Supporting Information

Wang et al. 10.1073/pnas.1212769110

SI Materials and Methods

Vectors. DNA fragments corresponding to the premiR-200a, premiR-200b, premiR-200c, premiR-141, and premiR-429 were amplified by PCR from genomic DNA and cloned into the retroviral vector by using BamHI and SalI restriction sites. The cDNA of zinc finger E-box binding homeobox 2 (ZEB2) was inserted into pBabe-puro retrovirus vector by using SalI restriction site, and ZEB2-shRNA was inserted in pMko.1 vector. The DNA encoding ZEB1 3′ untranslated region (UTR), ZEB2 3′ UTR, mir-141/200c promoter with octamer-binding protein 4 (Oct4) binding region, and mir-200a/b/429 promoter with SRY-box containing gene 2 (Sox2) binding site were cloned into the pGL3 reporter vector by using SacI and XbaI restriction sites, whereas the corresponding mutant vectors were obtained by using QuikChange Lightning Multi-Site Directed Mutagenesis Kit (Agilent Technologies catalog no. 210515-5). All of the constructed plasmids were verified by DNA sequencing.

Luciferase Reporter Assay. NIH 3T3 cells grown in 24-well plates were transfected with 50 nM micr*ON* miRNA mimics or 150 nM micr*OFF* miRNA inhibitors synthesized by Ribobio, 0.4 μg of pGL-3 luciferase reporter vector containing 3′ UTR, and 0.01 μg of Renilla luciferase vector (pRL-TK; Promega) in the presence of the Lipofectamine 2000 reagent (Invitrogen). In the transcription activity assay of Oct4/Sox2 on mir-200s promoter, we used 0.1 μg of pGL-3 luciferase reporter vector, in combination with 0.1 μg, 0.2 μg, and 0.4 μg of pMx-Oct4 or pMx-Sox2. Firefly and Renilla luciferase activities in cell lysates were assayed with a Dual-Luciferase Reporter Assay System (Promega) at 48 h aftertransfection. Firefly luciferase activity was normalized by that of Renilla luciferase.

FACS Analysis. Mouse embryonic fibroblast (MEF) cells after infection with Oct4, Sox2, Klf4, c-Myc and individual member of the miR-200 family for 7 d were used to detect the proportion of E-cadherin (E-cad)–positive cells. For surface staining, cells were trypsinized, washed once in PBS, fixed in 4% (wt/vol) paraformaldehyde in PBS, washed twice in PBS, and incubated in 100 μL of 1% FBS in PBS containing a rabbit anti-mouse E-cadherin antibody (1:200 dilution; Cell Signaling Technology) for 1 h and then washed three times in PBS. The cells were resuspended in 100 μL of 1% FBS in PBS containing fluorophore-labeled donkey anti-rabbit IgG (H+L) antibodies Alexa Fluor 488 (1:500 dilution) and incubated for 30 min. The cells were washed three times and resuspended in 500 μL of PBS. The proportion of E-cad–positive cells was analyzed by flow cytometry by using Guava EasyCyte 8HT (Millipore).

PCR for Transgenes Genetic Integration and Exogenous Transgenes Silencing. Genomic DNA was extracted from induced pluripotent stem cells (iPSCs) by using DNA extraction kit (TianGen) to detect exogenous *Oct4*, *Sox2*, *Klf4*, *c-Myc*, mir-200a, mir-200b, mir-200c, mir-141, and mir-429 genetic integration. Primers for exo-Oct4, Sox2, Klf4, c-Myc, miR-200a, miR-200b, miR-200c, miR-141, and miR-429 integration are listed in Table S1. Primers for analysis of exogenous Oct4, Sox2, Klf4, and c-Myc transgene silencing are listed in Table S2.

Promoter Methylation Analyses. CpG methylation of the *Nanog* and *Oct4* (*Pou5f1*) promoters was analyzed as the following procedures. Briefly, genomic DNA was extracted by using genomic DNA extraction kit (TianGen). Totally, 1 μg of DNA was used for genome modification analysis. The bisulfate-PCR primers for mouse *Oct4* and *Nanog* promoter are listed in Table S3, and the products were cloned into pMD19-T vector (Takara) followed by sequencing.

Embryonic bodies (EBs) Formation and in Vitro Differentiation Assay. Mouse embryonic stem cells and iPSCs were harvested by trypsinization (0.25% trypsin) and transferred 5×10^5 to bacterial culture dishes (60-mm) in 5 mL of KOSR medium including Knockout DMEM (Gibco) + 20% Knockout Serum replacement (KOSR) + NEAA + L-Glutamine without Leukemia Inhibitory Factor (LIF) to generate EBs. RNA was extracted from EBs and used for differentiation analyses by qPCR.

