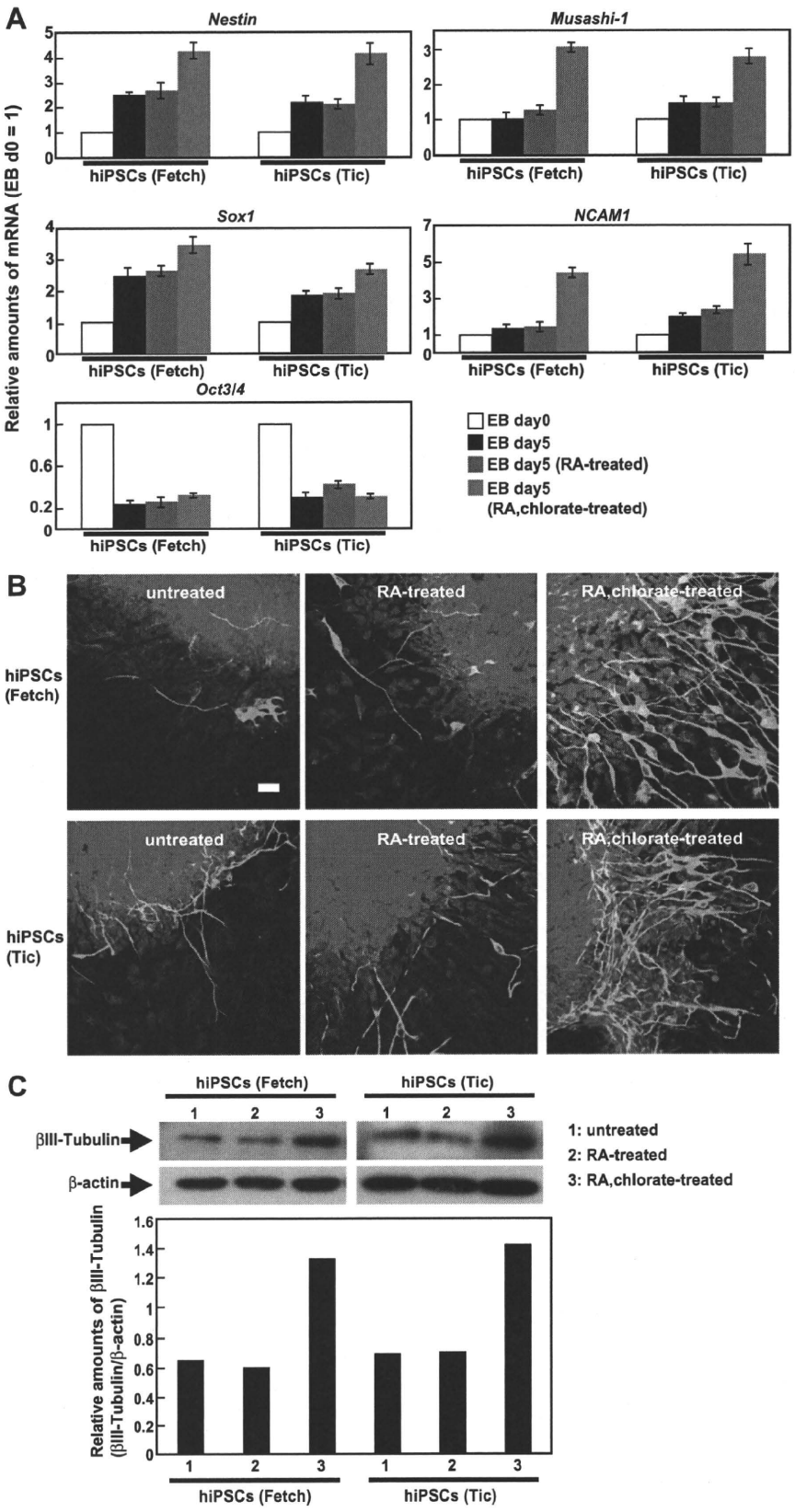


Fig. 4. Reduction of sulfation accelerates neural differentiation of hiPSCs. (A) Real-time PCR analysis of neural markers (*Nestin*, *Musashi-1*, *NCAM1* and *Sox1*) and an undifferentiated cell marker (*Oct3/4*) 5 days after EB formation. The relative amounts of each mRNA were normalized by *GAPDH* mRNA. The results are shown after normalization against the values obtained with undifferentiated hiPSCs (value = 1). The values shown are the means ± SD of duplicate measurements from representative experiments. Two independent experiments were performed. (B) Immunocytochemical staining 2 days after replating of EBs. Representative confocal images from two independent experiments are shown (βIII-tubulin, green; PI, red). Scale bar, 20 μm. (C) Western blotting 2 days after replating of EBs. Representative immunoblots from two independent experiments are shown. The histogram shows mean densitometric readings ± SD of βIII-tubulin/loading controls.

Fig. 3, Wnt/β-catenin, BMP/Smad, and FGF/ERK signaling, which are involved in the mesodermal differentiation of mESCs [24], were reduced in chlorate-treated EBs. Therefore, the inhibition of these signaling pathways by chlorate treatment seemed to cause defects in mesodermal differentiation and induced ectodermal differentiation. Chlorate treatment from the start of EB formation inhibited initial differentiation, as reported previously (data not shown, [25]). Thus, chlorate treatment from the appropriate time induces differentiation into a specific lineage, in particular neural differentiation.

Next, we investigated further the neural differentiation of chlorate-treated EBs. Firstly, we examined the expression of neural differentiation markers (Fig. 2B). Even at 5 days after EB formation, the expression levels of these markers were significantly higher in chlorate-treated EBs than in EBs only treated with RA, which indicated that differentiation into neural stem/neural progenitor cells and neural precursor cells was accelerated by chlorate treatment. These results were thought to be due to the reduction of Wnt and BMP signaling as shown in Fig. 3. Because, signaling by BMP and Wnt inhibits the neurogenesis of mESCs via EB formation [26,27]. Secondly, we examined the ability of chlorate-treated EBs to differentiate into mature neurons. Immunocytochemical staining for the mature neuron marker βIII-tubulin showed that only 2 days after plating, chlorate-treated EBs generated extremely dense networks of neurite outgrowths, which were not seen with EBs only treated with RA (Fig. 2C). We confirmed by western blotting that the level of βIII-tubulin in chlorate-treated cells was significantly higher than that in cells only treated with RA (Fig. 2D). Taken together, these results demonstrated that the reduction of sulfation by chlorate treatment accelerated the neurogenesis of mESCs.

As described above, the reduction of sulfation on HS and CS by chlorate treatment (Fig. 1) caused defects in Wnt/β-catenin, BMP/Smad, and FGF/ERK signaling during EB formation (Fig. 3), which was followed by the acceleration of neural differentiation (Fig. 2). These results are consistent with previous report [13], in which reduction of PAPST-dependent sulfation promotes neurogenesis of mESCs due to the decreased sulfation on HS and CS.

To date, several signaling pathways have been shown to contribute to the neural differentiation of hESCs and hiPSCs [18,28]. The BMP inhibitor Noggin and the inhibitor of the Lefty/Activin/TGFβ pathways SB431542 enhance neural differentiation [18]. Thus, we examined whether the inhibition of signaling pathways by chlorate treatment also enhanced the neural differentiation of hiPSCs. Real-time PCR analysis 5 days after EB formation showed that the expression of several neural stem/progenitor cell markers was higher in chlorate-treated EBs than in EBs only treated with RA, whereas expression of the marker of undifferentiated cells, *Oct3/4*, was reduced equally in both types of EBs (Fig. 4A). These results indicate that chlorate treatment accelerated the neural differentiation of hiPSCs. Furthermore, as shown in Fig. 4B and C, differentiation into mature neurons was induced markedly in chlorate-treated EBs.

In conclusion, we have demonstrated that the down-regulation of sulfation by chlorate treatment could accelerate neural differentiation of hiPSCs as well as mESCs. This acceleration was induced by the addition of chlorate at the correct time, and involved the inhibition of signaling pathways involved in the induction of mesodermal differentiation (Wnt, BMP, and FGF) and inhibition of neural differentiation (Wnt and BMP). This study as well as our previous report [13] have demonstrated that reduction of sulfation is effective to induce neural differentiation compared with the standard method using RA. In our previous report [13], we used RNAi method using retrovirus vector for long-term neural differentiation. This method requires great care and may give damages to the transfected cells. Moreover, the reduction of sulfation is more drastic in chlorate-treated cells than in *PAPST*-knockdown cells. This reflects more rapid induction of neural differentiation in chlorate-treated cells as shown in this report than in *PAPST*-knockdown cells. Furthermore, in hiPSCs including hESCs, efficiency of gene transfer is very low and for clinical application gene transfer should be avoided. Therefore, chlorate is useful for rapid, simple, and safety reduction of sulfation for rapid neural differentiation of hiPSCs, possibly including hESCs. Taken together, we propose that chlorate, which is available at low cost, could be used to achieve efficient and rapid neural induction of hiPSCs as well as mESCs in place of expensive signaling inhibitors.

Acknowledgments

Our research was partially supported by funds from Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Grant-in-Aid for Scientific Research (B) to SN, 20370051, 2008–2010, and from MEXT, the Matching Fund for Private Universities, S0901015, 2009–2014.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2010.09.085.

References

[1] M.J. Evans, M.H. Kaufman, Establishment in culture of pluripotential cells from mouse embryos, Nature 292 (1981) 154–156.
[2] G.R. Martin, Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells, Proc. Natl. Acad. Sci. USA 78 (1981) 7634–7638.
[3] J.A. Thomson, J. Itskovitz-Eldor, S.S. Shapiro, M.A. Waknitz, J.J. Swiergiel, V.S. Marshall, J.M. Jones, Embryonic stem cell lines derived from human blastocysts, Science 282 (1998) 1145–1147.
[4] G. Keller, Embryonic stem cell differentiation: emergence of a new era in biology and medicine, Genes Dev. 19 (2005) 1129–1155.
[5] M. Bernfield, M. Gotte, P.W. Park, O. Reizes, M.L. Fitzgerald, J. Lincecum, M. Zako, Functions of cell surface heparan sulfate proteoglycans, Annu. Rev. Biochem. 68 (1999) 729–777.
[6] N. Sasaki, K. Okishio, K. Ui-Tei, K. Saigo, A. Kinoshita-Toyoda, H. Toyoda, T. Nishimura, Y. Suda, M. Hayasaka, K. Hanaoka, S. Hitoshi, K. Ikenaka, S. Nishihara, Heparan sulfate regulates self-renewal and pluripotency of embryonic stem cells, J. Biol. Chem. 283 (2008) 3594–3606.
[7] D.C. Kraushaar, Y. Yamaguchi, L. Wang, Heparan sulfate is required for embryonic stem cells to exit from self-renewal, J. Biol. Chem. 285 (2010) 5907–5916.
[8] S. Besset, J.B. Vincourt, F. Amalric, J.P. Girard, Nuclear localization of PAPS synthetase 1: a sulfate activation pathway in the nucleus of eukaryotic cells, FASEB J. 14 (2000) 345–354.
[9] S. Kamiyama, T. Suda, R. Ueda, M. Suzuki, R. Okubo, N. Kikuchi, Y. Chiba, S. Goto, H. Toyoda, K. Saigo, M. Watanabe, H. Narimatsu, Y. Jigami, S. Nishihara, Molecular cloning and identification of 3′-phosphoadenosine 5′-phosphosulfate transporter, J. Biol. Chem. 278 (2003) 25958–25963.
[10] F. Luders, H. Segawa, D. Stein, E.M. Selva, N. Perrimon, S.J. Turco, U. Hacker, Slalom encodes an adenosine 3′-phosphate 5′-phosphosulfate transporter essential for development in Drosophila, EMBO J. 22 (2003) 3635–3644.
[11] S. Kamiyama, N. Sasaki, E. Goda, K. Ui-Tei, K. Saigo, H. Narimatsu, Y. Jigami, R. Kannagi, T. Irimura, S. Nishihara, Molecular cloning and characterization of a novel 3′-phosphoadenosine 5′-phosphosulfate transporter, PAPST2, J. Biol. Chem. 281 (2006) 10945–10953.
[12] E. Goda, S. Kamiyama, T. Uno, H. Yoshida, M. Ueyama, A. Kinoshita-Toyoda, H. Toyoda, R. Ueda, S. Nishihara, Identification and characterization of a novel Drosophila 3′-phosphoadenosine 5′-phosphosulfate transporter, J. Biol. Chem. 281 (2006) 28508–28517.
[13] N. Sasaki, T. Hirano, T. Ichimiya, M. Wakao, K. Hirano, A. Kinoshita-Toyoda, H. Toyoda, Y. Suda, S. Nishihara, The 3′-phosphoadenosine 5′-phosphosulfate transporters, PAPST1 and 2, contribute to the maintenance and differentiation of mouse embryonic stem cells, PLoS One 4 (2009) e8262.
[14] T.C. Ullrich, R. Huber, The complex structures of ATP sulfurylase with thiosulfate, ADP and chlorate reveal new insights in inhibitory effects and the catalytic cycle, J. Mol. Biol. 313 (2001) 1117–1125.
[15] P.A. Baeuerle, W.B. Huttner, Chlorate – a potent inhibitor of protein sulfation in intact cells, Biochem. Biophys. Res. Commun. 141 (1986) 870–877.
[16] F. Safaiyan, S.O. Kolset, K. Prydz, E. Gottfridsson, U. Lindahl, M. Samivirta, Selective effects of sodium chlorate treatment on the sulfation of heparan sulfate, J. Biol. Chem. 274 (1999) 36267–36273.

[17] K. Takahashi, S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors, Cell 126 (2006) 663–676.
[18] S.M. Chambers, C.A. Fasano, E.P. Papapetrou, M. Tomishima, M. Sadelain, L. Studer, Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling, Nat. Biotechnol. 27 (2009) 275–280.
[19] A. Nagy, J. Rossant, R. Nagy, W. Abramow-Newerly, J.C. Roder, Derivation of completely cell culture-derived mice from early-passage embryonic stem cells, Proc. Natl. Acad. Sci. USA 90 (1993) 8424–8428.
[20] A.G. Smith, M.L. Hooper, Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells, Dev. Biol. 121 (1987) 1–9.
[21] H. Makino, M. Toyoda, K. Matsumoto, H. Saito, K. Nishino, Y. Fukawatase, M. Machida, H. Akutsu, T. Uyama, Y. Miyagawa, H. Okita, N. Kiyokawa, T. Fujino, Y. Ishikawa, T. Nakamura, A. Umezawa, Mesenchymal to embryonic incomplete transition of human cells by chimeric OCT4/3 (POU5F1) with physiological co-activator EWS, Exp. Cell Res. 315 (2009) 2727–2740.
[22] M.F. Finley, S. Devata, J.E. Huettner, BMP-4 inhibits neural differentiation of murine embryonic stem cells, J. Neurobiol. 40 (1999) 271–287.
[23] Y. Yoshikawa, T. Fujimori, A.P. McMahon, S. Takada, Evidence that absence of Wnt-3a signaling promotes neuralization instead of paraxial mesoderm development in the mouse, Dev. Biol. 183 (1997) 234–242.
[24] D.A. Loebel, C.M. Watson, R.A. De Young, P.P. Tam, Lineage choice and differentiation in mouse embryos and embryonic stem cells, Dev. Biol. 264 (2003) 1–14.
[25] F. Lanner, K.L. Lee, M. Sohl, K. Holmborn, H. Yang, J. Wilbertz, L. Poellinger, J. Rossant, F. Farnebo, Heparan sulfation-dependent fibroblast growth factor signaling maintains embryonic stem cells primed for differentiation in a heterogeneous state, Stem Cells 28 (2010) 191–200.
[26] L. Haegele, B. Ingold, H. Naumann, G. Tabatabai, B. Ledermann, S. Brandner, Wnt signaling inhibits neural differentiation of embryonic stem cells by controlling bone morphogenetic protein expression, Mol. Cell Neurosci. 24 (2003) 696–708.
[27] J. Aubert, H. Dunstan, I. Chambers, A. Smith, Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation, Nat. Biotechnol. 20 (2002) 1240–1245.
[28] M. Denham, M. Dottori, Signals involved in neural differentiation of human embryonic stem cells, Neurosignals 17 (2009) 234–241.

