Genes to Cells

Efficient reprogramming of human and mouse primary extra-embryonic cells to pluripotent stem cells

Shogo Nagata[1,2], Masashi Toyoda[3], Shinpei Yamaguchi[1], Kunio Hirano[1], Hatsune Makino[3], Koichiro Nishino[3], Yoshitaka Miyagawa[4], Hajime Okita[4], Nobutaka Kiyokawa[4], Masato Nakagawa[5], Shinya Yamanaka[5], Hidenori Akutsu[3], Akihiro Umezawa[3] and Takashi Tada[1,2*]

[1]Stem Cell Engineering, Institute for Frontier Medical Sciences, Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan
[2]JST, CREST, 4-1-8 Hon-cho, Kawaguchi-shi, Saitama 332-0012, Japan
[3]Department of Reproductive Biology, National Research Institute for Child Health and Development, 2-10-1 Ookura, Setagaya-ku, Tokyo 157-8535, Japan
[4]Department of Developmental Biology, National Research Institute for Child Health and Development, 2-10-1 Ookura, Setagaya-ku, Tokyo 157-8535, Japan
[5]Center for iPS Cell Research and Application (CiRA), Institute for Integrated Cell-Material Sciences, Kyoto University, 53 Kawaharacho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan

Practical clinical applications for current induced pluripotent stem cell (iPSC) technologies are hindered by very low generation efficiencies. Here, we demonstrate that newborn human (h) and mouse (m) extra-embryonic amnion (AM) and yolk-sac (YS) cells, in which endogenous KLF4•Klf4, c-MYC•c-Myc and RONIN•Ronin are expressed, can be reprogrammed to hiPSCs and miPSCs with efficiencies for AM cells of 0.02% and 0.1%, respectively. Both hiPSC and miPSCs are indistinguishable from embryonic stem cells in colony morphology, expression of pluripotency markers, global gene expression profile, DNA methylation status of OCT4 and NANOG, teratoma formation and, in the case of miPSCs, generation of germline transmissible chimeric mice. As copious amounts of human AM cells can be collected without invasion, and stored long term by conventional means without requirement for in vitro culture, they represent an ideal source for cell banking and subsequent 'on demand' generation of hiPSCs for personal regenerative and pharmaceutical applications.

Introduction

Induced pluripotent stem cells (iPSCs) have been generated through nuclear reprogramming of somatic cells via retrovirus or lentivirus-mediated transduction of exogenous reprogramming factors Oct4, Sox2, Klf4 and C-Myc (Yamanaka 2007). This has led to greatly enhanced promise for exploring the causes of, and potential cures for, many genetic diseases, as well as increased promise for regenerative medicine. Improvements in delivery methodology have further facilitated iPSC generation by minimizing the requirement for genetic modification (Feng et al. 2009). Notably, generation of genetic modification-free iPSCs with reprogramming proteins (Kim et al. 2009; Zhou et al. 2009) suggests regenerative medicine with personal iPSCs could soon be realized. However, the markedly low efficiency of iPSC generation, with all adult somatic cell types tested to date, remains problematic (Wernig et al. 2008). Technological advancements in this field have mainly been achieved using mouse embryonic fibroblasts (MEFs), in which the efficiency of iPSC generation is 10–100 times higher than that with adult somatic cells (Yu et al. 2007; Wernig et al. 2008). Therefore, current methods would appear to be less than ideal for generating iPSCs from adult somatic cells.

Communicated by: Fuyuki Ishikawa
*Correspondence: ttada@frontier.kyoto-u.ac.jp

DOI: 10.1111/j.1365-2443.2009.01356.x
• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

Here, to find nuclear reprogramming-sensitive cells collectable with no risk by physical invasion, we generated iPSCs from human and mouse newborn extra-embryonic membranes, amnion (AM) and yolk sac (YS), which consist huge amounts of discarded cells after birth. Interestingly, the efficiency of mouse iPSC (miPSC) generation from the AM was comparable to that of MEFs by retroviral transduction with Oct4, Sox2, Klf4 and c-Myc. Importantly, human iPSC (hiPSC) is also efficiently generated from human AM cells. Expression of the endogenous KLF4•Klf4, c-MYC•c-Myc and RONIN•Ronin in human•mouse AM cells may function in facilitating the generation efficiency of iPSCs. The human AM cell, which is conventionally freeze-storable, could be a useful cell source for the generation of pluripotent stem cells including iPSCs mediated by nuclear reprogramming in the purpose of personal regenerative and pharmaceutical cure in the future of infants.

Results

Generation of iPSCs from mouse AM and YS cells

Extra-embryonic membranes, AM (amniotic ectoderm and mesoderm layers) and YS (visceral yolk sac endoderm and mesoderm layers) express a high level of proto-oncogene (Curran et al. 1984) which function, at least in part, to maintain and protect the fetus in utero. In E18.5 mouse embryos just before birth, AM and YS can be easily recognized microscopically (Fig. 1a). The membranes were dissected from Oct4-GFP (OG)•Neo-LacZ (Rosa26) embryos as approximately 5–10 mm^2 sections and digested with collagenase. Isolated cells were cultured for 4–5 days resulting in morphologically heterogeneous populations (Fig. 1a) in which OG expression was undetectable. Approximately $1 \cdot 10^5$ cells were then retrovirally transfected with exogenous Oct4, Sox2, Klf4 and c-Myc (OSKM). After approximately 3 weeks, OG-positive embryonic stem cell (ESC)-like miPSC colonies were picked and expanded without drug selection. All AM (female) and YS (male)-miPSC lines generated here, which closely resembled ESCs in morphology (Fig. 1a), had a 2n = 40 normal karyotype (data not shown).

Characterization of AM and YS-miPSCs

As with ESCs, all AM- and YS-miPSC colonies were positive for alkaline phosphatase (ALP) (Fig. 1b). Immunohistochemical analyses also demonstrated that the cells were positive for pluripotent cell-specific nuclear proteins Oct4 and Nanog, and the surface glycoprotein SSEA1 (Fig. 1b). Thus, the expression profile of all marker proteins tested in AM and YS-miPSCs was similar to that observed in ESCs.

To examine the global transcription profile of these cells, comparative Affymetrix gene expression microarray analyses were performed between AM cells, YS cells, YS-miPSCs and R1 ESCs (Fig 1c). The global gene expression profile of YS-miPSCs was significantly different from that of YS cells. We detected a similar behavior between AM-miPSCs and AM cells (data not shown). Notably, the profile was similar to that of ESCs (Fig. 1c). Together, the data indicate that significant global nuclear reprogramming had occurred in these cells in response to OSKM transfection. We next applied RT-PCR analysis to gain a more focused transcriptional profile of pluripotent cell-specific marker genes in the induced cells. We found that Nanog, Rex1, ERas, Gdf3, Zfp296 and Ronin were expressed in both AM and YS-miPSCs, whereas the AM and YS genes, Igf1 and Ccl6 were silenced (Fig. 1c). Notably, Ronin was expressed not only in AM and YS-miPSCs but also in the precursor AM and YS cells. To investigate whether the exogenous Oct4, Sox2, Klf4 and c-Myc genes were silenced by DNA methylation as reported for other iPSCs (Jaenisch & Young 2008) in the AM and YS-miPSCs, we examined expression using gene-specific primer sets designed to distinguish endogenous and exogenous transcripts. In all miPSC lines, the expression of endogenous Oct4, Sox2, Klf4 and c-Myc was similar to that in R1 ESCs, whereas the exogenous c-Myc and Klf4 were fully silenced in some YS-miPSC clones but not in others (Fig. 1c). Notably, high-level expression of endogenous Klf4 and c-Myc was detected even in AM and YS cells, consistent with the expression of proto-oncogene (Curran et al. 1984). Endogenous expression of Klf4, c-Myc and Ronin genes that are involved in maintaining pluripotency may play a key function in enhancing the generation efficiency of miPSCs from AM and YS cells.