In immunostaining analysis for in vitro differentiation of iPSCs, iPSCs were trypsinized into single-cell suspension and the hanging drop method was used to generate EBs. For each drop, 1×10^3 iPSCs in 20 μL of KOSR medium without LIF were used. EBs were cultured in hanging drops for 3 d before being reseeded onto gelatin-coated 48-well plates for another 6 d.

Alkaline Phosphatase (AP) Staining and Immunostaining. AP staining was carried out by using the FastRed Alkaline Phosphatase Kit (Sigma). For immunostaining, primary antibodies of anti-Oct4 (Santa Cruz), anti-Nanog (Abcam), anti–SSEA-1 (Santa Cruz), anti–E-cadherin (Abcam), anti-Tuj1 (Covance), anti-Gata4 (Santa Cruz), and anti–HNF-3β (Santa Cruz) were used.

Teratoma Formation and H&E Staining. To generate teratomas, iPSCs were trypsinized and resuspended at a concentration of 2×10^6 cells per 150 μL, and then injected into Athymus nude mice [non-obese diabetic (NOD)-severe combined immune deficiency (SCID)] approved by the Institutional Animal Care and Use Committee of Tongji University, which were obtained from National Resource Center of Mutant Mice (NRCMM), Model Animal Research Center of Nanjing University. Mice were checked for tumors every week for 4 wk. Tumors were harvested and fixed in formaldehyde solution for 24 h at room temperature before paraffin embedding and then for) H&E staining.

Western Blotting. Cells were lysed in SDS lysis buffer. An equal amount of cell lysis was separated on SDS/PAGE, blotted on nitrocellulose filter membrane, and probed with following primary antibodies: anti-ZEB2/SIP1 (Santa Cruz), and anti-GAPDH (Sigma). GAPDH was used as the loading control. After incubation with secondary antibodies, signals were visualized by enhanced chemiluminescence.

Statistical Analysis. The data in this study are means ± SD from three independent experiments; $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$, (Student's *t* test).