Defining Cell Identity by Comprehensive Gene Expression Profiling

M. Toyoda[1], T. Hamatani[2], H. Okada[3], K. Matsumoto[4], H. Saito[4] and A. Umezawa*,[1]

Department of [1]Reproductive Biology and [4]Allergy and Immunology, National Institute for Child Health and Development, Tokyo, 157-8535, Japan; [2]Department of Obstetrics and Gynecology, Keio University School of Medicine, Tokyo, 160-8582, Japan; [3]Departments of Nephrology, Faculty of Medicine, Saitama Medical University, Saitama, 350-0451, Japan

Abstract: The human body is composed of 60 trillion cells, which have their origin in a fertilized egg. During development, the potential of a cell or tissue can be achieved by environmental manipulation. Then, what molecular determinants underlie or accompany the potential of the cells? To obtain a broader understanding of these problems, it is important to analyze all transcripts / genes in a wide selection of cell types. The development of microarray technologies, which allow us to undertake parallel analyses of many genes, has led to a new era in medical science. In this review, we show that the global expression data have clearly elucidated discernible major trends of the phenomenon in preimplantation development and epithelial-mesenchymal transition, and of the character of marrow stromal cells, which are attracting a great deal of attention as they represent a valuable source of cells for regenerative medicine. One of the interesting results is obtained from microarray data of marrow stromal cells: OP9 cells that have been recognized as a type of niche-constituting preadipocyte derived from marrow stroma, are found to be chondroblasts. We also describe what effect each type of expression data would bring to reproductive and regenerative medicine, as well as offering an excellent model of cell differentiation in biology.

Keywords: Gene chip array, cell potency, early embryogenesis, transdifferentiation, cellular plasticity, hyaline cartilage formation, endochondral ossification, bioinformatics.

INTRODUCTION

What type of "human" do you like? – do you judge by his appearance and/or his character? When the word "human" is replaced by "cell" in this question, the present situation surrounding regenerative medicine emerges. With the derivation of pluripotent embryonic stem (ES) and somatic stem cells that can differentiate into many different cell types, excitement has increased for the prospect of replacing dysfunctional or failing cells and organs. Somatic stem cells have been identified in hematopoietic [1], hepatic [2], epidermal [3], gastrointestinal [4], neural [5, 6], muscle [7], and bone marrow [6-8] tissues. Many researchers have since demonstrated the developmental pluripotency of these cells. What is not yet clear, however, are the critical molecular mechanisms that can harness or manipulate the potential of cells to foster therapeutic applications targeted to specific tissues.

Then, what are the appearance and the character of these cells? This question is: can the diverse morphology of the cell and/or the differential activities of genes provide the distinction between totipotent cells, pluripotent cells and terminally differentiated cells? One approach to this question is through markers that appear on the surfaces of cells by flow cytometric analysis. Hematopoietic stem cells (HSCs) are somatic stem cells found in the bone marrow and the precursor cells that give rise to all the types of both the myeloid and lymphoid lineages. This includes monocytes and macrophages, neutrophils, basophils, eosinophils, T cells, B cells, NK cells, erythrocytes, megakaryocytes, and dendritic cells. How do researchers find the desired cell populations and stem cells at a specific hierarchical stage? Multipotent HSCs present various clusters of differentiation markers on their surface: CD34, CD38, CD90, CD133, Lin, Thy1, and CD45. Understanding the cell lineage of HSCs will eventually allow the generation of expanded populations of HSCs *ex vivo* that can be used therapeutically.

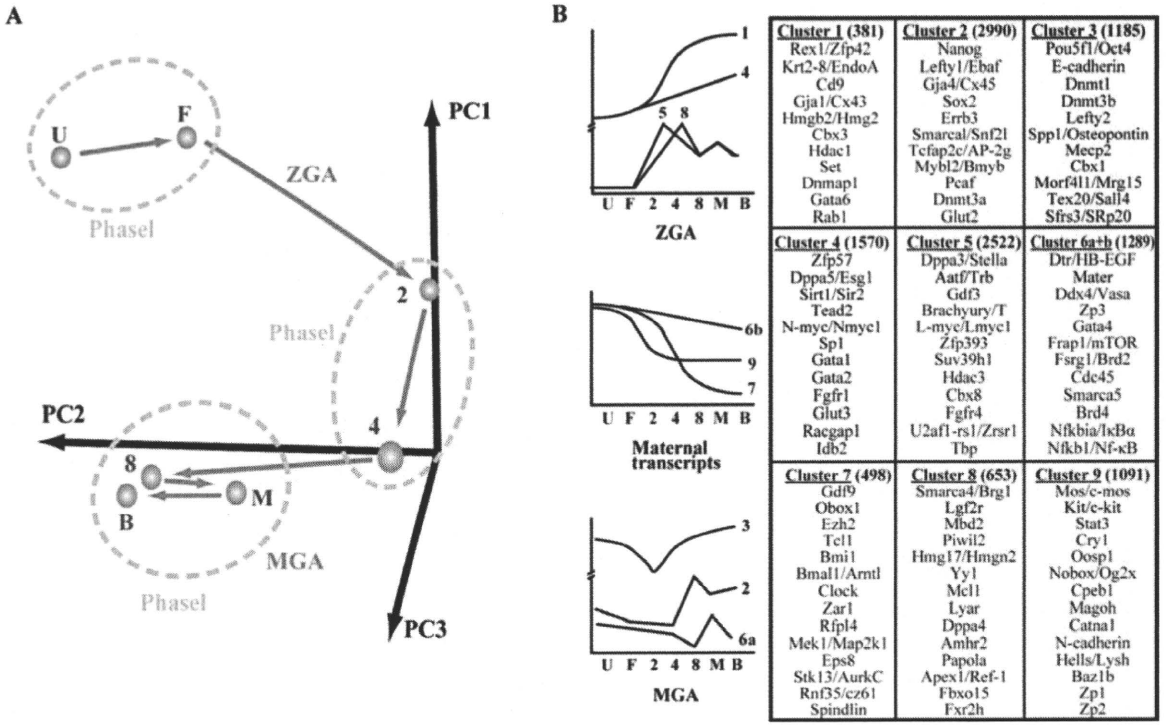
Another powerful approach to these questions is that of systematic genomic methodologies [9]. One of these methods, cDNA microarray/chip technology, is providing useful information [10-13]. Because of the logical connection between the function of a gene and its pattern of expression, the correlation of gene expression patterns with the variation in the phenotype of the cell can begin the process by which the function of a gene can be inferred. Similarly, the patterns of expression of known genes can reveal novel phenotypic aspects of the cell and tissues studied [14-16]. In this review we describe the use of microarray technology to determine cell identity based on gene expression pattern, with applications in regenerative medicine, especially preimplantation embryos, epithelial-mesenchymal transition and the mesenchymal stem cells.

THE BIG WAVE IN PREIMPLANTATION EMBRYO DEVELOPMENT

Preimplantation development encompasses the period from fertilization to implantation, and is marked by a number of critical sequential events. Understanding preimplantation development is important both for basic reproductive biology and for practical applications, including regenerative medicine and stock breeding. Preimplantation development is marked by 4 major events: the transition of maternal transcripts to zygotic transcripts, compaction, the first lineage differentiation into inner cell mass (ICM) and trophectoderm (TE), and implantation. The scarcity of the materials of preimplantation embryos, both in size (diameter <100 um) and in quantity (only a few to tens of oocytes from each ovulation), has hampered molecular analysis of preimplantation

*Address correspondence to this author at the Department of Reproductive Biology, National Institute for Child Health and Development, 2-10-1, Okura, Setagaya, Tokyo, 157-8535, Japan; Tel: +81-3-5494-7047; Fax: ????????????????; E-mail: umezawa@1985.jukuin.keio.ac.jp

0929-8673/10 $55.00+.00



Cluster 1 (381)	Cluster 2 (2990)	Cluster 3 (1185)
Rex1/Zfp42	Nanog	Pou5f1/Oct4
Krt2-8/EndoA	Lefty1/Ebaf	E-cadherin
Cd9	Gja4/Cx45	Dnmt1
Gja1/Cx43	Sox2	Dnmt3b
Hmgb2/Hmg2	Errb3	Lefty2
Cbx3	Smarcal/Snf2l	Spp1/Osteopontin
Hdac1	Tcfap2c/AP-2g	Mecp2
Set	Mybl2/Bmyb	Cbx1
Dnmap1	Pcaf	Morf4l1/Mrg15
Gata6	Dnmt3a	Tex20/Sall4
Rab1	Glut2	Sfrs3/SRp20
Cluster 4 (1570)	**Cluster 5 (2522)**	**Cluster 6a+b (1289)**
Zfp57	Dppa3/Stella	Dtr/HB-EGF
Dppa5/Esg1	Aatf/Trb	Mater
Sirt1/Sir2	Gdf3	Ddx4/Vasa
Tead2	Brachyury/T	Zp3
N-myc/Nmyc1	L-myc/Lmyc1	Gata4
Sp1	Zfp393	Frap1/mTOR
Gata1	Suv39h1	Fsrg1/Brd2
Gata2	Hdac3	Cdc45
Fgfr1	Cbx8	Smarca5
Glut3	Fgfr4	Brd4
Racgap1	U2af1-rs1/Zrsr1	Nfkbia/IκBα
Idb2	Tbp	Nfkb1/Nf-κB
Cluster 7 (498)	**Cluster 8 (653)**	**Cluster 9 (1091)**
Gdf9	Smarca4/Brg1	Mos/c-mos
Obox1	Lgf2r	Kit/c-kit
Ezh2	Mbd2	Stat3
Tcl1	Piwil2	Cry1
Bmi1	Hmg17/Hmgn2	Oosp1
Bmal1/Arntl	Yy1	Nobox/Og2x
Clock	Mcl1	Cpeb1
Zar1	Lyar	Magoh
Rfpl4	Dppa4	Catna1
Mek1/Map2k1	Amhr2	N-cadherin
Eps8	Papola	Hells/Lysh
Stk13/AurkC	Apex1/Ref-1	Baz1b
Rnf35/cz61	Fbxo15	Zp1
Spindlin	Fxr2h	Zp2

Fig. (1). Expression Profiling of Preimplantation Embryos.

A. Principal component analysis based on the expression data by gene chip array.

A matrix of scatter plots. U, F, 2, 4, 8, M, and B denote unfertilized egg, fertilized egg, 2-cell embryo, 4-cell embryo, 8-cell embryo, morula, and blastocyst, respectively. Each scatter plot shows the comparison of gene expression between embryo stages.

B. Expression changes of individual genes analyzed by the k-means non-hierarchical clustering method. Gene expression patterns can be assigned to three main groups. Group 1 (Cluster 1, 4, 5 and 8) appears to represent ZGA genes that are first activated from the zygotic genome. Group 2 (Cluster 7 and 9) represents maternal transcripts with distinctive patterns of degradation during preimplantation development. Group 3 (Cluster 2 and 3) appears to represent genes that follow a combination of these two patterns.

(Modified from Hamatani *et al.* Dev Cell, 2004, 6, 117 [18]).

embryos. Recent progress in RNA amplification methods using *in vitro* transcription and microarray platforms, including genes unique to preimplantation embryos, allows us to apply global gene expression profiling to the study of preimplantation embryos [17]. Hamatani *et al.* reported, for the first time, the global gene expression profiles of preimplantation embryos at all stages [18]. More than half of 21,939 gene features show statistically significant changes during preimplantation development. Pair-wise comparison, hierarchical clustering analysis, and principal component analysis (PCA) reveal two major transient waves of de novo transcription as follows: the first wave corresponds to zygotic genome activation (ZGA); the second wave, mid-preimplantation gene activation (MGA), contributes dramatic morphological changes during late preimplantation development (Fig. (**1A**)). Unsupervised methods such as principal component analysis (PCA) can transform the original features into new features (principal components (PC)), each PC representing a linear combination of the original features. PCA reduces input dimensionality by providing a subset of components that captures most of the information in the original data. For example, those genes that are highly correlated with the most informative PCs could be selected as classifier inputs, rather than a large dimension of original variables containing redundant features. To trace the expression changes of individual genes, statistically significant genes are analyzed by the k-means non-hierarchical clustering method. Gene expression patterns of these clusters can be assigned to three main groups (Fig. (**1B**)). The first group appears to represent ZGA genes that are first activated from the zygotic genome (Fig. (**1B**) Clusters 1, 4, 5, and 8). According to Gene Ontology (GO) terms [19] by MAPPFinder [20, 21], ZGA is suggested not to be promiscuous and to contribute mainly to the preparation of basic cellular machinery during the 2-cell and the 4-cell stages. The second group represents maternal transcripts with distinctive patterns of degradation during preimplantation development (Fig. (**1B**) Clusters 7 and 9). The third group appears to represent genes that follow a combination of these two patterns (Fig. (**1B**) Clusters 2 and 3). Genes whose expression first significantly increases from the 4-cell to 8-cell stage are identified as MGA genes, of which there are 4,216. The functional assignment of these genes by MAPPFinder characterizes the function of the MGA genes by the following three representative GO terms: "endopeptidase inhibitor," "intercellular junction," and "DNA (cytosine-5-)-methyltransferase." The implication of these GO terms and the timing of MGA seems consistent with the proposed role of MGA in compaction, cavitation, and the first differentiation of ICM and TE. Expression profiling of embryos treated

with inhibitors of transcription and translation reveals that the translation of maternal RNAs is required for the initiation of ZGA, suggesting a cascade of gene activation from maternal RNA/protein sets to ZGA gene sets and thence to MGA gene sets.

Decreasing oocyte competence with maternal aging is a major concern in human infertility because the rate of late childbearing is increasing even though reproductive capacity in women declines dramatically with advancing age. Studies of molecular mechanisms involved in the decline of oocyte quality with maternal age could have important implications for the efficacy and safety of clinical ooplasmic donation to rejuvenate aging oocytes. Hamatani *et al.* and Pan *et al.* also reported age-associated alteration of gene expression patterns in mouse oocytes, which has implications for aging research [22, 23]. Genes related to oxidative stress (e.g., Sod1, Apacd and Txn1), mitochondrial function (e.g., Sdha, Pdhb and Cyb5), chromatin structure (e.g., Hdac2, Hmgb3 and Bmi1), DNA methylation (e.g., Dnmt1, Dnmt3b, and Dnmt3L), and genome stability (e.g., Tert, Exo1, and Msh3) are altered with aging. Furthermore, kinetochore components of the spindle assembly checkpoint (e.g., Bub1, BubR1, Aurora kinase) and Cdc20, a critical activator of the Anaphase Promoting Complex, may contribute to aneuploidy in aged oocytes [23].

