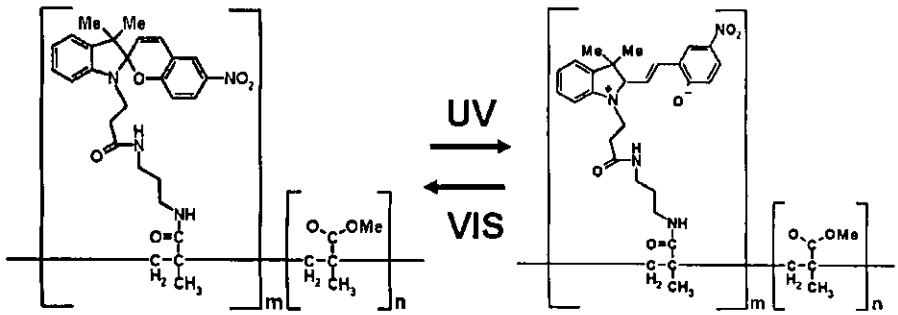

Figure 1. Schematic of transition of poly(NSP-*co*-MMA) upon exposure to UV irradiation.

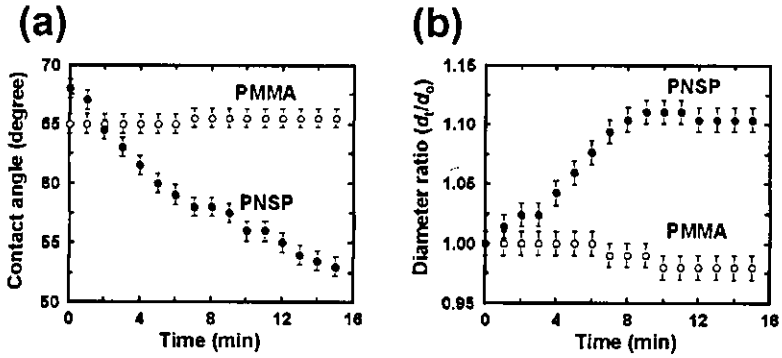

Figure 2. Hydrophobicity–hydrophilicity change on poly(NSP-*co*-MMA)-coated and PMMA-coated glass plates induced by UV irradiation. (a) Time dependence of water contact angle of water drop on poly(NSP-*co*-MMA)-coated and PMMA-coated glass plates. (b) Time dependence of diameter ratio (d/d_0) of water drop on poly(NSP-*co*-MMA) and PMMA coated g'ass plates. PNSP indicates poly(NSP-*co*-MMA)-coated glass plates.

After the water contact angles on poly(NSP-*co*-MMA)-coated glass plates were measured at first time, the poly-(NSP-*co*-MMA)-coated glass plates were dried under vacuum and in the dark place at ambient temperature for 24 h. After this treatment, the color of poly(NSP-*co*-MMA)-coated glass plates changed from purple color to colorless, which indicated the spiropyran in poly(NSP-*co*-MMA)-coated glass plates returned to the first form of nonionic spiropyran. The water contact angles of the poly(NSP-*co*-MMA)-coated glass plates were again measured. Exactly the same results to Figure 2 were obtained within the experimental error.

Therefore, the reversibility between the form of nonionic spiropyran and that of zwitterionic merocyanine isomer in the poly(NSP-*co*-MMA)-coated glass plates was observed in this study.

Light-Induced Detachment of Cells. Light-induced detachment of platelets and mesenchymal stem (KUSA-A1) cells was also examined. Figure 3 shows KUSA-A1 cells on poly(NSP-*co*-MMA)-coated glass plates before and after UV irradiation. After UV irradiation to the poly(NSP-*co*-MMA)-coated glass plates, KUSA-A1 cells were rarely observed on the surface of poly(NSP-*co*-MMA)-coated glass plates. On the other hand, KUSA-A1 cells remained attached to the surface of poly(NSP-*co*-MMA)-coated glass plates submitted to the same procedures, but not to be exposed to UV irradiation. The cell density of KUSA-A1 cells and platelets on the surface of poly(NSP-*co*-MMA)-coated glass plates before and after UV irradiation was examined and summarized in Figure 4a. In addition, light-induced detachment of KUSA-A1 cells and platelets was examined on

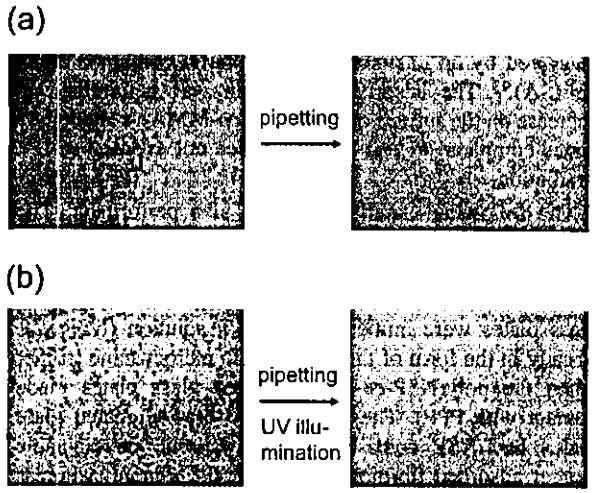

Figure 3. Light-induced detachment of KUSA-A1 cells. (a) KUSA-A1 cells on poly(NSP-*co*-MMA)-coated glass plates. (b) KUSA-A1 cells on PMMA-coated glass plates.

PMMA-coated glass plates, as a nonphotosensitive surface (Figure 4b). Light-induced detachment of KUSA-A1 cells was clearly not observed on PMMA-coated glass plates, indicating that the cell detachment on poly(NSP-*co*-MMA)-coated glass plates upon UV irradiation is not due to the stimulation of the cells by UV light irradiation. Thus, it is thought to be caused by the change in the surface energy and/or the change in the switching movement of closed nonpolar spiropyran to the polar zwitterionic merocyanine isomer upon UV irradiation.

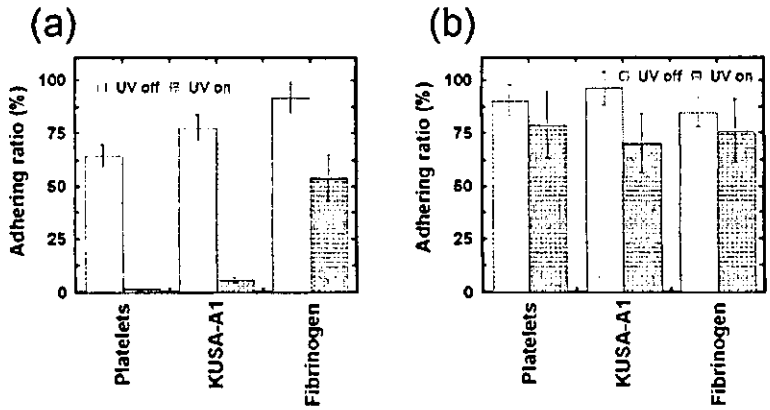

Figure 4. Light-induced detachment of platelets, KUSA-A1 cells and fibrinogen. (a) Platelets, KUSA-A1 cells, and fibrinogen on poly(NSP-*co*-MMA)-coated glass plates. (b) Platelets, KUSA-A1 cells, and fibrinogen on PMMA-coated glass plates.

The cell viability of the detached cells by UV irradiation was found to be more than 98% based on the trypan blue exclusion test.

Light-Induced Detachment of Fibrinogen. Suppression of platelet adhesion is generally believed to be due to a reduction of protein adsorption, particularly fibrinogen, which binds to the platelet membrane glycoprotein GP IIb-IIIa.[13,14] Light-induced detachment of fibrinogen adsorbed on poly(NSP-*co*-MMA) coated glass plates was, therefore, examined (Figure 4a). After the poly(NSP-*co*-MMA) coated glass plates were immersed in platelet-poor plasma (PPP) solution for 30 min, UV light was illuminated on the surface for 4 min, and the fibrinogen adsorbed on the surface was directly measured by an enzyme-immunoglobulin conjugate assay (ELISA).[13] The adsorbed fibrinogen was determined to decrease on the surface of poly(NSP-*co*-MMA)-coated glass plates compared to that on the surface not exposed to UV irradiation.

The following experiments were also performed: The adsorbed amount of fibrinogen was measured on the poly(NSP-*co*-MMA)-coated glass plates where UV light was irradiated for 4 min before the poly(NSP-*co*-MMA)-coated glass plates were immersed into PPP solution (NSP was already in the form of the zwitterionic merocyanine isomer when the poly(NSP-*co*-MMA)-coated glass plates made contact with PPP). The poly(NSP-*co*-MMA) coated glass plates had NSP consisting of zwitterionic merocyanine isomer adsorbed fibrinogen at 6.2 ± 0.5 μg/cm^2, whereas poly(NSP-*co*-MMA) coated glass plates had NSP consisting of nonionic spiropyran adsorbed at 5.0 ± 0.4 μg/cm^2. These findings indicate that the amount of adsorbed fibrinogen on the glass plates with the zwitterionic merocyanine isomer NSP was 1.2 times higher than that with the nonionic spiropyran NSP. These finding were not in agreement with the results shown in Figure 4a. This contradiction suggests that the fibrinogen was detached by means of the switching movement of closed nonpolar spiropyran to the polar zwitterionic merocyanine isomer, and the surface energy (hydrophobicity−hydrophilicity) does not directly contribute to the amount of adsorbed fibrinogen on the poly(NSP-*co*-MMA) coated glass plates. Light-induced detachment of fibrinogen was also examined on the PMMA-coated glass plates, as a nonphotosensitive surface and is shown in Figure 4b. No light-induced detachment of fibrinogen was observed on the PMMA-coated glass plates. This further indicates that fibrinogen detachment by UV irradiation is not due to the stimulation of fibrinogen by UV light, but a result of the change in the switching movement of closed nonpolar spiropyran to the polar zwitterionic merocyanine isomer upon UV irradiation.

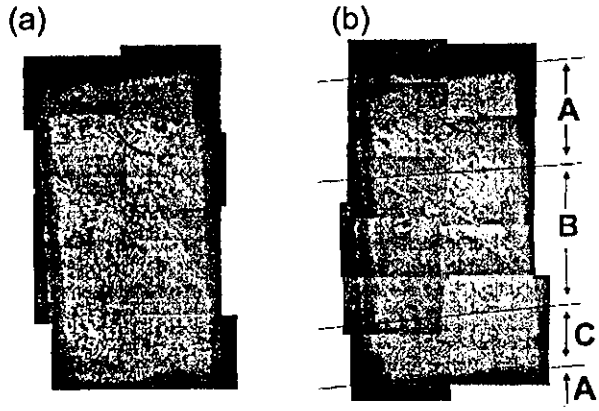

Figure 5. Light-induced detachment of KUSA-A1 cells with patterned light irradiation using striped pattern mask with 2.5 mm widths. (a) KUSA-A1 cells on poly(NSP-*co*-MMA)-coated glass plates before UV irradiation. (b) KUSA-A1 cells on poly(NSP-*co*-MMA)-coated glass plates after striped pattern UV irradiation for 4 min. Regions A and B indicate the area under non-UV irradiation (region A) and UV irradiation (region B), respectively. Region C indicates the border region.

Cell Detachment by Patterned Light Irradiation. Light-induced detachment of KUSA-A1 cells was achieved by simple patterned light irradiation using a striped pattern mask (width; 2.5 mm). Figure 5 shows KUSA-A1 cells on poly(NSP-*co*-MMA)-coated glass plates before and after UV irradiation with the striped pattern. KUSA-A1 cells were clearly observed to detach in the region exposed to UV irradiation (region B), whereas the cells remained attached in the masked area, not exposed to UV irradiation (region A). Therefore, patterned light detachment of KUSA-A1 cells was successively performed by the patterned light irradiation using the striped pattern mask.

Conclusion

KUSA-A1 cells, platelets, and the fibrinogen were detached by means of the switching movement of closed nonpolar spiropyran to the polar zwitterionic merocyanine isomer, and the surface energy (hydrophobicity−hydrophilicity) does not directly contribute to the amount of adsorbed fibrinogen on the poly(NSP-*co*-MMA)-coated glass plates.