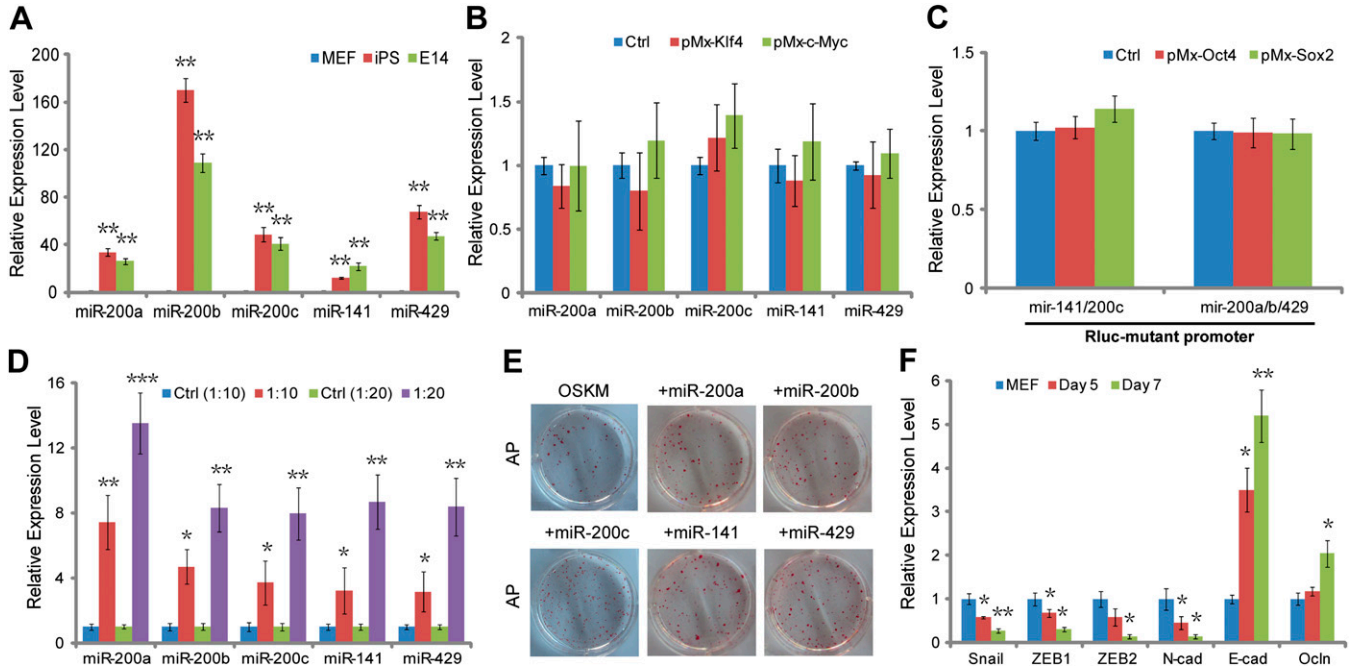

Fig. S1. Effects of the miR-200 family in OSKM-induced iPSC generation. (*A*) Relative expression level of miR-200a, miR-200b, miR-200c, miR-141, and miR-429 in mouse iPS (generated with OSKM) and ES cells (E14) by compared with that in MEF cells. (*B*) The expression level of miR-200a, miR-200b, miR-200c, miR-141, and miR-429 in MEFs, and MEFs after infection with pMx-Klf4 and pMx-c-Myc retroviruses. (*C*) Luciferase reporter assay for Oct4/Sox2 bindings at the promoter regions with mutant binding sites by transfection with 400 ng of pMx-Oct4 and pMx-Sox2. The pMx-GFP vector was used as a negative control (Ctrl). (*D*) qRT-PCR analyses for ectopic expression of the retroviral vectors containing miR-200a, miR-200b, miR-200c, miR-141, and miR-429. Cells after infection with the control vector viruses were used as the negative control (Ctrl), 1:10 and 1:20 represent the number of cells:viruses. U6 was used as an internal control. (*E*) Representative images of alkaline phosphatase-positive (AP^+) colonies on day 8 after retroviral infection of OSKM and OSKM in combination with miR-200a, miR-200b, miR-200c, miR-141, and miR-429. Cultures were fixed and stained for alkaline phosphatase activity. (*F*) qRT-PCR analyses for the expression levels of the mesenchymal genes (Snail, ZEB1, ZEB2, N-cad) and epithelial genes (E-cad, Ocln) in MEF cells after infection with OSKM viruses on day 5 and day 7 compared with that in MEFs. GAPDH was used as an internal control. Error bars denote the SD derived from three independent experiments. (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; Student's *t* test) ($n = 3$). (Scale bars: 100 μm.)

