| 213002_at   | MARCKS | myristoylated alanine-rich protein kinase C substrate                        | 1.20E-10 |
| 228798_x_at | MAZ    | MYC-associated zinc finger protein (purine-binding transcription factor)     | 6.06E-10 |
| 201987_at   | MED13  | mediator complex subunit 13  | 4.08E-10 |
| 212209_at   | MED13L | mediator complex subunit 13-like   | 4.75E-11 |
| 204349_at   | MED7   | mediator complex subunit 7   | 1.32E-09 |
| 212713_at   | MFAP4  | microfibrillar-associated protein 4  | 1.80E-09 |
| 211026_s_at | MGLL   | monoglyceride lipase   | 7.01E-10 |
| 226100_at   | MLL5   | myeloid/lymphoid or mixed-lineage leukemia 5 (trithorax homolog, Drosophila) | 1.38E-09 |
| 225125_at   | MMGT1  | membrane magnesium transporter 1   | 3.72E-09 |
| 204475_at   | MMP1   | matrix metalloproteinase 1 (interstitial collagenase)                        | 1.67E-10 |
| 218202_x_at | MRPL44 | mitochondrial ribosomal protein L44  | 2.63E-10 |
| 203200_s_at | MTRR   | 5-methyltetrahydrofolate-homocysteine methyltransferase reductase            | 1.83E-09 |
| 32811_at    | MYO1C  | myosin IC  | 6.78E-10 |
| 227400_at   | NFIX   | nuclear factor I/X (CCAAT-binding transcription factor)                      | 2.70E-09 |
| 222589_at   | NLK    | nemo-like kinase   | 8.17E-11 |
| 89476_r_at  | NPEPL1 | aminopeptidase-like 1  | 2.86E-09 |
| 218215_s_at | NR1H2  | nuclear receptor subfamily 1, group H, member 2                              | 2.17E-09 |
| 225592_at   | NRM    | nurim (nuclear envelope membrane protein)                                    | 2.70E-09 |

|              |                          |  |          |
|--------------|--------------------------|--|----------|
| 1569030_s_at | NUB1                     | negative regulator of ubiquitin-like proteins 1  | 4.16E-11 |
| 207545_s_at  | NUMB                     | numb homolog (Drosophila)  | 1.56E-09 |
| 218196_at    | OSTM1                    | osteopetrosis associated transmembrane protein 1   | 3.70E-10 |
| 38710_at     | OTUB1                    | OTU domain, ubiquitin aldehyde binding 1   | 2.06E-09 |
| 226140_s_at  | OTUD1                    | OTU domain containing 1  | 3.33E-09 |
| 220962_s_at  | PADI1                    | peptidyl arginine deiminase, type I  | 1.52E-09 |
| 206791_s_at  | PDE4C                    | phosphodiesterase 4C, cAMP-specific (phosphodiesterase E1 dunce homolog, Drosophila)   | 9.00E-10 |
| 225830_at    | PDZD8                    | PDZ domain containing 8  | 1.19E-09 |
| 1568986_x_at | PIGT                     | phosphatidylinositol glycan anchor biosynthesis, class T   | 2.23E-10 |
| 1558292_s_at | PIGW                     | phosphatidylinositol glycan anchor biosynthesis, class W   | 3.76E-09 |
| 218942_at    | PIP4K2C                  | phosphatidylinositol-5-phosphate 4-kinase, type II, gamma  | 1.59E-09 |
| 221854_at    | PKP1                     | plakophilin 1 (ectodermal dysplasia/skin fragility syndrome)   | 3.97E-09 |
| 209873_s_at  | PKP3                     | plakophilin 3  | 2.75E-09 |
| 205125_at    | PLCD1                    | phospholipase C, delta 1   | 2.24E-10 |
| 223370_at    | PLEKHA3                  | pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 3   | 2.31E-10 |
| 235360_at    | PLEKHM3                  | pleckstrin homology domain containing, family M, member 3  | 1.49E-09 |
| 1558897_at   | PLK5P                    | polo-like kinase 5 pseudogene  | 1.81E-09 |
| 208890_s_at  | PLXNB2                   | plexin B2  | 2.96E-09 |
| 209355_s_at  | PPAP2B                   | phosphatidic acid phosphatase type 2B  | 4.15E-11 |
| 203735_x_at  | PPFIBP1                  | PTPRF interacting protein, binding protein 1 (liprin beta 1)   | 1.49E-10 |
| 217136_at    | PP1A4A ///<br>PP1A4B /// | peptidylprolyl isomerase A (cyclophilin A)-like 4A /// peptidylprolyl isomerase A (cyclophilin A)-like 4B ///<br>peptidylprolyl isomerase A (cyclophilin A)-like 4C /// peptidylprolyl isomerase A (cyclophilin A)-like 4G | 4.15E-09 |

