Brief Report

DOI: 10.5582/bst.2017.01082

Comparative transcriptomic analysis identifies reprogramming and differentiation genes differentially expressed in UiPSCs and ESCs

Liang Shi[1,2,§], Yazhou Cui[2,§], Xiaoyan Zhou[2], Jing Luan[2], Linli Wang[3], Jinxiang Han[2,*]

[1] *School of Medicine and Life Sciences, University of Jinan-Shandong Academy of Medical Sciences, Ji'nan, Shandong, China;*
[2] *Key Laboratory for Rare Disease Research of Shandong Province, Key Laboratory for Biotech Drugs of the Ministry of Health, Shandong Medical Biotechnological Center, Shandong Academy of Medical Sciences, Ji'nan, Shandong, China;*
[3] *Guangzhou Biocare Institute of Cancer, Guangzhou, China.*

Summary **Embryonic stem cells (ESCs) technology has garnered worldwide attention for its therapeutic applications *in vivo*. Researchers have previously shown that non-viral induced pluripotent stem cells (iPSCs) can be generated from urine. As a potential alternative, Urinary iPSCs (UiPSCs) are highly similar to embryonic stem cells (ESCs) in many aspects such as morphology, expression of pluripotency markers and the capacity to develop into three germ layers *in vitro* and *in vivo*. However, the degree of gene expression similarity between iPSCs and ESCs has not been completely elucidated. In the present study, we performed a comparative study on the gene expression profile between UiPSCs and ESCs using microarray technology, and identified 19 differentially expressed genes. Furthermore, four genes associated with reprogramming and differentiation including neuronatin (*NNAT*), piwi-like RNA-mediated gene silencing 2 (*PIWIL2*), early growth response 1 (*EGR1*) and TATA-box binding protein associated factor 9b (*TAF9B*) were validated by quantitative real-time PCR (qRT-PCR) assays. Our results indicate that compared with ESCs, UiPSCs demonstrated a different pathway in reprogramming and differentiation preference from ESCs, and can be used as a potential tool in disease modeling, drug discovery and regenerative medicine.**

Keywords: Urinary iPSCs, embryonic stem cells, transcriptome

1. Introduction

The advent of human induced pluripotent stem cells (iPSCs) in 2007 (*1,2*), has ushered in an era of considerable excitement about the prospects of using these cells to develop new opportunities for healthcare, from their potential for regenerative medicine to their use as tools for studying the cellular basis of many diseases and the discovery of new drugs. In recent years, research on iPSCs technologies in mice and humans has progressed greatly. Human iPSCs can be generated from multiple donor sources, such as neural cells (*3*), hepatocytes (*4*), and amniocytes (*5*). Generation of urinary iPSCs (UiPSCs) may be a better choice since the isolation of urinary cells is simple, and safely, affordably, and frequently obtained (*6,7*). This approach has been widely used for modeling disorders and offering proof of principle for basic biological research and clinical applications (*8,9*). Previous methods used to derive iPSCs are not "footprint-free" and random integration may alter the transcriptional signature, a serious obstacle to comprehensive transcriptional analysis. Recently different integration-free methods have been used to reprogram these cells, which greatly improve the prospects for iPSCs applications (*10*). Although human iPSCs are shown to mimic ESCs, global transcriptional comparison of human ESCs and iPSCs derived from other sources has revealed some significant differences. Several studies have identified as many as 1267 to 3947 genes with varying levels of deviation (*11,12*). Nevertheless, previous study suggests that UiPSCs are "nearly identical" to ESCs,

Released online in J-STAGE as advance publication May 20, 2017.

[§]These authors contributed equally to this work.

**Address correspondence to:*
Dr. Jinxiang Han, Key Laboratory for Rare Disease Research of Shandong Province, Key Laboratory for Biotech Drugs of the Ministry of Health, Shandong Medical Biotechnological Center, Shandong Academy of Medical Sciences, Ji'nan, Shandong 250062, China.
E-mail: jxhan9888@aliyun.com

but it remains unclear whether the small percentage of genes that are differentially expressed between iPSCs and ESCs is biologically significant.