Light-induced detachment of cells on poly(NSP-*co*-MMA) surfaces will lead to mild isolation of cells. Furthermore, this will be a powerful tool for surface marker analysis using flow cytometry. This is because this method does not require the addition of trypsin for cell detachment, which degrades the extracellular matrix and cell adhering molecules between cells and the culture flask. Furthermore, patterned light detachment of nerve cells will provide an alternative method to create micro-patterned neuronal networks,[15] which are typically performed using a micro-contact printing method.[16]

Acknowledgment. We thank Dr. M. Kameda, Dr. K. Sumaru, and Dr. T. Kanamori (AIST) for their gift of poly(NSP-*co*-MMA). This research was partially supported from a Grant-in-Aid for Exploratory Research (No. 14655136) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. This research was also supported from Asahi Glass Foundation.

References and Notes

(1) Lee, H. J.; Fermin, D. J.; Corn, R. M. *Electrochem. Commun.* **1998**, *1*, 190.
(2) Digilov, R. *Langmuir* **2000**, *16*, 6719.
(3) Liang, L.; Shi, M.; Viswanathan, V. V.; Peurrung, L. M.; Young, J. S. *J. Membr. Sci.* **2000**, *177*, 97.
(4) Yakushiji, T.; Sakai, K.; Kikuchi, A.; Aoyagi, T.; Sakurai Y.; Okano T. *Langmuir* **1998**, *14*, 4657.
(5) Garcia, A. A.; Cherian, S.; Park, J.; Gust, D.; Jahnle, F.; Rosario, R. *J. Phys. Chem. A* **2000**, *104*, 6103.
(6) Ichimura, K.; Oh, S.; Nakagawa, M. *Science* **2000**, *288*, 1624.
(7) Rosario, R.; Gust, D.; Hayes, M.; Jahnke, F.; Springer, J.; Garcia, A. A. *Langmuir* **2002**, *18*, 8062.
(8) Lee, S.; Laibinis, P. E. *J. Am. Chem. Soc.* **2000**, *122*, 5395.
(9) Vo-Dinh, T.; Alarie, J. P.; Isola, N.; Landis, D.; Wintenberg, A. L.; Ericson, M. N. *Anal. Chem.* **1999**, *71*, 358.
(10) Maims, C.; Hulme, J.; Fielden, P. R.; Goddard, N. J. *Actuator B, Chem.* **2001**, *77*, 671.
(11) Hirose, M.; Kwon, O. H.; Yamato, M.; Kikuchi, A.; Okano, T. *Biomacromolecules* **2000**, *1*, 377.
(12) Ratner, B. R.; Horbett, T.; Hoffman, A. S.; Hauschka, S. D. *J. Biomed. Mater. Res.* **1975**, *9*, 407.
(13) Higuchi, A.; K. Sugiyama, K.; Yoon, B. O.; Sakurai, M.; Hara, M.; Sumita, M.; Sugahara, S.; Shirai, T. *Biomaterials* **2003**, *24*, 3235.
(14) Phillips, D. R.; Charo, I. F.; Parise, L. V.; Fitzgerald, L. A. *Blood* **1988**, *71*, 831.
(15) Higuchi, A.; Kitamura, H.; Shishimine, K.; Konishi, S.; Yoon, B. O.; Hara, M. *J. Biomat. Sci., Polym. Ed.* **2003**, *14*, 1377.
(16) Kam, L.; Shain, W.; J. N. Turner, J. N.; Bizios, R. *Biomaterials* **2001**, *22*, 1049.
(17) Imai, Y.; Adachi, K.; Naka, K.; Chujo Y. *Polym. Bull.* **2000**, *44*, 9.
(18) Labsky J.; Koropecky, I.; Nespurek, S.; Katal. J. *Polym. J.* **1981**, *17*, 309.
(19) Kameda, M.; Sumaru, K.; Kanamori, T. in preparation.
(20) Higuchi, A.; Tamiya, S.; Tsubomura, T.; Katoh, A.; Cho, C. S.; Akaike, T.; Hara, M. *J. Biomater. Sci. Polymer Ed.* **2000**, *11*, 149.
(21) Umezawa, A.; Maruyama, T.; Segawa, K.; Shadduck, R. K.; Waheed, A.; Hata, J. *J. Cell. Physiol.* **1992**, *151*, 197.

BM049737X

BIOLOGY OF REPRODUCTION 70, 415–418 (2004)
Published online before print 1 October 2003.
DOI 10.1095/biolreprod.103.020271

Nuclear Transfer of Adult Bone Marrow Mesenchymal Stem Cells: Developmental Totipotency of Tissue-Specific Stem Cells from an Adult Mammal[1]

Yoko Kato,[3] Hideaki Imabayashi,[4] Taisuke Mori,[4] Tetsuya Tani,[3] Masanori Taniguchi,[5] Mikihiko Higashi,[5] Michio Matsumoto,[5] Akihiro Umezawa,[4] and Yukio Tsunoda[2,3]

Laboratory of Animal Reproduction,[3] College of Agriculture, Kinki University, Nara 631-8505, Japan
Department of Pathology,[4] Keio University School of Medicine, Tokyo 160-8582, Japan
Kumamoto Prefecture Livestock Stations and Animal Public Health Center,[5] Kumamoto 861-1113, Japan

ABSTRACT

Recent studies have demonstrated that somatic stem cells have a flexible potential greater than previously expected when they are transplanted into different tissues. On the other hand, recent studies also have revealed that these potentials might occur because of spontaneous cell fusion with recipient cells. The nuclei of somatic cells could have been reprogrammed when they were artificially or spontaneously fused with mouse embryonic stem (ES) cells. The resultant hybrid cells acquired a developmental pluripotency that the original somatic cells did not have but that ES cells did. LaBarge and Blau (Cell 2002; 111: 589–601) demonstrated that adult bone marrow-derived cells contributed to muscle tissue in a stepwise biological progression. This means that bone marrow-derived cells became satellite cells of mononucleate muscle stem cells after the first irradiation-induced damage to the mouse, and after the second irradiation-induced damage, multinucleate myofibers appeared from the bone marrow-derived cells. Considered together, the differentiation potential of the somatic stem cell nucleus itself remains unclear. Although the pluripotency of somatic stem cell populations has been evaluated, the developmental totipotency of the nuclei of somatic stem cells, whether or not they fused with other cells, has not been shown, except in only one study concerning fetal neural cells (never in adult stem cells). Here, we showed the developmental totipotency of adult bovine mesenchymal stem cells by nuclear transfer.

assisted reproductive technology, developmental biology, early development, embryo, gamete biology

INTRODUCTION

Somatic stem cells were recently identified in hematopoietic [1], hepatic [2], epidermal [3], gastrointestinal [4], neural [5, 6], muscular [6], and bone marrow [6–8] tissues. Many researchers have demonstrated the developmental pluripotency of somatic stem cells of other tissues. Bone marrow-derived stem cells can be trans-differentiated into multilineage cells, such as muscle [9] of mesoderm, lung [10] and liver [10, 11] of endoderm, and brain [12–15] and skin [10] of ectoderm. Moreover, adult neural stem cells generated all germ layers in chimera [16]. Thus, somatic stem cells were evaluated to determine the extent of their potential for trans-differentiation in other tissues. It has been suggested that somatic stem cells are more desirable than embryonic stem (ES) cells for cell therapeutics because of ethical considerations and possible immunologic rejection of ES cells. The developmental totipotency of somatic stem cells, however, remains unknown.

Recent studies demonstrated that somatic stem cells have a more flexible potential than expected when put into different tissues. On the other hand, recent studies also revealed the possibility that this high potential might develop because of spontaneous cell fusion with recipient ES cells [17, 18]. Somatic cell nuclei can be reprogrammed when artificially [17] or spontaneously [17, 18] fused with mouse ES cells. The resulting hybrid cells acquire a developmental pluripotency that the original somatic cells do not have, but that ES cells do. LaBarge and Blau [19], however, demonstrated that adult bone marrow-derived cells contributed to muscle tissue in a stepwise biological progression. This finding indicates that bone marrow-derived cells became satellite cells of mononucleate muscle stem cells after the first irradiation-induced damage, and after the second irradiation-induced damage, multinucleate myofibers appeared from the bone marrow-derived cells.

Thus, the differentiation potential of somatic stem cell nuclei remains unclear. Although the pluripotency of somatic stem cells has been evaluated by fusion with other cells, the developmental totipotency of somatic stem cell nuclei has not been demonstrated in adult stem cells; however, it has been demonstrated in one report using fetal neural cells [20].

In the present study, we report the developmental totipotency of adult bovine mesenchymal stem cells by nuclear transfer. To our knowledge, this is the first report of developmental totipotency of tissue-specific stem cells derived from the adult mesodermal cell lineage. The present study clearly demonstrates that pluripotent bovine mesenchymal stem cells derived from adult bone marrow have developmental totipotency by nuclear transfer.

MATERIALS AND METHODS

All experiments and protocols were carried out in strict accordance with the Guiding Principles for the Care and Use of Research Animals promulgated by the Kinki University Committee on Animal Research and Bioethics.

Bone marrow mesenchymal stem cells were cultivated from female and male bovine stromal cells isolated from a femur, sternum, and rib [14, 21–23] by flushing the bone. Collected cells were cultured in 20% fetal bovine serum (FBS) supplemented with Dulbecco modified Eagle medium

[1]Supported by grants from the Programme for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN) and partly supported by a grant from the Ministry of Education, Science, and Culture (14034259, 12358014, 13308050, 15039233); the Ministry of Agriculture, Forestry and Fisheries (Cloning); Ito Kinen Foundation; and Nakajima International Interchange Foundation.
[2]Correspondence. FAX: 81 7 4243 115; e-mail: tsunoda@nara.kindai.ac.jp

Received: 12 June 2003.
First decision: 1 July 2003.
Accepted: 30 September 2003.

ISSN: 0006-3363. http://www.biolreprod.org

TABLE 1. In vitro development of nuclear transferred oocytes with adult bovine mesenchymal stem cells.

Used oocytes (n)	Fused oocytes (n [%])	Cultured oocytes (n)	Oocyte development (n [%])		
			2-Cell stage	Morula stage	Blastocyst stage
168[a]	152(90)	152	128(84)	61(40)	59(39)
161[b]	122(76)	122	71(58)	12(10)	8(7)
Total					
329	274(83)	274	199(73)	73(27)	67(24)

[a] Oocytes were collected from fresh ovaries.
[b] Oocytes were collected from ovaries preserved for 24 h at 10°C.

(DMEM; high glucose) for several days in vitro and subcultured several times before use.

Recipient ooplasm was obtained from enucleated metaphase II-phase oocytes that were derived from in vitro culture for 22–24 h (TCM-199, catalog no. 31100-027, supplemented with 10% FBS; Gibco BRL, Invitrogen, Carlsbad, CA) of immature germinal vesicle-phase oocytes residing in ovarian follicles. In some case, ovaries that had been preserved overnight at a low temperature were used for collection of oocytes because of a national guideline for avoiding bovine spongiform encephalopathic infection from slaughterhouses in Japan. The best temperature at which to preserve ovaries overnight was previously determined. Judging by the in vitro maturation rate of immature oocytes and the in vitro developmental potential after in vitro fertilization of in vitro-matured oocytes, 10°C preserves ovaries overnight in saline solution (Matsushita et al., unpublished observation). Thus, we preserved the ovaries at 10°C until the infection test for bovine spongiform encephalopathy was completed.

Nuclear transfer methods were performed as described previously [24–26]. Quiescent donor cells, which were produced by serum starvation or contact inhibition, were used. A single donor cell was electrically fused with oocytes immediately after enucleation using two direct current (DC) pulses of 150 V/mm for 25 μsec in Zimmerman fusion medium [24]. Two fusion pulses were administered at 15-min intervals until fusion was achieved. Fused oocytes were further stimulated by electrical pulses (DC pulses of 20 V/mm for 20 μsec) to confirm activation. Reconstituted oocytes were immediately treated with cycloheximide (10 μg/ml) in CR1-aa medium with 3 mg of bovine serum albumin (fatty acid-free) for 5–6 h. After treatment, oocytes were cultured in cycloheximide-free medium. On the third day (first day = day of nuclear transfer), embryos were transferred to CR1-aa medium supplemented with 10% FBS and cocultured with mouse embryonic fibroblast cells, which are routinely used for mouse ES cell culture [27]. On the seventh day of culture, they were moved to modified DMEM supplemented with 10% FBS.