|             | PPIAL4C /// |   |          |
|-------------|-------------|---|----------|
|             | PPIAL4G     |   |          |
| 213225_at   | PPM1B       | protein phosphatase 1B (formerly 2C), magnesium-dependent, beta isoform             | 1.07E-10 |
| 201603_at   | PPP1R12A    | protein phosphatase 1, regulatory (inhibitor) subunit 12A                           | 1.32E-09 |
| 202014_at   | PPP1R15A    | protein phosphatase 1, regulatory (inhibitor) subunit 15A                           | 2.20E-09 |
| 225984_at   | PRKAA1      | protein kinase, AMP-activated, alpha 1 catalytic subunit                            | 2.89E-10 |
| 213518_at   | PRKCI       | protein kinase C, iota  | 8.17E-10 |
| 38269_at    | PRKD2       | protein kinase D2   | 3.95E-09 |
| 211711_s_at | PTEN        | phosphatase and tensin homolog  | 2.77E-09 |
| 212640_at   | PTPLB       | protein tyrosine phosphatase-like (proline instead of catalytic arginine), member b | 3.79E-09 |
| 204469_at   | PTPRZ1      | protein tyrosine phosphatase, receptor-type, Z polypeptide 1                        | 1.82E-09 |
| 204021_s_at | PURA        | purine-rich element binding protein A   | 5.89E-10 |
| 1556262_at  | PWRN1       | Prader-Willi region non-protein coding RNA 1  | 1.76E-09 |
| 227999_at   | PWWP2B      | PWWP domain containing 2B   | 2.97E-09 |
| 203132_at   | RB1         | retinoblastoma 1  | 9.55E-10 |
| 209885_at   | RHOD        | ras homolog gene family, member D   | 1.28E-10 |
| 229648_at   | RICS        | Rho GTPase-activating protein   | 1.44E-09 |
| 227366_at   | RILP        | Rab interacting lysosomal protein   | 5.72E-10 |
| 35685_at    | RING1       | ring finger protein 1   | 4.18E-09 |
| 205805_s_at | ROR1        | receptor tyrosine kinase-like orphan receptor 1                                     | 2.07E-09 |
| 210822_at   | RPL13P5     | ribosomal protein L13 pseudogene 5  | 1.80E-09 |
| 200908_s_at | RPLP2       | ribosomal protein, large, P2  | 2.60E-09 |

|             |          |   |          |
|-------------|----------|---|----------|
| 213350_at   | RPS11    | Ribosomal protein S11   | 1.22E-10 |
| 1568613_at  | RSPH3    | radial spoke 3 homolog (Chlamydomonas)  | 2.46E-09 |
| 225574_at   | RWDD4A   | RWD domain containing 4A  | 2.26E-09 |
| 235849_at   | SCARA5   | scavenger receptor class A, member 5 (putative)   | 1.31E-10 |
| 209206_at   | SEC22B   | SEC22 vesicle trafficking protein homolog B ( <i>S. cerevisiae</i> )  | 1.53E-09 |
| 209176_at   | SEC23IP  | SEC23 interacting protein   | 1.18E-10 |
| 223610_at   | SEMA5B   | sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5B | 3.11E-09 |
| 226619_at   | SENP1    | SUMO1/sentrin specific peptidase 1  | 4.69E-10 |
| 1554616_at  | SERPINB8 | serpin peptidase inhibitor, clade B (ovalbumin), member 8   | 1.60E-09 |
| 204362_at   | SKAP2    | src kinase associated phosphoprotein 2  | 3.61E-09 |
| 209681_at   | SLC19A2  | solute carrier family 19 (thiamine transporter), member 2   | 5.38E-10 |
| 212683_at   | SLC25A44 | solute carrier family 25, member 44   | 2.73E-09 |
| 202614_at   | SLC30A9  | solute carrier family 30 (zinc transporter), member 9   | 1.92E-10 |
| 213251_at   | SMARCA5  | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5   | 3.79E-11 |
| 222270_at   | SMEK2    | SMEK homolog 2, suppressor of mek1 ( <i>Dictyostelium</i> )   | 1.40E-10 |
| 205622_at   | SMPD2    | sphingomyelin phosphodiesterase 2, neutral membrane (neutral sphingomyelinase)  | 1.46E-09 |
| 207390_s_at | SMTN     | smoothelin  | 3.68E-09 |
| 1559343_at  | SNRPN    | small nuclear ribonucleoprotein polypeptide N   | 2.28E-09 |
| 1566342_at  | SOD2     | Superoxide dismutase 2, mitochondrial   | 1.28E-09 |
| 221844_x_at | SPCS3    | signal peptidase complex subunit 3 homolog ( <i>S. cerevisiae</i> )   | 3.36E-10 |
| 226765_at   | SPTBN1   | Spectrin, beta, non-erythrocytic 1  | 3.55E-10 |