In this study, we investigated the differences in gene expression profiles of UiPSCs and ESCs, and identified a set of differentially expressed genes for the first time. After bioinformatic analysis, four genes related to reprogramming and differentiation were further validated by qRT-PCR. The results of the present study extended our understanding of the transcriptional profiles in ESCs and UiPSCs and highlighted that the substantial gene expression differences between these cell populations can be helpful to direct the utility of UiPSCs in the future.

2. Materials and Methods

2.1. *Cell origin and culturing*

The experimental procedures were approved by the ethnic committee of Shandong Medical Biotechnological Center. Human ESCs was derived from discarded human embryos and co-cultured with Guangzhou Biocare Cancer Institute (GBCI). UiPSCs were generated from urine using an integration-free reprogramming method provided by Guangzhou Institutes of Biomedicine and Health. All the above cells were maintained in defined medium BioCISO (Biocare Biotech., Ltd., Guangzhou, China) on matrigel (Corning, NewYork, USA). The culture medium was changed daily and cells were passaged with 0.5mM EDTA (Gibco, Carlsbad, CA, USA) when the culture grew confluent.

2.2. *Bioinformatic analysis*

Genome-wide expression profiling analysis was performed using Affymetrix GeneChip Human Transcriptome Array 2.0 between ESCs and UiPSCs according to the manufacturer's instructions. CEL-files of the raw data uploaded to the website of Gene Cloud of Biotechnology Information (GCBI Platform, Shanghai, China) (www.gcbi.com.cn) for further data mining, including differences in mRNA profiles, and other bioinformatic analysis. We selected the differentially expressed mRNAs based on the P-value, Q-value and at least a 2-fold change. To determine the interactions among differentially expressed genes, gene co-expression networks were built according to the normalized signal intensity of specific expressed genes. In a network analysis, degree is the most important measure of an mRNA centrality within a network. A higher degree of a gene indicates that it plays a more important role in the signaling network. A GO analysis was applied to analyze the main functions of the differentially expressed mRNAs (*13*). Pathway analysis of differentially expressed genes was performed based on the Kyoto encyclopedia of genes and genomes (KEGG) (*14*).

Table 1. Targeted gene sequences of the primers used for qRT-PCR

Gene	Sequences5'-3'
TAF9B	Forward, GGATGACGAGTGGCTGGATA
	Reverse, GCCAGTCTCACATCATCTGC
NNAT	Forward, ACCGCATTCTGATCTGGACA
	Reverse, ACCCTCCTTCCTCAACTGTG
PIWIL2	Forward, TTGTGGACAGCCTGAAGCTA
	Reverse, CCATCAGACACTCCATCACG
EGR1	Forward, CCACCACGTACTCCTCTGTT
	Reverse, GAACCCTCCTCTCCTATGGC
ACTIN	Forward, CCCAGAGCAAGAGAGG
	Reverse, GTCCAGACGCAGGATG

2.3. *qRT-PCR*

Alteration of targeted gens at the mRNA level was confirmed by qRT-PCR analysis. Total RNA was extracted using Trizol reagent (Gibco, Carlsbad, CA, USA) and the purified total RNA was used for cDNA synthesis with a first-strand cDNA synthesis kit (Toyobo, Osaka, Japan). After the reverse transcription reaction, cDNA was used as the template for qRT-PCR of *NNAT*, *EGR1*, *PIWIL2*, and *TAF9B*. The sequences of the primers used are listed in Table 1. qRT-PCR was performed (LightCycler 480 thermocycler, Roche Applied Science, Mannheim, Germany) using a SYBR Green qPCR Kit (Toyobo, Osaka, Japan). *ACTIN* was used as an internal control to determine the relative expression of target mRNA. All reactions were performed in triplicate.

2.4. *Statistical analysis*

Data are shown as the mean ± S.D., and student's *t*-test (two-tailed) was used to determine the statistical significance of quantitative data. $P < 0.05$ was considered to be statistically significant.