Timing and efficiency of miPSC generation

The molecular mechanisms that govern OKSM-induced nuclear reprogramming of somatic cells to iPSCs are poorly understood. It has been demonstrated that activation of endogenous Oct4 may be a landmark for irreversible epigenetic transition toward

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

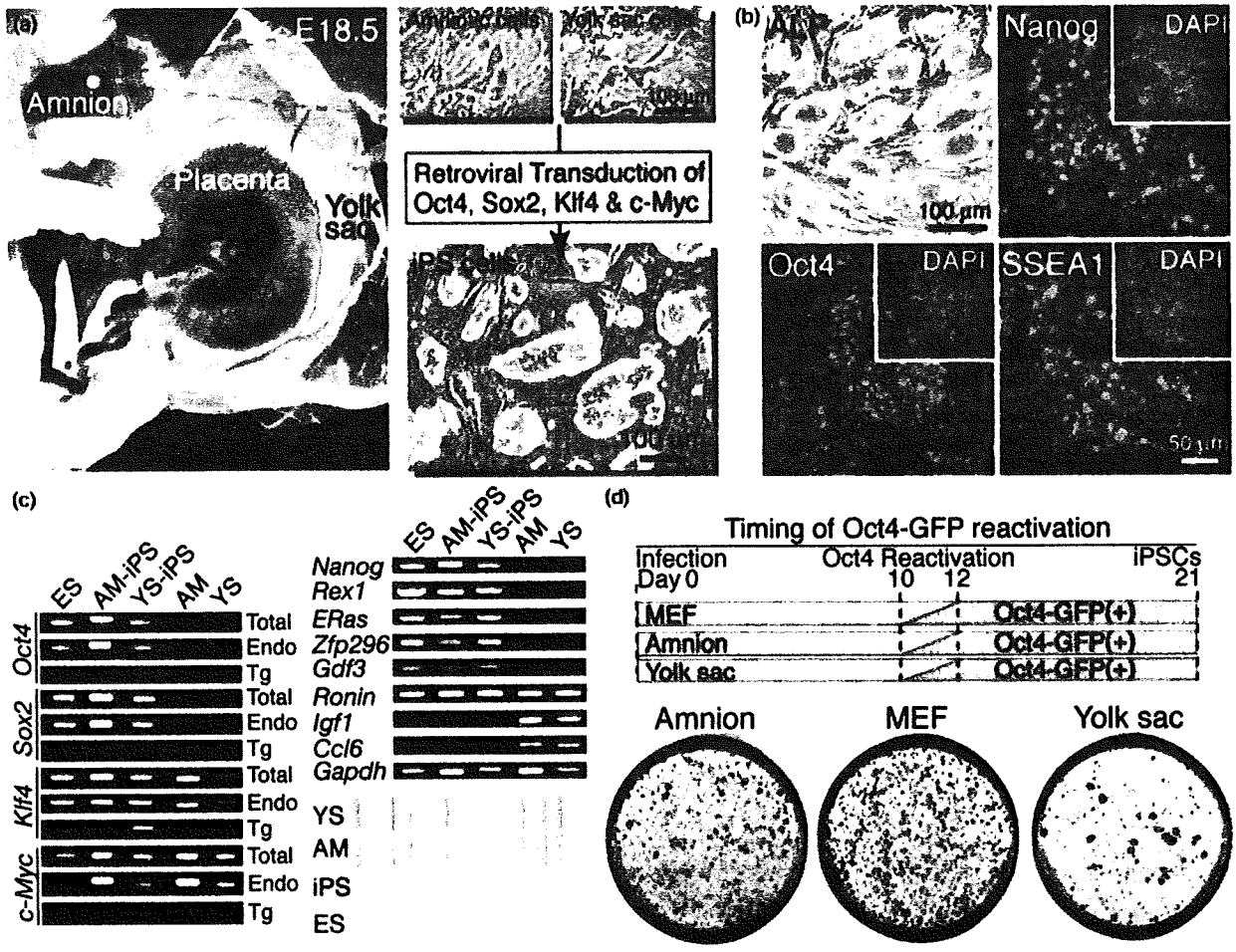

Figure 1 Generation of iPSCs from mouse AM and YS cells. (a) Isolation of AM and YS cells from the extra-embryonic tissues of newborn mice and generation of miPSCs through epigenetic reprogramming by retroviral infection-mediated expression of Oct4, Sox2, Klf4 and c-Myc. (b) Expression of pluripotent cell marker proteins, alkaline phosphatase (ALP), Nanog, Oct4 and SSEA1. Cell nuclei were visualized with DAPI. (c) Transcriptional activation and silencing of pluripotent and somatic cell marker genes by miPSC induction. RT-PCR analyses revealed that pluripotent marker genes were activated, somatic marker genes were silenced, and Klf4, c-Myc and Ronin were expressed even in AM and YS cells. Gapdh is a positive control. Microarray analyses demonstrated global alteration in gene expression profile between YS cells and YS-miPSCs, which more closely resemble mESCs. Relative level of gene expression is illustrated as red > yellow > green. (d) The generation efficiency of ALP-positive colonies and timing of GFP detection demonstrating Oct4-GFP reporter gene reactivation. ALP-positive colonies (red) in a 10-cm culture dish was shown when $1.0 \cdot 10^5$ of AM cells, YS cells and MEFs were exposed to OSKM reprogramming factors and reseeded at day 4.

fully reprogrammed iPSCs (Sridharan & Plath 2008). Thus, the timing of reactivation of OG is closely linked with the efficiency of reprogramming. Activation of exogenous OG was detected in some cell populations in every colony around 10 days after OSKM transfection of AM and YS cells, similar to control MEFs examined here and those reported previously (Fig. 1d) (Brambrink et al. 2008). The reprogramming efficiency of AM and YS cells was estimated by ALP-staining 21 days after OSKM transfection with reseeding at day 4. Notably, the number of ALP-positive colonies was similar between AM cells (4373 ± 983; mean ± SEM, n = 3) and MEFs (4997 ± 1049, n = 3), and • 50% in YS cells (2293 ± 487, n = 3). Thus, the efficiency of AM reprogramming by OSKM is comparable to that of MEFs, and far exceeds that of adult somatic cells (Fig. 1d).

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

Germline-transmissible chimeras with AM and YS-miPSCs

To address in vivo differentiation potential of the AM and YS-miPSCs, approximately 10 agouti miPSCs were microinjected into C57BL•6J • BDF1 blastocysts (black), and transferred into white ICR foster mothers to generate chimeras. Three male YS-miPSC and two female AM-miPSC lines were tested for chimera formation. X-gal staining analysis on sections of E15.5 embryos demonstrated successful generation of normally developing chimeric embryos with OG•Neo-LacZ miPSC contribution to the majority of tissues in all miPSC lines examined (data not shown). We next examined the miPSC potential for normal growth to sexual maturity and germline transmission. Two high-degree chimeric mice with a YS-miPSC line and three high-degree chimeric mice with two AM-miPSC lines, characterized by the >50% contribution of agouti coat color (Fig. 2a), developed normally into adulthood. However, an adult YS-miPSC chimera developed a neck tumor around 8–10 weeks after birth, which may be due to reactivation of the exogenous c-Myc as reported previously (Nakagawa et al. 2008). Testes isolated from affected males were bisected and one-half was X-gal-stained for LacZ activity whereas the other half was cryosectioned. Blue staining in the seminiferous tubule indicated that YS-miPSCs could contribute to germ cell development. To confirm this, testis cryosections immunohistochemically stained with antibodies against LacZ (iPSC-derived cell marker) and TRA98 (spermatogonia and spermatocyte marker) (Fig. 2b). Germ cells in all tubules were positive for TRA98, whereas germ cells in only some seminiferous tubules were positive for LacZ, clearly demonstrating that YS-miPSCs are capable of contributing to the differentiating germ line in chimeras. Finally, to examine whether the genetic information of YS-miPSCs was transmissible to the next generation, DNA isolated from progeny of the remaining YS-miPSC chimera was analyzed by genomic PCR with a primer set specific to Neo. Seven of the thirty-five pups examined were positive, demonstrating that YS-miPSCs are able to differentiate into fully functional germ cells (Fig. 2c). In one of three female AM-miPSC chimeric mice, competence for contribution to germ cells was detected by X-gal staining analysis of ovaries (data not shown).