These comprehensive expression profiles of the majority of genes should give a baseline for analysis of the complex gene regulatory networks in normal mouse preimplantation and for comparative analysis for other mammalian species, including humans.

WHAT'S GOING ON IN AN EPITHELIAL-MESENCHYMAL TRANSITION (EMT)?

The conversion of an epithelial cell to a mesenchymal cell is critical to vertebrate embryogenesis and a defining structural feature of organ development, such as forming fibroblasts in injured tissues [24, 25], or in initiating metastases in epithelial cancer [26-29]. From a general perspective, EMT is about disaggregating epithelial units and reshaping epithelia for movement. Epithelia in transition lose polarity, tight junctions, adherens junctions, desmosomes and cytokeratin intermediate filaments in order to rearrange their F-actin stress fibers and express lamellopodia.and filopodia. This phenotypic conversion requires the molecular reprogramming of epithelia with new biochemical instructions. It is known that commonly used molecular markers for EMT include increased expression of N-cadherin and vimentin, nuclear localization of beta-catenin, and increased production of the transcription factors such as Snail, Twist, and SIP1/ZEB2. Much of this conversion, however, has been studied during experiments that expose new transduction and signaling pathways in epithelia, and more recently in fibrogenic tissues. It is not yet clear whether the fibroblast transition of EMT is an expected middle phase of transdifferentiating epithelia, or whether EMT producing fibroblasts is an arrested form of transdifferentiation.

EMT is easily engaged by a combination of cytokines associated with proteolytic digestion of basement membranes upon which epithelia reside. We analyzed PCA and hierarchical clustering method of the gene expression pattern of the renal tubular cells and mammary gland cells. If PC1 were used to identify genes that are differentially expressed between phenotypes, then genes that are strongly associated with PC1 would be selected. If both PC axes are used, then genes strongly associated with two groups would be selected. We then identified the genes which discriminate between the renal tubular and the mammary gland epithelial cells (PC1), or EMT-induced and non-induced cells (PC3) (Fig. (**2**)). Undergoing EMT identifies the genes that discriminate between the renal tubular and the mammary gland epithelial cells (PC1), or EMT-induced and non-induced cells (PC3) (Table **1**).

The advanced study of the genes identified by PCA would yield new insight regarding EMT, and achieve a breakthrough in understanding the molecular mechanisms of

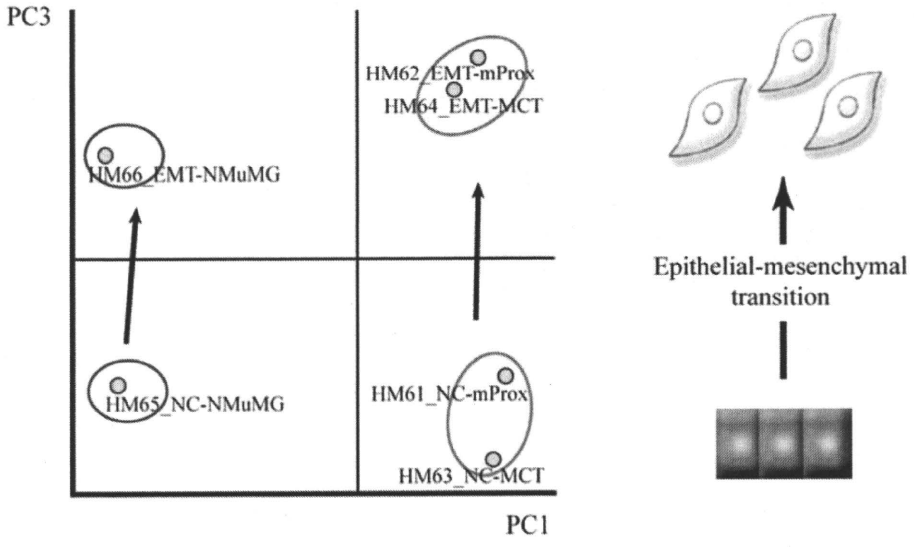


Fig. (2). EMT related genes by gene chip analysis.

Principal component analysis based on the gene expression pattern of the renal tubular cells and mammary gland cells.

Table 1. List of Genes that Had Up-Regulated (Positive) and Down-Regulated (Negative) Expression Related to EMT (PC3 Axis)

Positive			Negative		
Gene Symbol	**Gene Title**	**Representative Public ID**	**Gene Symbol**	**Gene Title**	**Representative Public ID**
Ccng2	cyclin G2	U95826	Bach1	BTB and CNC homology 1	NM_007520
Ccni	cyclin I	NM_017367	Cdc42bpa	Cdc42 binding protein kinase alpha	BM117074
Ctgf	connective tissue growth factor	NM_010217	Dnm1	dynamin 1	L31397
Dock7	Dedicator of cytokinesis 7	BB463580	Foxb1	forkhead box B1	U90538
Dok1	docking protein 1	BC006868	Gprc5c	G protein-coupled receptor, family C, group 5, member C	BC008228
Fgfr1	Fibroblast growth factor receptor 1	M33760	Il13ra1	interleukin 13 receptor, alpha 1	S80963
Gja1	gap junction membrane channel protein alpha 1	BB039269	Kcnk5	potassium channel, subfamily K, member 5	AF319542
Gtpbp4	GTP binding protein 4	AI987834	Kif13a	kinesin family member 13A	AB037923
Hgfac	hepatocyte growth factor activator	NM_019447	Kif17	kinesin family member 17	AW492270
Hoxa3	homeo box A3	BB496114	Mark2	MAP/microtubule affinity-regulating kinase 2	BI686265
Hoxb8 /b7	homeo box B8 / homeo box B7	X13721	Mef2d	myocyte enhancer factor 2D	NM_133665
Il15ra	interleukin 15 receptor, alpha chain	NM_008358	Mrpl51	mitochondrial ribosomal protein L51	AI594880
Irx2	Iroquois related homeobox 2	AF295369	Mxd4	Max dimerization protein 4	BE291523
Itga5	integrin alpha 5	BB493533	Neu1	neuraminidase 1	AI649303
Itgb1	integrin beta 1	BM120341	Rel	reticuloendotheliosis oncogene	NM_009044
Mapkbp 1	Mitogen activated protein kinase binding protein 1	BQ174980	Rgnef	Rho-guanine nucleotide exchange factor	BG069493
Mdm2	transformed mouse 3T3 cell double minute 2	X58876	Rps6kb2	ribosomal protein S6 kinase, polypeptide 2	NM_021485
Ncam1	neural cell adhesion molecule 1	NM_010875	Slc24a1	solute carrier family 24, member 1	BC016094
Pdgfa	platelet derived growth factor, alpha	BB371842	Slc25a19	solute carrier family 25, member 19	AV338420
Prkcc	protein kinase C, gamma	NM_011102	Slc25a22	solute carrier family 25 , member 22	AK018760
Rab23	RAB23, member RAS oncogene family	NM_008999	Slc40a1	solute carrier family 40 , member 1	AF226613
Rasa3	RAS p21 protein activator 3	NM_009025	Stat1	signal transducer and activator of transcription 1	AW214029
Rb1	retinoblastoma 1	NM_009029	Tgfa	transforming growth factor alpha	M92420
Sbno1	Sno, strawberry notch homolog 1	BC023136	Ubp1	upstream binding protein 1	NM_013699
Slc1a4	solute carrier family 1, member 4	BB277461	Usp12	ubiquitin specific protease 12	AF441835
Slc34a1	solute carrier family 34 , member 1	AI788646			
Slc4a7	Solute carrier family 4, member 7	AW555750			
Slc7a2	solute carrier family 7, member 2	M62838			
Ube1y1	ubiquitin-activating enzyme E1, Chr Y 1	X62581			
Vegfa	vascular endothelial growth factor A	NM_009505			
Wnt6	wingless-related MMTV integration site 6	NM_009526			
Wnt7b	wingless-related MMTV integration site 7B	W29605			

drug delivery for specific anti-cancer drugs, especially those affecting metastasis. Progress in understanding EMT has been an exercise in coming to appreciate the level of complexity required for changing cellular identity. The mecha-

nism of EMT highlights an integration of nuclear regulation and network signaling with alterations in the microenvironment to create a moving cell; in this sense, basic concepts based on EMT mechanisms would thus hold great promise for regenerative medicine.

GET OP9 CELLS DEAD TO RIGHTS

The concept of regenerative medicine refers to the cell-mediated restoration of damaged or diseased tissue. Candidate cell sources for tissue regeneration include ES cells, fetal cells, and adult cells such as marrow stromal cells, each of which has both advantages and drawbacks. Clinical trials with marrow stromal cells have been performed in patients with osteogenesis imperfecta and osteoporosis, and marrow stromal cells are expected to be a good source of cell therapy.

Bone marrow-derived stem cells can be transdifferentiated into multilineage cells, such as muscle [30] from mesoderm, lung [31] and liver [31, 32] from endoderm, and brain [33-36] and skin [31] from ectoderm. Somatic stem cells are more desirable than ES cells for cell therapeutics because of ethical considerations and the possible immunologic rejection of ES cells. Mesenchymal stem cells have become the most popular somatic stem cells in medicine and biology, not least because of their high reproductive capability *in vitro*.

Chondrocytes differentiate from mesenchymal cells during embryonic development [37], and the phenotype of the differentiated chondrocyte is characterized by the synthesis, deposition, and maintenance of cartilage-specific extracellular matrix molecules, including type II collagen and aggrecan [38-40]. The phenotype of differentiated chondrocytes is rapidly lost since it is unstable in culture [41-44]. This process is referred to as 'dedifferentiation' and is a major impediment to the use of mass cell populations for therapy or tissue engineering of damaged cartilage. When isolated chondrocytes are cultured in a monolayer at low density, the typical round chondrocytes morphologically transform into flattened fibroblast-like cells, with profound changes in biochemical and genetic characteristics, including reduced synthesis of type II collagen and cartilage proteins [45].

We established several stromal cells from murine bone marrow cultures [46]. One of them, KUSA-A1 cells, displays osteogenetic characteristics *in vitro* and *in vivo*. In order to clarify the specific gene expression profile of KUSA-A1, other established stromal cells, KUM5, 9-15c, KUSA-O, H-1/A [47], and mouse embryonic fibroblasts, we compared the expression levels of approximately 23,000 genes by using the Affymetrix gene chip oligonucleotide arrays. Of the 23,000 genes represented on the gene chip, chondrocyte-specific or -associated genes such as type II collagen α1, type XI collagen α1, Sox9, proline arginine-rich end leucine-rich repeat, and cartilage oligomeric matrix protein are more strongly expressed in KUM5 cells than in other marrow-derived mesenchymal cells. Does a gene expression pattern reflect the character of the cells *in vitro* and/or *in vivo*? – The answer is yes: KUM5 cells generate hyaline cartilage and exhibit endochondral ossification *in vitro* (Fig. (**3B**)) and *in vivo* [48].

Surprisingly, OP9 cells [49] also express these chondrocyte-specific or -associated genes at higher levels: the type II collagen α1, and type XI collagen α1 genes are expressed in OP9 cells at more than 10-fold higher levels than in 9-15c, KUSA-O, H-1/A, primary embryonic fibroblasts, or even KUM5 chondroblasts. In addition, expression of 'structural proteins' on Gene Ontology, including the extracellular matrix, is much higher by OP9 and KUM5 cells than by non-chondrogenic cells such as KUSA-A1, H-1/A, and 9-15c, implying that the OP9 and KUM5 cells are mainly engaged in synthesizing extracellular matrix. We also performed hierarchical clustering and PCA, based on the microarray data (Fig. (**3A**)). KUM5 and OP9 cells are grouped into the same subcategory and can clearly be separated from other stromal cells based on the expression data of cell surface markers and cell-type-specific genes, implying that KUM5 and OP9 cells have chondrogenic potential.

Are OP9 cells chondroblasts *in vitro* and/or *in vivo*? - the answer, again, is yes; OP9 cells are induced into the chondrogenic lineage by the pellet culture method (Fig. (**3C**)), and the OP9 pellets (micromasses) implanted in mice form the type II collagen-positive hyaline cartilage [48]. OP9 cells are derived from macrophage colony-stimulating factor-deficient osteopetrotic mice, and have also been used as feeder cells for embryonic stem cells [50-52]. The cells identified as a key participant in regulating the number of adult stem cells or hematopoietic stem cells are now considered to be of an osteoblastic lineage [53, 54]. OP9 cells have been recognized as a niche-constituting preadipocyte; however their true face is a chondroblast. We have two different types of cells, osteoblasts (KUSA-A1) and chondroblasts (OP9 and KUM5), showing distinctive *in vivo* characteristics. The unique characteristics of these cells provide an opportunity to analyze the process of membranous ossification and endochondral ossification. These cells are useful candidate cell sources, in addition to dedifferentiated chondrocytes obtained from cartilage for transplantation in osteoarthritis and rheumatoid arthritis.