3. Results

3.1. *Bioinformatic analysis of microarray data between ESCs and UiPSCs*

To identify differentially expressed genes between the ESCs and the UiPSCs, we selected the differentially expressed mRNAs according to the *P*-value and *Q*-value by using the GCBI platform. *P*-values < 0.01 and $Q < 0.01$ were considered significant. The list of significant genes was further filtered using fold change (FC) > 2 (Table 2). Although the hierarchical clustering analysis showed that UiPSCs were similar to ESCs in expression levels, there were still 19 mRNAs with the largest differences in each of the two cell populations (Figure 1A). Of these, 5 showed higher expression in UiPSCs than in ESCs, and 14 were more highly expressed in ESCs than in UiPSCs. To further evaluate

Table 2. The 19 differentially expressed genes

Gene Symbol	Gene Description	Fold Change	Gene Feature
TAF9B	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor, 31kDa	9.64011	Up
CLC	Charcot-Leyden crystal protein	9.120564	Down
NNAT	Neuronatin , transcript variant 1	5.586741	Down
CCDC152	Coiled-coil domain containing 152	4.242718	Up
LGALS14	Lectin, galactoside-binding, soluble, 14, transcript variant 1, mRNA	4.231168	Down
SERPINB9	Serpin peptidase inhibitor, clade B (ovalbumin), member 9	3.28019	Down
ZNF676	Zinc finger protein 676	2.949182	Down
CAPN6	Calpain 6	2.870614	Down
ZNF208	Zinc finger protein 208	2.797106	Down
LUZP2	Leucine zipper protein 2, transcript variant 1, mRNA	2.564051	Down
PIWIL2	Piwi-like RNA-mediated gene silencing 2, transcript variant 1	2.45924	Down
GRPR	Gastrin-releasing peptide receptor	2.356274	Down
GSTT1	Glutathione S-transferase theta 1	2.309566	Up
LCP1	Lymphocyte cytosolic protein 1 (L-plastin)	2.250113	Down
ZNF248	Zinc finger protein 248, transcript variant 1	2.223748	Up
SDR42E1	Short chain dehydrogenase/reductase family 42E, member 1	2.216083	Up
ZNF729	Zinc finger protein 729	2.074822	Down
EGR1	Early growth response 1	2.042328	Down
AASS	Aminoadipate-semialdehyde synthase, nuclear gene encoding mitochondrial protein	2.004129	Down