Teratoma formation with AM and YS-miPSCs

The differentiation competence of AM and YS-miPSCs was further tested by teratoma formation induced by injection of cells into the inguinal region of immunodeficient SCID mice. Teratomas were isolated 5–8 weeks after for histological analysis and for gene expression analysis. Hematoxylin–eosin (HE) staining of paraffin sections demonstrated that the three primary layers were generated as morphologically shown by ectodermal glia and neuroepithelium, mesodermal muscle and endodermal ciliated epithelium and cartilage (Fig. 2d). Multi-lineage differentiation of miPSCs was verified by transcription of endodermal, mesodermal and ectodermal genes in the majority of teratomas (Fig. 2e).

Generation of iPSCs from human AM cells

To examine whether hiPSCs could be efficiently generated from primary AM cells isolated from the amniotic membrane (• 100 cm^2) of the placenta of newborn human (Fig. 3a), the reprogramming factors OCT4, SOX2, KLF4 and c-MYC were introduced by vesicular stomatitis virus G glycoprotein (VSV-G) retroviral transduction. About 20 AM-hiPSC lines were established from 1.0 • 10^5 AM cells infected (0.02%). The efficiency of AM-hiPSC generation is markedly high relative to that with cells from human adult tissues (Yu et al. 2007). AM-hiPSCs were morphologically similar to human ESCs (hESCs) (Fig. 3a). Immunohistochemical analyses demonstrated expression of the pluripotent cell-specific nuclear proteins OCT4, SOX2 and NANOG, and the keratan sulfate proteoglycan TRA-1-60 (Fig. 3b) consistent with the profile observed in hESCs. To extend this analysis, we examined the expression profile of genes by RT-PCR. The endogenous reprogramming factor genes OCT4, SOX2, KLF4 and c-MYC were all activated in AM-hiPSCs, whereas the transgenes were fully silenced (Fig. 3c). Expression of pluripotent cell-specific genes NANOG, REX1, GDF3, ESG1, FGF4, TERT and RONIN were also activated in all AM-hiPSC clones consistent with the profile of control hESCs (Fig. 3c). Notably, transcription of KLF4, c-MYC, and RONIN was detected not only in AM-hiPSCs but also AM cells. Similar to mouse AM and YS cells, endogenous expression of KLF4, c-MYC and RONIN in human AM cells may facilitate acquisition of reprogramming competency for efficient generation of hiPSCs.

DNA methylation of OCT4 and NANOG in AM-hiPSCs

To further characterize the pluripotent nature of AM-hiPSCs, the promoter CpG methylation status

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

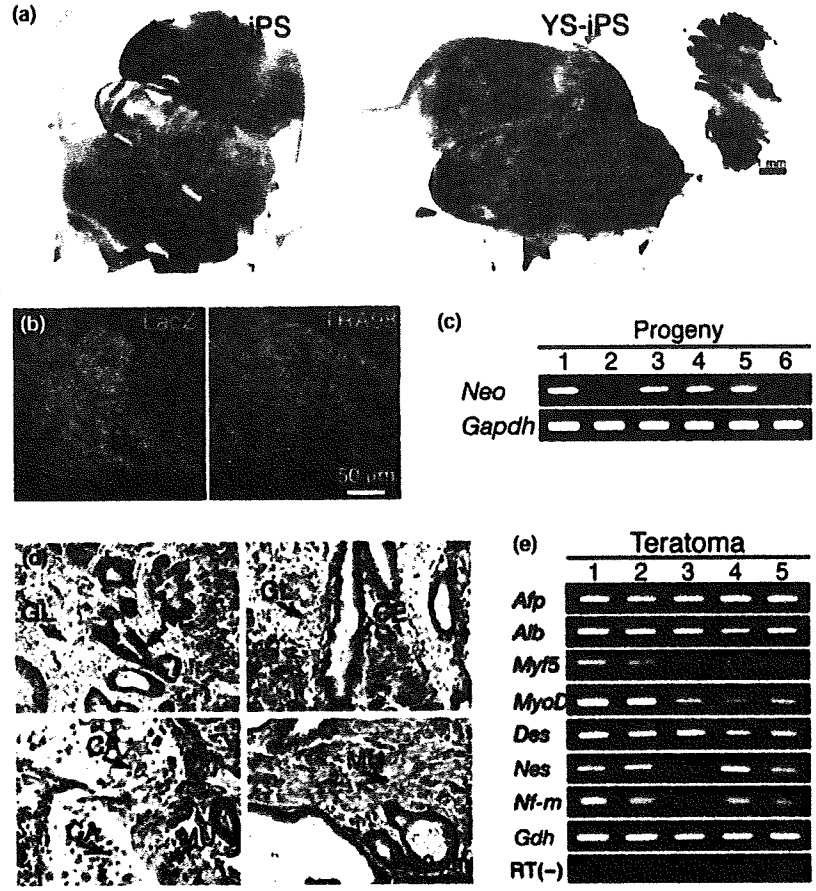

Figure 2 Pluripotency of AM and YS-miPSCs. (a) Chimeric mice with female AM-miPSCs and male YS-miPSCs. Inset: X-gal staining of testis collected from an adult YS-miPSC chimera (blue cells are YS-miPSC derivatives). (b) Immunohistochemical double staining of testis cryosections from a YS-miPSC chimera with anti-LacZ (YS-miPSC-derived germ cells) and anti-TRA98 (spermatogonia and spermatocytes) antibodies. (c) Genotyping of progeny obtained by backcrossing with YS-miPSC chimeras. Neo positive demonstrates germline transmission of YS-miPSC genetic information. Gapdh is positive control. (d) Hematoxylin-eosin staining of teratoma sections generated by AM and YS-miPSC implantation. GL, glia (ectoderm); NE, neuroepithelium (ectoderm); CE, ciliated epithelium (endoderm); CA, cartilage (ectoderm); MU, muscle (mesoderm). (e) Transcription analysis of lineage-specific genes in teratomas generated with AM and YS-miPSCs. Gray rectangle: endoderm makers; purple rectangle: mesoderm markers; pink rectangle: ectoderm markers. Afp, a-Fetoprotein; Alb, albumin; Des, desmin; Nes, Nestin; Nf-m, neurofilament-M; Gdh, Gapdh (positive control).

of key pluripotency genes was examined by bisulfite-modified DNA sequencing. Promoters of both OCT4 and NANOG were found to highly methylated in hAM cells, consistent with transcriptional silencing in these cells. Conversely, both promoter regions were hypo-methylated in AM-hiPSCs consistent with the observed reactivation (Fig. 3d). These data demonstrate that human AM cells are capable of being epigenetically reprogrammed to AM-hiPSCs through forced expression of reprogramming factors.

Teratoma formation with AM-hiPSCs

To address whether the AM-hiPSCs have competence to differentiate into specific tissues, teratoma formation was induced by implantation under the kidney capsule of immunodeficient nude mice. Twenty-one out of twenty-four AM-hiPS independent clones induced teratoma formation within 6–10 weeks of implantation ($1.0 \cdot 10^7$ cells·site). Histological analysis by HE staining of paraffin-embedded sections demonstrated that the three