GENE EXPRESSION PROFILING AND MEDICAL SCIENCE

Recently, gene expression profiling has been successfully used to predict outcomes in some types of malignant diseases [55-61] and, additionally, to assess drug discovery screening [62]. In reproductive and regenerative medicine, it is important to identify biomarkers that will establish the isolation, selection and expansion of stem cells *in vitro* to allow their use for cell therapy. On the road map for translational medicine-- often referred to as bench to bedside research--, stem cell therapy is a prime destination. Stem cells have not taken on the identity of any specific cell type and are not yet committed to any dedicated function; they can divide extensively or indefinitely, and may be induced to give rise to one or more specialized cell types. Stem cells derived from bone marrow can replace heart muscle lost as a result of a heart attack, and can improve cardiac function. Injecting bone-marrow stem cells into an injured heart potentially represented new therapy, triggering the launch of numerous clinical studies to investigate the effect of directly injecting these cells into the damaged heart muscle of patients following a

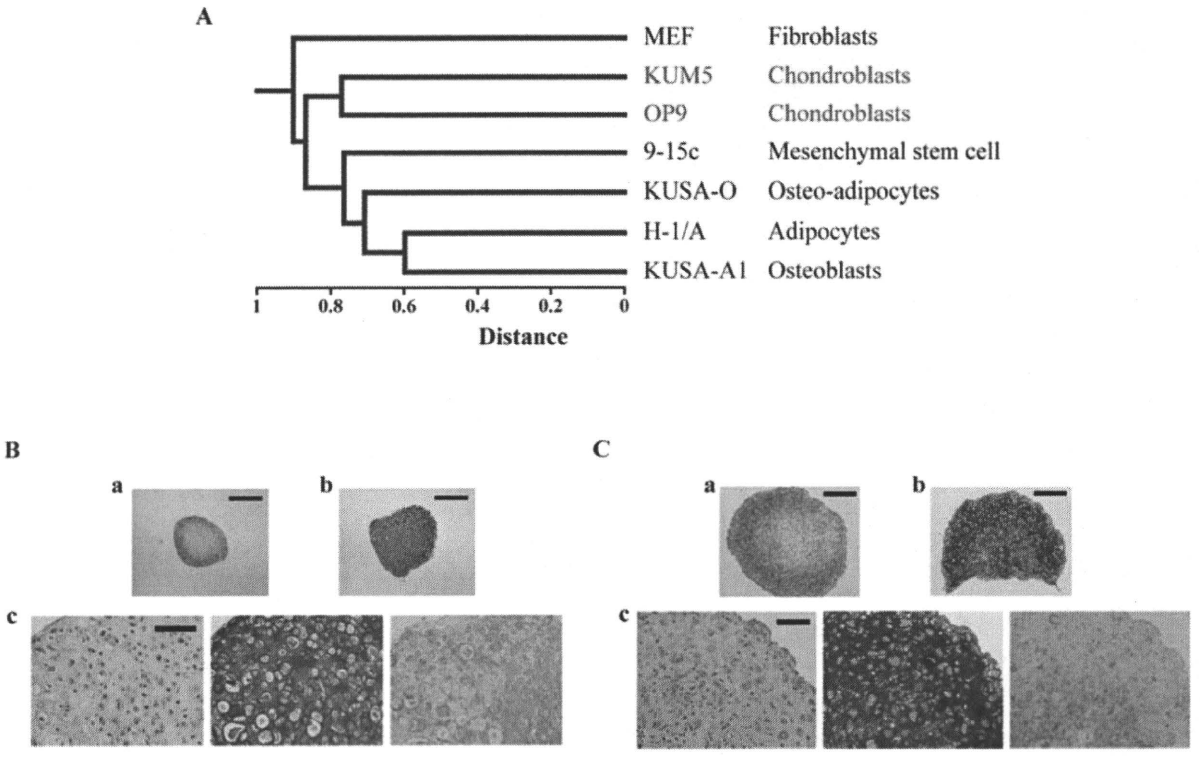


Fig. (3). Expression profiling and *in vitro* chondrogenesis of KUM5 and OP9 cells.

A. Dendrogram revealing clustering profile of six marrow stromal cells and mouse embryonic fibroblast (MEF) using 244 surface marker genes.

B. a, b: Toluidine blue stained section of KUM5 chondrogenic nodules in the pellet culture exposed TGF-β3 and BMP-2 for 1 (a) or 3 (b) weeks. Scale bars: 500 μm. c: Higher magnification of KUM5 chondrogenic pellet exposed to TGF-β3 and BMP-2 for 3 weeks. Left panel: hematoxylin and eosin stain; center panel: toluidine blue stain; right panel: alcian blue stain. Scale bars: 100 μm.

C. a, b: Toluidine blue stained section of OP9 chondrogenic nodules in the pellet culture exposed TGF-β3 and BMP-2 for 1 (a) or 3 (b) weeks. Scale bars: 500 μm. c: Higher magnification of OP9 chondrogenic pellet exposed to TGF-β3 and BMP-2 for 3 weeks. Left panel: hematoxylin and eosin stain; center panel: toluidine blue stain; right panel: alcian blue stain. Scale bars: 100 μm.

(Modified from Sugiki *et al.* J Cell Biochem, 2007, 100, 1240 [48]).

heart attack [63]. The scientific underpinnings of the ongoing human studies have been established. Now is the time to search for the presence of naturally occurring, authentic pluripotent cells and to identify and dissect the signals that guide their migration, self-renewal and differentiation. Furthermore, we need to commit the necessary time and resources to identify the best stem cells for cell therapy to translate.

CONCLUSION

Here, the expression pattern has been correlated with molecular structure descriptor; this consistency indicates that the expression profiling is valid. Consequently, understanding the global gene network that governs the pluripotency and self-renewal of stem cells is an important first step towards the experimental manipulation of cellular developmental potency. The cell potency is a fundamental concept in developmental biology and stem cell biology, providing a conceptual framework of sequential transition from totipotent fertilized eggs to pluripotent embryonic stem cells and stem cells to terminal differentiated cells. The global expression profiling can help to delineate the global architecture and dynamics of a gene regulatory network such as Oct4-regulated gene networks in mouse ES cells [64].

ACKNOWLEDGEMENTS

The authors wish to thank Yoriko Takahashi for helpful discussions, and Kayoko Saito for secretarial assistance.

This study was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan; the Ministry of Health, Labour and Welfare Sciences Research Grants; by a Research Grant on Health Science focusing on Drug Innovation from the Japan Health Science Foundation; by the program for promotion of Fundamental Studies in Health Science of the Pharmaceuticals and Medical Devices Agency; by a Research Grant for Cardiovascular Disease from the ministry of Health, Labour and Welfare; and by a Grant for Child Health and Development from the Ministry of Health, Labour and Welfare.

FOOTNOTE

Fig. (**1**) is prepared from ref. [18] with permission from Elsevier.

Fig. (**3**) is prepared from ref. [48] with permission from Wiley-Liss, Inc.

REFERENCES

[1] Bryder, D.; Rossi, D.J.; Weissman, I.L. Hematopoietic stem cells: the paradigmatic tissue-specific stem cell. *Am. J. Pathol.,* **2006**, *169*, 338-346.

[2] Alison, M.; Sarraf, C. Hepatic stem cells. *J. Hepatol.,* **1998**, *29*, 676-682.

[3] Watt, F.M. Epidermal stem cells: markers, patterning and the control of stem cell fate. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.,* **1998**, *353*, 831-837.

[4] Potten, C.S. Stem cells in gastrointestinal epithelium: numbers, characteristics and death. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.,* **1998**, *353*, 821-830.

[5] Gage, F.H. Mammalian neural stem cells. *Science,* **2000**, *287*, 1433-1438.

[6] Jiang, Y.; Vaessen, B.; Lenvik, T.; Blackstad, M.; Reyes, M.; Verfaillie, C.M. Multipotent progenitor cells can be isolated from postnatal murine bone marrow, muscle, and brain. *Exp. Hematol.,* **2002**, *30*, 896-904.

[7] Jiang, Y.; Jahagirdar, B.N.; Reinhardt, R.L.; Schwartz, R.E.; Keene, C.D.; Ortiz-Gonzalez, X.R.; Reyes, M.; Lenvik, T.; Lund, T.; Blackstad, M.; Du, J.; Aldrich, S.; Lisberg, A.; Low, W.C.; Largaespada, D.A.; Verfaillie, C.M. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature,* **2002**, *418*, 41-49.

[8] Fridenshtein, A. [Stromal bone marrow cells and the hematopoietic microenvironment]. *Arkh. Patol.,* **1982**, *44*, 3-11.

[9] Ko, M.S. Embryogenomics: developmental biology meets genomics. *Trends Biotechnol.,* **2001**, *19*, 511-518.

[10] Ivanova, N.B.; Dimos, J.T.; Schaniel, C.; Hackney, J.A.; Moore, K.A.; Lemischka, I.R. A stem cell molecular signature. *Science,* **2002**, *298*, 601-604.

[11] Ramalho-Santos, M.; Yoon, S.; Matsuzaki, Y.; Mulligan, R.C.; Melton, D.A. "Stemness": transcriptional profiling of embryonic and adult stem cells. *Science,* **2002**, *298*, 597-600.

[12] Schena, M.; Shalon, D.; Davis, R.W.; Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science,* **1995**, *270*, 467-470.

[13] Tanaka, T.S.; Kunath, T.; Kimber, W.L.; Jaradat, S.A.; Stagg, C.A.; Usuda, M.; Yokota, T.; Niwa, H.; Rossant, J.; Ko, M.S. Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. *Genome Res.,* **2002**, *12*, 1921-1928.

[14] Brown, P.O.; Botstein, D. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.,* **1999**, *21*, 33-37.

[15] DeRisi, J.L.; Iyer, V.R.; Brown, P.O. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science,* **1997**, *278*, 680-686.

[16] Iyer, V.R.; Eisen, M.B.; Ross, D.T.; Schuler, G.; Moore, T.; Lee, J.C.; Trent, J.M.; Staudt, L.M.; Hudson, J., Jr.; Boguski, M.S.; Lashkari, D.; Shalon, D.; Botstein, D.; Brown, P.O. The transcriptional program in the response of human fibroblasts to serum. *Science,* **1999**, *283*, 83-87.

[17] Carter, M.G.; Hamatani, T.; Sharov, A.A.; Carmack, C.E.; Qian, Y.; Aiba, K.; Ko, N.T.; Dudekula, D.B.; Brzoska, P.M.; Hwang, S.S.; Ko, M.S. In situ-synthesized novel microarray optimized for mouse stem cell and early developmental expression profiling. *Genome Res.,* **2003**, *13*, 1011-1021.

[18] Hamatani, T.; Carter, M.G.; Sharov, A.A.; Ko, M.S. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell,* **2004**, *6*, 117-131.

[19] Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; Harris, M.A.; Hill, D.P.; Issel-Tarver, L.; Kasarskis, A.; Lewis, S.; Matese, J.C.; Richardson, J.E.; Ringwald, M.; Rubin, G.M.; Sherlock, G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.,* **2000**, *25*, 25-29.

[20] Dahlquist, K.D.; Salomonis, N.; Vranizan, K.; Lawlor, S.C.; Conklin, B.R. GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. *Nat. Genet.,* **2002**, *31*, 19-20.

[21] Doniger, S.W.; Salomonis, N.; Dahlquist, K.D.; Vranizan, K.; Lawlor, S.C.; Conklin, B.R. MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biol,* **2003**, *4*, R7.

[22] Hamatani, T.; Falco, G.; Carter, M.G.; Akutsu, H.; Stagg, C.A.; Sharov, A.A.; Dudekula, D.B.; VanBuren, V.; Ko, M.S. Age-associated alteration of gene expression patterns in mouse oocytes. *Hum. Mol. Genet.,* **2004**, *13*, 2263-2278.

[23] Pan, H.; Ma, P.; Zhu, W.; Schultz, R.M. Age-associated increase in aneuploidy and changes in gene expression in mouse eggs. *Dev. Biol.,* **2008**, *316*, 397-407.

[24] Strutz, F.; Okada, H.; Lo, C.W.; Danoff, T.; Carone, R.L.; Tomaszewski, J.E.; Neilson, E.G. Identification and characterization of a fibroblast marker: FSP1. *J. Cell Biol.,* **1995**, *130*, 393-405.

[25] Iwano, M.; Plieth, D.; Danoff, T.M.; Xue, C.; Okada, H.; Neilson, E.G. Evidence that fibroblasts derive from epithelium during tissue fibrosis. *J. Clin. Invest.,* **2002**, *110*, 341-350.

[26] Janda, E.; Lehmann, K.; Killisch, I.; Jechlinger, M.; Herzig, M.; Downward, J.; Beug, H.; Grunert, S. Ras and TGF[beta] cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. *J. Cell Biol.,* **2002**, *156*, 299-313.

[27] Kiemer, A.K.; Takeuchi, K.; Quinlan, M.P. Identification of genes involved in epithelial-mesenchymal transition and tumor progression. *Oncogene,* **2001**, *20*, 6679-6688.

[28] Vincent-Salomon, A.; Thiery, J.P. Host microenvironment in breast cancer development: epithelial-mesenchymal transition in breast cancer development. *Breast Cancer Res,* **2003**, *5*, 101-106.

[29] Xue, C.; Plieth, D.; Venkov, C.; Xu, C.; Neilson, E.G. The gatekeeper effect of epithelial-mesenchymal transition regulates the frequency of breast cancer metastasis. *Cancer Res.,* **2003**, *63*, 3386-3394.

[30] Ferrari, G.; Cusella-De Angelis, G.; Coletta, M.; Paolucci, E.; Stornaiuolo, A.; Cossu, G.; Mavilio, F. Muscle regeneration by bone marrow-derived myogenic progenitors. *Science,* **1998**, *279*, 1528-1530.

[31] Krause, D.S.; Theise, N.D.; Collector, M.I.; Henegariu, O.; Hwang, S.; Gardner, R.; Neutzel, S.; Sharkis, S.J. Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell,* **2001**, *105*, 369-377.

[32] Petersen, B.E.; Bowen, W.C.; Patrene, K.D.; Mars, W.M.; Sullivan, A.K.; Murase, N.; Boggs, S.S.; Greenberger, J.S.; Goff, J.P. Bone marrow as a potential source of hepatic oval cells. *Science,* **1999**, *284*, 1168-1170.

[33] Sanchez-Ramos, J.; Song, S.; Cardozo-Pelaez, F.; Hazzi, C.; Stedeford, T.; Willing, A.; Freeman, T.B.; Saporta, S.; Janssen, W.; Patel, N.; Cooper, D.R.; Sanberg, P.R. Adult bone marrow stromal cells differentiate into neural cells *in vitro*. *Exp. Neurol.,* **2000**, *164*, 247-256.

[34] Woodbury, D.; Schwarz, E.J.; Prockop, D.J.; Black, I.B. Adult rat and human bone marrow stromal cells differentiate into neurons. *J. Neurosci. Res.,* **2000**, *61*, 364-370.

[35] Kohyama, J.; Abe, H.; Shimazaki, T.; Koizumi, A.; Nakashima, K.; Gojo, S.; Taga, T.; Okano, H.; Hata, J.; Umezawa, A. Brain from bone: efficient "meta-differentiation" of marrow stroma-derived mature osteoblasts to neurons with Noggin or a demethylating agent. *Differentiation,* **2001**, *68*, 235-244.

[36] Kim, B.J.; Seo, J.H.; Bubien, J.K.; Oh, Y.S. Differentiation of adult bone marrow stem cells into neuroprogenitor cells *in vitro*. *Neuroreport,* **2002**, *13*, 1185-1188.

[37] Solursh, M. Differentiation of cartilage and bone. *Curr. Opin. Cell Biol.,* **1989**, *1*, 989-994.

[38] Hauselmann, H.J.; Fernandes, R.J.; Mok, S.S.; Schmid, T.M.; Block, J.A.; Aydelotte, M.B.; Kuettner, K.E.; Thonar, E.J. Phenotypic stability of bovine articular chondrocytes after long-term culture in alginate beads. *J. Cell Sci.,* **1994**, *107 (Pt 1)*, 17-27.

[39] Reginato, A.M.; Iozzo, R.V.; Jimenez, S.A. Formation of nodular structures resembling mature articular cartilage in long-term primary cultures of human fetal epiphyseal chondrocytes on a hydrogel substrate. *Arthritis Rheum.,* **1994**, *37*, 1338-1349.

[40] Archer, C.W.; McDowell, J.; Bayliss, M.T.; Stephens, M.D.; Bentley, G. Phenotypic modulation in sub-populations of human articular chondrocytes *in vitro*. *J. Cell Sci.,* **1990**, *97 (Pt 2)*, 361-

371.

[41] Benya, P.D.; Padilla, S.R.; Nimni, M.E. Independent regulation of collagen types by chondrocytes during the loss of differentiated function in culture. *Cell,* **1978**, *15*, 1313-1321.