To determine the pluripotency of the donor cells, the cells were induced to differentiate into osteogenic precursor cells or adipocytes [17, 28]. Cells were stained with von Kossa and alkaline phosphatase for the osteogenic precursor cells, Oil-Red-O for adipocytes, and antibodies against nestin and glial fibrillary acidic protein (GFAP) for neuroectodermal differentiation. Moreover, specific antibodies were used to check the differentiation status and stem cell specificity, such as CD29, CD44, CD166, CD14, CD31, CD34, CD45, and CD117, which are routinely used as markers for human mesenchymal stem cells.

Data were analyzed using the chi-square and Student *t*-tests.

RESULTS

Some cells induced for specific differentiation were positive for von Kossa or alkaline phosphatase, indicating differentiation into osteogenic progenitor cells. In another culture, cells were positive for Oil-Red-O, indicating that they differentiated into adipocytes. Some cells were positive for nestin and GFAP antibodies, demonstrating successful neuroectodermal differentiation [18, 28].

When several antibodies used as markers of mesenchymal stem cells in human and mouse were examined, cells were positive for CD29, CD44, and CD166 and negative for CD14, CD31, CD34, CD45, and CD117. These results were consistent with those from human mesenchymal stem cells. The cells were negative, however, against CD90, CD105, and CD140a, which are usually positive for human and mouse mesenchymal stem cells.

TABLE 2. In vivo development after transfer to recipient females.

Preservation time of ovaries (h)	Embryos transferred (n)	Recipient females (n)	Pregnancies (n)[a]	Young (n)
0	2	2	1	0
24	11	8	2	1
Total	13	10	3	1

[a] On Day 40.

We preliminarily examined whether oocytes collected from preserved ovaries have the same ability to mature in vitro to support development after in vitro fertilization and nuclear transfer with cumulus cells. We concluded that oocytes from ovaries preserved at 10°C overnight had an ability similar to that of those in the nonpreserved group (Matsushita et al., unpublished observation). When bovine mesenchymal stem cells were fused with enucleated metaphase II-phase oocytes with or without preservation and then electrically activated, they developed to the blastocyst stage in vitro after 8–9 days (Table 1). The proportion of blastocysts after nuclear transfer of mesenchymal stem cells was significantly decreased in the preserved group (7% vs. 39%), but after transfer to the foster mother, two recipients with embryos from preserved oocytes and one from nonpreserved oocytes were pregnant at Day 40. One female each from the preserved and nonpreserved groups was aborted between 80 and 120 days and between 40 and 80 days, respectively. One healthy offspring was obtained from the preserved group (Table 2). The results of the microsatellite marker analysis verified that donor cells were the source of the genetic material used to produce the cloned bovine.

DISCUSSION

Recently, Jiang et al. [8] reported that mesenchymal stem cells derived from adult mouse bone marrow have pluripotency in vitro. Those authors also demonstrated that mesenchymal stem cells differentiated into almost all tissues of adult mouse after injection of a single cell into blastocysts. This finding that tissue-specific stem cells could act like ES cells was extremely important. On the other hand, two different groups [17, 18] recently demonstrated that the high potential of somatic stem cells residing in other tissues might develop because of spontaneous cell fusion with other cells. These authors demonstrated that these stem cells have the ability to fuse with ES cells spontaneously in vitro. Once fused, the resulting hybrid cells acquire the developmental pluripotency of the ES cells. A similar observation was reported when somatic cells were artificially fused with ES cells [29]. Thus, it has remained unclear whether somatic stem cell nuclei have the ability to trans-differentiate. The results of the present study, however, clearly indicate that adult bovine mesenchymal stem cells have developmental totipotency after nuclear transfer.

Recently developed animal cloning technology using adult somatic cells by nuclear transfer [30] has several advantages, one of which is that the resulting animals are exactly the same as the provider of the donor cells. Compared with nuclear transfer of germ line cells, somatic cells are easily collected and cultured without serious injury to animals. Also, somatic cells do not undergo the essential modification of imprinted genes that occurs during the production of germ cells. When primordial germ cells (PGCs)

were used as donors for nuclear transfer, the resulting fetuses were inviable at Day 10.5 because of the erasing and abnormal expression of imprinted genes that occur when PGCs differentiate into functional germ cells [31, 32]. In contrast, the results of nuclear transfer using ES cells largely depend on the cell lines, number of passages, and background of the mouse strains of the ES cells used [33–36]. The ES cells derived from inner-cell-mass cells of a blastocyst, which include a subset of cells that are PGCs, should develop into a fetus. Therefore, some ES cells must also undergo modification of imprinted genes during in vitro culture. Thus, somatic cells have advantages for nuclear transfer, and a solid nuclear transfer system using somatic cells must be established. For this reason, it is essential to determine the somatic cell type best suited for nuclear transfer to produce normal young.

Because the efficiency of somatic cell cloning remains low and peri- and postnatal death of young remains high [37], it is still too early to apply this method in human clinics or farm animal production without further basic research to solve the problems of abnormal development. Previously [25], we evaluated the cell type most appropriate for efficient cloning in bovines. We demonstrated that among the cell types tested, cumulus cells are good candidate nuclear donors, judging by the incidence of peri- and postnatal death of young. Even in cumulus cells, however, the potential after nuclear transfer depends on the cell line [25]. It is unclear both why and how cell lines affect developmental totipotency after nuclear transfer. Although cumulus cells are good cloning candidates, they can be obtained only from female animals. On the other hand, mesenchymal stem cells can be obtained from both female and male living animals. In the present study, we examined whether this cell type is a suitable nuclear donor candidate. As the results show, adult mesenchymal stem cells are good cloning candidates (one healthy clone was obtained).

Yamazaki et al. [20] demonstrated the developmental totipotency of embryonic neural cells in the cerebral cortex of mouse postimplantation embryos after nuclear transfer. They reported that the cloning efficiency of premature and early postmitotic neural cells from the ventricular side of the cortex was higher than that of postmitotic differentiated neurons from the pial side of the cortex and also higher than that of other somatic cells. Those authors concluded that for cloning, premature somatic stem cells in fetuses are better than differentiated somatic stem cells. This hypothesis is consistent with the previous success of Hochedlinger and Jaenisch [38]. Those authors obtained cloned mice derived from terminally differentiated, mature B and T cells using a two-step method, but they never succeeded with a simple nuclear transfer. In the present study, we demonstrated that adult bovine stem cells have preserved developmental totipotency.

Moreover, in most studies, the derivation of somatic cells used as donors was not clear. For example, most cells cultivated from tissues such as skin and liver were probably fibroblast cells, the viability of which might be higher than that of other specific cell types in tissues. In the present study, we examined the characteristics of bone marrow cells in detail using specific antibodies and stains. Bovine bone marrow mesenchymal stem cells had developmental pluripotency in vitro. They differentiated into not only osteogenic cells but also adipocytes, as examined by the in vitro culture experiments. Because no markers have been established for bovine mesenchymal stem cells [28], we used several markers for human mesenchymal stem cells. We do not know why CD90, CD105, and CD140a, which are positive markers for human mesenchymal stem cells, were negative in this investigation, but species-specific effects of these antibodies might be involved. Whether the cloned offspring were derived from such pluripotent cells is unclear, because it is technically impossible to determine precisely. However, as a group, our cell population was pluripotent, and the cloned animal was born from pluripotent mesenchymal stem cells.

In conclusion, the present study clearly demonstrates that bone marrow mesenchymal stem cells can be isolated and identified from adult bovine and that adult mesenchymal stem cells have developmental totipotency after nuclear transfer.

ACKNOWLEDGMENTS

We thank all staff members at the Kumamoto Prefecture Livestock Stations and Animal Public Health Center for embryo transfer, assistance, and management of recipient animals.

REFERENCES

1. Weissman IL. Translating stem and progenitor cell biology to the clinic: barriers and opportunities. Science 2000; 287:1442–1446.
2. Alison M, Sarraf C. Hepatic stem cells. J Hepatol 1998; 29:676–682.
3. Watt F. 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. Stem cells in gastrointestinal epithelium: numbers, characteristics and death. Philos Trans R Lond B Biol Sci 1998; 353:821–830.
5. Gage FH. Mammalian neural stem cells. Science 2000; 287:1433–1438.
6. Jiang Y, Vaessen B, Lenvik T, Blackstad M, Reyes M, Verfaillie CM. Multipotent progenitor cells can be isolated from postnatal murine bone marrow, muscle, and brain. Exp Hematol 2002; 30:896–904.
7. Fridenshtein A. Stromal bone marrow cells and the hematopoietic microenvironment. Arkh Patol 1982; 44:3–11.
8. Jiang Y, Jahagirdar BN, Reinhardt RL, Schwartz RE, Keenek CD, Ortiz-Gonzalezk XR, Reye M, Lenvik T, Lund T, Blackstad M, Du J, Aldrich S, Lisberg A, Lowk WC, Largaespada DA, Verfaillie CM. Pluripotency of mesenchymal stem cells derived from adult marrow. Nature 2002; 418(6893):41–49.
9. 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.
10. Krause DS, Theise ND, Collector MI, Henegariu O, Hwang S, Gardner R, Neutzel S, Sharkis SJ. Multiorgan, multilineage engraftment by a single bone marrow-derived stem cell. Cell 2001; 105:369–377.
11. Petersen BE, Bowe WC, Patrene KD, Mars WM, Sullivan AK, Murase N, Boggs SS, Greenberger JS, Goff JP. Bone marrow as a source of hepatic oval cells. Science 1999; 284:1168–1170.
12. Sanchez-Ramos J, Song S, Cardozo-Pelaez F, Hazzi C, Stedeford T, Willing A, Freeman TB, Saporta S, Janssen W, Patel N, Cooper DR, Sanberg PR. Adult bone marrow stromal cells differentiate into neural cells in vitro. Exp Neurol 2000; 164:247–256.
13. Woodbury D, Schwarz EJ, Prockop DJ, Black IB. Adult rat and human bone marrow stromal cells differentiate into neurons. J Neurosci Res 2000; 61:364–370.
14. 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 demethylation agent. Differentiation 2001; 68:235–244.
15. Kim BJ, Seo JH, Bubien JK, Oh YS. Differentiation of adult bone marrow stem cells into neuroprogenitor cells in vitro. Regeneration and Transplantation 2002; 13:1185–1188.
16. Clarke DL, Johansson CB, Wilbertz J, Veress B, Nilsson E, Karlstrom H, Lendahl U, Frisen J. Generalized potential of adult neural stem cells. Nature 2000; 288:1660–1663.
17. Terada N, Hamazaki T, Oka M, Hoki M, Mastalerz DM, Nakano Y, Meyer EM, Morel L, Petersen BE, Scott EW. Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. Nature 2002; 416:542–545.