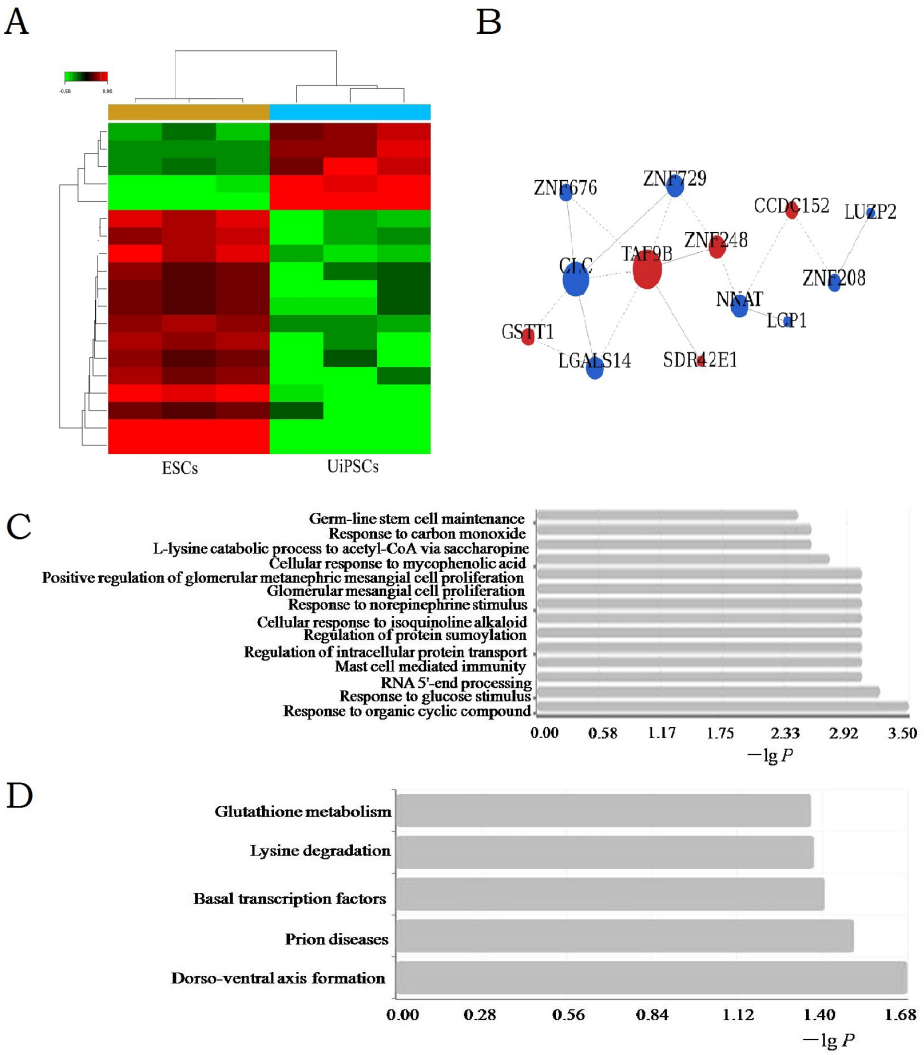

Figure 1. Comparative transcriptomic analysis between ESCs and UiPSCs. (A) Hierarchical clustering analysis of the differentially expressed genes in ESCs and UiPSCs. Red color indicates upregulated genes and green color indicates downregulated genes. **(B)** Gene co-expression network analysis based on the differentially expressed genes. The red circles represent upregulated genes, and blue circles represent downregulated genes. The size of the circle represents the degree value. **(C)** Histogram of changed GO analysis based on the differentially expressed genes. **(D)** Histogram of signaling pathways based on the differentially expressed genes.

the interactions among the differentially expressed genes and to locate core regulatory genes in the network, we constructed a gene co-expression network by GCBI. The higher degree of a gene indicated that it was regulating or being regulated by a greater number of genes, implying it had a more important role in the signaling network. A P value < 0.05 was considered statistically significant. As shown in Figure 1B, *TAF9B* was identified as the core regulatory node with the highest degree. By using the GCBI platform, significantly altered cell functions were generated. We focused on GOs with a P value of < 0.05 and a FDR of < 0.05 (top 14 affected cell functions are listed in Figure 1C). High-enrichment GOs of biological processes included the organic cyclic compound response, glucose stimulus response, glomerular mesangial cell proliferation, positive regulation of glomerular metanephric mesangial cell proliferation, and germ-line stem cell maintenance. Among these differentially expressed genes, *PIWIL2* participated in germ-line stem cell maintenance, and *EGR1* is involved in glomerular mesangial cell proliferation. Pathway analyses were used to determine the significantly enriched pathways of the differentially expressed genes. As shown in Figure 1D, significant signaling pathways between ESCs and UiPSCs groups included dorso-ventral axis formation, prion diseases, basal transcription factors, lysine degradation, and glutathione metabolism.

3.2. *Validation of differently expressed genes by qRT-PCR*

In order to validate differential mRNA expression patterns, 4 selected genes, related to reprogramming and differentiation, were analyzed by qRT-PCR. *NNAT*, *PIWIL2*, *EGR1*were found to be more highly expressed in ESCs, and *TAF9B* was found to be more highly expressed in UiPSCs (Figure 2), which is consistent with our findings using the microarray platform.

UiPSCs are shown to mimic human ESCs (*15,16*), the degree of molecular similarity between UiPSCs derived from urine by transcriptional reprogramming and those of embryo-derived human ESCs has not been completely elucidated. In this study, we performed a comparison of gene expression profiling between the UiPSCs and ESCs. Our data suggest that, although the global transcriptional profiles of human ESCs and UiPSCs were globally similar, small but significant differences indeed exist. A total of 19 differentially expressed genes were identified and 4 genes (*TAF9B*, *NNAT*, *EGR1*, *PIWIL2*) were further validated.

Among down-regulated genes in UiPSCs, *NNAT*, *EGR1* have been demonstrated to be involved in generation and maintenance of stem cell properties as negative regulators (*17,18)*. Teichroeb *et al*. found that *NNAT* was consistently silenced in iPSCs compared with its isogenic ESCs, and suppression of *NNAT* could