[42] Benya, P.D.; Shaffer, J.D. Dedifferentiated chondrocytes reexpress the differentiated collagen phenotype when cultured in agarose gels. *Cell,* **1982**, *30*, 215-224.

[43] Lefebvre, V.; Peeters-Joris, C.; Vaes, G. Production of collagens, collagenase and collagenase inhibitor during the dedifferentiation of articular chondrocytes by serial subcultures. *Biochim. Biophys. Acta,* **1990**, *1051*, 266-275.

[44] Bonaventure, J.; Kadhom, N.; Cohen-Solal, L.; Ng, K.H.; Bourguignon, J.; Lasselin, C.; Freisinger, P. Reexpression of cartilage-specific genes by dedifferentiated human articular chondrocytes cultured in alginate beads. *Exp. Cell Res.,* **1994**, *212*, 97-104.

[45] Yoon, Y.M.; Kim, S.J.; Oh, C.D.; Ju, J.W.; Song, W.K.; Yoo, Y.J.; Huh, T.L.; Chun, J.S. Maintenance of differentiated phenotype of articular chondrocytes by protein kinase C and extracellular signal-regulated protein kinase. *J. Biol. Chem.,* **2002**, *277*, 8412-8420.

[46] Matsumoto, S.; Shibuya, I.; Kusakari, S.; Segawa, K.; Uyama, T.; Shimada, A.; Umezawa, A. Membranous osteogenesis system modeled with KUSA-A1 mature osteoblasts. *Biochim. Biophys. Acta,* **2005**, *1725*, 57-63.

[47] Umezawa, A.; Maruyama, T.; Segawa, K.; Shadduck, R.K.; Waheed, A.; Hata, J. Multipotent marrow stromal cell line is able to induce hematopoiesis *in vivo*. *J. Cell. Physiol.,* **1992**, *151*, 197-205.

[48] Sugiki, T.; Uyama, T.; Toyoda, M.; Morioka, H.; Kume, S.; Miyado, K.; Matsumoto, K.; Saito, H.; Tsumaki, N.; Takahashi, Y.; Toyama, Y.; Umezawa, A. Hyaline cartilage formation and enchondral ossification modeled with KUM5 and OP9 chondroblasts. *J. Cell. Biochem.,* **2007**, *100*, 1240-1254.

[49] Nakano, T.; Kodama, H.; Honjo, T. *In vitro* development of primitive and definitive erythrocytes from different precursors. *Science,* **1996**, *272*, 722-724.

[50] Kodama, H.; Nose, M.; Niida, S.; Nishikawa, S. Involvement of the c-kit receptor in the adhesion of hematopoietic stem cells to stromal cells. *Exp. Hematol.,* **1994**, *22*, 979-984.

[51] Nakano, T.; Kodama, H.; Honjo, T. Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science,* **1994**, *265*, 1098-1101.

[52] Nakano, T. Lymphohematopoietic development from embryonic stem cells *in vitro*. *Semin. Immunol.,* **1995**, *7*, 197-203.

[53] Calvi, L.M.; Adams, G.B.; Weibrecht, K.W.; Weber, J.M.; Olson, D.P.; Knight, M.C.; Martin, R.P.; Schipani, E.; Divieti, P.; Bringhurst, F.R.; Milner, L.A.; Kronenberg, H.M.; Scadden, D.T. Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature,* **2003**, *425*, 841-846.

[54] Zhang, J.; Niu, C.; Ye, L.; Huang, H.; He, X.; Tong, W.G.; Ross, J.; Haug, J.; Johnson, T.; Feng, J.Q.; Harris, S.; Wiedemann, L.M.; Mishina, Y.; Li, L. Identification of the haematopoietic stem cell niche and control of the niche size. *Nature,* **2003**, *425*, 836-841.

[55] Alizadeh, A.A.; Eisen, M.B.; Davis, R.E.; Ma, C.; Lossos, I.S.; Rosenwald, A.; Boldrick, J.C.; Sabet, H.; Tran, T.; Yu, X.; Powell, J.I.; Yang, L.; Marti, G.E.; Moore, T.; Hudson, J., Jr.; Lu, L.; Lewis, D.B.; Tibshirani, R.; Sherlock, G.; Chan, W.C.; Greiner, T.C.; Weisenburger, D.D.; Armitage, J.O.; Warnke, R.; Levy, R.; Wilson, W.; Grever, M.R.; Byrd, J.C.; Botstein, D.; Brown, P.O.; Staudt, L.M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature,* **2000**, *403*, 503-511.

[56] Beer, D.G.; Kardia, S.L.; Huang, C.C.; Giordano, T.J.; Levin, A.M.; Misek, D.E.; Lin, L.; Chen, G.; Gharib, T.G.; Thomas, D.G.; Lizyness, M.L.; Kuick, R.; Hayasaka, S.; Taylor, J.M.; Iannettoni, M.D.; Orringer, M.B.; Hanash, S. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat. Med.,* **2002**, *8*, 816-824.

[57] Pomeroy, S.L.; Tamayo, P.; Gaasenbeek, M.; Sturla, L.M.; Angelo, M.; McLaughlin, M.E.; Kim, J.Y.; Goumnerova, L.C.; Black, P.M.; Lau, C.; Allen, J.C.; Zagzag, D.; Olson, J.M.; Curran, T.; Wetmore, C.; Biegel, J.A.; Poggio, T.; Mukherjee, S.; Rifkin, R.; Califano, A.; Stolovitzky, G.; Louis, D.N.; Mesirov, J.P.; Lander, E.S.; Golub, T.R. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature,* **2002**, *415*, 436-442.

[58] Rosenwald, A.; Wright, G.; Chan, W.C.; Connors, J.M.; Campo, E.; Fisher, R.I.; Gascoyne, R.D.; Muller-Hermelink, H.K.; Smeland, E.B.; Giltnane, J.M.; Hurt, E.M.; Zhao, H.; Averett, L.; Yang, L.; Wilson, W.H.; Jaffe, E.S.; Simon, R.; Klausner, R.D.; Powell, J.; Duffey, P.L.; Longo, D.L.; Greiner, T.C.; Weisenburger, D.D.; Sanger, W.G.; Dave, B.J.; Lynch, J.C.; Vose, J.; Armitage, J.O.; Montserrat, E.; Lopez-Guillermo, A.; Grogan, T.M.; Miller, T.P.; LeBlanc, M.; Ott, G.; Kvaloy, S.; Delabie, J.; Holte, H.; Krajci, P.; Stokke, T.; Staudt, L.M. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.,* **2002**, *346*, 1937-1947.

[59] Shipp, M.A.; Ross, K.N.; Tamayo, P.; Weng, A.P.; Kutok, J.L.; Aguiar, R.C.; Gaasenbeek, M.; Angelo, M.; Reich, M.; Pinkus, G.S.; Ray, T.S.; Koval, M.A.; Last, K.W.; Norton, A.; Lister, T.A.; Mesirov, J.; Neuberg, D.S.; Lander, E.S.; Aster, J.C.; Golub, T.R. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat. Med.,* **2002**, *8*, 68-74.

[60] van 't Veer, L.J.; Dai, H.; van de Vijver, M.J.; He, Y.D.; Hart, A.A.; Mao, M.; Peterse, H.L.; van der Kooy, K.; Marton, M.J.; Witteveen, A.T.; Schreiber, G.J.; Kerkhoven, R.M.; Roberts, C.; Linsley, P.S.; Bernards, R.; Friend, S.H. Gene expression profiling predicts clinical outcome of breast cancer. *Nature,* **2002**, *415*, 530-536.

[61] Yeoh, E.J.; Ross, M.E.; Shurtleff, S.A.; Williams, W.K.; Patel, D.; Mahfouz, R.; Behm, F.G.; Raimondi, S.C.; Relling, M.V.; Patel, A.; Cheng, C.; Campana, D.; Wilkins, D.; Zhou, X.; Li, J.; Liu, H.; Pui, C.H.; Evans, W.E.; Naeve, C.; Wong, L.; Downing, J.R. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell,* **2002**, *1*, 133-143.

[62] Scherf, U.; Ross, D.T.; Waltham, M.; Smith, L.H.; Lee, J.K.; Tanabe, L.; Kohn, K.W.; Reinhold, W.C.; Myers, T.G.; Andrews, D.T.; Scudiero, D.A.; Eisen, M.B.; Sausville, E.A.; Pommier, Y.; Botstein, D.; Brown, P.O.; Weinstein, J.N. A gene expression database for the molecular pharmacology of cancer. *Nat. Genet.,* **2000**, *24*, 236-244.

[63] Chien, K.R. Stem cells: lost in translation. *Nature,* **2004**, *428*, 607-608.

[64] Matoba, R.; Niwa, H.; Masui, S.; Ohtsuka, S.; Carter, M.G.; Sharov, A.A.; Ko, M.S. Dissecting Oct3/4-regulated gene networks in embryonic stem cells by expression profiling. *PLoS One,* **2006**, *1*, e26.

ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Production of a Non-Triple Helical Collagen α Chain in Transgenic Silkworms and Its Evaluation as a Gelatin Substitute for Cell Culture

Takahiro Adachi,[1,2] Xiaobiao Wang,[1] Tomoko Murata,[1] Masanobu Obara,[3] Hidenori Akutsu,[4] Masakazu Machida,[4] Akirihiro Umezawa,[4] Masahiro Tomita[1]

[1]Neosilk Co., Ltd., 3-13-26 Kagamiyama, Higashihiroshima, Hiroshima 739-0046, Japan; telephone: +81 82 431 0652; fax: +81 82 431 0654; e-mail: mtomita@neosilk.co.jp

[2]Hiroshima Prefectural Institute of Industrial Science and Technology, Higashihiroshima, Hiroshima, Japan

[3]Developmental Biology Laboratory, Department of Biological Science, Graduate School of Science, Hiroshima University, Higashihiroshima, Hiroshima, Japan

[4]National Center for Child Health and Development, Setagaya-ku, Tokyo, Japan

Received 16 December 2009; revision received 25 March 2010; accepted 29 March 2010

Published online 8 April 2010 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22752

ABSTRACT: We generated transgenic silkworms that synthesized human type I collagen α1 chain [α1(I) chain] in the middle silk glands and secreted it into cocoons. The initial content of the recombinant α1(I) chain in the cocoons of the transgenic silkworms was 0.8%. The IE1 gene, a trans-activator from the baculovirus, was introduced into the transgenic silkworm to increase the content of the chain. We also generated silkworms homozygous for the transgenes. These manipulations increased the α1(I) chain content to 8.0% (4.24 mg per cocoon). The α1(I) chain was extracted and purified from the cocoons using a very simple method. The α1(I) chain contained no hydroxyprolines due to the absence of prolyl-hydroxylase activity in the silk glands. Circular dichroism analysis showed that the secondary structure of the α1(I) chain is similar to that of denatured type I collagen, demonstrating the absence of the triple helical structure. Human skin fibroblasts were seeded on the α1(I) chain-coated dishes. The cells attached and spread, although at decreased chain concentrations the spreading rate was lower than that of the collagen and gelatin. Cynomolgus monkey embryonic stem cells cultured on the α1(I) chain-coated dishes maintained an undifferentiated state after 30 passages, and their pluripotency was confirmed by teratoma formation in severe combined immunodeficient mice. These results show that the recombinant human α1(I) chain is a promising candidate biomaterial as a high-quality and safe gelatin substitute for cell culture.

Biotechnol. Bioeng. 2010;106: 860–870.

KEYWORDS: transgenic silkworm; cocoon; collagen; gelatin; cell culture

Correspondence to: M. Tomita

Introduction

Collagen, a protein with a triple helical structure composed of three α chains, is used as a biomaterial for a variety of medical and cosmetic applications (Lee et al., 2001; Miyata et al., 1992). Gelatin, the denatured and fragmented form of collagen, is applied in medicine as a material for oral drug delivery and parenteral formulations (Gill and Feinberg, 2001; Tabata and Ikada, 1998). These materials come mainly from animal sources, which carry a risk of pathogen contamination and can also cause allergic reactions (Bradley, 1993; Mullins et al., 1996). One promising approach to overcome such problems is to produce them by recombinant means using appropriate hosts. For example, collagens have been produced successfully using host cells possessing prolyl-hydroxylase activity, which is required for the formation of the triple helical structure (Berg and Prockop, 1973; Fichard et al., 1997; Geddis and Prockop, 1993). Collagens have also been produced in host cells that did not possess sufficient prolyl-hydroxylase activity for triple helix formation but were transfected with prolyl-hydroxylase genes (John et al., 1999; Lamberg et al., 1996; Merle et al., 2002; Vuorela et al., 1997). In contrast, cells lacking hydroxylase activity have been utilized as hosts for the synthesis of recombinant gelatins (Olsen et al., 2003; Werten et al., 1999).

We previously generated transgenic silkworms that synthesized a recombinant fusion protein containing the collagen sequence in their silk glands and secreted it into cocoons. Due to the absence of prolyl-hydroxylase activity in the glands, prolyl residues in the collagen sequence were not hydroxylated (Tomita et al., 2003, 2005). To produce

prolyl-hydroxylated protein, prolyl-hydroxylase genes were simultaneously introduced to the silkworms along with the recombinant protein gene. The resulting silkworms produced a protein containing a prolyl-hydroxylated collagen sequence in their cocoons (Adachi et al., 2006). While these studies demonstrated the possibility of mass production of recombinant collagens in transgenic silkworms, we recognized an important issue that needed to be addressed: given that the recombinant protein was incorporated into insoluble silk fibers, the protein could not be extracted without using strong chaotropic reagents.

Silk fibers are composed mainly of two types of protein: fibroin and sericin. The former comprises 70–80% of all silk proteins, constitutes the core of insoluble silk fibers, and is synthesized in the posterior silk glands (PSGs). Sericin, which accounts for 20–30% of silk proteins, refers to a group of hydrophilic glue proteins that surround the fibroin core and are synthesized in the middle silk glands (MSGs; Garel et al., 1997; Grzelak, 1995). The above-described recombinant fusion protein containing the collagen sequence was expressed in the PSGs, resulting in production of the recombinant protein into the insoluble fibroin core (Tomita et al., 2005). Recently, we also developed a sericin promoter-driven recombinant protein expression system, in which recombinant proteins are expressed in the MSGs and secreted into hydrophilic sericin layers of silk fibers. Green fluorescent protein (Tomita et al., 2007), human serum albumin (Ogawa et al., 2007), and mouse IgG (Iizuka et al., 2009) were successfully produced in cocoons. These proteins were extractable from cocoons by soaking them in mild neutral aqueous solutions such as phosphate-buffered saline (PBS) or Tris-buffered saline.