18. Ying Q, Nichols J, Evans EP, Smith AG. Changing potency by spontaneous fusion. Nature 2002; 416:545–548.
19. LaBarge MA, Blau HM. Biological progression from adult bone marrow to mononucleate muscle stem cell to multinucleate muscle fiber in response to injury. Cell 2002; 111:589–601.
20. Yamazaki Y, Makino H, Hamaguchi-Hamada K, Hamada S, Sugino H, Kawase E, Miyata T, Ogawa M, Yanagimachi R, Yagi T. Assessment of the developmental totipotency of neural cells in the cerebral cortex of mouse embryo by nuclear transfer. Proc Nat Acad Sci U S A 2001; 98:14022–14026.
21. Umezawa A, Murayama T, Segawa K, Shadduck RK, Waheed A, Hata J. Multipotent marrow stromal cell line is able to induce hematopoiesis in vivo. J Cell Physiol 1992; 151:197–205.
22. Makino S, Fukuda K, Miyoshi S, Konishi F, Kodama H, Pan J, Sano M, Takahashi T, Hori S, Abe H, Hata J, Umezawa A, Ogawa S. Cardiomyocytes can be generated from marrow stromal cells in vitro. J Clin Invest 1999; 103:697–705.
23. Hakone D, Fukuda K, Makino S, Konishi F, Tomita Y, Manabe T, Suzuki Y, Umezawa A, Ogawa S. Bone marrow-derived regenerated cardiomyocytes (CMG cells) express functional adrenergic and muscarinic receptors. Circulation 2002; 105:380–386.
24. Kato Y, Tani T, Sotomaru Y, Kurokawa K, Kato J, Doguchi H, Yasue H, Tsunoda Y. Eight calves cloned from somatic cells of a single adult. Science 1998; 282:2095–2098.
25. Kato Y, Tani T, Tsunoda Y. Cloning of calves from various somatic cell types of male and female adult, newborn and fetal cows. J Reprod Fertil 2000; 120:231–237.
26. Tani T, Kato Y, Tsunoda Y. Direct exposure of chromosomes to non-activated ovum cytoplasm is effective for bovine somatic cell nucleus reprogramming. Biol Reprod 2001; 64:324–330.
27. Robertson EJ. Embryo-derived stem cell lines. In: Robertson EJ (ed.), Teratocarcinomas and Embryonic Stem Cells. Oxford: IRL Press; 1987:71–112.
28. Deans RJ, Moseley AB. Mesenchymal stem cells: biology and potential clinical uses. Exp Hematol 2000; 28:875–884.
29. Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol 2001; 11:1553–1558.
30. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KHS. Viable offspring derived from fetal and adult mammalian cells. Nature 1997; 385:810–813.
31. Kato Y, Rideout WM III, Hilton K, Barton SC, Tsunoda Y, Surani MA. Developmental potential of mouse primordial germ cells. Development 1999; 126:1823–1832.
32. Lee J, Inoue K, Ono R, Ogonuki N, Kohda T, Kaneko-Ishino T, Ogura A, Ishino F. Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. Development 2002; 129:1807–1817.
33. Amano T, Tani T, Kato Y, Tsunoda Y. Mouse cloned from embryonic stem (ES) cells synchronized in metaphase with nocodazole. J Exp Zool 2001; 289:139–145.
34. Amano T, Kato Y, Tsunoda Y. The developmental potential of the inner cell mass of blastocysts that were derived from mouse ES cells using nuclear transfer technology. Cell Tissue Res 2002; 307:367–370.
35. Eggan K, Akutsu H, Loring J, Jackson-Grusby L, Klemm M, Rideout WM III, Yanagimachi R, Jaenisch R. Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. Proc Natl Acad Sci U S A 2001; 98:6209–6214.
36. Humpherys D, Eggan K, Akutsu H, Hochedlinger K, Rideout WM III, Biniszkiewicz D, Yanagimachi R, Jaenisch R. Epigenetic instability in ES cells and cloned mice. Science 2001; 293(5527):95–97.
37. Tsunoda Y, Kato Y. Recent progress and problems in animal cloning. Differentiation 2002; 69:158–161.
38. Hochedlinger K, Jaenisch R. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. Nature 2002; 415:1035–1038.

Transcriptome Analysis of Mouse Stem Cells and Early Embryos

Alexei A. Sharov[1], Yulan Piao[1], Ryo Matoba[1], Dawood B. Dudekula[1], Yong Qian[1], Vincent VanBuren[1], Geppino Falco[1], Patrick R. Martin[1], Carole A. Stagg[1], Uwem C. Bassey[1], Yuxia Wang[1], Mark G. Carter[1], Toshio Hamatani[1], Kazuhiro Aiba[1], Hidenori Akutsu[1], Lioudmila Sharova[1], Tetsuya S. Tanaka[1], Wendy L. Kimber[1], Toshiyuki Yoshikawa[1], Saied A. Jaradat[1], Serafino Pantano[1], Ramaiah Nagaraja[1], Kenneth R. Boheler[1], Dennis Taub[1], Richard J. Hodes[1,2], Dan L. Longo[1], David Schlessinger[1], Jonathan Keller[3], Emily Klotz[2], Garnett Kelsoe[4], Akihiro Umezawa[5], Angelo L. Vescovi[6], Janet Rossant[7], Tilo Kunath[7], Brigid L. M. Hogan[4] Anna Curci[8], Michele D'Urso[8], Janet Kelso[9], Winston Hide[9], Minoru S. H. Ko[1]*

1 National Institute on Aging, Baltimore, Maryland, United States of America, 2 National Cancer Institute, Bethesda, Maryland, United States of America, 3 Basic Research Program, SAIC-Frederick, National Cancer Institute at Frederick, Frederick, Maryland, United States of America, 4 Duke University Medical Center, Durham, North Carolina, United States of America, 5 National Research Institute for Child Health and Development, Tokyo, Japan, 6 Institute for Stem Cell Research, Ospedale San Raffaele, Milan, Italy, 7 Mount Sinai Hospital, Toronto, Ontario, Canada, 8 Institute of Genetics and Biophysics, Consiglio Nazionale delle Ricerche, Naples, Italy, 9 South African National Bioinformatics Institute, University of the Western Cape, Bellville, South Africa

Understanding and harnessing cellular potency are fundamental in biology and are also critical to the future therapeutic use of stem cells. Transcriptome analysis of these pluripotent cells is a first step towards such goals. Starting with sources that include oocytes, blastocysts, and embryonic and adult stem cells, we obtained 249,200 high-quality EST sequences and clustered them with public sequences to produce an index of approximately 30,000 total mouse genes that includes 977 previously unidentified genes. Analysis of gene expression levels by EST frequency identifies genes that characterize preimplantation embryos, embryonic stem cells, and adult stem cells, thus providing potential markers as well as clues to the functional features of these cells. Principal component analysis identified a set of 88 genes whose average expression levels decrease from oocytes to blastocysts, stem cells, postimplantation embryos, and finally to newborn tissues. This can be a first step towards a possible definition of a molecular scale of cellular potency. The sequences and cDNA clones recovered in this work provide a comprehensive resource for genes functioning in early mouse embryos and stem cells. The nonrestricted community access to the resource can accelerate a wide range of research, particularly in reproductive and regenerative medicine.

Introduction

With the derivation of pluripotent human embryonic stem (ES) (Thomson et al. 1998) and embryonic germ (EG) (Shamblott et al. 1998) cells that can differentiate into many different cell types, excitement has increased for the prospect of replacing dysfunctional or failing cells and organs. Very little is known, however, about critical molecular mechanisms that can harness or manipulate the potential of cells to foster therapeutic applications targeted to specific tissues.

A related fundamental problem is the molecular definition of developmental potential. Traditionally, potential has been operationally defined as "the total of all fates of a cell or tissue region which can be achieved by any environmental manipulation" (Slack 1991). Developmental potential has thus been likened to potential energy, represented by Waddington's epigenetic landscape (Waddington 1957), as development naturally progresses from "totipotent" fertilized eggs with unlimited differentiation potential to terminally differentiated cells, analogous to a ball moving from high to low points on a slope. Converting differentiated cells to pluripotent cells, a key problem for the future of any stem cell-based therapy, would thus be an "up-hill battle," opposite the usual direction of cell differentiation. The only current way to do this is by nuclear transplantation into enucleated oocytes, but the success rate gradually decreases according to developmental stages of donor cells, providing yet another operational definition of developmental potential (Hochedlinger and Jaenisch 2002; Yanagimachi 2002).

What molecular determinants underlie or accompany the potential of cells? Can the differential activities of genes provide the distinction between totipotent cells, pluripotent cells, and terminally differentiated cells? Systematic genomic methodologies (Ko 2001) provide a powerful approach to these questions. One of these methods, cDNA microarray/chip technology, is providing useful information (Ivanova et al. 2002; Ramalho-Santos et al. 2002; Tanaka et al. 2002),

Received August 4, 2003; Accepted October 13, 2003; Published December 22, 2003
DOI: 10.1371/journal/pbio.0000074

Abbreviations: ATCC, American Type Culture Collection; 2D, two dimensional; 3D, three dimensional; EG, embryonic germ (cell); ES, embryonic stem (cell); EST, expressed sequence tag; FDR, false discovery rate; GAP-DH, glyceraldegyde-3-phosphate dehydrogenase; HS, hematopoietic stem/progenitor (cell); LIF, leukemia inhibitory factor; MS, mesenchymal stem (cell); NCBI, National Center for Biotechnology Information; NIA, National Institute on Aging; NS, neural stem/progenitor (cell); ORF, open reading frame; PC1, first principal component; PCA, principal component analysis; PGC, primordial germ cell; TS, trophoblast stem (cell); VRML, virtual reality modeling language

Academic Editor: Patrick Tam, University of Sydney

*To whom correspondence should be addressed. E-mail: KoM@grc.nia.nih.gov

although analyses have been restricted to a limited number of genes and cell types. To obtain a broader understanding of these problems, it is important to analyze all transcripts/genes in a wide selection of cell types, including totipotent fertilized eggs, pluripotent embryonic cells, a variety of ES and adult stem cells, and terminally differentiated cells. Despite the collection of a large number of expressed sequence tags (ESTs) (Adams et al. 1991; Marra et al. 1999) and full-insert cDNA sequences (Okazaki et al. 2002), systematic collection of ESTs on these hard-to-obtain cells and tissues has been done previously only on a limited scale (Sasaki et al. 1998; Ko et al. 2000; Solter et al. 2002).

Accordingly, we have attempted to (i) complement other public collections of mouse gene catalogs and cDNA clones by obtaining and indexing the transcriptome of mouse early embryos and stem cells and (ii) search for molecular differences among these cell types and infer features of the nature of developmental potential by analyzing their repertoire and frequency of ESTs. Here we report the collection of approximately 250,000 ESTs, enriched for long-insert cDNAs, and signature genes associated with the potential of cells, various types of stem cells, and preimplantation embryos.

Results and Discussion

Novel Genes Derived from Early Mouse Embryos and Stem Cells

Twenty-one long-insert-enriched cDNA libraries with insert ranges from 2–8 kb (Piao et al. 2001) were generated from preimplantation embryos (unfertilized egg, fertilized egg, two-cell embryo, four-cell embryo, eight-cell embryo, morula, and blastocyst), ES cells (Anisimov et al. 2002) and EG cells (Matsui et al. 1992), trophoblast stem (TS) cells (Tanaka et al. 1998), adult stem cells (e.g., neural stem/progenitor [NS] cells) (Galli et al. 2002), mesenchymal stem (MS) cells (Makino et al. 1999), osteoblasts (Ochi et al. 2003), and hematopoietic stem/progenitor (HS) cells (Ortiz et al. 1999), their differentiated cells, and newborn organs (e.g., brain and heart) (see Protocol S1 and Dataset S1 for methods, full list of libraries, and references). In total, 249,200 ESTs (170,059 cDNA clones: 114,437 5′ ESTs and 134,763 3′ ESTs) were generated and assembled together with public data into a gene index (see Materials and Methods; Protocol S1).

Of 29,810 mouse genes identified in our gene index (Figure 1; Dataset S2; Dataset S3), 977 were not present as either known or predicted transcripts in other major transcriptome databases, such as RefSeq (Pruitt and Maglott 2001), Ensembl (Hubbard et al. 2002), and RIKEN (Okazaki et al. 2002) (see Dataset S3 for details and Dataset S4 for sequences). These genes represent possible novel mouse genes, as they either encode open reading frames (ORFs) greater than 100 amino acids or have multiple exons. In particular, 554 of the 977 genes remained novel with high confidence even after more thorough searches against GenBank and other databases. Comparisons of these 977 genes against all National Center for Biotechnology Information (NCBI) UniGene representative sequences showed that 377 genes did not match even fragmentary ESTs and are therefore unique to the National Institute on Aging (NIA) cDNA collection (see Dataset S3). A random subset of 19 cDNA clones representing these genes was sequenced completely to confirm their novelty (Figure 2). Protein domain searches using InterPro (Mulder et al. 2003)