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

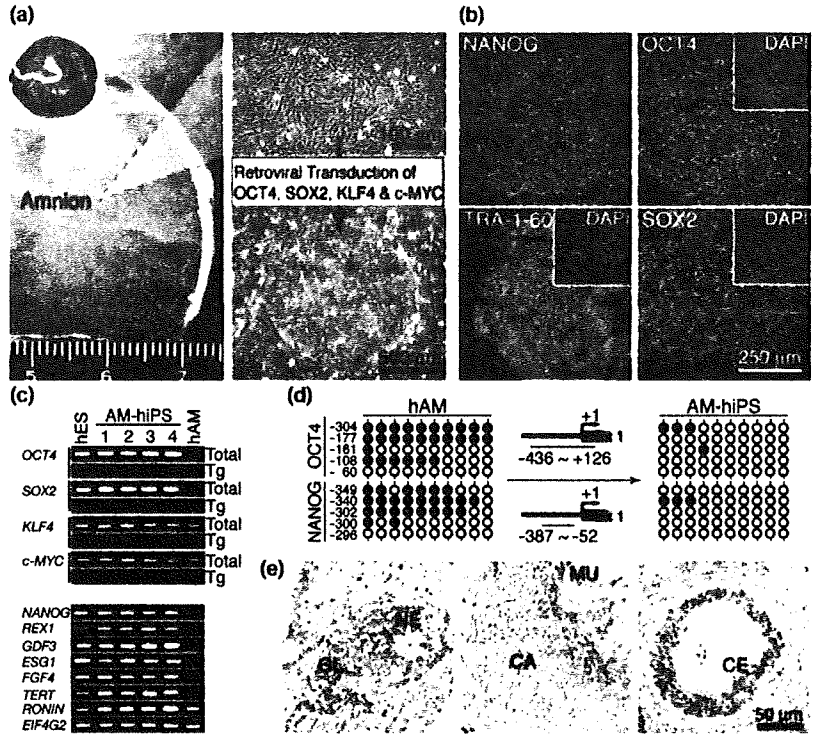

Figure 3 Generation of iPSCs from human AM cells. (a) Isolation of hAM cells from extra-embryonic tissues of human newborns and generation of hiPSCs through epigenetic reprogramming by retroviral infection-mediated expression of OCT4, SOX2, KLF4 and c-MYC. (b) Expression of pluripotent cell marker proteins, NANOG, OCT4, TRA-1-60 and SOX2. Cell nuclei were visualized with DAPI. (c) Transcriptional activation of pluripotent marker genes by hiPSC induction. RT-PCR analyses revealed that the exogenous OCT4, SOX2, KLF4 and c-MYC genes were silenced and the endogenous pluripotent marker genes were activated in AM-hiPSCs. KLF4, c-MYC and RONIN were expressed even in hAM cells before reprogramming. EIF4G2 (eukaryotic translation initiation factor 4 gamma 2) is included as a positive control. (d) Epigenetic reprogramming of the OCT4 and NANOG promoter regions. Bisulfite-modified DNA sequence analysis demonstrated a transition from hyper-methylation in AM cells (black circles) to hypo-methylation in AM-hiPSCs (white circles). (e) Hematoxylin-eosin staining of teratoma sections of teratoma generated by AM-hiPSC implantation. GL, glia (ectoderm); NE, neuroepithelium (ectoderm); CE, ciliated epithelium (endoderm); CA, cartilage (ectoderm); MU, muscle (mesoderm).

primary layers were generated as shown by ectodermal glia and neuroepithelium, mesodermal muscle and endodermal ciliated epithelium and cartilage morphologically (Fig. 3e). Thus, the majority of AM-hiPSC clones have potential for multi-lineage differentiation in vivo.

Discussion

We here demonstrated that hiPSCs and miPSCs were efficiently generated from newborn AM cells, in which endogenous Klf4, c-Myc and Ronin were highly expressed. The generation efficiency of miPSCs from AM cells was comparable to that from MEFs in mice and was notably high to that from adult somatic cells in humans. The properties of AM-hiPSCs and AM or YS-miPSCs resemble those of fully reprogrammed iPSCs from other tissues and ESCs.

iPSCs are generated through epigenetic reprogramming of somatic cells. Information on the base sequence of DNA in nuclei is unchanged through the reprogramming, although the gene expression profile is altered through the reprogramming from the somatic cell to the iPSC type. Developmentally rewound iPSCs retain aged DNA base sequence information inherited from somatic cells. The base sequence of DNA accumulates mutations through aging with cell division and mis-repair. Young somatic cells are suitable for iPSC generation rather than aged somatic cells. Therefore, it is suggested that the AM cells accumulating less genetic mutation are safer than the adult somatic cells as a cell source for iPSC generation.

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

The generation efficiency of OG-positive colonies was approximately four times lower than that of ALP-positive colonies and it is likely that miPSC generation will be further reduced (Wernig et al. 2008). Furthermore, when pre-iPSCs are reseeded, the generation efficiency of iPSC outcome could be roughly estimated as $1 \cdot 2^X$ (X = reseeded day after infection or transfection; doubling time of pre-iPSC is estimated as 24 h). Recently, iPSC generation technology has been developed and improved with MEFs and human embryonic or newborn fibroblasts (HNFs) as representative somatic cells. Even with these types of cells, application of the current technology resulted in a marked decrease in iPSC generation efficiency. The retroviral transduction-mediated miPSC generation efficiency is 0.05–0.1% with MEFs (Takahashi et al. 2007; Wernig et al. 2007). The generation efficiency of hiPSCs (• 0.01% in ALP-positive colony and 0.0025% in hiPSC outcome) (Yu et al. 2007; Wernig et al. 2008) is • 10 times lower than that of miPSCs. The generation efficiency of genetic modification-free hiPSCs from HNFs by direct delivery of reprogramming proteins is estimated at about 0.001% in outcome (Kim et al. 2009). Notably, it is evident that the generation of hiPSCs from adult somatic cells is much harder than that from MEFs. In fact, analysis with a secondary dox-inducible transgene system shows that the efficiency varies between different somatic cell types (Wernig et al. 2008). Thus, for practical application of iPSC technology to medical care, identification of reprogramming-sensitive cell types is a key issue. Human primary keratinocytes are one candidate cell type for efficient generation of hiPSCs from adult patients (the efficiency of ALP-positive colony = 1.0%) (Aasen et al. 2008). Here, we have shown that human and mouse AM cells, in which the endogenous KLF4 •Klf4, c-MYC •c-Myc and RONIN •Ronin are naturally expressed, are highly reprogramming-sensitive (hiPSC generation efficiency was approximately 0.02% in outcome). An important point is that relatively huge amounts of human AM cells can be collected from discarded AM membranes at birth with no risk to the individual. Furthermore, these cells can be kept in long-term storage without requirement for amplification by in vitro cell culture.

Our findings illustrate that human AM cells are a strong candidate cell source for collection and banking that could be retrieved on demand and used for generating personalized genetic modification-free iPSCs applicable for clinical treatment and drug screening.

Experimental procedures

Amnion and yolk sac cells

In mice, AM and YS membranes collected from E18.5 embryos from GOF-18 •delta PE •GFP (Oct4-GFP) transgenic females (Yoshimizu et al. 1999) mated with 129 •Rosa26 transgenic males (Friedrich & Soriano 1991) were digested with 0.1% collagenase (Wako, Osaka, Japan) and 20% fetal bovine serum (FBS) at 37 •C for 1 h, and then repeatedly passed through a 26-gauge needle. The cell suspension was cultured with mES medium (DMEM •F12 (Dulbecco's modified Eagle's medium •Ham's F12) (Wako) supplemented with 15% FBS, $10^{)4}$M 2-mercaptoethanol (Sigma) and 1000 U •mL of recombinant leukemia inhibitory factor (Chemicon, Temecula, CA, USA) containing 5 ng •mL basic fibroblast growth factor (bFGF) (Peprotech, Rocky Hill, NJ, USA). Following culture for 2–3 days, the adherent AM and YS cells growing to near-confluence were applied for iPSC experiments.

In humans, the AM membrane was cut into tiny pieces with dissection scissors. The AM membrane pieces were cultured in DMEM with 10% FBS for 7–10 days. The adherent AM cells growing to near-confluence were applied for iPSC experiments. Primary AM cells were provided from the cell bank of RIKEN Bioresource Center, Japan.

Generation of iPSCs

In mouse, each of pMXs-Oct4, Sox2, Klf4, c-Myc and DsRed (an indicator of retroviral silencing) was transfected into the Plat-E cells using the FuGENE6 Transfection Reagent (Roche Diagnostics, Indianapolis, IN, USA). A 1 : 1 : 1 : 1 : 4 mixture of Oct4, Sox2, Klf4, c-Myc and DsRed retroviruses in supernatants with 4 lg •mL polybrene (Nacalai Tesque, Kyoto, Japan) was added to AM and YS cells at $1.0 \cdot 10^5$ cells per 3 cm well. At day 4 after infection, the cells were reseeded into a 10 cm culture dish on feeder cells with mES medium. Colonies were picked around day 20.