In this study, we generated transgenic silkworms that expressed the full-length triple helical region of the human type I collagen α1 chain [α1(I) chain] in the MSGs. Type I collagen is the most abundant fibril-forming collagen in the human body. A common form of type I collagen is a heterotrimer composed of two α1(I) chains and one α2(I) chain. It is also known that α1(I) chains are able to form a homotrimer in the absence of α2(I) chain (Nicholls et al., 1979). The recombinant α1(I) chain expressed in the MSGs was efficiently secreted into cocoons and was easily recoverable. Given that the silk glands had no prolyl-hydroxylase activity, the chains contained no hydroxyproline residues. In addition, the recombinant α1(I) chain expressed in this study lacks the telopeptide and propeptide promoting triple helix formation (Doege and Fessler, 1986; Olsen et al., 2001; Rosenbloom et al., 1976). Therefore, the α1(I) chains did not form the homotrimer with the triple helical structure, and their physio-chemical properties were similar to those of gelatin rather than collagen. We tested the possibility of using the recombinant α1(I) chain as a cell culture scaffold, and found that cells could be cultured on the chain as on marketed gelatin. As the recombinant α1(I) chain has uniform molecular weight and no risks of animal-derived pathogen contamination, the recombinant chain may be useful as a high-quality and safe substitute for marketed gelatin.

Materials and Methods

Experimental Animals

A strain of silkworm, pnd-w1, was obtained from the National Institute of Agrobiological Science (Tsukuba, Japan). Larvae were reared at 25°C on an artificial diet (Nosan Corporation, Yokohama, Japan).

Vector Construction for Transgenic Silkworms

The vector COL1A1/pMSG for transgenic silkworms was constructed using a full-length cDNA coding for the pro α1 chain of human type I procollagen (GenBank Accession No. Z74615) obtained by RT-PCR from human placenta total RNA (Clontech, Palo Alto, CA). The PCR-amplified product was cloned into the pCR4-TOPO vector (Invitrogen, Carlsbad, CA) using a TOPO cloning system (Invitrogen), which yielded pCR4COL1A1. To obtain the DNA fragment containing the baculovirus polyhedrin 5′-untranslated region (UTR; nt 140–189; GenBank Accession No. M30925; Iizuka et al., 2008), the signal sequence of human calreticulin (nt 109–159; GenBank Accession No. M84739), and the cDNA for the triple helical region of the α1(I) chain (nt 654–3695; Fig. 1A), PCR was performed using pCR4COL1A1 as a template. A set of primers, calSP/COL1A1-F (5′-ATGCTGCTATCCGTGCCGTTGCTGCTC-GGCCTCCTCGGCCTGGCCGTCGCCGGCCCCATGGGT-CCCTCT-3′) and *Nru*I/STP/COL1A1-R (5′-TCGCGATT-AGGGAGGACCAGGG-GGACC-3′), was used for the first PCR. A second PCR was performed using the product of the first PCR as a template and *Nru*I/Bm5UTR/calSP-F (5′-TCGCGAAAGTATTTTACTGTTTTCGTAACAGTTTT-GTAATAAAAAAACCTATAAATATGCTGCTATCCGTG-CCG-3′) and *Nru*I/STP/COL1A1-R as a primer set. The amplified product was cloned into the pCR4-TOPO vector to generate pCR4BmUTRcalSPCOL1A1. The DNA fragment was re-excised from pCR4BmUTRcalSPCOL1A1 by digestion with *Nru*I, and inserted between the *Bombyx mori* sericin 1 promoter and the fibroin light chain 3′-flanking sequence of pMSG1.1MG (Iizuka et al., 2009), giving rise to COL1A1/pMSG (Fig. 1B).

A vector carrying the gene of baculovirus trans-activator IE1 (Tomita et al., 2007) was prepared as follows. To obtain the DNA fragment consisting of the polyhedrin 5′-UTR and the IE1 gene, PCR was performed using pIE1 (Tomita et al., 2007) as a template. A set of primers, *Eco*RV/Bm5UTR/IE1-F (5′-GATATCAAGTATTTTACTGTTTTC-GTAACAGTTTTGTAATAAAAAAACCTATAAATATGA-CGCAAATTAATTTTAACGCGTCG-3′) and *Bgl*/IE1-R (5′-AGATCTTTAATTAAATTCAAATTTTTTATATTTACAA-TTTAG-3′), was used for the PCR. The amplified product

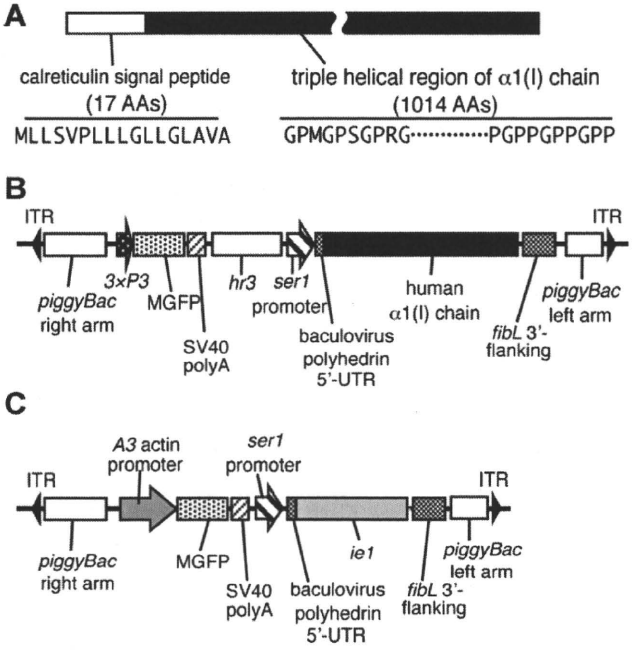


Figure 1. The structure of the transformation vector. **A**: The structure of the amino acid sequence of the α1(I) chain. The sequence of the α1(I) chain consists of a 17 amino acid (AA)-long calreticulin signal peptide (open box) and a 1,014 AA-long triple helical region of the α1(I) chain (filled box). **B**: The structure of the transformation vector COL1A1/pMSG. COL1A1/pMSG contained two expression units between the right and left arms of *piggyBac* as follows: *3xP3* promoter-driven MGFP cDNA with the SV40 polyA signal sequence (SV40 polyA) and the baculovirus-derived enhancer *hr3*-linked sericin 1 (*ser1*) promoter-driven α1(I) chain cDNA with the fibroin light chain gene polyA signal sequence (*fibL* 3′-flanking). ITR, inverted terminal repeat. **C**: Transformation vector pIM1. pIM1 contained two expression units between the right and left arms of *piggyBac* as follows: *Bombyx mori A3* actin promoter-driven MGFP cDNA with the SV40 polyA and *ser1* promoter-driven *ie1* gene with the *fibL* 3′-flanking.

was cloned into pCRII-TOPO vector (Invitrogen) to generate pCRBmUTRie1. The DNA fragment was released from pCRBmUTRie1 by digestion with *Eco*RV, and inserted between the *B. mori* sericin1 promoter and the fibroin light chain 3′-flanking sequence of pMSG1.1A3-MG, which had been created by replacing the *3xP3* promoter with the *B. mori A3* actin promoter (nt 1757–2595, GenBank Accession No. U49854) in pMSG1.1MG. The resulting vector was termed pIM1 (Fig. 1C).

Generation of Transgenic Silkworms

COL1A1/pMSG was injected with the helper vector pHA3PIG (Tamura et al., 2000) into pre-blastoderm embryos as described previously (Tomita et al., 2003). Hatched G_0 larvae were reared at 25°C to the moth stage. G_1 embryos obtained by sibling mating were screened for the expression of Monster Green Fluorescent Protein (MGFP; Promega, San Luis Obispo, CA) in the eyes and nervous tissues to obtain transgenic silkworms.

pIM1 was injected into eggs, and transgenic silkworms were created as described above, except that screening was performed by observing MGFP fluorescence in the larval body. The resulting transgenic silkworm (IM1) was crossed with silkworms carrying the α1(I) chain to obtain silkworms hemizygous for both the IE1 and the α1(I) chain genes.

To generate silkworms homozygous for the IE1 and the α1(I) chain genes, the above hemizygous silkworms were crossed with one another. From the next generations, homozygous silkworms for both genes were screened by genomic PCR using DNA extracted from blood cells as a template. Primers used in this PCR were designed from the genomic sequences flanking the transgene insertions, which were determined with an APA Transgene Locator Kit (BIO S&T Inc., Montreal, QC, Canada) according to the manufacturer's instructions.

SDS–PAGE and Western Blotting

Proteins in the sericin layer of silk fibers were extracted by incubating cocoons at 80°C for 5 min in a buffer consisting of 8 M urea, 2% 2-mercaptoethanol, and 50 mM Tris–HCl, pH 8.0. After centrifugation at 23,500*g* for 5 min, the obtained supernatant was electrophoresed on 0.1% SDS/5–20% polyacrylamide gradient gels (Atto, Tokyo, Japan) in running buffer (0.1% SDS, 12.5 mM Tris, and 125 mM glycine). The gels were stained with CBB-R250, and densitometric analyses were performed using the image-processing program, ImageJ (http://rsb.info.nih.gov/ij/). For Western blotting, the proteins on the gels were transferred onto PVDF membranes (Immobilon-P; Millipore, Billerica, MA), reacted with antihuman/bovine type I collagen antibodies (LB-1190; Cosmo Bio, Tokyo, Japan) and then with horseradish peroxidase (HRP)-linked anti-(rabbit IgG) antibodies (Cell Signaling Technology, Danvers, MA). The antigen–antibody complexes were visualized using the ECL Western Blotting Detection System (GE Healthcare, Little Chalfont, Buckinghamshire, UK).

Purification of the Recombinant α1(I) Chain From Transgenic Silkworm Cocoons

Cocoons were crushed into powder using a mill, suspended in a solution of 8 M urea and 0.5 M CH_3COOH, pH 2.0, at a concentration of 10 mg powdered cocoons/mL, and incubated at 4°C for 2 days with stirring. The resulting solution was filtered through filter paper and 70-μm nylon mesh. The solution was concentrated approximately threefold with an ultrafiltration membrane (Millipore) and urea was removed from the solution by adding 0.5 M CH_3COOH, pH 2.0, and subsequent ultrafiltration. NaCl was then added to the solution at a concentration of 0.5 M, and precipitated sericin was removed by centrifugation. Next, the recombinant α1(I) chain in the solution was collected by the addition of 3.5 M NaCl. The precipitate was

solubilized in 0.5 M CH_3COOH and subjected to dialysis against water.

N-Terminal Sequencing and Measurement of Amino Acid Composition

The N-terminal sequences of the purified α1(I) chain were determined with a G1000A protein sequencer (Hewlet Packard, Palo Alto, CA) using the standard protocol of Edman degradation. For the degradation reaction, 36 pmol of the chain were loaded onto the sequencer and the reaction was carried out for five cycles.

The purified α1(I) chain was hydrolyzed in 6 N HCl for 22 h at 110°C under vacuum, and analyzed for amino acid composition using a Hitachi 835 amino acid analyzer (Hitachi, Tokyo, Japan).

Determination of Melting and Gelation Points

The purified α1(I) chain was dissolved in distilled water at a concentration of 50 mg/mL, and the solution was subjected to determination of gelation and melting points as follows. The α1(I) chain solution was gradually cooled from 35 to −5°C at a rate of 1°C/min with a thermal cycler (Atto). The gelation point was determined by reading the temperature of the sample when its fluidity disappeared. For the analysis of the melting point, the α1(j) chain solution was incubated in icy water for 30 min to create a gel. The α1(I) chain gel was gradually heated from 0 to 40°C at a rate of 1°C/min with the thermal cycler. The melting point was determined by reading the temperature of the sample when the bubble at the bottom of the tube reached the surface of the sample solution.

Measurement of Circular Dichroism (CD) Spectra

CD spectra were recorded for the recombinant α1(I) chain using a spectropolarimeter J-820 (Jasco, Tokyo, Japan). The purified α1(I) chain was dissolved in 50 mM CH_3COOH at a concentration of 100 μg/mL for measurement in far ultraviolet (190–240 nm) regions, and the solution was placed in a cuvette with 2-mm pass length. The temperature was kept at 4°C during the procedures. Measurements were performed three times and the averaged values were plotted. Thermal transition curves were recorded by measuring molar ellipticity at 224 nm between 5 and 60°C at a rate of 30°C/h. The denaturing temperature was calculated with J-820 software (Jasco).

Endotoxin Test

An endotoxin level of the recombinant α1(I) chain was quantified using the limulus amebocyte lysate (LAL) as per the manufacturer's instructions (Endospecy ES-50M and Toxicolor DIA-MP; Seikagaku Biobusiness, Tokyo, Japan). Briefly, 50 μL standards or samples diluted with endotoxin-free water were mixed with LAL and incubated at 37°C for 30 min. After the substrate solution was added, the absorbance at 545 nm was measured. A standard curve was constructed by using the standards in the range of 0.02–0.1 EU/mL, and the concentration of endotoxin in each sample was determined by plotting the absorbance to the standard curve.

Quantifying the Spread of Human Dermal Skin Fibroblasts

The cell adhesivity to the recombinant α1(I) chain was analyzed by a quantitative cell-spreading assay using human dermal skin fibroblasts (HSFs; Kurabo, Osaka, Japan) as described previously (Yamada and Kennedy, 1984). In brief, a 96-well tissue culture plate (Becton Dickinson Labware, Franklin Lakes, NJ) was incubated with the recombinant α1(I) chain at various concentrations (0.31, 0.62, 1.25, 2.5, and 5.0 μg/mL) at 37°C for 1 h, treated with heat-denatured 0.5% bovine serum albumin to block the direct interaction between cells and the plate surface, and extensively washed with PBS. HSFs were seeded on the wells at a density of 1×10^4 cells per well, cultured in serum-free Dulbecco's modified Eagle's medium for 1 h, fixed in 4% paraformaldehyde, and viewed through a phase-contrast microscope (Nikon, Tokyo, Japan) to calculate the ratio of fully spreading cells to all cells observed in five randomly selected fields. HSFs were also cultured on the 96-well tissue culture plate coated with 10 μg/mL of the α1(I) chain as above to observe the cell morphology.

Culture of Cynomolgus Monkey Embryonic Stem (ES) Cells on Dishes Coated With the Recombinant α1(I) Chain

Cynomolgus monkey ES cells were cultured on murine embryonic fibroblast (MEF) feeder cells in Petri dishes (90 mm in diameter) according to previously established protocols (Cameron et al., 2006). Briefly, dishes were coated with the recombinant α1(I) chain by incubation with a solution of 1 mg/mL α1(I) chain for 30 min at room temperature. γ-Irradiated MEF cells were then cultured on the coated dishes as feeder cells for 1 day. Subsequently, monkey ES cells were seeded on MEF feeder layers and maintained by changing the medium. Immunostaining of monkey ES cell colonies was performed using NANOG, TRA1-81, SSEA-4, SOX2, and OCT4-specific primary antibodies (Millipore) according to the manufacturer's protocol. Nuclei were visualized by DAPI staining (Lin et al., 1976).

Monkey ES cells were injected subcutaneously into the hind leg of severe combined immunodeficient (SCID) mice (Suemori et al., 2001). Two months after injection, the mice were killed to remove teratomas. The teratomas

were then fixed with 4% paraformaldehyde in PBS, paraffin-embedded, sectioned, and histologically analyzed following staining with hematoxylin and eosin.