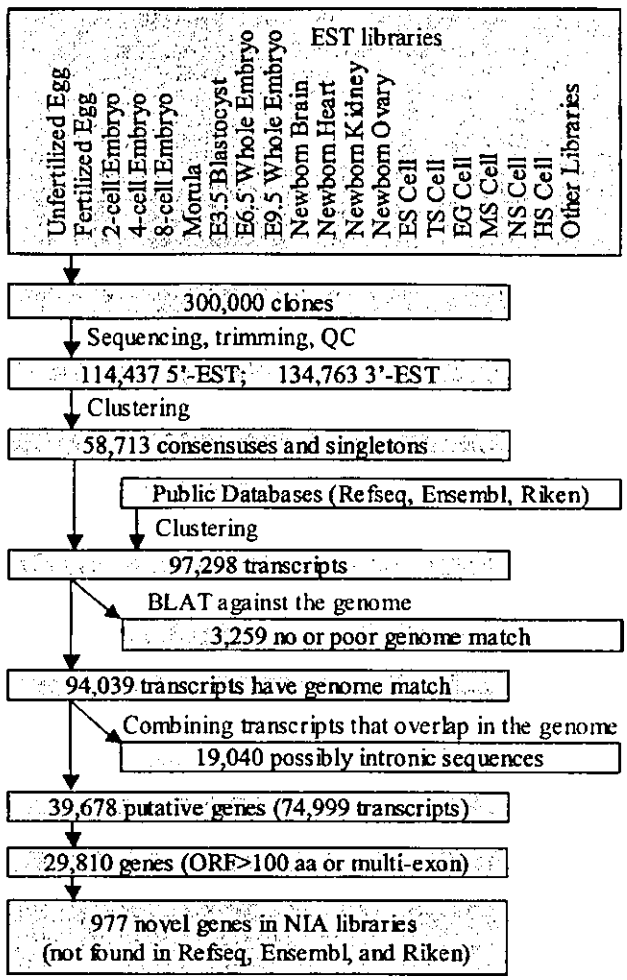

Figure 1. Flow Chart of Sequence Data Analysis

Using TIGR gene indices clustering tools (Pertea et al. 2003), 249,200 ESTs were clustered, generating 58,713 consensuses and singletons. NIA consensuses and singletons were further clustered with Ensembl transcripts (Hubbard et al. 2002), RIKEN transcripts (Okazaki et al. 2002), and RefSeq transcripts and transcript predictions (Pruitt and Maglott 2001). Alignments of these sequences to the mouse genome (UCSC February 2002 freeze data, available from ftp://genome.cse.ucsc.edu/goldenPath/mmFeb2002) (Waterston et al. 2002) using BLAT (Kent 2002) helped to avoid false clustering of similar sequences at nonmatching genome locations. Erroneous clusters were reassembled based on the analysis of genome alignment. A total 94,039 putative transcripts were thus generated and then grouped into 39,678 putative genes based on their overlap in the genome on the same chromosome strand and on clone-linking information. Using criteria of an ORF greater than 100 amino acids or of multiple exons (excluding sequences that are potentially located in a wrong strand), 29,810 mouse genes were identified. Finally, 977 genes unique to the NIA database were identified.

DOI: 10.1371/journal/pbio.0000074.g001

revealed that one of them, *U004160*, is an orthologue of human gene Midasin (*MDN1*), but the remaining 18 genes do not encode any known protein motifs. However, they were split into multiple exons in the alignment to the mouse genome sequences, and we therefore considered them genes. As these sequences are mainly derived from early embryos and stem cells, they most likely represent new candidates for genes specific to particular types of stem cells. RT-PCR analysis revealed that they are expressed in specific cell types

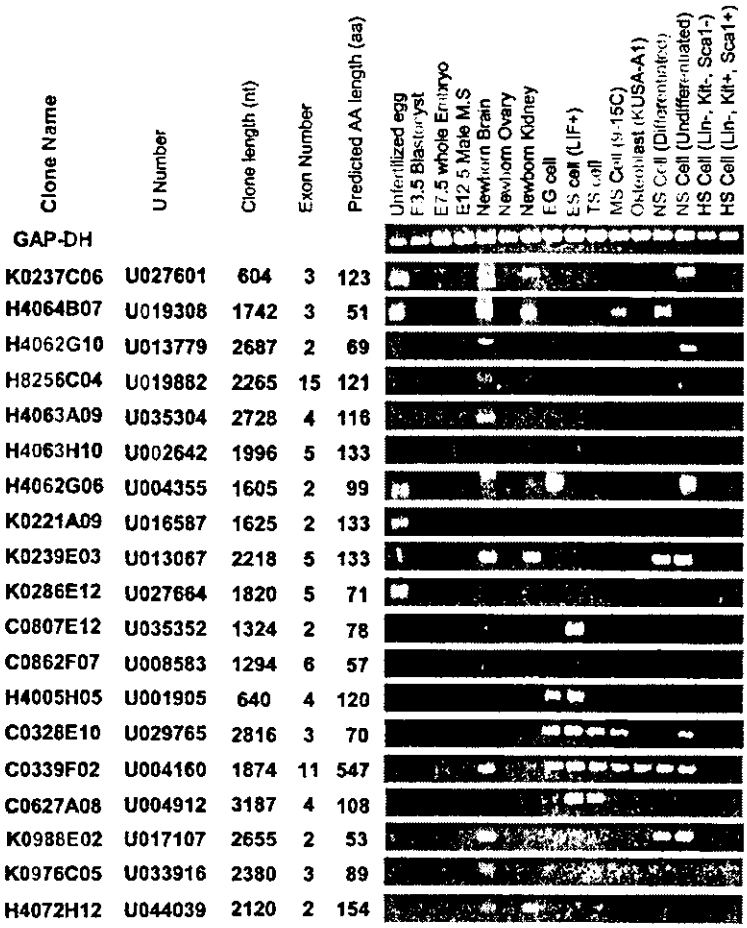

Figure 2. Examples of NIA-Only cDNA Clones and RT–PCR Results

Expression pattern of 19 novel cDNA clones in 16 different cell lines or tissues: unfertilized egg, E3.5 blastocyst, E7.5 whole embryo (embryo plus placenta), E12.5 male mesonephros (gonad plus mesonephros), newborn brain, newborn ovary, newborn kidney, embryonic germ (EG) cell, embryonic stem (ES) cell (maintained as undifferentiated in the presence of LIF), trophoblast stem (TS) cell, mesenchymal stem (MS) cell, osteoblast, neural stem/progenitor (NS) cell, NS differentiated (differentiated neural stem/progenitor cells), and hematopoietic stem/progenitor (HS) cells. Glyceraldegyde-3-phosphate dehydrogenase (GAP-DH) was used as a control. A U number is assigned to each gene in the gene index (see Dataset S2). The exon number was predicted from alignment with the mouse genome sequence, and the amino acid sequence was predicted with the ORF finder from NCBI.

DOI: 10.1371/journal/pbio.0000074.g002

(Figure 2; Dataset S5). For example, the expression of gene *U035352* was unique to ES cells, expression of *U004912* unique to ES and TS cells, and expression of *U001905* unique to ES and EG cells. In addition, one gene showed apparent specific expression in several stem cells and is thus a potential pan-stem cell marker (*U029765*). Taken together, these data suggest that most of the putative genes represented only in the NIA cDNA collection are bona fide genes that have not been previously identified.

Signature Genes That Characterize Preimplantation Embryos and Stem Cells

To identify genes that were consistently overrepresented in a given set of cDNA libraries when compared with other libraries, we performed the correlation analysis of log-transformed EST frequency combined with the false discovery rate (FDR) method (Benjamini and Hochberg 1995) (FDR = 0.1) (Figure 3; Dataset S6; Dataset S7).

First, we analyzed various combinations of preimplantation stages and identified the following genes: (i) 196 genes specific to unfertilized eggs (oocytes) and fertilized eggs (Group A in Figure 3), (ii) 122 genes specific to two- to four-cell embryos (Group B in Figure 3), (iii) 119 genes specific to eight-cell embryos, morula, and blastocyst (Group C in Figure 3), (iv) 81 genes specific to all preimplantation embryos (Group D in Figure 3), and (v) 143 genes specific to all preimplantation embryos except for blastocysts (Group E in Figure 3) (see also Dataset S7). Blastocyst EST frequencies are unique even among preimplantation embryos, most likely reflecting the switch of the transcriptome from the maternal genetic program to the zygotic genetic program (Latham and Schultz 2001; Solter et al. 2002) or to the differentiation of the trophectoderm. At least 35 out of 196 genes in the egg signature gene list (Group A in Figure 3) have ATP-related protein domains. Genes in the following categories were also enriched in this gene list: the ubiquitin–proteasome pathway, the energy pathway, cell signaling (kinase and membrane) proteins, ribosomal proteins, and zinc finger proteins. Two *SWI/SNF*-related genes (*5930405J04Rik*, the homologue of human *SMARCC2*, and *Smarcf1*) and two *Polycomb* genes (*Scmh1* and *Sfmbt*) overrepresented in eggs may be candidate genes for strong chromatin remodeling activity of eggs during nuclear transplantation of somatic cell nuclei.

Addition of ES and EG cells to preimplantation embryos (143 genes; Group E in Figure 3) yielded only 54 signature genes (Group F in Figure 3). Addition of adult stem cells, MS and NS, or MS, NS, and HS (Lin^-, Kit^+, $Sca1^+$ and Lin^-, Kit^-, $Sca1^+$) cells further reduced the number of signature genes to five and one, in Groups G and H, respectively (Dataset S7). Taken together, these results seem to indicate that preim-