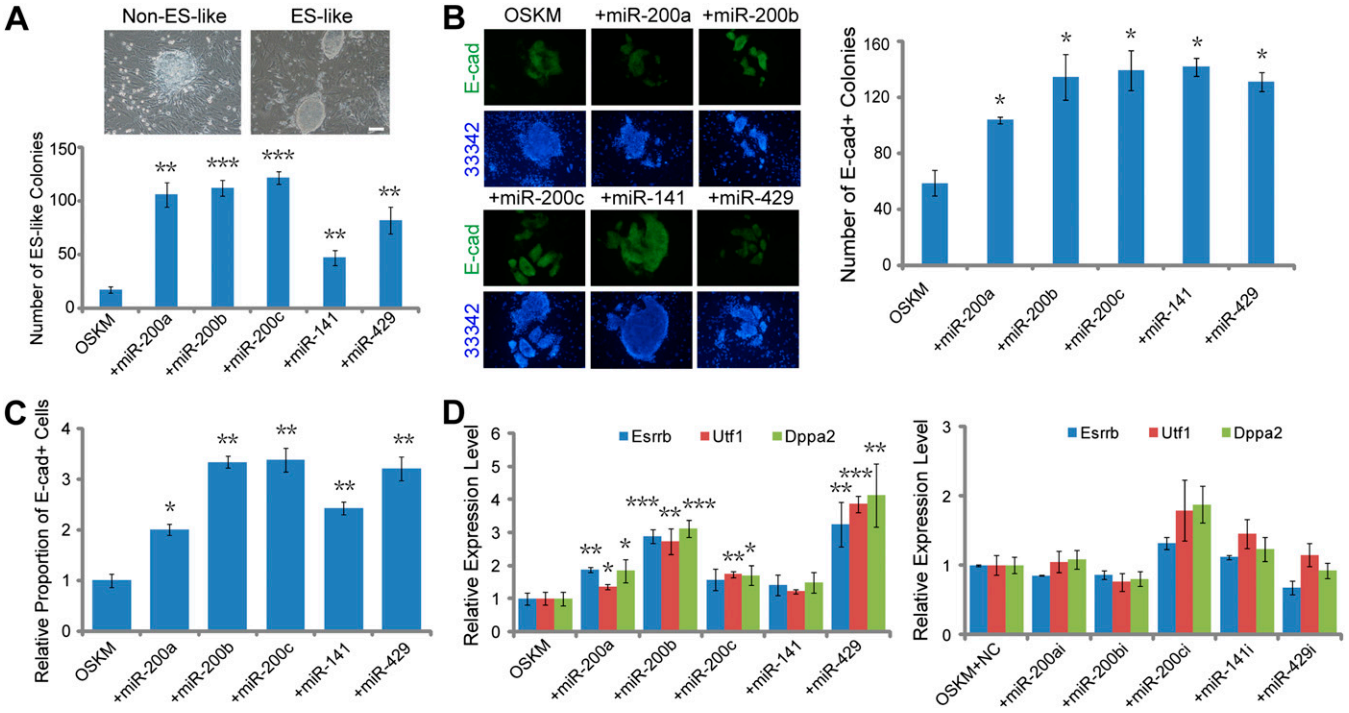

Fig. S2. Members of the miR-200 family promote early events of iPSC generation. (*A*) Representative images of ES-like colony and non–ES-like colony and the numbers of ES-like colonies in OSKM, and OSKM in combination with individual miRNA of the miR-200 family were quantified and shown. (*B*) Representative immunostaining images of the epithelial marker E-cadherin (E-cad, green) with Hochest33342 (33342, blue) for nucleus staining in MEF cells after infection with viruses containing OSKM, and OSKM in combination with miR-200a, miR-200b, miR-200c, miR-141, and miR-429 on day 7. The number of E-cad–positive (E-cad$^+$) colonies were also quantified. (*C*) Quantification for the proportion of E-cad$^+$ cells in MEF cells after infection with OSKM and individual miRNA on day 7, the relative proportion level was compared with that of OSKM group. (*D*) qRT-PCR analyses for the expression levels of the early reprogramming predictors (Esrrb, Utf1, Dppa2) in MEF cells after infection with OSKM and in combination with miR-200s or inhibitor for each miRNA compared with the control group. Error bars denote the SD derived from three independent experiments (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; Student's *t* test) ($n = 3$). (Scale bars: 100 μm.)