In humans, pMXs-OCT4, SOX2, KLF4 or c-MYC, pCL-GagPol, and pHCMV-VSV-G vectors were transfected into 293FT cells (Invitrogen, Carlsbad, CA, USA) using the TransIT-293 reagent (Mirus). A 1 : 1 : 1 : 1 mixture of OCT4, SOX2, KLF4 and c-MYC viruses in supernatant with 4 lg •mL polybrene were added to AM cells at $1.0 \cdot 10^5$ cells per 3 cm well. The cells were subcultured on feeder cells into a 10 cm dish with the iPSellon medium (Cardio) supplemented with 10 ng •mL bFGF (Wako) (hES medium). Colonies were picked up around day 28.

Immunocytochemistry

Human and mouse cells were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) for 10 min at 4 •C. After washing with 0.1% Triton X-100 in PBS (PBST), the cells were prehybridized with blocking buffer for 1–12 h at 4 •C and then incubated with primary antibodies; anti-SSEA4

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

Table 1 Primers for RT-PCR and PCR

Gene name	5•-Forward-3•	5•-Reverse-3•
Mice		
Oct4 (total)	CTGAGGGCCAGGCAGGAGCACGAG	CTGTAGGGAGGGCTTCGGGCACTT
Oct4 (endogenous)	TCTTTCCACCAGGCCCCCGGCTC	TGCGGGCGGACATGGGGAGATCC
Oct4 (transgene)	CCCATGGTGGTGGTACGGGAATTC	AGTTGCTTTCCACTCGTGCT
Sox2 (total)	GGTTACCTCTTCCTCCCACTCCAG	TCACATGTGCGACAGGGGCAG
Sox2 (transgene)	CCCATGGTGGTGGTACGGGAATTC	TCTCGGTCTCGGACAAAAGT
Klf4 (total)	CACCATGGACCCGGGCGTGGCTGCCAGAAA	TTAGGCTGTTCTTTTCCGGGGCCACGA
Klf4 (endogenous)	GCGAACTCACACAGGCGAGAAACC	TCGCTTCCTCTTCCTCCGACACA
Klf4 (transgene)	CCCATGGTGGTGGTACGGGAATTC	GTCGTTGAACTCCTCGGTCT
c-Myc (total)	CAGAGGAGGAACGAGCTGAAGCGC	TTATGCACCAGAGTTTCGAAGCTGTTCG
c-Myc (endogenous)	CAGAGGAGGAACGAGCTGAAGCGC	AAGTTTGAGGCAGTTAAAATTATGGCTGAAGC
c-Myc (transgene)	CTCCTGGCAAAAGGTCAGAG	GACATGGCCTGCCCGGTTATTATT
Nanog	ATGAAGTGCAAGCGGTGGCAGAAA	CCTGGTGGAGTCACAGAGTAGTTC
Eras	CAAAGATGCTGGCAGGCAGCTACC	GACAAGCAGGGCAAAGGCTTCCTC
Gdf3	AGTTTCTGGGATTAGAGAAAGC	GGGCCATGGTCAACTTTGCCT
Rex1	GACATCATGAATGAACAAAAAATG	CCTTCAGCATTTCTTCCCTG
Zfp296	AAGCACCCAGATCTGTTGACCT	GAGCCTCTGGGGTATCTAGG
Ronin	GCCTCAGAGCTAGAGGCTGCTACG	TGGAAGGAGTCACGAATTCTGCAG
Igf1	GGACCAGAGACCCTTTGCGGGG	GGCTGCTTTTGTAGGCTTCAGTGG
Ccl6	CCTAAGCACCCTGAAGCAAG	ACAACTGGGAACCCACAAAGC
Gapdh	CCCACTAACATCAAATGGGG	CCTTCCACAATGCCAAAGTT
a-Fetoprotein	TCGTATTCCAACAGGAGG	CACTCTTCCTTCTGGAGATG
Albumin	AAGGAGTGCTGCCATGGTGA	CCTAGGTTTCTTGCAGCCTC
Myf-5	TGCCATCCGCTACATTGAGAG	CCGGGTAGCAGGCTGTGAGTTG
MyoD	GCCCGCGCTCCAACTGCTCTGAT	CCTACGGTGGTGCGCCCTCTGC
Desmin	TTGGGGTCGCTGCGGTCTAGCC	GGTCGTCTATCAGGTTGTCACG
Nestin	GGAGTGTCGCTTAGAGGTGC	TCCAGAAAGCCAAGAGAAGC
Neurofilament-M	GCCGAGCAGACCAAGGAGGCCATT	CTGGATGGTGTCCTGGTAGCTGCT
Neo	CGGCAGGAGCAAGGTGAGAT	CAAGATGGATTGCACGCAGG
Humans		
OCT4 (total)	GCCGTATGAGTTCTGTGG	TCTCCTTCTCCAGCTTCAC
SOX2 (total)	TAAGTACTGGCGAACCATCT	AAATTACCAACGGTGTCAAC
KLF4 (total)	ACTCGCCTTGCTGATTGTCT	GAACGTGGAGAAAGATGGGA
c-MYC (total)	GCGTCCTGGGAAGGGAGATCCGGAGC	TTGAGGGGCATCGTCGCGGGAGGCTG
NANOG	ATTATGCAGGCAACTCACTT	GATTCTTTACAGTCGGATGC
REX1	CAGATCCTAAACAGCTCGCAGAAT	GCGTACGCAAATTAAAGTCCAGA
GDF3	CTTATGCTACGTAAAGGAGCGGG	GTGCCAACCCAGGTCCCGGAAGTT
ESG1	ATATCCCGCCGTGGGTGAAAGTTC	ACTCAGCCATGGACTGGAGCATCC
FGF4	CTACAACGCCTACGAGTCCTACA	GTTGCACCAGAAAAGTCAGAGTTG
TERT	CCTGCTCAAGCTGACTCGACACCGTG	GGAAAAGCTGGCCCTGGGGTGGAGC
RONIN	CACTGTAGACAGCAGTCAGG	TGCCTTTCATCTCTTTCATC
EIF4G2	AAGGAAAGGGACTGAGTTTC	CCAAGAAAGCTTCTTCTTCA
Bis-OCT4	GATTAGTTTGGGTAATATAGTAAGGT	ATCCCACCCACTAACCTTAACCTCTA
Bis-NANOG	TGGTTAGGTTGGTTTTAAATTTTTG	AACCCACCCTTATAAATTCTCAATTA

(1 : 300) (Chemicon), anti-TRA-1–60 (1 : 300) (Chemicon), anti-Oct4 (1 : 50) (Santa Cruz Biotechnology, Santa Cruz, CA, USA), anti-Nanog (1 : 300) (ReproCELL, Tokyo, Japan), anti-Sox2 (1 : 300) (Abcam, Cambridge, UK) and•or anti-SSEA1 (1 : 1000) (DSHB) antibodies for 6–12 h at 4 •C. They were incubated with secondary antibodies; anti-rabbit IgG, anti-mouse IgG or anti-mouse IgM conjugated with Alexa 488 or 546 (1 : 500) (Molecular Probes, Eugene, OR, USA) in blocking buffer for 1 h at room temperature. The cells were counterstained with 4,6-diamidino-2-phenylindole (DAPI) and then mounted with a SlowFade light antifade kit (Molecular Probes). To examine germline competence,

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

cryosections of a half of a testis of 4- to 5-week-old chimeric mice were fixed with 4% paraformaldehyde in PBS for overnight at 4 °C, and then prehybridized with blocking buffer. The sections were double-stained with primary antibodies; anti-LacZ antibody (1 : 500) (Promega, Madison, WI, USA) specific to miPSC-derived cells and with anti-TRA98 antibody (1 : 500) specific to spermatogonia and spermatocytes. The remaining testis and ovaries were stained with X-gal.

RT-PCR

Total RNAs were isolated from mouse and human cells using the TRIzol (Invitrogen) and the RNeasy Plus Mini Kit (Qiagen, Valencia, CA, USA), respectively. cDNAs were synthesized from 1 lg total RNAs using Superscript III reverse transcriptase (Invitrogen) with random hexamers according to the manufacturer's instructions. Template cDNA was PCR-amplified with gene-specific primer sets (Table 1).