|              |           |  |          |
|--------------|-----------|--|----------|
| 242748_at    | SREBF2    | sterol regulatory element binding transcription factor 2         | 5.89E-10 |
| 223103_at    | STARD10   | STAR-related lipid transfer (START) domain containing 10         | 1.80E-09 |
| 212625_at    | STX10     | syntaxin 10  | 1.65E-09 |
| 228853_at    | STYX      | serine/threonine/tyrosine interacting protein                    | 8.41E-10 |
| 222715_s_at  | SYNRG     | synergin, gamma  | 3.42E-09 |
| 227134_at    | SYTL1     | synaptotagmin-like 1   | 8.24E-12 |
| 223013_at    | TBL1XR1   | transducin (beta)-like 1 X-linked receptor 1                     | 5.60E-10 |
| 203743_s_at  | TDG       | thymine-DNA glycosylase  | 1.27E-10 |
| 202720_at    | TES       | testis derived transcript (3 LIM domains)                        | 1.55E-09 |
| 219292_at    | THAP1     | THAP domain containing, apoptosis associated protein 1           | 1.77E-10 |
| 40837_at     | TLE2      | transducin-like enhancer of split 2 (E(sp1) homolog, Drosophila) | 3.69E-09 |
| 1552426_a_at | TM2D3     | TM2 domain containing 3  | 9.54E-10 |
| 217758_s_at  | TM9SF3    | transmembrane 9 superfamily member 3                             | 6.95E-11 |
| 1553636_at   | TMCO5A    | transmembrane and coiled-coil domains 5A                         | 2.17E-09 |
| 1558281_a_at | TMEM184A  | transmembrane protein 184A                                       | 4.73E-11 |
| 221512_at    | TMEM222   | transmembrane protein 222  | 4.46E-10 |
| 222391_at    | TMEM30A   | transmembrane protein 30A  | 1.02E-09 |
| 231697_s_at  | TMEM49    | Transmembrane protein 49   | 3.60E-10 |
| 225302_at    | TMX3      | thioredoxin-related transmembrane protein 3                      | 1.21E-10 |
| 209295_at    | TNFRSF10B | tumor necrosis factor receptor superfamily, member 10b           | 1.12E-09 |
| 238627_at    | TRAPPC2L  | trafficking protein particle complex 2-like                      | 2.97E-09 |
| 202504_at    | TRIM29    | tripartite motif-containing 29                                   | 2.20E-10 |