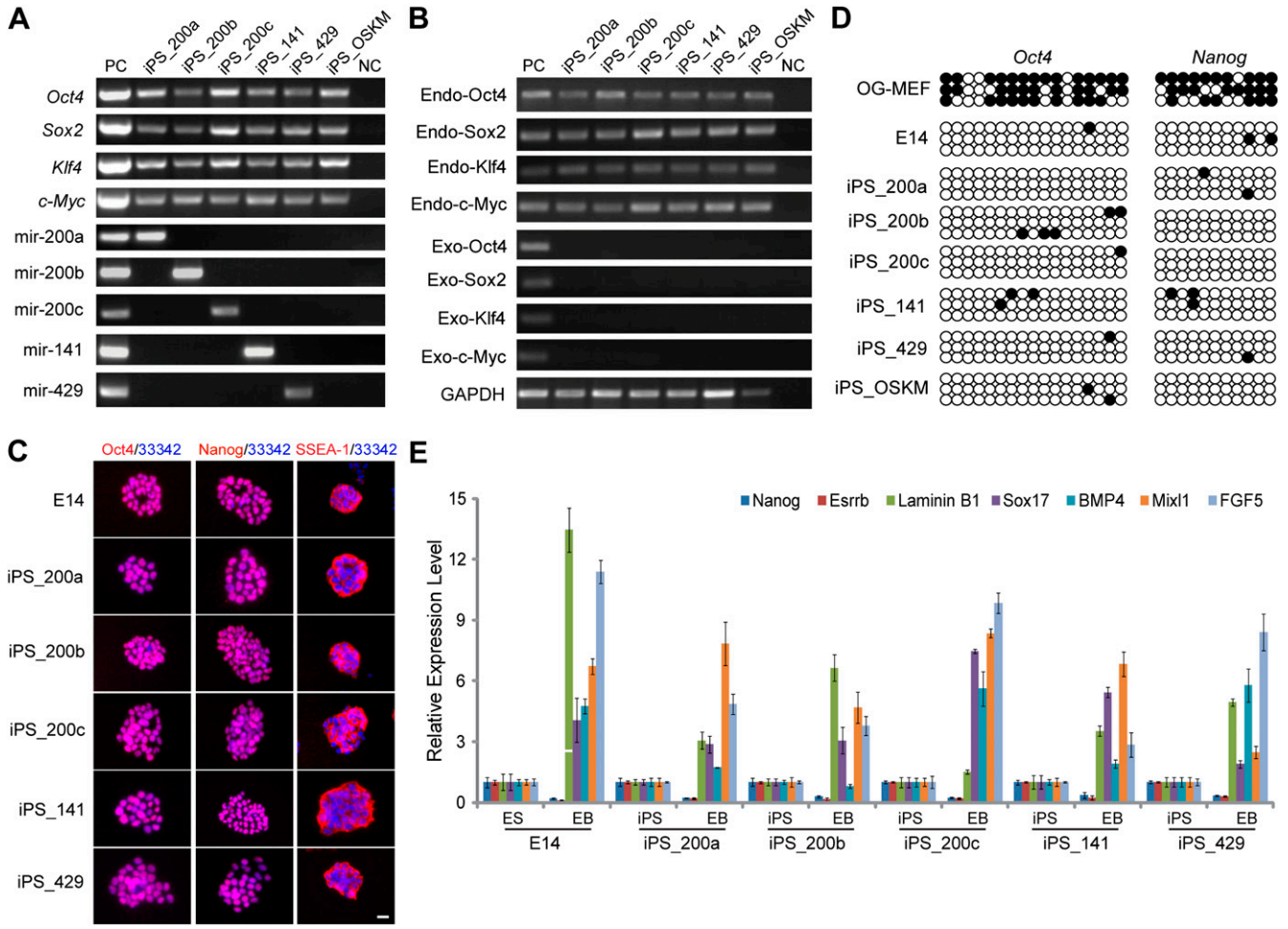

Fig. S3. Identification of OSKM+miRNA-derived iPSCs. (*A*) Genomic integration of *Oct4*, *Sox2*, *Klf4*, *c-Myc*, mir-200a, mir-200b, mir-200c, mir-141, and mir-429 transgenes was confirmed by PCR using specific primers in OSKM+miRNA-derived iPSC lines (iPS_miR-200a, iPS_miR-200b, iPS_miR-200c, iPS_miR-141, iPS_miR-429), with MEFs as the negative control (NC) and pMX vectors as the positive control (PC). (*B*) RT-PCR analysis for the expression of endogenous OSKM, and silencing of exogenous OSKM in OSKM+miRNA-derived iPSC lines. GAPDH was used as an internal control. Mouse ES cells (E14) and MEF cells after infection with retroviruses of OSKM were used as the positive control (PC) for endogenous OSKM and exogenous OSKM. (*C*) Immunostaining assay of mouse pluripotency markers (Oct4, Nanog, SSEA-1) (red) in E14 and OSKM+miRNA-derived iPSC lines. (*D*) Methylation analysis of *Oct4* and *Nanog* promoters by bisulfate sequencing in OG-MEFs, E14, OSKM+miRNA-derived iPSC lines, and OSKM-derived iPSC line (iPS_OSKM). Each horizontal row of circles represents an individual sequencing result from one amplicon. Open and filled circles indicate unmethylated and methylated CpGs, respectively. (*E*) qRT-PCR analysis for pluripotency markers (Nanog, Esrrb) and all three germ layer markers (Laminin B1, Sox17, BMP4, Mixl1, FGF5) in EBs derived from E14 and OSKM+miRNA-derived iPSCs. The corresponding pluripotent stem cells were used as the calibrators, and GAPDH was used as the internal control. (Scale bars: 100 μm.)