Gene expression microarray

Total RNA was extracted from mouse cells using the TRIzol Reagent. Double-stranded cDNA synthesized from the total RNA was amplified and labeled using the One-Cycle Target Labeling and Control Regents (Affymetrix, Santa Clara, CA, USA). Global gene expression was examined with the GeneChip Mouse Genome 430 2.0 Array (Affymetrix). The fluorescence intensity of each probe was quantified by using the GeneChip Analysis Suite 5.0 computer program (Affymetrix). The level of gene expression was determined as the average difference (AD). Specific AD levels were then calculated as percentages of the mean AD level of probe sets for housekeeping genes Actin and Gapdh. To eliminate changes within the range of background noise and to select the most differentially expressed genes, data were used only if the raw data values were less than 50 AD. Further data were analyzed with GeneSpring GX 7.3.1 (Agilent Technologies, Santa Clara, CA, USA).

Reprogramming efficiency

The reprogramming efficiency of mouse YS and AM cells was estimated by counting the number of ALP-positive colonies 21 days after retroviral infection. The cells in 10 cm culture dish were fixed with 4% paraformaldehyde in PBS for 15 min at room temperature and washed with PBS. After treating with ALP stain (pH 9.0) for 30 min at room temperature, the number of ALP-positive cells was counted.

Chimera

AM-miPSCs (2n = 40, XX) and YS-miPSCs (2n = 40, XY) were microinjected into blastocysts (C57BL•6J • BDF1). The blastocysts were transferred into the uterus of pseudopregnant ICR female mice. Chimeric mice were mated with C57BL•6J for examining germline transmission. The genotype of the progeny was determined with tail tip DNA by genomic PCR with a Neo-specific primer set (Table 1). All animal experiments were performed according to the guidelines of animal experiments of Kyoto University, Japan.

Teratoma

In mice, cell suspension of $1.0 \cdot 10^6$ AM or YS-miPSCs•100 lL DMEM•F12 was subcutaneously injected into the inguinal region of immunodeficient SCID mice (CLEA). In humans, the 1 : 1 mixture of the AM-hiPSC suspension and Basement Membrane Matrix (BD Biosciences, San Jose, CA, USA) were implanted at $1.0 \cdot 10^7$ cells•site under the kidney capsule of immunodeficient nude mice (CLEA). Teratomas surgically dissected out 5–8 weeks in mice and 6–10 weeks in human after implantation, were fixed with 4% paraformaldehyde in PBS, and embedded in paraffin. Sections at 10 lm in thickness were stained with HE.

Bisulfite-modified DNA sequencing

Genomic DNAs (1 lg) extracted from AM-hiPSCs and hAM cells were bisulfite-treated with EZ DNA methylation-Gold Kit (ZYMO Research, Orange, CA, USA) according to the manufacturer's instruction. The promoter regions of the human NANOG and OCT4 genes were PCR-amplified with specific primer sets (Table 1). Ten clones of each PCR product were gel-purified, sub-cloned and sequenced with the SP6 universal primer.

Acknowledgements

We thank Dr Gen Kondoh and Miss Hitomi Watanabe for generating chimeras, and Dr Justin Ainscough for critical comments on the manuscript.

References

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G., Edel, M., Boue, S. & Belmonte, J.C. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat. Biotechnol. 26, 1276–1284.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H. & Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell 2, 151–159.

Curran, T., Miller, A.D., Zokas, L. & Verma, I.M. (1984). Viral and cellular fos proteins: a comparative analysis. Cell 36, 259–268.

Feng, B., Ng, J.H., Heng, J.C. & Ng, H.H. (2009). Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. Cell Stem Cell 4, 301–312.

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

Friedrich, G. & Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev. 5, 1513–1523.

Jaenisch, R. & Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 132, 567–582.

Kim, D., Kim, C.H., Moon, J.I., Chung, Y.G., Chang, M.Y., Han, B.S., Ko, S., Yang, E., Cha, K.Y., Lanza, R. & Kim, K.S. (2009). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell Stem Cell 4, 472–476.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N. & Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat. Biotechnol. 26, 101–106.

Sridharan, R. & Plath, K. (2008). Illuminating the black box of reprogramming. Cell Stem Cell 2, 295–297.

Takahashi, K., Okita, K., Nakagawa, M. & Yamanaka, S. (2007). Induction of pluripotent stem cells from fibroblast cultures. Nat. Protoc. 2, 3081–3089.

Wernig, M., Lengner, C.J., Hanna, J., Lodato, M.A., Steine, E., Foreman, R., Staerk, J., Markoulaki, S. & Jaenisch, R. (2008). A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nat. Biotechnol. 26, 916–924.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E. & Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448, 318–324.

Yamanaka, S. (2007). Strategies and new developments in the generation of patient-specific pluripotent stem cells. Cell Stem Cell 1, 39–49.

Yoshimizu, T., Sugiyama, N., De Felice, M., Yeom, Y.I., Ohbo, K., Masuko, K., Obinata, M., Abe, K., Scholer, H.R. & Matsui, Y. (1999). Germline-specific expression of the Oct-4•green fluorescent protein (GFP) transgene in mice. Dev. Growth Differ. 41, 675–684.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I. & Thomson, J.A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917–1920.

Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., Trauger, S., Bien, G., Yao, S., Zhu, Y., Siuzdak, G., Scholer, H.R., Duan, L. & Ding, S. (2009). Generation of induced pluripotent stem cells using recombinant proteins. Cell Stem Cell 4, 381–384.

Received: 20 August 2009
Accepted: 16 September 2009

2009年(平成21年)11月17日 火曜日 朝日新聞

羊膜から効率よくiPS細胞

京都大などのグループ発表

胎児を包む羊膜を使って効率よくiPS（人工多能性幹）細胞をつくることに、京都大再生医科学研究所の多田高准教授らが成功し、16日に専門誌に発表した。ヒトでもマウスでも確認された。羊膜は長く保存できるため、将来構想されるiPS細胞バンクで使用する細胞の候補になる可能性がある。同研究所の山中伸弥教授や国立成育医療センター研究所との共同研究。

iPS細胞は、体細胞に四つの遺伝子を入れてつくる。多田准教授らは、羊膜の細胞では2遺伝子がすでに働いていることに注目した。羊膜の細胞に4遺伝子を組み込んでiPS細胞をつくったところ、大人の体の細胞からつくるより、iPS細胞ができる率は10倍以上となり、効率よくできた。

iPS細胞の作製効率を高める方法としては、特定の遺伝子の働きを抑えたり、化学物質を使ったりするなどの報告があるが、羊膜からつくる方法はより操作が少なくて済むという。多田准教授は「出産時に胎盤とともに出てくる羊膜を保存し、必要になったときにiPS細胞をつくるような応用が考えられる」と話している。（瀬川茂子）

• 2009 The Authors
Journal compilation • 2009 by the Molecular Biology Society of Japan/Blackwell Publishing Ltd.

BMC Research Notes

BioMed Central

Data Note

Open Access

Collection of *Macaca fascicularis* cDNAs derived from bone marrow, kidney, liver, pancreas, spleen, and thymus

Naoki Osada*[1], Makoto Hirata[1], Reiko Tanuma[1], Yutaka Suzuki[2], Sumio Sugano[2], Keiji Terao[3], Jun Kusuda[1], Yosuke Kameoka[1], Katsuyuki Hashimoto[1] and Ichiro Takahashi[1]

Address: [1]Department of Biomedical Resources, National Institute of Biomedical Innovation, 7-6-8 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan, [2]Department of Medical Genome Sciences, Graduate School of Frontier Sciences, University of Tokyo, 5-1-5 Kashiwano-ha, Kashiwa, Chiba 277-8561. Japan and [3]Tsukuba Primate Research Center, National Institute of Biomedical Innovation, 1 Hachimandai, Tsukuba 305-0843, Japan

Email: Naoki Osada* - nosada@nibio.go.jp; Makoto Hirata - mhirata@nibio.go.jp; Reiko Tanuma - tanumark@nibio.go.jp; Yutaka Suzuki - ysuzuki@hgc.jp; Sumio Sugano - ssugano@ims.u-tokyo.ac.jp; Keiji Terao - terao@nibio.go.jp; Jun Kusuda - jkusuda@nibio.go.jp; Yosuke Kameoka - ykameoka@nibio.go.jp; Katsuyuki Hashimoto - khashi@nih.go.jp; Ichiro Takahashi - ichiro-t@nibio.go.jp

* Corresponding author

Published: 29 September 2009

BMC Research Notes 2009, 2:199 doi:10.1186/1756-0500-2-199

Received: 15 May 2009
Accepted: 29 September 2009

This article is available from: http://www.biomedcentral.com/1756-0500/2/199

© 2009 Osada et al; licensee BioMed Central Ltd.
This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Consolidating transcriptome data of non-human primates is essential to annotate primate genome sequences, and will facilitate research using non-human primates in the genomic era. *Macaca fascicularis* is a macaque monkey that is commonly used for biomedical and ecological research.