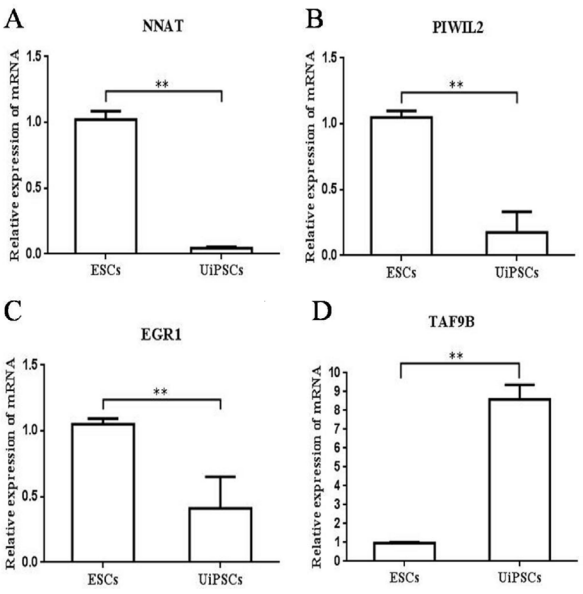

Figure 2. qRT-PCR analysis of 4 selected mRNAs expression. (A) *NNAT*. **(B)** *PIWIL2*. **(C)** *EGR1*. **(D)** *TAF9B*. Bars are shown as the mean ± S.D. ***p* < 0.01 *vs*. group ESCs.

be used as a biomarker for successful reprogramming (*17*). *EGR1* is a zinc-finger pro-differentiation factor that plays an important role in the regulation of differentiation and development in several contexts (*19*). Recently, Worringer *et al*. suggested that *EGR1* might be a barrier to reprogramming of *let-7*, and inhibition *EGR1* mRNA by RNA-binding protein (RBP) LIN-41 could promote reprogramming (*18*).

PIWI proteins have been found to play essential and conserved roles in germline stem cell maintenance, and are expressed in ESCs at higher levels (*20*). Moreover, a recent study reported that PIWI proteins are dispensable for reprogramming of mouse fibroblasts into iPSCs (*21*). Therefore, combining the above facts with our findings, it is reasonable to deduce that UiPSCs may be dependent on a different gene background to maintain stem cell properties from ESCs. According to current evidence, UiPSCs seem to be far more transcriptionally similar to iPSCs derived from other sources than ESCs.

Interestingly, in this study, we also found that a neuron differentiation regulator *TAF9B* as a top candidate gene increases in UiPSCs more than in ESCs (more than 9 fold). Previous studies reported that *TAF9B* was dispensable for global gene expression and pluripotency of murine ESCs, but *TAF9B* was required for the efficient *in vitro* differentiation of murine ESCs into motor neurons (*22*). Moreover, it has been shown that epithelial-like cells from human urine can be reprogrammed into UiPSCs, and even directly into human neural progenitor cells (*23,24*). Therefore, our result may indicate that the level of *TAF9B* in UiPSCs would account for its preference towards neuron differentiation. Compared with ESCs and even other iPSCs, UiPSCs may be able to be induced into neurons

more easily.

In conclusion, our above results further revealed that UiPSCs and ESCs had different gene expression profiles, in particular in stem cells properties maintainance mechanisms. The high basic level of *TAF9B* may partly account for the potential of UiPSCs in neuron differentiation. The molecular differences between UiPSCs and ESCs described here should drive intense efforts in the future aimed at uncovering UiPSCs as a potential tool for disease modeling, drug discovery and regenerative medicine.

Acknowledgements

This study was supported by The Innovation Project of Shandong Academy of Medical Sciences & the Key Projects in the National Science & Technology Pillar Program during the Twelfth Five-year Plan Period (2013BAI07B00).