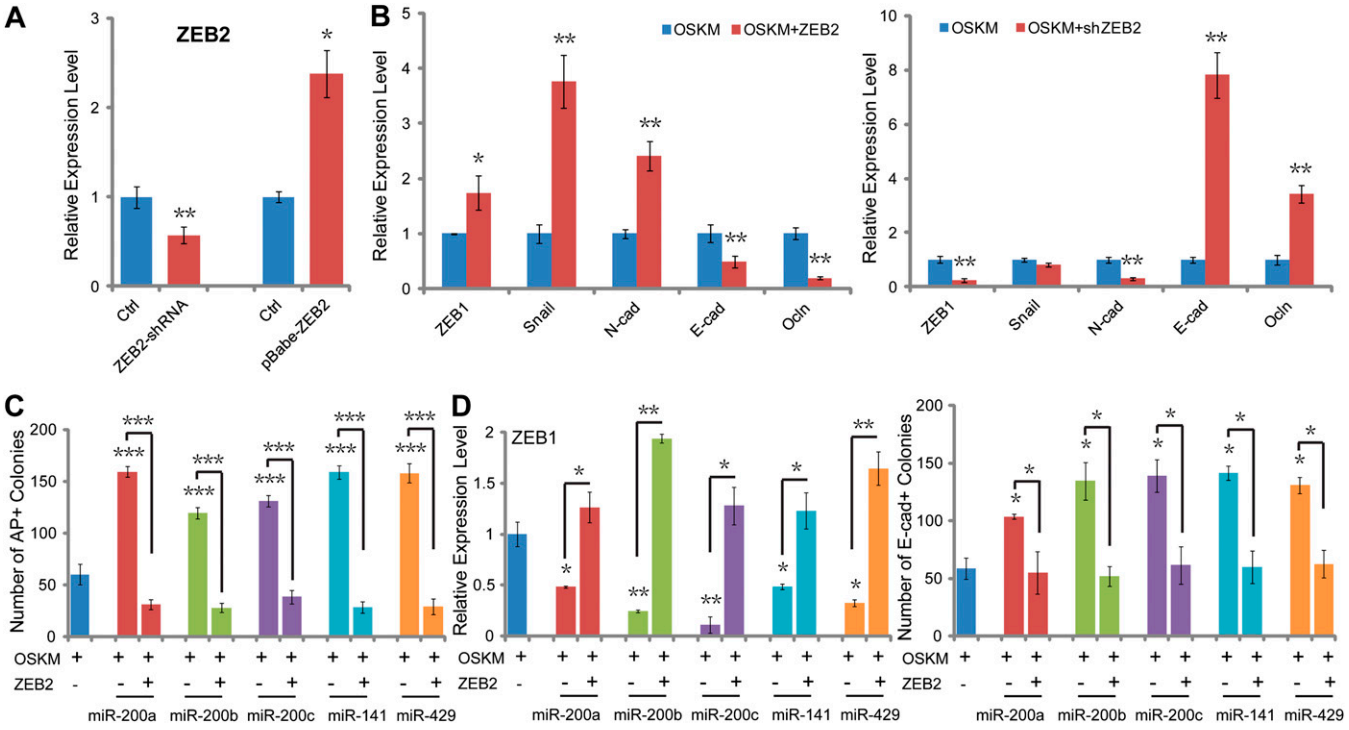

Fig. S4. Effects of ZEB2 on MET and iPSC generation. (*A*) qRT-PCR analysis of the ectopic expression and knockdown efficiency of ZEB2 for the pBabe-ZEB2 and pMko-ZEB2-shRNA vector, by using pBabe-puro and pMko.1 vector as the corresponding negative control (Ctrl) and GAPDH as an internal control. (*B*) Effects of ZEB2 overexpression and inhibition on MET process in OSKM-induced iPSC generation. qRT-PCR analysis of mesenchymal genes (ZEB1, Snail, N-cad) and epithelial genes (E-cad, Ocln) in MEF cells after infection on day 7. (*C*) Quantification for the number of AP^+ colonies after infection with OSKM, OSKM+miRNA, and OSKM+miRNA+ZEB2 on day 8. (*D*) qRT-PCR analysis of mesenchymal marker genes (ZEB1) (*Left*) and quantification of the number of E-cad^+ colonies on day 7 after infection with OSKM, OSKM in combination with miRNA, and OSKM in combination with miRNA and ZEB2 (*Right*). Error bars denote the SD derived from three independent experiments (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; Student's *t* test) ($n = 3$).

Table S1. Primers for genomic PCR genotyping

Primer names	Primer sequences	F/R
pMx-S1811	GCTTGGATACACGCCGC	F
Oct4-virus	TTCATGTCCTGGGACTCCTC	R
Sox2-virus	TTGCTGATCTCCGAGTTGTG	R
Klf4-virus	AACCGCTCCACATACAGTCC	R
c-Myc-virus	GGAAGACGAGGATGAAGCTG	R
mir-200a-virus	ACGCGTCGACGCATGAAGTAATAGACAAGGGTCAAAGTCC	R
mir-200b-virus	ACGCGTCGACTCAGAGGTGCCTGGTTTGGGGGAACTCTCC	R
mir-200c-virus	ACGCGTCGAC ATGGGTTGCTGCCAGATAAAAGTGCCTAAA	R
mir-141-virus	ACGCGTCGACTGCTGCTGTTCTGAAGGCAGCCATTTTGTC	R
mir-429-virus	ACGCGTCGACCAGTCAGGCTCAGCGGGGCCTGTATATTTA	R

F, forward; R, reverse.

Table S2. Primers for exogenous transgene silencing

Primer names	Primer sequences	F/R
pMXs-TgUS	GTGGTGGTACGGGAAATCAC	F
pMXs-Oct3/4-TgDS	TAGCCAGGTTCGAGAATCCA	R
pMXs-Klf4-TgDS	GGGAAGTCGCTTCATGTGAG	R
pMXs-Sox2-TgDS	GGTTCTCCTGGGCCATCTTA	R
pMXs-c-Myc-TgDS	AGCAGCTCGAATTTCTTCCA	R

F, forward; R, reverse.