Results

Generation of Transgenic Silkworms

COL1A1/pMSG was injected into 9,834 pre-blastoderm embryos, and the hatched 5,282 G_0 larvae were allowed to develop to moths. G_1 embryos from the G_0 moths were screened for MGFP expression in the eyes and nervous tissues to obtain transgenic silkworms. Genomic Southern blot analysis of the transgenic silkworms demonstrated the existence of 41 independent transgenic lines with respect to the chromosomal insertion positions and copy numbers of the transgenes. Among them, 34 lines of transgenic silkworms with a single-copy transgene were selected, and the cocoon proteins of the lines were analyzed by SDS–PAGE. The line with the highest level of α1(I) chain expression was crossed with wild-type silkworms, and offspring hemizygous for the α1(I) chain gene were used in the following experiments as the COL249 line.

The worms of the COL249 line spun cocoons that were normal in appearance, size, and weight. Proteins in the sericin layer of the silk fibers of COL249 and wild-type silkworms were separated by SDS–PAGE and stained with CBB (Fig. 2, lanes 1–2). A band with an apparent molecular weight of 120 kDa was seen only in the cocoon proteins of COL249 (Fig. 2, lane 1). This band reacted with antihuman/bovine type I collagen antibody (Fig. 2, lane 6), indicating that this was the recombinant product from the human α1(I) chain gene. The band intensity of the recombinant α1(I) chain on the CBB-stained gel was quantified by densitometry. The content of the α1(I) chain was estimated to be 0.8% of the dried cocoon.

To increase the α1(I) chain content in the cocoon, a COL249 moth was crossed with an IM1 moth bearing the gene of the baculovirus-derived trans-activator IE1. Approximately 25% of the offspring carried both the α1(I) chain and IE1 genes hemizygously. The α1(I) chain/IE1 hemizygous silkworms (COL249/IM1) were further crossed with each other and silkworms homozygous for both the α1(I) chain and IE1 genes $(COL249/IM)^2$ were screened from the offspring by genomic PCR. Proteins in the cocoon extracts of COL249, COL249/IM, and $(COL249/IM)^2$ were separated by SDS–PAGE and stained with CBB (Fig. 3). By measuring the band intensity, the contents of the α1(I) chain in cocoons of the silkworm lines COL249, COL249/IM, and $(COL249/IM)^2$ were estimated to be 0.8%, 4.8%, and 8.0%, respectively. The average weights of cocoons in the COL249, COL249/IM, and $(COL249/IM)^2$ lines were 72, 65, and 53 mg, respectively. Although the cocoon weight decreased slightly with increased transgene copy numbers, the synthesis of the recombinant α1(I) chain per larvae was improved by this procedure; the amounts of the chain per cocoon of the COL249, COL249/IM, and $(COL249/IM)^2$ lines were 0.58, 3.12, and 4.24 mg, respectively.

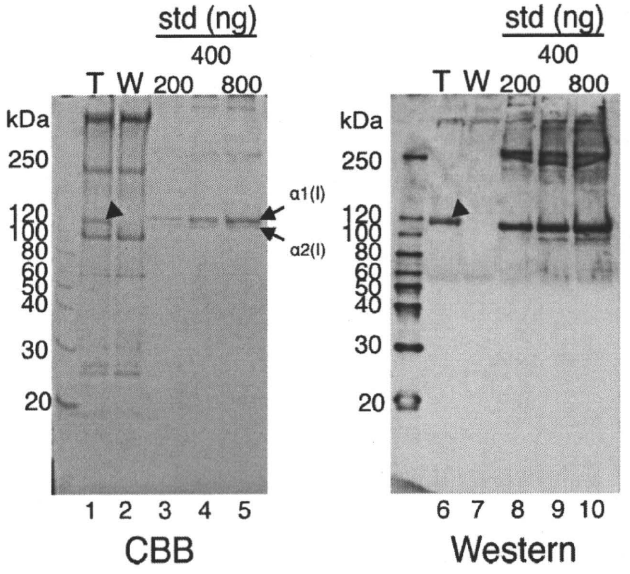


Figure 2. Expression of the α1(I) chain in cocoons of transgenic silkworms. The proteins in the cocoons of COL249 (T) and wild-type (W) silkworms were extracted with 8 M urea, 2% 2-mercaptoethanol, and 50 mM Tris–HCl, pH 8.0, separated by SDS–PAGE, and stained with CBB (**left panel**). Aliquots of the cocoon extracts were also assessed by Western blotting with antihuman/bovine type I collagen antibodies (**right panel**). Bovine pepsinized type I collagen in the amounts indicated was analyzed by CBB staining and Western blotting as a standard (std). The arrowheads in lanes 1 and 6 point to the band of the recombinant α1(I) chain. The arrows in lane 5 point to α1(I) and α2(I) chains of bovine type I collagen. Arabic numerals at the left side are molecular masses in kDa.

Extraction and Purification of the Recombinant α1(I) Chain From Cocoons

The extraction efficiency of the recombinant α1(I) chain from cocoons was examined. The powder of $(COL249/IM)^2$ cocoons was suspended in either PBS (Fig. 4A, lane 3), 0.5 M CH_3COOH, pH 3.0 (Fig. 4A, lane 4), 0.5 M CH_3COOH, pH 2.0 (Fig. 4A, lane 5), or 8 M urea and 0.5 M CH_3COOH, pH 2.0 (Fig. 4A, lane 6), at 4°C for 16 h, and the extracted proteins were analyzed by SDS–PAGE. Total proteins in the sericin layer of silk fibers of wild-type and transgenic silkworms were extracted by incubating cocoons at 80°C for 5 min in a buffer consisting of 8 M urea, 2% 2-mercaptoethanol, and 50 mM Tris–HCl, pH 8.0 (Fig. 4A, lanes 1 and 2). The ratios of the α1(I) chain extracted with the solutions to the total α1(I) chain in the sericin layer were calculated by quantifying the band intensities of the CBB-stained α1(I) chain. The efficiency of extraction was estimated to be 10% for PBS, 30% for 0.5 M CH_3COOH, pH 3.0, 50% for 0.5 M CH_3COOH, pH 2.0, and 80% for 8 M urea and 0.5 M CH_3COOH, pH 2.0. Thus, the use of CH_3COOH-solutions was more effective than PBS for extracting the recombinant α1(I) chain from

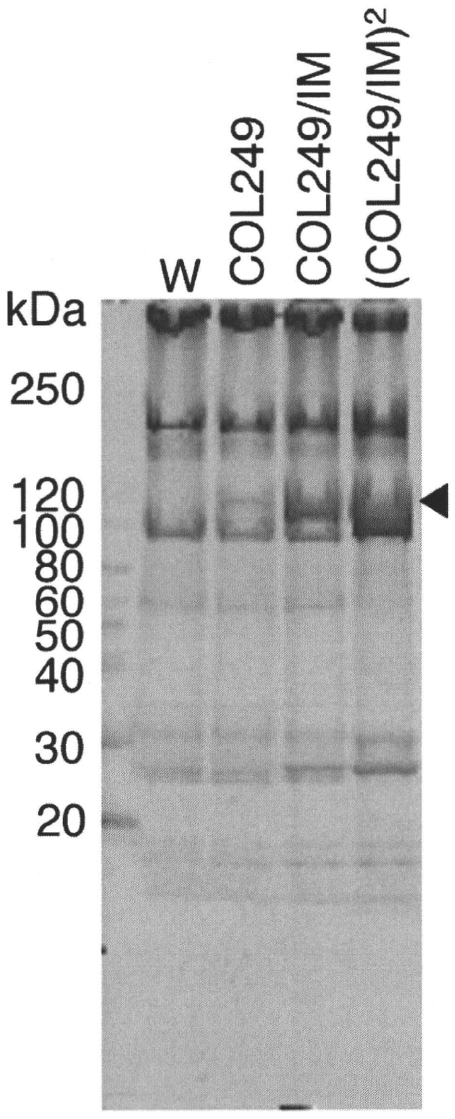


Figure 3. Increase in the α1(I) chain content in cocoons of transgenic silkworms. Transgenic silkworms were genetically modified to increase the amount of the α1(I) chain. Proteins were extracted from cocoons produced by wild-type (W), COL249, COL249/IM, and $(COL249/IM)^2$ silkworms, separated by SDS–PAGE, and stained with CBB. The arrowhead points to the band of the recombinant α1(I) chain. Arabic numerals at the left side are molecular masses in kDa.

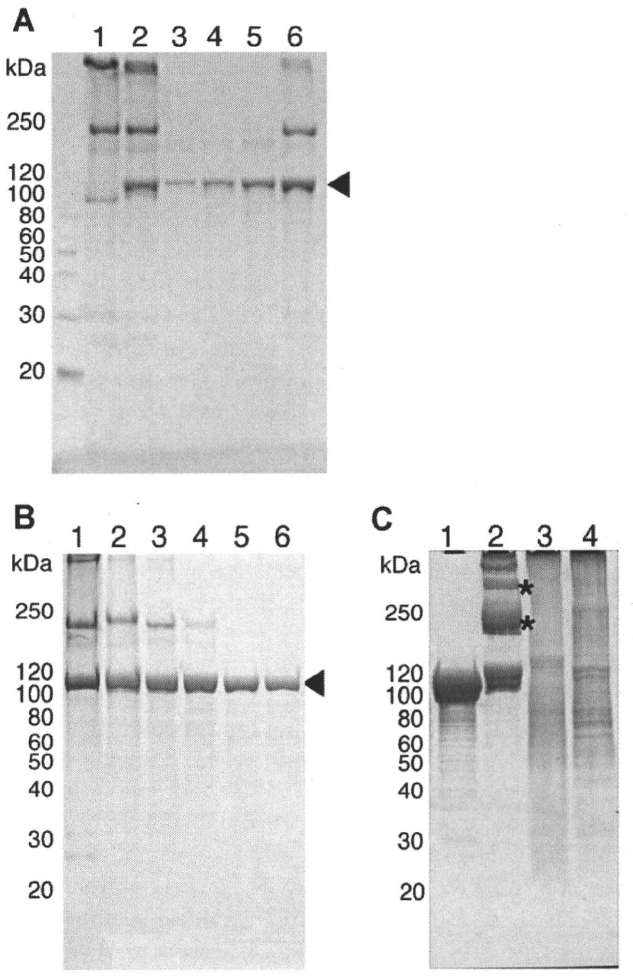


Figure 4. Extraction and purification of the α1(I) chain from cocoons of $(COL249/IM)^2$ silkworms. **A**: Extraction of the α1(I) chain from cocoons. Cocoon proteins were extracted with PBS (**lane 3**), 0.5 M CH_3COOH, pH 3.0 (**lane 4**), 0.5 M CH_3COOH, pH 2.0 (**lane 5**), and 8 M urea and 0.5 M CH_3COOH, pH 2.0 (**lane 6**). The extracted proteins were analyzed by SDS–PAGE. **Lanes 1** and **2** represent total proteins in the sericin layer of the silk fibers of wild-type and transgenic silkworms, respectively. The arrowhead points to the α1(I) chain. **B**: Analysis of the α1(I) chain in the purification processes. Cocoon proteins were extracted with 8 M urea and 0.5 M CH_3COOH, pH 2.0 (**lane 2**). The extracted proteins were concentrated by ultrafiltration (**lane 3**), and urea in the extract was removed by adding 0.5 M CH_3COOH, pH 2.0, and subsequent ultrafiltration (**lane 4**). After removing sericin by 0.5 M NaCl precipitation (**lane 5**), the α1(I) chain in the extract was collected by 4 M NaCl precipitation (**lane 6**). The proteins in each purification step were analyzed by SDS–PAGE and CBB staining. **Lane 1** represents total proteins in the sericin layer of the silk fibers. The arrowhead points to the α1(I) chain. **C**: Analysis of the purified α1(I) chain. Aliquots of 15 μg of proteins were electrophoresed and stained with CBB. **Lane 1**, the recombinant α1(I) chain; **lane 2**, bovine type I collagen; **lane 3**, alkaline-treated bovine gelatin; **lane 4**, acid-treated porcine gelatin. The asterisks in **lane 2** point to the dimer and trimer composed of the two and three α chains, respectively. Arabic numerals at the left side are molecular masses in kDa.

cocoons, and a major part of the α1(I) chain was extractive with 8 M urea and 0.5 M CH_3COOH, pH 2.0.

For purification of the α1(I) chain, 30 g of $(COL249/IM)^2$ cocoon powder, which was estimated to contain 2.4 g of the α1(I) chain, were suspended in 8 M urea and 0.5 M CH_3COOH, pH 2.0. The extracted α1(I) chain (Fig. 4B, lane 2) was concentrated by ultrafiltration (Fig. 4B, lane 3). The urea in the solution was removed by adding 0.5 M CH_3COOH, pH 2.0, and subsequent ultrafiltration. In this process, the endogenous sericin in the extract formed an insoluble aggregate, increasing the α1(I) chain content in the extract (Fig. 4B, lane 4). Small amounts of contaminated sericin were removed by 0.5 M NaCl precipitation (Fig. 4B, lane 5), and the α1(I) chain in the supernatant was then collected by precipitation with 4 M NaCl. The collected α1(I) chain was dissolved in 0.5 M CH_3COOH, pH 2.0, again, and the α1(I) chain solution was finally dialyzed against water (Fig. 4B, lane 6). The proteins in each

purification step and total proteins in the sericin layer (Fig. 4B, lane 1) were analyzed by SDS–PAGE, demonstrating that this simple purification process is sufficient to purify the α1(I) chain to apparent homogeneity. As a result, 990 mg of the α1(I) chain were purified from 30 g of cocoons; the recovery rate was estimated to be approximately 41%.

Biochemical Characterization

The purified recombinant α1(I) chain was analyzed by SDS–PAGE. Although small amounts of degradation products were found, the purified recombinant chain was composed of the polypeptide with a uniform molecular weight. The molecular weight of the chain was slightly smaller than the standard bovine α1(I) chain (Fig. 4C, lanes 1 and 2), indicating the possibility of insufficient prolyl-hydroxylation in the recombinant chain. The dimer (β chain) and trimer (γ chain) of the α chain, which were present in the standard collagen, were not detected from the purified recombinant chain, suggesting the absence of covalent cross-linking among the α1(I) chains. The molecular weight distribution of the recombinant α chain was quite different from that of the alkali-treated bovine (Fig. 4C, lane 3) or acid-treated porcine gelatins (Fig. 4C, lane 4). The gelatins gave broad molecular weight distributions because they were hydrolyzed products of collagens.

The α1(I) chain was subjected to an amino acid sequencer with five cycles of Edman degradation. The N-terminal amino acid sequencing of the α1(I) chain detected major and minor amino acid peaks in each cycle as shown in Table I. The sequence deduced from the minor peaks (GPM) was consistent with that of the predicted signal peptide cleavage (Fig. 1A) although peaks were not detected in the fourth and fifth cycles. The sequence from the major peaks (MGPSG) was probably derived from a cleavage at two amino acids downstream of the predicted site.

The amino acid composition of the purified α1(I) chain was determined after acid hydrolysis using a Hitachi L835 automated analyzer (Table II). The determined values were very similar to the predicted ones, except for the absence of hydroxyprolines and hydroxylysines.

The endotoxin levels of the α1(I) chain and the porcine gelatin were measured. The endotoxin level of the α1(I) chain was much lower (26 EU/g) than the gelatin (6,400 EU/g).