Findings: We constructed cDNA libraries of *Macaca fascicularis*, derived from tissues obtained from bone marrow, liver, pancreas, spleen, and thymus of a young male, and kidney of a young female. In total, 5'-end sequences of 56,856 clones were determined. Including the previously established cDNA libraries from brain and testis, we have isolated 112,587 cDNAs of *Macaca fascicularis*, which correspond to 56% of the curated human reference genes.

Conclusion: These sequences were deposited in the public sequence database as well as in-house macaque genome database http://genebank.nibio.go.jp/qfbase/. These data will become valuable resources for identifying functional parts of the genome of macaque monkeys in future studies.

Findings

Macaca fascicularis (cynomolgus, crab-eating, or long-tail macaque) is one of the most popular primate species used in biomedical research, and is closely related to *Macaca mulatta* (rhesus macaque). The draft sequence of the *Macaca mulatta* genome, which has an evolutionary important position, was published in 2007 [1].

Transcriptiome data broadens the application of genome sequences. Compared with several millions of human transcript sequences, macaque transcriptome data has only been analyzed in a limited numbers of studies [2-6]. A complete list of macaque genes will be beneficial for performing genetic studies using macaques in the future. We aim to elucidate all the macaque transcripts that cor-

respond to human genes, which have been widely accepted as reference sequences, such as the RefSeq sequences [7].

We have published expressed sequence tag (EST) and full-length sequences, which were obtained from cDNA libraries of brain and testis of *Macaca fascicularis*, using a variety of research subjects [5,8-13]. Here, we present 5'-EST sequences from six other tissues of *Macaca fascicularis*. Bone marrow, liver, pancreas, spleen, and thymus from a 4-year-old male Malaysian *Macaca fascicularis*, and kidney from a 3-year-old female Philippine *Macaca fascicularis* were harvested. These animals are bred and reared in the Tsukuba Primate Research Center (TPRC), National Institute of Biomedical Innovation (Ibaraki, Japan). The tissues were harvested in the P2 facility in TPRC, in accordance with the guidelines of the Laboratory Biosafety Manual, World Health Organization. The libraries for kidney (QreA and QreB) and liver (QlvC) were constructed using the vector-capping method [14], and those for bone marrow (QbmA), pancreas (QpaA), spleen (QspA), and thymus (QthA) were constructed using the oligo-capping method [15]. The sequences of 5'-EST were determined by Sanger sequencing using an ABI 3730 sequencer, and all vector sequences were filtered out [5]. Nucleotide calls with a quality value (QV) of less than 15 were masked as ambiguous. After the masking, the sequences were trimmed, such that they did not contain more than four ambiguous nucleotides in a 10-bp width window, and sequences shorter than 100 bp after the trimming were filtered out. After the trimming, the average sequence length was 886.9 bp.

In total, we obtained 56,856 EST sequences from the six tissues. The repeat sequences were masked by Repbase Update before the BLAST search [16]. The BLAST search (BLASTN) was performed with a cut-off value (*E*-value) of 1e-60 against human RefSeq data [7]. Since RefSeq sequences contain partially overlapped isoforms, we constructed non-redundant RefSeq sequences based on the Entrez Gene database [17]. Hereafter, we shall refer to the non-redundant RefSeq sequences as RefSeq genes. There were 23,236 RefSeq genes, including non-coding RNAs in the human genome at the time of investigation (Release 34) [7]. Out of the newly isolated 56,856 cDNA clones, 44,603 matched to 4940 human RefSeq genes. Of the 12,253 non-RefSeq clones, 40 consisted of repeat sequences, and the other 1631 did not show any homology to human transcript sequences in public databases using a lower cutoff value (1e-15). Meanwhile, 23,900 EST sequences were homologous to multiple RefSeq genes with the high cutoff value (1e-60). The average nucleotide sequence identity between the best BLAST hit pairs was 95.26%. The nucleotide sequence identity was slightly lower than that estimated using full-length cDNA sequences of high quality [5], and supposed to reflect some sequencing errors in the EST sequences. In some cases, the nucleotide sequence identity between the best and second best hit pairs were very close, which was probably due to gene duplications specific in the human lineage. The difference in nucleotide sequence identities between the best and second best BLAST hits were less than 0.5% in 8996 ESTs. In such cases, the best hit orthologs would not be regarded as unique orthologs of humans and macaques. In Figure 1, we classify the macaque ESTs according to the number of BLAST hits to RefSeq genes. The average nucleotide sequence identities were ordered by the rank of BLAST hits. For example, the nucleotide sequence identity in the second bin represents the identity between the second best hit pairs.

In conjunction with the previously sequenced cDNA clones, we obtained 112,587 EST sequences corresponding to 8262 human RefSeq genes, which correspond to 36% of all human RefSeq genes. When we restricted the analysis of the human RefSeq genes in the manually curated status (Reviewed or Validated status) [7], 56% (6,177/11,080) of the human RefSeq genes were covered by the macaque transcriptome.

As shown in Table 1, the number of RefSeq genes that were represented in the libraries was different in different tissues. In order to measure the unbiased transcript redun-

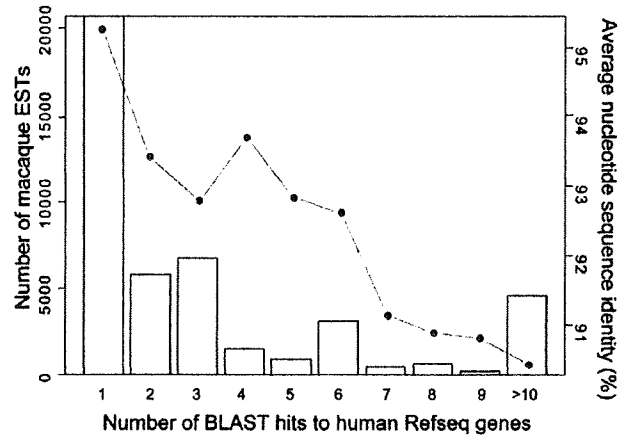

Figure 1
Number of BLAST hits (cutoff: 1e-60) against the human RefSeq genes. The grey bars represent the number of macaque ESTs matched to the human RefSeq genes. ESTs matched more than nine RefSeq genes were combined into a single bin. The red circles and lines represent the average nucleotide sequence identity between the macaque ESTs and RefSeq genes, ordered by the rank of BLAST hits. For example, the sequence identity in the second bin represents the sequence identity between the second best hits.