Table S3. Primers for promoter bisulfate sequencing

Gene symbols	Primer sequences	Forward/reverse
Oct4	GAGGATTGGAGGTGTAATGGTTGTT	Outside forward
	CTACTAACCCATCACCCCCACCTA	Outside reverse
	CAAGCTTTGGGTTGAAATATTGGGTTTATTT	Inside forward
	CGGATCCCTAAAACCAAATATCCAACCATA	Inside reverse
Nanog	AAGTATGGATTAATTTATTAAGGTAGTT	Outside forward
	AAAAAACCCACACTCATATCAATATA	Outside reverse
	AAGTATGGATTAATTTATTAAGGTAGTT	Inside forward
	CAACCAAATCAACCTATCTAAAAA	Inside reverse

Table S4. Primers used for vectors construction

Vector symbols	Primer sequences	F/R
premiR-200a	CGCGGATCCGGTCAACATCTCTGCTGGCTTATTGCGATG	F
	ACGCGTCGACGCATGAAGTAATAGACAAGGGTCAAAGTCC	R
premiR-200b	CGCGGATCCAAGGGGTGCCCTCCTTCTGCAATGCTCTGG	F
	ACGCGTCGACTCAGAGGTGCCTGGTTTGGGGGAACTCTCC	R
premiR-200c	CGCGGATCCATGGGAGCAGGAGATCTGCCGCTTCTCTTG	F
	ACGCGTCGACATGGGTTGCTGCCAGATAAAAGTGCCTAAA	R
premiR-141	CGCGGATCCAGAAGCAGACCTTGGCCTTGGGATTGTCTA	F
	ACGCGTCGACTGCTGCTGTTCTGAAGGCAGCCATTTTGTC	R
premiR-429	CGCGGATCCGGTCCTAGAAGGGTGAACCCCAAGAATGGC	F
	ACGCGTCGACCAGTCAGGCTCAGCGGGGCCTGTATATTTA	R
ZEB2	CACAGGTCGACATGAAGCAGCCGATCATGGCGGAT	F
	CACAGGTCGACTTATTCCATGCCATCTTCCAT	R
ZEB2-shRNA	CCGGCCACTAGACTTCAATGACTATCTCGAGATAGTCATTGAAGTCTAGTGGTTTTTG	F
	AATTCAAAAACCACTAGACTTCAATGACTATCTCGAGATAGTCATTGAAGTCTAGTGG	R
ZEB1 3′UTR	GGCGAGCTCTTCCTGCTCACTACTGTGTAATGTC	F
	GGCTCTAGAAATCTTAGCAACTTTCCAAAATTTC	R
ZEB2 3′UTR	GGCGAGCTCGGCAGTTCAGCCAAGACAGAGTTAG	F
	GGCTCTAGACACAATCTGGAATCAGGATCAGTTG	R
ZEB1 mutant 3′UTR primer_1	GATGGGCGTCCGGCATGGCTACGTCATAAATCACTCTTACGTTGGCTCATTC	F
	GAATGAGCCAACGTAAGAGTGATTTATGACGTAGCCATGCCGGACGCCCATC	R
ZEB1 mutant 3′UTR primer_2	GTAAACATATGCTAAATCCGCTTGTCATAATTATTATGTTTTTTAAAATG	F
	CATTTTAAAAAACATAATAATTATGACAAGCGGATTTAGCATATGTTTAC	R
ZEB2 mutant 3′UTR primer_1	GTGCCCGCACTACCATACATGACATAATTATTATTATTATTATTGTTATTC	F
	GAATAACAATAATAATAATAATAATTATGTCATGTATGGTAGTGCGGGCAC	R
ZEB2 mutant 3′UTR primer_2	GTGCCCGCACTACAGTGCATGTCATAAACGATTCCTCAACACTTTCC	F
	GGAAAGTGTTGAGGAATCGTTTATGACATGCACTGTAGTGCGGGCAC	R
ZEB2 mutant 3′UTR primer_3	GTTCTTTGAAGCACCCGTGTGTCATAAAGACGACTAGGCAGCAGTTCC	F
	GGAACTGCTGCCTAGTCGTCTTTATGACACACGGGTGCTTCAAAGAAC	R
ZEB2 mutant 3′UTR primer_4	CTTTTTTGTTCATTAATTTTGTGTCATAAACACCAAACTTTTTTTGCAAC	F
	GTTGCAAAAAAAGTTTGGTGTTTATGACACAAAATTAATGAACAAAAAAG	R
ZEB2 mutant 3′UTR primer_5	ATTTATTTCATACTGTAGTGTAGTCATAAATAGTTCTTCAATATATAG	F
	CTATATATTGAAGAACTATTTATGACTACACTACAGTATGAAATAAAT	R
ZEB2 mutant 3′UTR primer_6	CCTGAATTGTCTTTGTAAATCATAAATTGAATTTTCAATTTGTAATTTC	F
	GAAATTACAAATTGAAAATTCAATTTATGATTTACAAAGACAATTCAGG	R
ZEB2 mutant 3′UTR primer_7	GCATTATGCAAAAATTTTGTTGTCACAATAGGCCTAAAAACTGTGTGGTTC	F
	GAACCACACAGTTTTTAGGCCTATTGTGACAACAAAATTTTTGCATAATGC	R
ZEB2 mutant 3′UTR primer_8	GGTTACCTTTGCACCAGCTTGTCACAAAAGCTCACCGTGTTCTTTGAAG	F
	CTTCAAAGAACACGGTGAGCTTTTGTGACAAGCTGGTGCAAAGGTAACC	R
Rluc-mir-200a/b/429 promoter	CACAGCTCGAGGGTCATCCAAGTCCAGATCCG	F
	CACAGCTCGAGCAGCCATGCAGGAAGTACCTG	R
Rluc-mir-141/ 200c promoter	CACAGCTCGAGGCGATGCCCGTACGGAAG	F
	CACAGCTCGAGCGTGAGGCATGGCTCTTTTG	R