Group	Unfertilized Egg	Fertilized Egg	2-cell Embryo	4-cell	8-cell	Morula	E3.5 Blastocyst	E6.5 Whole Embryo	E7.5 Whole Embryo	E7.5 Embryonic Part	E7.5 Extraembryonic Part	E8.5 Whole embryo	E9.5 Whole Embryo	E12.5 Male Gonad/Mesonephros	E12.5 Female Gonad/Mesonephros	E13.5 VMB Dopamine Cell	Newborn Brain	Newborn Heart	Newborn Kidney	Newborn Ovary	ES Cell (LIF+)	ES Cell (LIF-)	TS Cell	EG Cell	MS Cell (9-15C)	Osteoblast (KUSA-A1)	NS Cell (Undifferentiated)	NS Cell (Differentiated)	HS Cell (Lin-, Kit+, Sca1+)	HS Cell (Lin-, Kit-, Sca1+)	HS Cell (Lin-, Kit+,Sca1-)	HS Cell (Lin-, Kit-, Sca1-)	gene number	Gene symbols
A	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	196	*Akap1; Alox12e; Apg5l; Arl6; Arl6ip2; Banp; Bcl2l10; Birc2; Bmp15; Bpgm; Btg4; Bub1b; Ccnb1rs1; Cdc25a; Cdc45l; Cryl1; Cyp11a; Dtx2; Eed; Epb4.1l2; Fbxw4; Fmn2; Folr4; Gdf9; Gtr2; H1fo; Hsd3b1; Ing1; Irf1; Itga9; Kcnh1; Kdel1; Krt2_16; Map2k6; Map4k5; Mapkbp1; Mater; Mdm4; Mitc1; Mmp23; Mrg1; Npc1; Oas1d; Oas1e; Obox3; Orc5l; Orc6l; P37nb; Pabpc4; Pcsk6; Phf1; Plat; Pld1; Pole3; Prkab1; Rbbp7; Rdx; Rfpl4; Rgs2; Rnf35; Rnpc1; Sh3d5; Sip1; Slc21a11; Smarcf1; Snrpb2; Tbn; Tcl1; Tcl1b1; Tcl1b3; Tes3; Tex14; Thrsp; Top1; Wbscr21; Xbp1; Zfp296; and 119 unknown genes.*
D	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	81	*Akp5; Bcl2l10; Bmp15; Bpgm; Btg4; Cdkap1; Ctsc; Dusp14; Fbxo15; Fxyd4; Gdf9; Hspa8; Klf5; Obox3; Ovgp1; Prkab1; Rfpl4; Rnf35; Rnpc1; Spin; Tcl1; Tcl1b3; Timd2; Ulk1; and 57 unknown genes.*
E	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	143	*Akp5; Arpc1b; Bcl2l10; Birc2; Bmp15; Bpgm; Btg4; Cd160; Cdkap1; Cry1; Daf1; Degs; Dusp14; E2f1; Fbxo15; Fkbp5; Fkbp6; Fmn2; Folr4; Fxyd4; Gdf9; Gja4; Gstm2; H1fo; Hspa8; Krt2_16; Lcn2; Magoh; Mater; Mbtd1; Mdm4; Mil1; Oas1d; Oas1e; Obox3; Ovgp1; P37nb; Prkab1; Rfpl4; Rgs2; Rnf33; Rnf35; Rnpc1; Rps8; Sat; Slc34a2; Spin; Tcl1; Tcl1b1; Tcl1b3; Timd2; Ybx3; and 91 unknown genes.*
F	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	54	*Bcl2l10; Bmp15; Btg4; Cdkap1; Degs; Dusp14; Fxyd4; Gdf9; Mitc1; Obox3; Ovgp1; Prkab1; Rfpl4; Rnf35; Rnpc1; Spin; Tcl1; Tcl1b3; Timd2; and 35 unknown genes.*
I	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	+	-	+	-	-	-	-	-	140	*Akap10; Akap12; Arnt; Ash2l; Atm; Birc6; Cask; Cbx5; Cdh11; CHD6; Crkol; Dnchc1; Ect2; Edr1; Enah; ENPP3; Fshprh1; Gabrg1; Galnt1; GPCR; Hells; Hmgcr; Hmmr; Impact; Kars_ps1; Kif10; Kif15; Knsl1; Lamc1; Lbr; Lox; Mad5; Mki67; Np95; Oazi; Odf2; Opa1; Pald; Plrg1; Pola1; Ppp4r1; Ptch; Ptch2; Rasa2; Rb1; Rex2; Rpl31; Sh3bp3; Shcbp1; Ski; Slc27a1; Sms; Spna2; Synj2; Tnc; TRB_2; Trif; Trps1; Zfp148; Zfp191; and 80 unknown genes.*
J	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	-	-	-	-	93	*Abce1; Akap10; Ccnf; Cdh11; Col12a1; Crkol; Dda3; Dmd; Enh; Fanca; GPCR; Hmmr; Img; Lrig1; Iqgap1; Kars_ps1; Kif1b; Lox; Morf; Nedd4; Opa1; Pald; Pola1; Ppp4r1; Ptprf; Rpl31; Sec23a; Sh3bp3; Slc27a1; Slc7a3; Snrp116; Tardbp; Thbs1; TRB-2; Trps1; Zfhx1a; Zfhx1b; Zfp191; and 55 unknown genes.*
K	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-	-	75	*Abhd2; Ccnf; Cldn4; Dda3; Dnmt3b; ENPP3; F11r; Fkbp4; Foxh1; Grb7; Grcc8; Helb; Hmga1; Jmj; Jub; Lsm10; Map4k1; Mif; Morf; Mta3; Mybl2; Ncl; Nek4; Nfyb; Nol5; Pabpn1; Pcnxl3; Pdcd4; Ppp4r1; Prker2a; Prss8; Rbbp6; Rfng; Rpl13; Rps2; Rps6ka1; Slc29a1; Slc7a3; Sntb2; Sox13; Tdh; Tsbp; Ubb; Unc13h1; Wee1; Xrcc2; Zfp42; and 28 unknown genes.*
L	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-	39	*Akap10; Cldn4; Dnmt3b; ENPP3; Foxh1; Galk1; Grb7; Kars-ps1; Lcn7; Nfyb; Ngfrap1; Pcbp1; Rgds; Ubb; Unc13h1; Wee1; Xrcc2; Zfp278; Zfp42; and 20 unknown genes.*
M	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	-	+	+	-	-	-	44	*Atf2; Cbl; Ccne2; Cd44; Elf1; Fshprh1; Fyn; G7e; Herc3; IFI 203; Itga4; Jak1; MALT-1; Mbnl; Nab1; Nfat5; Phc3; SENP6; Stxbp4; Tde1l; Tex2; Tm6sf1; Wwp4; and 34 unknown genes.*
N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-	+	-	-	-	-	-	-	-	-	108	*Abce1; Abhd2; Akap10; Arl6ip2; Ccnf; Col12a1; Dda3; Dstn; Edr1; Enah; ENPP3; Fkbp4; Foxh1; Gfpt2; Helb; Impact; Ing5; Jmj; Jub; Mad5; Map4k1; Mkrn1; Morf; Mov10; Mta3; Mtf2; Mybl2; Pask; Pcnxl3; Pola1; Pola2; Ppp2r5e; Ppp4r1; Ptch2; Ranbp17; Rbbp6; Rest; Rex2; Rw1; Slc29a1; Slc7a3; Smarca4; Sntb2; Sox13; Tacc3; Tcof1; Tdgf1; Tdh; Zfp42; and 59 unknown genes.*
O	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	+	-	-	-	-	-	-	-	-	113	*Akap10; Akap12; Ash2l; Atm; Cipp; Col18a1; Dda3; Ect2; Edr1; Enah; ENPP3; Ermelin; Ell1; Gab1; Gfpt2; Hic2; Hmmr; Impact; Ing5; Itga3; Jmj; Lamc1; Lyar; Mad5; Mkrn1; Mtf2; Mybl2; Ncoa3; Np95; Opa1; Pald; Pola1; Ppp2r5a; Ptch2; Ranbp17; Rex2; Rnf17; Rw1; Slc29a1; Slc7a3; Smcx; Sms; Taf7; Tdh; Tex20; Trif; Wee1; Zfp110; and 65 unknown genes.*

Figure 3. Signature Genes for Specific Groups of Early Embryos and Stem Cells
DOI: 10.1371/journal/pbio.0000074.g003

plantation embryos, particularly totipotent fertilized eggs and highly pluripotent cells (ES and EG cells), have quite distinct genetic programs, but that less pluripotent adult stem cells (MS, NS, and HS) have even more specialized genetic programs. This supports the notion of a gradual decrease of developmental potential from preimplantation embryos to stem cells to differentiated cells.

Additional analysis was done to determine genes that are enriched in stem cells, but not in preimplantation embryos and other tissues (see Figure 3; Dataset S6; Dataset S7). In this

analysis, 140 genes were identified as signature genes for pluripotent stem cells (ES, EG, NS, and MS in Group I in Figure 3), whereas 93 genes were identified as signature genes for these stem cells and their differentiated forms (cultured cells in Group J in Figure 3). Similarly, 75 and 39 genes, respectively, were identified as ES- and TS-specific (Group K in Figure 3), whereas 44 genes were identified as signature genes for adult stem cells (NS, MS, and HS in Group M in Figure 3). Lists of these genes showed that distinctive sets of genes are responsible for cell specificity (Figure 3).

FDR analysis revealed that 113 genes were specifically expressed in ES and EG cells in Group O (the most pluripotent stem cells), but not in all other cell types examined (Figure 3; Dataset S7). The most abundant group of these genes was transcription regulatory factors (about 30% of all specific genes), most of which were members of the zinc finger family, including *Mtf2*, *Ing5*, *Mkrn1*, *Hic2*, and the KRAB box zinc finger. Other abundant genes specifically expressed in ES and EG cells included matrix/cytoskeleton/membrane structural proteins such as *Itga3*, *Dstn*, *Smtn*, *Dctn1*, and *Col18a1* and the DNA remodeling proteins such as *Rcc1*, *Kars-ps1*, *Pola2*, *Mov10*, and *Rad54l*. These two groups of genes may be associated with the unique feature of ES/EG cell cycle structure, where greater than 70% of the cell population are in S phase (Savatier et al. 1996).

Previous studies have identified genes specific to particular stem cells or genes common to a group of stem cells, although there was little agreement about which transcripts are commonly enriched in these studies (e.g., Anisimov et al. 2002; Ivanova et al. 2002; Ramalho-Santos et al. 2002; Tanaka et al. 2002). The difference in the method and platform used could be a major reason for the difficulty in identifying a common gene set. The analysis of limited number of cell types could also contribute to differences in the resulting gene lists, because genes that appeared specific to certain cell types may also be expressed in other cells that were not included in the analysis. In contrast, the current study has analyzed a large number of different stem cells, preimplantation embryos, and newborn organs from our own EST collections as well as all publicly available ESTs that were derived from a few hundred cell types. Combined with stringent FDR statistics (see Materials and Methods), the analysis of this large number of cell types may provide broader perspectives on this issue. Comparison between the gene lists of the present study and the gene lists from the previously published studies identified areas of agreement (common genes), but also revealed that many genes previously reported as specifically expressed in one cell type or group of cells are actually expressed in other cell types and thus are not specific (see the details in Dataset S8). The signature genes identified in this study distinguish different stem cells, and this gene list may provide a way to recognize or purify specific stem cell types and provide insights into stem cell–specific functions.

Principal Component Analysis Identified Clusters of Cells/Tissues with Similar EST Frequency

The global expression patterns of 2,812 relatively abundant genes (see Materials and Methods; Dataset S9) were further analyzed by principal component analysis (PCA), which reduces high-dimensionality data into a limited number of principal components. The first principal component (PC1) captures the largest contributing factor of variation, which in this case corresponds to the average EST frequency in all tissues, and subsequent principal components correspond to other factors with smaller effects, which characterize the differential expression of genes. As we were interested in the differential gene expression component, we plotted the position of each cell type against the PC2, PC3, and PC4 axis in three-dimensional (3D) space by using virtual reality modeling language (VRML) (Figure 4A; Video S1; a full interactive view is available on http://lgsun.grc.nia.nih.gov/Supplemental-Information). Genes were also plotted in the same 3D space (a version of PCA called a biplot) (Chapman et al. 2002) to see their association with cell/tissue types. Close examination of the 3D model identified PC2 and PC3 as the most representative views of the 3D model (Figure 4B). A two-dimensional (2D) plot of PC2 and PC3 is therefore used for the following discussion, with references to the 3D model. It is important to keep in mind that the distance between cell types along principal components has a substantial error associated with randomness of clone counts in EST libraries. The estimated error range (2*SE) in the PC3 scale is about 7%–9% based on Poisson distribution (Figure 4B). Nonetheless, PCA identifies major trends and clusters in gene expression among these cell types.

The most conspicuous trend was that cells that differ in their developmental potential appeared well separated along the PC3 axis. In Figure 4A and 4B, preimplantation embryos (unfertilized egg, fertilized egg, two-cell, four-cell, eight-cell, morula, and blastocyst) are positioned at the top of the PC3 axis; embryos and extraembryonic tissues from early- to mid-gestation stage, such as E6.5, E7.5, E8.5, and E9.5, are positioned at the middle; and cells and tissues mostly from terminally differentiated cells (newborn ovary, newborn heart, and newborn brain) are positioned at the bottom. PCA is unsupervised (performed without using knowledge of developmental stages of each cell types), and so this ordering along the PC3 axis seems to reflect the structures of global gene expression patterns among the cells. The PC2 axis provided an additional dimension to separate cells into developmental stages, functional groups, or both. The correlation of the PC2 axis to known biological stages, functions, or both, however, remains unclear.

Interestingly, both ES cells and adult stem cells are positioned at the middle of the PC3 axis together with whole-embryo libraries from early- to mid-gestation stages (Figure 4B). ES and EG cells were derived from embryos, and thus their positions matched with their developmental timing. Although NS, MS, and HS cells were all derived from adult organs (brain, bone marrow, and bone marrow, respectively), their position along the PC3 axis corresponded to early embryonic tissues and embryo-derived stem cells (ES and EG). The results are consistent with the notion that adult stem cells acquire or retain the pluripotency with characters of less-differentiated cell types. This also suggests that the PC3 axis does not represent just developmental timing, but also indicates the developmental potential of cells, with totipotent eggs at the top, pluripotent embryonic cells and stem cells at the middle, and terminally differentiated cells at the bottom.