Table 1: Summary of *Macaca fascicularis* cDNA libraries

Tissue	Total clones	Covered RefSeq[d]	non-RefSeq[e]	Redundancy[f]
Brain cortex[a, c]	28679	4035	10259	2.32
Brain stem[b, c]	5758	1591	2050	2.40
Cerebellum[c]	11003	2340	4179	2.32
Testis[c]	8551	1833	3300	2.36
Liver	9188	1360	3853	3.21
Kidney	9558	2495	2630	1.91
Bone marrow	9472	1366	1317	3.26
Spleen	9783	1556	1527	3.15
Thymus	9566	1295	1491	2.96
Pancreas	9289	534	1435	9.83
All	112587	8262	32269	2.14

[a]Brain cortex includes parietal lobe (Qnp), temporal lobe (Qtr), occipital lobe (Qor), and frontal lobe (Qfl).
[b]Brain stem includes medulla oblongata (Qmo) and the other part of brain stem (Qbs).
[c]These sequences were determined by the previous studies [8-10,12].
[d]Number of human RefSeq genes that have macaque homologs in each library.
[e]TheNumber of macaque cDNA clones that do not have human RefSeq homologs.
[f]Estimated from randomly chosen 1000 macaque transcripts, averaged over 1000 simulations.

dancy in each tissue, we estimated the redundancy of the human RefSeq homologs in 1000 macaque transcripts in each tissue. We randomized the transcript data and selected 1000 transcripts to enumerate the human RefSeq genes covered by the transcripts. The redundancy was given by the number of transcripts (1000) divided by the number of human RefSeq genes covered by the transcripts. This procedure was repeated 1000 times for each tissue, and the average redundancy was estimated. The results are shown in the last column of Table 1. Pancreas showed the highest redundancy; while brain and testis showed low redundancy, indicating that the gene expression complexity in brain and testis is higher than that in the other tissues, as suggested previously [18]. We also found that the kidney library (QreA) had very low redundancy. It was constructed using the vector-capping method, which does not amplify the template cDNA by PCR and may reduce the redundancy of the library [14]. In order to test the effectiveness of the cloning methods, we compared the redundancy of the transcript in our liver library constructed using the vector-capping method, and the previously reported liver library constructed using the oligo-capping method [6]. The redundancy in the vector-capped liver library was 3.21 (Table 1). In contrast, the redundancy in the oligo-capped liver library was 5.19 [6], which was significantly higher than that in the vector-capped library ($P < 0.001$, permutation test).

We have developed an in-house database for the genome data of *Macaca fascicularis* (QFbase: http://genebank.nibio.go.jp/qfbase/) [5]. The *Macaca fascicularis* cDNA sequences described in this report were annotated and added to this database. They were also mapped on the rhesus macaque genome sequence using the BLAT program [19]. The results can be viewed in the *Macaca fascicularis* genome browser http://genebank.nibio.go.jp/cgi-bin/gbrowse/rheMac2/, which is implemented using GBrowse software [20]. The DDBJ/EMBL/Genbank accession numbers of these sequences are DC629777-DC639249 (bone marrow), DC639249-DC648806 (kidney), DC620589-DC629776 (liver), FS362802-FS372090 (pancreas), DC848487-DC858269 (spleen), and DK575154-DK584719 (thymus).

Availability and requirements

- **Project name**: *Macaca fascicularis* cDNA sequencing project
- **Project home page**: http://genebank.nibio.go.jp/qfbase/
- **Operating system(s)**: Platform independent
- **Programming language**: PERL
- **Other requirements**: Generic web browser
- **License**: GNU, GPL
- **Any restrictions to use by non-academics**: none

Abbreviations

EST: expressed sequence tag; QV: quality value;

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

NO, KT, JK, YK, KH, and IT contributed to the design of the research. NO analyzed the data. NO and KH wrote the

manuscript. MH performed the computational analysis. RT, YK, and IT were involved in the cDNA sequencing. YS and SS constructed the oligo-capped cDNA libraries. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by a Health Science Research grant from the Ministry of Health, Labor, and Welfare of Japan.

References

1. Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, *et al.*: **Evolutionary and biomedical insights from the rhesus macaque genome.** *Science* 2007, **316**:222-234.
2. Magness CL, Fellin PC, Thomas MJ, Korth MJ, Agy MB, Proll SC, Fitzgibbon M, Scherer CA, Miner DG, Katze MG, Iadonato SP: **Analysis of the Macaca mulatta transcriptome and the sequence divergence between Macaca and human.** *Genome Biol* 2005, **6**:R60.
3. Chen WH, Wang XX, Lin W, He XW, Wu ZQ, Lin Y, Hu SN, Wang XN: **Analysis of 10,000 ESTs from lymphocytes of the cynomolgus monkey to improve our understanding of its immune system.** *BMC Genomics* 2006, **7**:82.
4. Wallace JC, Korth MJ, Paeper B, Proll SC, Thomas MJ, Magness CL, Iadonato SP, Nelson C, Katze MG: **High-density rhesus macaque oligonucleotide microarray design using early-stage rhesus genome sequence information and human genome annotations.** *BMC Genomics* 2007, **8**:28.
5. Osada N, Hashimoto K, Kameoka Y, Hirata M, Tanuma R, Uno Y, Inoue I, Hida M, Suzuki Y, Sugano S, *et al.*: **Large-scale analysis of Macaca fascicularis transcripts and inference of genetic divergence between M. fascicularis and M. mulatta.** *BMC Genomics* 2008, **9**:90.
6. Uno Y, Suzuki Y, Wakaguri H, Sakamoto Y, Sano H, Osada N, Hashimoto K, Sugano S, Inoue I: **Expressed sequence tags from cynomolgus monkey (Macaca fascicularis) liver: a systematic identification of drug-metabolizing enzymes.** *FEBS Lett* 2008, **582**:351-358.
7. Pruitt KD, Tatusova T, Klimke W, Maglott DR: **NCBI Reference Sequences: current status, policy and new initiatives.** *Nucleic Acids Res* 2009, **37**:D32-36.
8. Osada N, Hida M, Kusuda J, Tanuma R, Hirata M, Hirai M, Terao K, Suzuki Y, Sugano S, Hashimoto K: **Prediction of unidentified human genes on the basis of sequence similarity to novel cDNAs from cynomolgus monkey brain.** *Genome Biol* 2002, **3**:RESEARCH0006.
9. Osada N, Hida M, Kusuda J, Tanuma R, Hirata M, Suto Y, Hirai M, Terao K, Sugano S, Hashimoto K: **Cynomolgus monkey testicular cDNAs for discovery of novel human genes in the human genome sequence.** *BMC Genomics* 2002, **3**:36.
10. Osada N, Hida M, Kusuda J, Tanuma R, Iseki K, Hirata M, Suto Y, Hirai M, Terao K, Suzuki Y, *et al.*: **Assignment of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes.** *Gene* 2001, **275**:31-37.
11. Osada N, Hirata M, Tanuma R, Kusuda J, Hida M, Suzuki Y, Sugano S, Gojobori T, Shen CK, Wu CI, Hashimoto K: **Substitution rate and structural divergence of 5'UTR evolution: comparative analysis between human and cynomolgus monkey cDNAs.** *Mol Biol Evol* 2005, **22**:1976-1982.
12. Osada N, Kusuda J, Hirata M, Tanuma R, Hida M, Sugano S, Hirai M, Hashimoto K: **Search for genes positively selected during primate evolution by 5'-end-sequence screening of cynomolgus monkey cDNAs.** *Genomics* 2002, **79**:657-662.
13. Wang HY, Chien HC, Osada N, Hashimoto K, Sugano S, Gojobori T, Chou CK, Tsai SF, Wu CI, Shen CK: **Rate of Evolution in Brain-Expressed Genes in Humans and Other Primates.** *PLoS Biol* 2007, **5**:e13.
14. Kato S, Ohtoko K, Ohtake H, Kimura T: **Vector-capping: a simple method for preparing a high-quality full-length cDNA library.** *DNA Res* 2005, **12**:53-62.
15. Maruyama K, Sugano S: **Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides.** *Gene* 1994, **138**:171-174.
16. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J: **Repbase Update, a database of eukaryotic repetitive elements.** *Cytogenet Genome Res* 2005, **110**:462-467.
17. Maglott D, Ostell J, Pruitt KD, Tatusova T: **Entrez Gene: gene-centered information at NCBI.** *Nucleic Acids Res* 2007, **35**:D26-31.
18. Chikaraishi DM, Deeb SS, Sueoka N: **Sequence complexity of nuclear RNAs in adult rat tissues.** *Cell* 1978, **13**:111-120.
19. Kent WJ: **BLAT--the BLAST-like alignment tool.** *Genome Res* 2002, **12**:656-664.
20. Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, Lewis S: **The generic genome browser: a building block for a model organism system database.** *Genome Res* 2002, **12**:1599-1610.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

BioMedcentral