F, forward; R, reverse.

Table S5. Primer sets used in qPCR assays

Gene symbols	Primer sequences	F/R
Endo-Oct4	TCTTTCCACCAGGCCCCCGGCTC	F
	TGCGGGCGGACATGGGGAGATCC	R
Endo-Sox2	TAGAGCTAGACTCCGGGCGATGA	F
	TTGCCTTAAACAAGACCACGAAA	R
Endo-Klf4	GCGAACTCACACAGGCGAGAAACC	F
	TCGCTTCCTCTTCCTCCGACACA	R
Endo-c-Myc	TGACCTAACTCGAGGAGGAGCTGGAAT C	F
	AAGTTTGAGGCAGTTAAAATTATGGCTGAAGC	R
Nanog	AGGGTCTGCTACTGAGATGCTCTG	F
	CAACCACTGGTTTTTCTGCCACCG	R
Esrrb	GTCCCTCTCCGCGTTAGC	F
	GGGGCAGGTTCGTCATTT	R
Utf1	TCTACTGGCCCTGGACG	F
	TGGAAGAACTGAATCTGAGCG	R
Dppa2	CTGCAATTTCTTGCTGGACA	F
	TTTTTCTTTGCACGGCTTCT	R
Lin28	TCGGCTTCCTGTCTATGACC	F
	CTTCGGAACCCTTCCATGT	R
ZEB1	TGCTCACCTGCCCGTATTGTGATA	F
	AGTGCACTTGAACTTGCGGTTTCC	R
ZEB2	ATTGCACATCAGACTTTGAGGAA	F
	ATAATGGCCGTGTCGTTCG	R
Snail	TTGTGTCTGCACGACCTGTGGAAA	F
	TCTTCACATCCGAGTGGGTTTGGA	R
N-cadherin	GGCTGAAAATAGACCCCGTG	F
	GCTGTGATGTTAATTGAGTTGGG	R
E-cadherin	CAGCCTTCTTTTCGGAAGACT	F
	GGTAGACAGCTCCCTATGACTG	R
Ocln	CCTCCAATGGCAAAGTGAATGGCA	F
	TGTTTCATAGTGGTCAGGGTCCGT	R
Laminin B1	CCCCAATCTCTGTGAACCATG	F
	GCAATTTGCACCGACACTGA	R
Sox17	CGAGCCAAAGCGGAGTCTC	F
	TGCCAAGGTCAACGCCTTC	R
BMP4	AAAAGTCGCCGAGATTCAG	F
	CGGTAAAGATCCCTCATGTAA	R
Mixl1	ACGCAGTGCTTTCCAAACC	F
	CCCGCAAGTGGATGTCTGG	R
FGF5	AAAGTCAATGGCTCCCACGAA	F
	GGCACTTGCATGGAGTTTTCC	R
GAPDH	AGGTCGGTGTGAACGGATTTG	F
	TGTAGACCATGTAGTTGAGGTCA	R
mir-200a/b/429 promoter binding region (<200 nt)	CTTGTCAGTTGAGCCATTTC	F
	CTTACAGATTCCCAGCCAC	R
mir-200a/b/429 promoter binding negative control region (>1,000 nt)	TTGCCTTAGAACTCCACAG	F
	AGAACACCATTATGCCTTG	R
mir-141/200c promoter binding region (<200 nt)	CTTCCGGTGCCCTTTCTCC	F
	GGCGTCCAGCTAAGTCCTTCA	R
mir-141/200c promoter binding negative control region (>1,000 nt)	GGACTATGAGGAGCGAGTG	F
	CCCAGTAACCCTGACCAA	R

F, forward; R, reverse.