Structural Characterization

Far-ultraviolet (190–240 nm) CD spectra were recorded for the recombinant α1(I) chain, the native bovine type I collagen, the heat-denatured bovine type I collagen, and the porcine gelatin (Fig. 5A). The positive peak at 224 nm that is characteristic of the triple helical structure of collagen (Miller and Gay, 1982) was observed in the type I collagen. The gelatin exhibited a positive low peak at this wavelength, suggesting that the gelatin contained a partly formed triple helical structure in the molecule. In contrast, the recombinant α1(I) chain did not show a positive peak at 224 nm. The peak of the type I collagen disappeared when it was heat-denatured. The spectra of the denatured collagen were almost identical to those of the recombinant α1(I) chain. These results suggest that the α1(I) chain contained no triple helical structure. A negative peak at 198 nm represents the triple helical structure (Miller and Gay, 1982). The peak intensity at this wavelength of the recombinant α1(I) chain was similar to that of the heat-denatured collagen rather than that of gelatin, confirming the absence of the triple helical structure in the recombinant chain.

The 224-nm spectra were recorded for the recombinant chain, the native collagen, the denatured collagen and the gelatin at temperatures from 4 to 60°C (Fig. 5B). Apparent structural transition of the native collagen was observed in the range 39–46°C, which is in accordance with a report that the denatured temperature of bovine type I collagen is 42.8°C (Peltonen et al., 1980). In contrast, the recombinant α1(I) chain, the denatured collagen and the gelatin showed slight structural changes in the range 25–45°C.

The melting and gelation points of the 5% α1(I) chain or the gelatin solution were measured as described in the Materials and Methods Section. The melting and gelation

Table I. N-terminal sequencing.

Amino acid number	1	2	3	4	5
Major peaks	M	G	P	S	G
Minor peaks	G	P	M	—	—

Table II. Measurement of amino acid composition.

Amino Acid	Composition (mol%)	
	Recombinant α1(I) chain	Human α1(I) chain (predicted)
Aspartic acid	4.48	4.14
Threonine	1.56	1.58
Serine	2.91	3.35
Glutamic acid	7.71	7.30
Glycine	34.23	33.63
Alanine	11.72	11.74
Valine	1.81	1.87
Cystein	0.00	0.00
Methionine	0.33	0.69
Isoleucine	0.70	0.59
Leucine	2.12	1.87
Tyrosine	0.00	0.00
Phenylalanine	1.27	1.18
Hydroxylysine	0.00	2.37*
Lysine	3.66	1.18
Histidine	0.00	0.20
Arginine	4.71	5.03
Tryptophan	0.00	0.00
Hydroxyproline	0.00	11.44*
Proline	22.80	11.83
Total	100.00	100.00

*Assuming that all of lysine and proline residues in Y-position are hydroxylated.

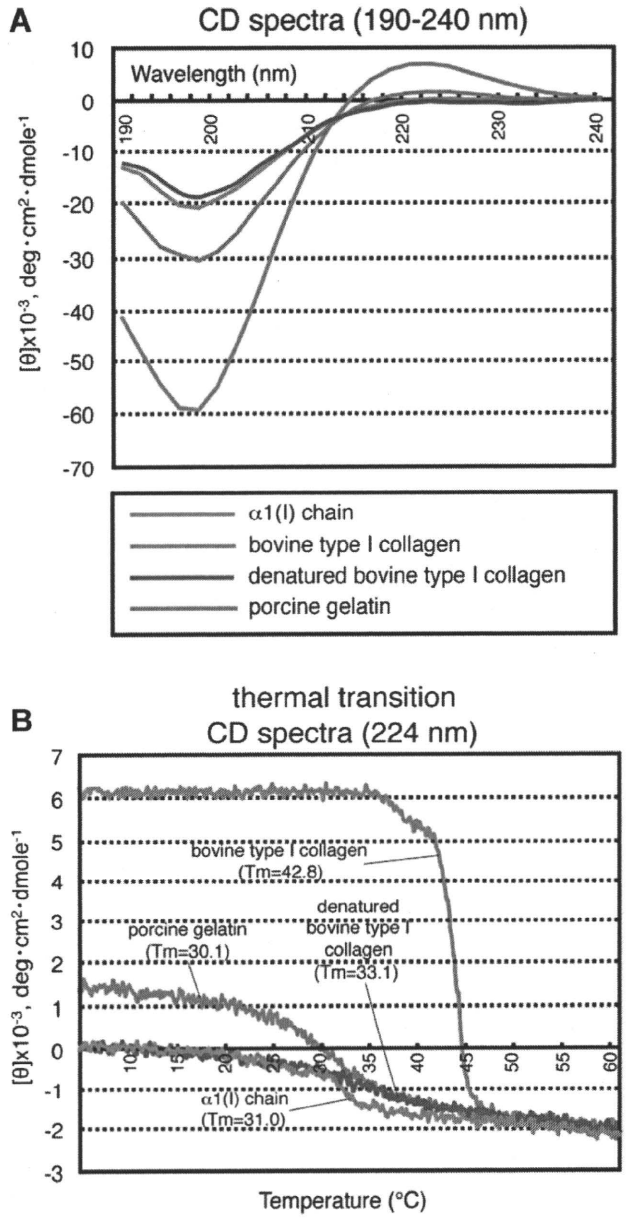


Figure 5. CD spectra of the α1(I) chain. **A:** Measurement of CD spectra of the α1(I) chain. Far-ultraviolet (190–240 nm) CD spectra were recorded for the recombinant α1(I) chain (red line), bovine type I collagen (pink line), denatured bovine type I collagen (blue line), and porcine gelatin (green line) at a concentration of 100 μg/mL. **B:** Thermal transition curves of the α1(I) chain. The CD spectra at 224 nm of the α1(I) chain, bovine type I collagen, denatured bovine type I collagen and porcine gelatin were monitored at temperatures from 4 to 60°C.

points of the α1(I) chain were 17 and 10°C, respectively, while the melting and gelation points of the gelatin were 30 and 26°C, respectively. Thus, both the melting and gelation points of the α1(I) chain were lower than the respective points of the gelatin, which may support the result from the measurement of CD spectra showing an absence of triple helical structures in the α1(I) chain. We also analyzed whether the α1(I) chain formed collagen fibrils under the physiological conditions as native collagen (Michalopoulos and Pitot, 1975) and found that the α1(I) chain did not form the fibrils (data not shown).

Cell Biological Properties

To investigate the cell biological properties of the recombinant α1(I) chain, HSFs were cultured on dishes coated with the α1(I) chain, the native collagen, or gelatin at various concentrations, and cell spreading on the materials was analyzed as described in the Materials and Methods Section (Fig. 6A). HSFs spread on all coating materials in a concentration-dependent manner. On collagen at concentrations of more than 0.63 μg/mL, HSFs spread at a rate of 100%. More than 80% of the cells spread on gelatin at concentrations >0.63 μg/mL, but 100% cell spreading was never observed even at the highest concentration tested (5.0 μg/mL). Cell spreading rates for the α1(I) chain at concentrations of <2.5 μg/mL were slightly low compared to those on gelatin at the same concentrations. However, HSFs spread on the α1(I) chain at a similar rate to on gelatin when inoculated at a concentration of 5.0 μg/mL. No differences were observed among cell morphologies when the cells were cultured on the α1(I) chain, the native collagen, or the gelatin at a concentration of 10 μg/mL (Fig. 6B, panels a–c). Cell-spreading was not observed on the uncoated dishes (Fig. 6B, panel d).

Cynomolgus monkey ES cells were cultured on feeder cells that had been cultured on dishes coated with the α1(I) chain or porcine gelatin. The ES cells cultured on dishes coated with the α1(I) chain formed tightly packed and flattened colonies (Fig. 7A, panel a). This morphology was the same as that of ES cell colonies cultured on dishes coated with porcine gelatin (Fig. 7A, panel b). Immunocytochemical studies confirmed that the monkey ES cell colonies on the α1(I) chain expressed the ES cell marker proteins NANOG, TRA1-81, SSEA-4, SOX2, and OCT4 (Fig. 7B). When the ES cells were subcutaneously injected into SCID mice after the passages on the α1(I) chain, the cells formed teratomas in the mouse tissues. Histological analyses of the teratomas showed formation of pigment epithelium, gastrointestinal epithelium, and cartilage (Fig. 7C). Thus, the α1(I) chain was confirmed to be useful for the maintenance of monkey ES cells.

Discussion

We generated transgenic silkworms that secreted the recombinant human α1(I) chain into the sericin layer of silk fibers. The content of the α1(I) chain in the cocoons of the established line COL249 was estimated to be 0.8%. By introducing the gene of the trans-activator IE1 into the silkworm as in our previous studies (Ogawa et al., 2007; Tomita et al., 2007), the expression of the α1(I) chain was enhanced to 4.8%. We then generated silkworms (COL249/

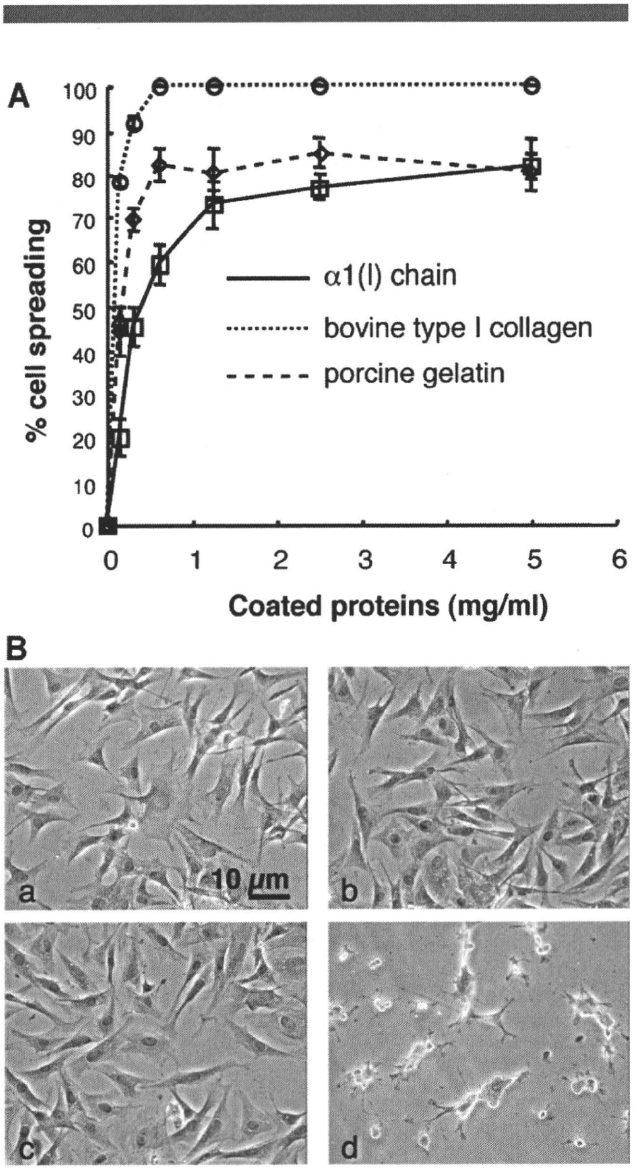


Figure 6. Spreading of HSFs on the α1(I) chain-coated dishes. **A**: Cell spreading assay using HSFs. The wells of tissue culture plates were coated with the α1(I) chain (black line), bovine type I collagen (dotted line), and porcine gelatin (dashed line) at various concentrations, and treated with heat-denatured bovine serum albumin to block the direct interaction between cells and the plate. HSFs were seeded on these wells and cultured for 1 h. The cells were then fixed, and the ratio of spreading cells to all cells in observed fields was calculated. **B**: Cell morphology of HSFs cultured on the α1(I) chain. HSFs were cultured on dishes coated with 10 μg/mL of the α1(I) chain (**a**), bovine type I collagen (**b**), porcine gelatin (**c**). The cells were also cultured on the uncoated but the albumin-treated dish (**d**). Scale bar, 10 μm.

IM)^2 homozygous for both the α1(I) chain and IE1 genes. This manipulation increased the α1(I) chain content up to 8.0%. Given that the average weight of a (COL249/IM)^2 cocoon was 53 mg, the α1(I) chain content per cocoon was calculated to be 4.24 mg. If 1,000 (COL249/IM)^2 silkworms were reared, we could produce 4.24 g of the α1(I) chain. The (COL249/IM1)^2 silkworm was generated from the pnd-w1 strain, which produces small cocoons (50–70 mg). Our preliminary experiment revealed that the cocoon weight could be increased to approximately 150 mg by crossing it with typical silkworm strains that produce 300- to 500-mg cocoons, leading to elevation of the α1(I) content to more than 10 mg per cocoon (data not shown). We also demonstrated the superiority of the transgenic silkworm system for the purification of the recombinant α1(I) chain. The α1(I) chain was highly purified from the cocoon extract by a simple method consisting of ultrafiltration and salt precipitation. Thus, this study offers experimental evidence for the viability of using transgenic silkworms in the production of the human recombinant α1(I) chain on an industrial scale.

Our previous study demonstrated that prolyl-hydroxylase activity is absent in silk glands (Adachi et al., 2005), and the recombinant fusion protein containing the collagen sequence expressed in the glands includes no hydroxyprolines (Tomita et al., 2005). The recombinant α1(I) chain produced in this study also contained no hydroxyprolines as predicted. The presence of hydroxyprolines is a prerequisite for forming the stable collagen triple helix (Berg and Prockop, 1973). In addition, the α1(I) chain did not contain the C-telopeptide and C-propeptide, which are known to promote triple helix formation (Doege and Fessler, 1986; Rosenbloom et al., 1976). Therefore, we postulated that the α1(I) chain is not capable of forming the triple helix. Indeed, CD spectra of the chain showed a complete absence of the triple helical structure. The importance of the telopeptide and propeptide for the triple helix formation was also shown in the previous studies. Unhydroxylated type I collagen with the telopeptide, and unhydroxylated α1(I) chain with the telopeptide and propeptide were synthesized as correctly folded triple helices in yeast (Olsen et al., 2001) and tobacco (Ruggiero et al., 2000), respectively. On the other hand, this study revealed that the animal-derived gelatin contained a partially folded triple helix, suggesting the significance of hydroxyprolines in the stability of triple helix. To further clarify the difference of physiological properties among the recombinant α1(I) chain, the gelatin and the collagen, we investigated the gelation and fibril-forming properties of the α1(I) chain. Unlike the collagen, the α1(I) chain did not form collagen fibrils under the physiological conditions examined. In contrast, the α1(I) chain, as well as the gelatin, gelated at lower temperatures than physiological ones. The melting and gelation points of the α1(I) chain were lower than those of gelatin. Thus, the physico-chemical properties of the recombinant α1(I) chain were similar to gelatin rather than collagen. However, due to the complete absence of the triple helical structure, the properties of the α1(I) chain differed slightly from those of gelatin.

The α1(I) chain promoted cell attachment and the spread of HSFs, but the cell-spreading rates for the α1(I) chain as well as gelatin were lower than those for collagen at all concentrations tested. At decreased concentrations of the materials, fewer cells spread on the α1(I) chain than on gelatin. Thus, HSFs were likely able to distinguish among these three materials. The cell–collagen interaction is mediated via integrins. Integrins α1β1 and α2β1 recognize collagens as collagen receptors (Hynes, 2002), and integrin