This hypothesis seems to be consistent with another interesting observation that the differentiated forms of stem cells were always positioned lower than their stem cell

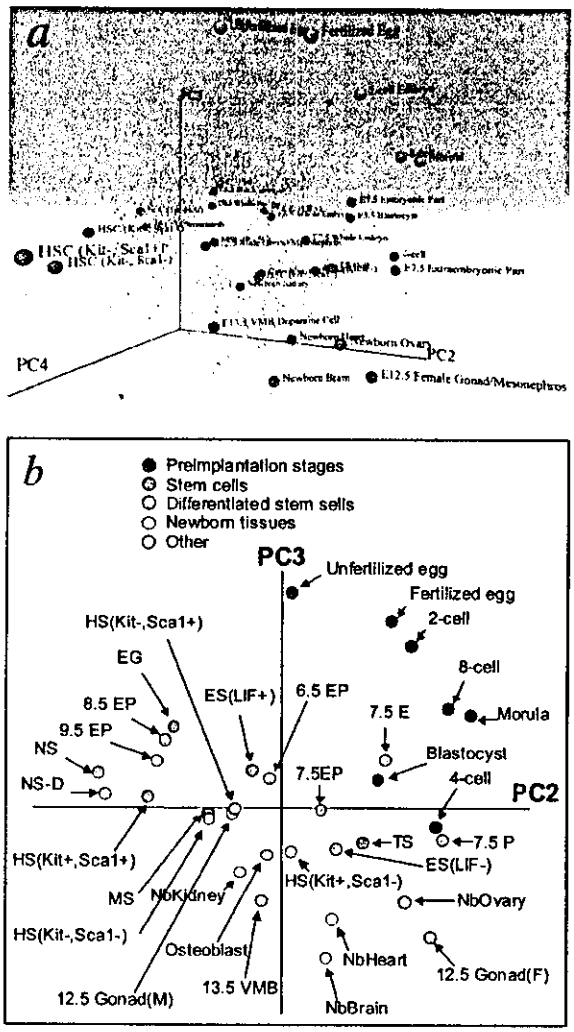

Figure 4. PCA Analysis of EST Frequency

The results were obtained by analyzing 2,812 genes that exceeded 0.1% in at least one library. (A) 3D biplot that shows both cell types (red spheres) and genes (yellow boxes). (B) 2D PCA of cell types. EST frequencies were log-transformed before the analysis. Names of some cells and tissues are abbreviated as follows: 6.5 EP, E6.5 whole embryo (embryo plus placenta); 7.5 EP, E7.5 whole embryo (embryo plus placenta); 8.5 EP, E8.5 whole embryo (embryo plus placenta); 9.5 EP, E9.5 whole embryo (embryo plus placenta); 7.5 E, E7.5 embryonic part only; 7.5 P, E7.5 extraembryonic part only; NbOvary, newborn ovary; NbBrain, newborn brain; NbHeart, newborn heart; NbKidney, newborn kidney; 13.5 VMB, E13.5 ventral midbrain dopamine cells; 12.5 Gonad (F), E12.5 female gonad/mesonephros; 12.5 Gonad (M), E12.5 male gonad/mesonephros; HS (Kit^-, $Sca1^-$), hematopoietic stem/progenitor cells (Lin^-, Kit^-, $Sca1^-$); HS (Kit^-, $Sca1^+$), hematopoietic stem/progenitor cells (Lin^-, Kit^-, $Sca1^+$); HS (Kit^+, $Sca1^-$), hematopoietic stem/progenitor cells (Lin^-, Kit^+, $Sca1^-$); HS (Kit^+, $Sca1^+$), hematopoietic stem/progenitor cells (Lin^-, Kit^+, $Sca1^+$); and NS-D, differentiated NS cells.

DOI: 10.1371/journal/pbio.0000074.g004

counterparts (undifferentiated forms) in the PC3 axis (Figure 4A and 4B). For example, the position of NS (differentiated) cells, a mixture of neuron and glia obtained after culturing NS cells in the differentiation conditions, was lower and nearer to the terminally differentiated cells than were NS cells. Osteoblast cells, which are more differentiated than the MS cells from which they are derived, were again positioned lower than the MS cells. The same holds true for ES (LIF^-) cells (lower PC3 position), which were obtained by culturing ES cells in the absence of leukemia inhibitory factor (LIF), allowing ES cells to differentiate into many different cell types, and ES (LIF^+) cells (higher PC3 position), which were maintained as highly pluripotent by culturing them in the presence of LIF. For HS cells, all four cell types were selected first as lineage marker-negative cells, and thus they were all relatively undifferentiated cells. These cells were then sorted by c-Kit^+ and $Sca1^+$ into four separate fractions. The most pluripotent cells (Lin^-, c-Kit^+, $Sca1^+$) were again positioned higher than other three cell types in the PC3 axis. Finally, TS cells were positioned at the least-potent place among stem cells, which seemed to fit to their known characteristics. It has previously been shown that TS cells are already committed to the extraembryonic lineage and are less pluripotent than ES and EG cells, because TS cells injected back to mouse blastocysts only differentiate into extraembryonic trophoblast lineages (Tanaka et al. 1998). The microarray analysis of TS cells also shows that they already express many placenta-specific genes, which is a sign of lineage-committed cells (Tanaka et al. 2002).

Finally, it is interesting to note that EG cells were positioned closely to E8.5 whole embryos and E9.5 whole embryos, whereas ES cells were positioned closely to blastocysts, E6.5, and E7.5 whole embryos (Figure 4). Because ES cells are derived from E3.5 blastocysts and EG cells are derived from primordial germ cells (PGCs) of E8.5 (in this particular line), these results indicate that the expression patterns of relatively abundant genes in ES and EG cells reflect their developmental stages of origin. Although ES and EG cells were established from different sources, EG cells are often considered to be ES cells and the distinction of their origin is ignored. However, the result here suggests potentially significant differences between the genetic programs of EG cells and ES cells.

Genes Correlated with the Developmental Potential of Cells

To identify a group of genes associated with the PC3 axis, we first fixed the coordinate of each cell type on PC3 and searched for genes whose log-transformed frequencies correlated with this coordinate in each cell type. Correlation analysis combined with the FDR method (FDR = 0.1) revealed 88 genes whose expression levels were significantly associated with PC3 (Dataset S10). To test how well these genes represent PC3, we plotted the sum of log-transformed EST frequencies for these 88 genes versus PC3 projections of the same cell types (Figure 5). Most cells were positioned diagonally relative to the original PC3 coordinates, indicating that the average expression levels of these 88 genes can roughly represent cell type position along the PC3 coordinate. Because the PC3 axis does not have a unit and cannot be directly translated to variables measured by molecular biological techniques, the possible use of 88 genes as a surrogate for the PC3 axis will help to test this working hypothesis in the future.

What are the characteristics of these 88 potential correlating genes? Based on the available protein domain information, Gene Ontology (GO) annotation (Ashburner et al. 2000; http://www.geneontology.org/doc/GO.annotation.html), and literature, 58 genes can be classified into putative functional categories (Dataset S10). For example, signature genes in the

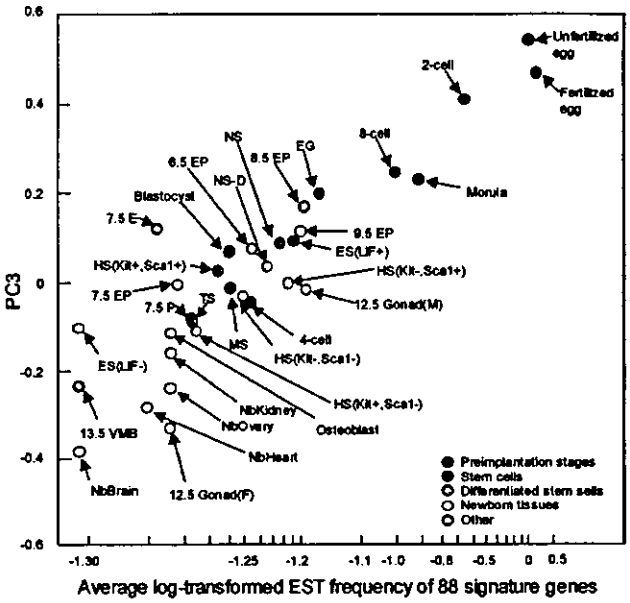

Figure 5. Relationship between PC3 and Average Expression Levels of 88 Signature Genes

A list of 88 genes associated with developmental potential: *Birc2, Bmp15, Btg4, Cdc25a, Cyp11a, Dtx2, E2f1, Fmn2, Folr4, Gdf9, Krt2–16, Mitc1, Oas1d, Oas1e, Obox3, Prkab1, Rfpl4, Rgs2, Rnf35, Rnpc1, Slc21a11, Spin, Tcl1, Tcl1b1, Tcl1b3, 1810015H18Rik, 2210021E03Rik, 2410003C07Rik, 2610005B21Rik, 2610005H11Rik, 3230401D17Rik, 4833422F24Rik, 4921528E07Rik, 4933428G09Rik, 5730419I09Rik, A030007L17Rik, A930014I12Rik, E130301L11Rik, AA617276, Bcl2l10, MGC32471, MGC38133, MGC38960, D7Ertd784e*, and 44 genes with only NIA U numbers (see Dataset S10).

DOI: 10.1371/journal/pbio.0000074.g005

"transcriptional control" category include eight genes, such as MAD homologue 4 interacting transcription coactivator 1 (*Mitc1*), *Drosophila* Deltex 2 homologue (*Dtx2*), and oocyte-specific homeobox 5 (*Obox5*); the "RNA binding" category includes five genes such as RNA-binding region containing 1 (*Rnpc1*) and 2′-5′-oligoadenylate synthetase 1D (*Oas1d*); the "signal transduction" category includes ten genes, such as AMP-activated protein kinase (*Prkab1*) and regulator of G-protein signaling 2 (*Rgs2*); and the "proteolysis" category includes six genes, such as Ret finger protein-like 4 (*Rfpl4*) and ring finger protein 35 (*Rnf35*). These categories were diverse, and the domination of any specific categories was not observed.

Although all 88 genes shared the general trend of continuous decrease of expression levels from eggs to terminally differentiated tissues, these genes can be further subdivided by their expression patterns. First, 53 genes were those identified as preimplantation specific, particularly unfertilized and fertilized egg-specific genes, which include already well-known genes for their functions in oogenesis and zygotic gene activation, such as *Gdf9, Bmp15, Rfpl4, Fmn2, Tcl1, Obox5*, and *Oospl*. Second, ten genes were represented as ESTs in both preimplantation embryos and postimplantation embryos, including *Cyp11a* and *D7Ertd784e*. Third, 25 genes were represented well as ESTs in preimplantation embryos, postimplantation embryos, and stem cells, including *Mitc1*, actin-binding Kelch family protein, *Dtx2, Cdc25a, Spin, Rgs2, Prkab1*, and *Birc2*. The seemingly continuous decrease of the expression of these genes is therefore not caused by passive dilution of transcripts that are abundant in oocytes, but is most likely caused by a specific mechanism that actively regulates the expression levels of these genes.

Concluding Remarks

The sequence information and cDNA clones collected in this work provide the most comprehensive database and resources for genes functioning in early mouse embryos and stem cells. All cDNA clones developed in this project have been made available through the American Type Culture Collection (ATCC). The subset of these cDNA clones have been rearrayed into the condensed clone sets, the NIA Mouse 15K cDNA Clone Set (Tanaka et al. 2000; Kargul et al. 2001) and the 7.4K cDNA Clone Set (VanBuren et al. 2002), which have been made available through designated academic distribution centers. Many genes that are uniquely or predominantly expressed in mouse early embryos and stem cells have been recently incorporated into a 60mer oligonucleotide microarray (Carter et al. 2003). Sequence information has been made available at public sequence databases (e.g., dbEST [Boguski et al. 1993]). Finally, all the information discussed here, as well as the graphical interfaces of the Mouse Gene Index, is available on our Web site at http://lgsun.grc.nia.nih.gov/cDNA/cDNA.html.

Although the full appreciation of these resources is yet to be realized, the initial assessment of the first comprehensive transcriptome of early mouse embryos and stem cells has already provided three major points presented in this report.

First, approximately 1,000 putative genes that were newly identified using our cDNA collection most likely represent mouse genes unidentified previously, as they either encode ORFs greater than 100 amino acids or have multiple exons. The RT–PCR analysis of 19 selected genes confirmed the notion that novel cDNAs from our libraries tend to be expressed specifically in cells and tissues that we used in this project. These gene candidates will be a rich source of genes that are expressed at low levels, but play major roles in ES cells and adult stem cells as well as in early embryos.

Second, the analysis provided lists of genes specific to particular embryonic stages or stem cells and not expressed in other cell types. For example, we have identified signature genes for the individual preimplantation stages, all preimplantation stages, ES cells, and adult stem cells.

Finally, the PCA of 2,812 genes with relatively abundant expression revealed 88 genes with average expression levels that correlate well to the developmental potentials of cells. These genes may provide the first scale to characterize the developmental potential of cells and tissues at the molecular level.

The developmental potential of cells is a fundamental concept in developmental biology, providing a conceptual framework of sequential transition from totipotent fertilized eggs to pluripotent embryonic cells and stem cells to terminally differentiated cells. It is worth noting that genes associated with developmental potential can be identified only by simultaneous analysis of preimplantation embryos and a variety of stem cells. The analyses of stem cells alone could not provide these broader perspectives (Ivanova et al. 2002; Ramalho-Santos et al. 2002; Tanaka et al. 2002). The 88 genes we have identified here may provide a set of marker

genes for scaling the potential of cells. It is important to note that this scale is an operational construct. As such, further studies of the genes in the list will be required to test whether they provide critical clues to resolve the classic problem of the relation of stem cells to development. But the list could have immediate practical utility in assessing the effectiveness of treatments, gene manipulation, or both to convert differentiated cells such as fibroblasts into more potent cells such as ES—one of the most important goals required to achieve stem cell–based therapy.