|             |         |   |          |
|-------------|---------|---|----------|
| 218617_at   | TRIT1   | tRNA isopentenyltransferase 1   | 3.65E-10 |
| 223314_at   | TSPAN14 | tetraspanin 14  | 8.02E-10 |
| 203008_x_at | TXNDC9  | thioredoxin domain containing 9   | 2.32E-09 |
| 244121_at   | UBR3    | Ubiquitin protein ligase E3 component n-recogin 3 (putative)                | 1.60E-10 |
| 224827_at   | UBTD2   | ubiquitin domain containing 2   | 9.99E-10 |
| 41856_at    | UNC5B   | Unc-5 homolog B (C. elegans)  | 1.68E-09 |
| 225583_at   | UXS1    | UDP-glucuronate decarboxylase 1   | 2.42E-10 |
| 202205_at   | VASP    | vasodilator-stimulated phosphoprotein                                       | 4.69E-10 |
| 224562_at   | WASF2   | WAS protein family, member 2  | 2.51E-09 |
| 228953_at   | WHAMM   | WAS protein homolog associated with actin, golgi membranes and microtubules | 1.37E-10 |
| 225814_at   | XRN1    | 5'-3' exoribonuclease 1   | 1.70E-11 |
| 224953_at   | YIPF5   | Yip1 domain family, member 5  | 1.41E-09 |
| 212342_at   | YIPF6   | Yip1 domain family, member 6  | 4.01E-09 |
| 221741_s_at | YTHDF1  | YTH domain family, member 1   | 1.71E-09 |
| 228402_at   | ZBED3   | zinc finger, BED-type containing 3  | 1.11E-09 |
| 226284_at   | ZBTB2   | zinc finger and BTB domain containing 2                                     | 1.02E-09 |
| 225629_s_at | ZBTB4   | zinc finger and BTB domain containing 4                                     | 2.97E-10 |
| 231393_x_at | ZBTB43  | Zinc finger and BTB domain containing 43                                    | 1.69E-10 |
| 235145_at   | ZBTB7B  | zinc finger and BTB domain containing 7B                                    | 2.89E-11 |
| 231899_at   | ZC3H12C | zinc finger CCCH-type containing 12C  | 1.98E-09 |
| 221193_s_at | ZCCHC10 | zinc finger, CCHC domain containing 10                                      | 1.14E-09 |
| 221613_s_at | ZFAND6  | zinc finger, AN1-type domain 6  | 2.05E-09 |

|             |        |                         |          |
|-------------|--------|-------------------------|----------|
| 238711_s_at | ZNF148 | zinc finger protein 148 | 1.01E-09 |
| 233992_x_at | ZNF445 | zinc finger protein 445 | 2.54E-09 |
| 241348_at   | ZNF654 | zinc finger protein 654 | 4.64E-10 |

**Supplementary Table 4. Clinicopathological characteristics of 118 patients with lung adenocarcinoma**

| <b>Patient characteristics</b> | <b>Total of 118 patients</b>      |                        |
|--------------------------------|-----------------------------------|------------------------|
|                                | <b>Surgery + adjuvant therapy</b> | <b>Definitive CCRT</b> |
| <b>Numbers of patients</b>     | <b>50</b>                         | <b>68</b>              |
| <b>Age</b>                     |                                   |                        |
| <b>Range</b>                   | <b>48-79</b>                      | <b>45-74</b>           |
| <b>Median</b>                  | <b>60</b>                         | <b>58</b>              |
| <b>Cancer stage</b>            |                                   |                        |
| <b>I-II</b>                    | <b>7</b>                          | <b>10</b>              |
| <b>III-IV</b>                  | <b>43</b>                         | <b>58</b>              |
| <b>OCT4 staining</b>           |                                   |                        |
| <b>Negative</b>                | <b>17</b>                         | <b>23</b>              |
| <b>Positive</b>                | <b>33</b>                         | <b>45</b>              |
| <b>Nanog staining</b>          |                                   |                        |
| <b>Negative</b>                | <b>20</b>                         | <b>20</b>              |
| <b>Positive</b>                | <b>30</b>                         | <b>48</b>              |
| <b>Slug staining</b>           |                                   |                        |
| <b>Negative</b>                | <b>15</b>                         | <b>25</b>              |
| <b>Positive</b>                | <b>35</b>                         | <b>43</b>              |

CCRT: concurrent chemoradiotherapy

## Supplementary Materials and Methods

### Quantitative Real-Time PCR.

RNA was extracted from cultured cells using TRIzol (Invitrogen) according to the manufacturer's protocol. Extracted total RNA was then reverse transcribed into cDNA using the oligo(dT)12-18 primer with Superscript III reverse transcriptase (Invitrogen). Twenty-fold dilutions of each cDNA were prepared for subsequent PCR amplification with the SYBR Green I Master Kit using Lightcycler480 instrument (Roche Applied Science). Primer sequences designed to detect specific genes are listed in Supplementary Table 1.