References

1. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007; 131:861-872.
2. Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin, II, Thomson JA. Induced pluripotent stem cell lines derived from human somatic cells. Science. 2007; 318:1917-1920.
3. Kim JB, Greber B, Arauzo-Bravo MJ, Meyer J, Park KI, Zaehres H, Scholer HR. Direct reprogramming of human neural stem cells by *OCT4*. Nature. 2009; 461:649-643.
4. Liu H, Ye Z, Kim Y, Sharkis S, Jang YY. Generation of endoderm-derived human induced pluripotent stem cells from primary hepatocytes. Hepatology. 2010; 51:1810-1819.
5. Ruiz S, Brennand K, Panopoulos AD, Herrerias A, Gage FH, Izpisua-Belmonte JC. High-efficient generation of induced pluripotent stem cells from human astrocytes. PLoS One. 2010; 5:e15526.
6. Zhou T, Benda C, Duzinger S, *et al*. Generation of induced pluripotent stem cells from urine. J Am Soc Nephrol. 2011; 22:1221-1228.
7. Xue Y, Cai X, Wang L, Liao B, Zhang H, Shan Y, Chen Q, Zhou T, Li X, Hou J, Chen S, Luo R, Qin D, Pei D, Pan G. Generating a non-integrating human induced pluripotent stem cell bank from urine-derived cells. PLoS One. 2013; 8:e70573.
8. Zhang SZ, Li HF, Ma LX, Qian WJ, Wang ZF, Wu ZY. Urine-derived induced pluripotent stem cells as a modeling tool for paroxysmal kinesigenic dyskinesia. Biol Open. 2015; 4:1744-1752.
9. Hildebrand L, Rossbach B, Kühnen P, Gossen M, Kurtz A, Reinke P, Seemann P, Stachelscheid H. Generation of integration free induced pluripotent stem cells from fibrodysplasia ossificans progressiva (FOP) patients from urine samples. Stem Cell Res. 2016; 16:54-58.
10. Seo BJ, Hong YJ, Do JT. Cellular reprogramming using protein and cell-penetrating peptides. Int J Mol Sci. 2017; 18. pii:E552.
11. Chin MH, Mason MJ, Xie W, *et al*. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell. 2009; 5:111-123.
12. Marchetto MC, Yeo GW, Kainohana O, Marsala M, Gage FH, Muotri AR. Transcriptional signature and memory retention of human-induced pluripotent stem cells. PLoS One. 2009; 4:e7076.
13. Ashburner M, Ball CA, Blake JA, *et al*. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25:25-29.
14. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. Nucleic Acids Res. 2004; 32:D277-280.
15. Drozd AM, Walczak MP, Piaskowski S, Stoczynska-Fidelus E, Rieske P, Grzela DP. Generation of human iPSCs from cells of fibroblastic and epithelial origin by means of the oriP/EBNA-1 episomal reprogramming system. Stem Cell Res Ther. 2015; 6:122.
16. Zhou T, Benda C, Dunzinger S, *et al*. Generation of human induced pluripotent stem cells from urine samples. Nat Protoc. 2012; 7:2080-2089.
17. Teichroeb JH, Betts DH, Vaziri H. Suppression of the imprinted gene *NNAT* and X-chromosome gene activation in isogenic human iPS cells. PLoS One. 2011; 6:e23436.
18. Worringer KA, Rand A, Hayashi Y, Sami S, Takahashi K, Tanabe K, Narita M, Srivastava D, Yamanaka S. The *let-7/LIN-41* pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes. Cell Stem Cell. 2014; 14:40-52.
19. de Klerk N, Saroj SD, Wassing GM, Maudsdotter L, Jonsson AB. The host cell transcription factor *EGR1* is induced by bacteria through the EGFR–ERK1/2 pathway. Front Cell Infect Microbiol. 2017; 7:16.
20. Juliano C, Wang J, Lin H. Uniting germline and stem cells: The function of piwi proteins and the piRNA pathway in diverse organisms. Annu Rev Genet. 2011; 45:447-469.
21. Cheng EC, Kang D, Wang Z, Lin H. Piwi proteins are dispensable for mouse somatic development and reprogramming of fibroblasts into pluripotent stem cells. PLoS One. 2014; 9:e97821.
22. Herrera FJ, Yamaguchi T, Roelink H, Tjian R. Core promoter factor *TAF9B* regulates neuronal gene expression. Elife. 2014; 3:e02559.
23. Wang L, Wang L, Huang W, Su H, Xue Y, Su Z, Liao B, Wang H, Bao X, Qin D, He J, Wu W, So KF, Pan G, Pei D. Generation of integration-free neural progenitor cells from cells in human urine. Nat Methods. 2013; 10:84-89.
24. Wang L, Li X, Huang W, Zhou T, Wang H, Lin A, Hutchins AP, Su Z, Chen Q, Pei D, Pan G. TGFβ signaling regulates the choice between pluripotent and neural fates during reprogramming of human urine derived cells. Sci Rep. 2016; 6:22484.

(Received March 30, 2017; Accepted May 12, 2017)