### Immunoblotting.

Western blots were performed as described (1). Primary antibodies included anti-Oct4, anti-Nanog, anti-Snail, anti-Slug, anti-Cytokeratin 18 (Cell Signaling), anti-E-cadherin, anti-Vimentin, anti-N-cadherin (Santa Cruz Biotechnology), and anti- $\beta$ -actin (Sigma). HRP-conjugated secondary antibodies were obtained from Santa Cruz Biotechnology (1:5000).

### Slug-promoter assay

$5 \times 10^5$  A549 lung cancer cells were seeded in 24-well plates and allowed to attach for 24 hours. Cells were transfected with 0.2  $\mu$ g of pGL3 plasmids containing Slug promoter region either 1.9K (-1.9K) or 1.2K (-1.2K) upstream to its start codon, together with plasmids expressing *Oct4* and *Nanog* (0.2  $\mu$ g) as indicated. Fifty nanograms of Renilla luciferase plasmid (pRL-SV40) was co-transfected as an internal control. Cell lysate was collected 48 hours after transfection, and the dual-luciferase assay was performed according to the manufacturer's protocol (Promega). The intensity of luciferase activity was measured and quantified by luminal meter.

### Immunofluorescence Staining.

Cells were seeded on glass coverslips at ~30% confluence 24 hr before transfection. Forty-eight hours after transfection, cells were fixed with 4% formaldehyde in PBS for 15 min at room temperature and permeabilized with 0.5% Triton X-100 in PBS for 10 min. After blocking with 3% skim milk in PBS, cells were incubated with primary antibody diluted in PBS overnight at 4°C and then incubated with secondary antibody for 2 hr at room temperature. Cells were mounted with Prolong Gold antifade reagent (Invitrogen) and imaged with a fluorescence microscope.

### Tumor tissues acquirement and preparation

All procedures of tissues acquirements are following the tents of the Declaration of Helsinki and were got inform consents from the enrolled lung cancer patients. Resected tissues from 118 lung cancer patients (Supplementary Table 3), who gave informed consent for the use of their tissue, were harvested at surgery. Human primary lung carcinoma (P), normal paired non-cancerous matched tissues (N), along

with as well as available lymph node metastatic lesions (LN) and metastatic/secondary lung tumors (M) were obtained from surgical procedures send to the pathology lab for frozen section diagnosis. Tumor tissues with cryosections were subjected to H&E stain for morphological inspection and Ki-67 immunohistological staining for proliferation activity before microscopically screened to have > 75% of their areas occupied by tumor cells under two pathologists' survey (2). The specimen (tumor, normal counterpart, and lymph node metastatic lesions) were snap frozen in liquid nitrogen and stored at -80°C.

### **Isolation and cultivation of CD133-positive cells**

All procedures of tissues acquisitions are following the tents of the Declaration of Helsinki and were got inform consents from the enrolled lung cancer patients. In brief, the dissociated cells from the samples were labeled with 1 mL CD133/1 micromagnetic beads per 1 million cells using the CD133 cell isolation kit (Miltenyi Biotech, auburn, CA). CD133<sup>+</sup> cells were cultured in a medium consisting of serum-free DMEM/F12 (Gibco-BRL, Gaitherburg, MD), N2 supplement (R&D Systems Inc., Minneapolis), 10 ng/ml human recombinant bFGF (R&D Systems) and 10 ng/ml EGF (R&D System) (3).

## Reference List

1. Hong CF, Chou YT, Lin YS, Wu CW. MAD2B, a novel TCF4-binding protein, modulates TCF4-mediated epithelial-mesenchymal transdifferentiation. *J.Biol.Chem.* 2009;284:19613-22.
2. Dong QZ, Zhao Y, Liu Y et al. Overexpression of SCC-S2 correlates with lymph node metastasis and poor prognosis in patients with non-small-cell lung cancer. *Cancer Sci.* 2010;101:1562-9.
3. Chiou SH, Yu CC, Huang CY et al. Positive correlations of Oct-4 and Nanog in oral cancer stem-like cells and high-grade oral squamous cell carcinoma. *Clin.Cancer Res.* 2008;14:4085-95.