Materials and Methods

cDNA library construction, clone handling, and sequencing. Sources of tissue materials and RNA extraction methods are available as associated documents in the GenBank DNA sequence records (see also http://lgsun.grc.nia.nih.gov/cDNA/cDNA.html). cDNA libraries were constructed as described elsewhere (Piao et al. 2001). More details are available in Protocol S1.

Assembling of a gene index. See description in the legend to Figure 1 and in Protocol S1.

Analysis of 19 cDNA clones. Sequencing of full-length cDNA clones and RT-PCR analysis were done by the standard methods. More details are available as Protocol S1.

Identification of differentially expressed genes. Most methods for selecting differentially expressed genes from EST frequencies are based on the assumption that each cDNA clone is a random sample from the mRNA pool in the cell and hence that EST frequencies correspond to the Poisson distribution (Audic and Claverie 1997). Real EST libraries, however, do not satisfy this assumption because even small changes in experimental conditions may affect the stability of particular species of mRNA, which in turn will cause a bias in EST frequency. Thus, a reliable detection of differentially expressed genes requires either library replications or comparison of classes of libraries. Because our EST libraries do not have true replications, we selected the latter approach, which yields genes that are specifically expressed in one class of tissues/stages and do not express in other tissues/stages. Some cDNA clones were represented by 5′ EST, some were by 3′ EST, and some were by both 5′ EST and 3′ EST. To avoid counting the same cDNA clone twice by 5′ EST and 3′ EST, all EST frequency analysis was done at the cDNA clone level.

To detect genes specific to a particular group of libraries, we first estimated the correlation between log-transformed clone frequencies, $\log(1000*n_i/N + 0.05)$, where n_i is the abundance of clone i in the library and N is the total number of clones, with membership indicated (0 or 1) in a particular group (see Dataset S6). The first three group classifications are targeted on oocytes. The next two classifications include all preimplantation stages with and without blastocysts. There are four classifications attempting to differentiate between pluripotent cells and other tissues. The final nine classifications capture various groups of stem cells. Results of these analyses are given in Dataset S7 and a subset of the data is shown in Figure 3. We analyzed only positive correlations because we were interested in genes that are overexpressed in tissues of interest, and P-values were estimated using a one-tailed t-test. Because P-values cannot be used for simultaneous assessment of multiple hypotheses, we determined significant genes using the FDR method (Benjamini and Hochberg 1995). The FDR was set to 0.1, which corresponds to the average proportion of false positives equal to 10%.

As this study is focused on embryo- and stem cell–specific genes, we analyzed EST frequencies in public databases (Boguski et al. 1993) to exclude those genes that are predominantly expressed in adult tissues. A total of 3,338,847 public ESTs have been grouped into the following categories: NIA Collection, Preimplantation, Embryo, Embryonic Stem Cells, Fetus, Neonate, Adult, Adult Gonad, Adult Stem Cells, Adult Tumor, and Unclassified/Pooled Tissues (Dataset S11). Of 29,810 mouse genes, 5,425 genes were not represented by ESTs, 11,574 genes were expressed predominantly in adult tissues (EST frequency in adult tissues exceeds one-third of the maximum EST frequencies in all tissues), and 12,811 were genes expressed in embryos or in gonads, tumors, and stem cells. By removing 2,055 gonad-specific and 56 tumor-specific genes (20 times more ESTs in gonad or tumors than in other tissues), we obtained 10,700 genes that are predominantly expressed in embryos and stem cells (Dataset S12). Only ESTs matching to these genes were analyzed for differential expression.

PCA of clone frequencies. For the PCA shown in Figure 4, we selected 2,812 genes that had transcript frequencies of greater than or equal to 0.1% in at least one library (see Dataset S9). Clone/EST frequencies were log-transformed as $\log(1000*n_i/N + 0.05)$, where n_i is the number of clones in U-cluster i in the library, and N is the total number of all clones in this library.

Statistical significance of gene contribution to PC3 (see Figure 5) was evaluated using correlation between log-transformed clone frequencies in various libraries and library position on the PC3 axis. P-values, estimated using a one-tailed t-distribution, characterize the significance of correlation for a single clone. To control the proportion of false positives, we used FDR, which was set to 0.1.

Supporting Information

To view this Supporting Information with dynamic Web links, see http://lgsun.grc.nia.nih.gov/Supplemental-Information/.

The NIA Mouse Gene Index has recently made available to the public (http://lgsun.grc.nia.nih.gov/geneindex/). The Web interface provides a view of transcripts and genes on the mouse genome sequence. Unique IDs (U plus 6 digits, e.g., U018631) have been assigned to individual genes in the gene index. "U numbers" in the following datasets have direct links to corresponding genes in the NIA Mouse Gene Index. Clicking the "U number" in the datasets will lead to a Web page of the NIA public Web site.

Dataset S1. List of NIA Mouse cDNA Libraries and the Number of ESTs Generated

View online at DOI: 10.1371/journal/pbio.0000074.sd001 (22 KB XLS).

Dataset S2. Summary of Gene Counts in the NIA Mouse Gene Index

In addition to the list here, the Web interface at http://lgsun.grc.nia.nih.gov/geneindex provides a view of transcripts and genes on the mouse genome sequence.

View online at DOI: 10.1371/journal/pbio.0000074.sd002 (36 KB XLS).

Dataset S3. List of 977 Genes Unique to the NIA Mouse cDNA Collection

These are not found in RefSeq, Ensembl, and RIKEN. For sequence information, see Dataset S4.

View online at DOI: 10.1371/journal/pbio.0000074.sd003 (268 KB XLS).

Dataset S4. Sequence Information of 977 Genes in the FASTA Format

View online at DOI: 10.1371/journal/pbio.0000074.sd004 (685 KB TXT).

Dataset S5. Primer Sequences for RT-PCR Analysis

View online at DOI: 10.1371/journal/pbio.0000074.sd005 (30 KB DOC).

Dataset S6. Classification of cDNA Libraries for the Analysis of Differentially Expressed Genes

This table describes how cDNA libraries were logically grouped for further EST analysis, where membership to a group is indicated with a 1 and nonmembership is indicated with a 0.

View online at DOI: 10.1371/journal/pbio.0000074.sd006 (19 KB XLS).

Dataset S7. List of Genes Overexpressed in Preimplantation Embryos and Stem Cells

This table identifies the genes overexpressed in each group of cells/tissues described in Dataset S6.

View online at DOI: 10.1371/journal/pbio.0000074.sd007 (510 KB XLS).

Dataset S8. Comparison of the Gene Lists Identified in Dataset S7 with the Published Data

View online at DOI: 10.1371/journal/pbio.0000074.sd008 (23 KB DOC).

Dataset S9. List of 2,812 Genes Used for PCA to Investigate the Global Feature of Gene Expression Patterns

View online at DOI: 10.1371/journal/pbio.0000074.sd009 (633 KB XLS).

Dataset S10. List of 88 Genes Correlated with Developmental Potential of Cells

View online at DOI: 10.1371/journal/pbio.0000074.sd010 (72 KB XLS).

Dataset S11. Comprehensive Data about EST Frequencies of Genes in NIA Mouse cDNA Libraries and in Public Sequence Databases

View online at DOI: 10.1371/journal/pbio.0000074.sd011 (13.9 MB XLS).

Dataset S12. List of 10,699 Genes Predominantly Expressed in Embryos and Stem Cells

These genes were identified by the analysis of NIA EST and public EST datasets.

View online at DOI: 10.1371/journal/pbio.0000074.sd012 (3.2 MB XLS).

Protocol S1. Supplemental Materials and Methods

View online at DOI: 10.1371/journal/pbio.0000074.sd013 (59 KB DOC).

Video S1. 3D View of Results Obtained by PCA of Log-Transformed EST Frequencies in NIA Mouse cDNA Libraries

Red spheres represent libraries and yellow boxes represent genes. Gene names can be legible at closer distance. (For Windows, Media Player or Real Player is required to view. For Macintosh, Quicktime Player is required.) A virtual reality modeling language (VRML) formatted version is also available on our Web site (http://lgsun.grc.nia.nih.gov/Supplemental-Information). The VRML version allows users to freely rotate and zoom the image in 3D space. Genes are also hyperlinked to the NIA Mouse Gene Index Web site (mentioned in Dataset S2).

View online at DOI: 10.1371/journal/pbio.0000074.sv001 (3.9 MB AVI).

Accession Numbers

The LocusLink (http://www.ncbi.nih.gov/LocusLink/) accession numbers for the genes discussed in this paper are *1810015H18Rik* (LocusLink ID 69104), *2210021E03Rik* (LocusLink ID 52570), *2410003C07Rik* (LocusLink ID 66977), *2610005B21Rik* (LocusLink ID 72119), *2610005H11Rik* (LocusLink ID 72114), *3230401D17Rik* (LocusLink ID 66680), *4833422F24Rik* (LocusLink ID 74614), *4921528E07Rik* (LocusLink ID 114874), *4933428G09Rik* (LocusLink ID 66768), *5730419I09Rik* (LocusLink ID 74741), *5930405J04Rik* (LocusLink ID 68094), *A030007L17Rik* (LocusLink ID 68252), *A930014I12Rik* (LocusLink ID 77805), *AA617276* (LocusLink ID 100012), actin-binding Kelch family protein (LocusLink ID 246293), *Bcl2l10* (LocusLink ID 12049), *Birc2* (LocusLink ID 11796), *Bmp15* (LocusLink ID 12155), *Btg4* (LocusLink ID 56057), *Cdc25a* (LocusLink ID 12530), *Col18a1* (LocusLink ID 12822), *Cyp11a* (LocusLink ID 13070), *D7Ertd784e* (LocusLink ID 52428), *Dctn1* (LocusLink ID 13191), *Dstn* (LocusLink ID 56431), *Dtx2* (LocusLink ID 74198), *E130301L11Rik* (LocusLink ID 78733), *E2f1* (LocusLink ID 13555), *Fmn2* (LocusLink ID 54418), *Folr4* (LocusLink ID 64931), *Gdf9* (LocusLink ID 14566), *Hic2* (LocusLink ID 58180), *Ing5* (LocusLink ID 66262), *Itga3* (LocusLink ID 16400), *Kars-ps1* (LocusLink ID 85307), KRAB box zinc finger (LocusLink ID 170763), *Krt2-16* (LocusLink ID 16680), *MGC32471* (LocusLink ID 212980), *MGC38133* (LocusLink ID 243362), *MGC38960* (LocusLink ID 235493), *Mitc1* (LocusLink ID 75901), *Mkrn1* (LocusLink ID 54484), *Mov10* (LocusLink ID 17454), *Mtf2* (LocusLink ID 17765), *Oas1d* (LocusLink ID 100535), *Oas1e* (LocusLink ID 231699), *Obox3* (LocusLink ID 246791), *Obox5* (LocusLink ID 252829), *Oosp1* (LocusLink ID 170834), *Pola2* (LocusLink ID 18969), *Prkab1* (LocusLink ID 19079), *Rad54l* (LocusLink ID 19366), *Rcc1* (LocusLink ID 100088), *Rfpl4* (LocusLink ID 192658), *Rgs2* (LocusLink ID 19735), *Rnf35* (LocusLink ID 260296), *Rnpc1* (LocusLink ID 56190), *Scmh1* (LocusLink ID 29871), *Sfmbt* (LocusLink ID 54650), *Slc21a11* (LocusLink ID 108116), *SMARCC2* (LocusLink ID 6601), *Smarcf1* (LocusLink ID 93760), *Smtn* (LocusLink ID 29856), *Spin* (LocusLink ID 20729), *Tcl1* (LocusLink ID 21432), *Tcl1b1* (LocusLink ID 27379), and *Tcl1b3* (LocusLink ID 27378).

The GenBank (http://www.ncbi.nih.gov/Genbank/index.html) accession numbers of new ESTs reported in this paper are AA406988–AA407326, AA409386–AA409982, AA409984–AA410173, AW536060–AW536143, AW537733–AW537828, AW545917–AW545921, BE824469–BE825132, BI076411–BI076872, BM114148–BM121445, BM121647–BM125459, BM194710–BM203257, BM203259–BM214569, BM214575–BM251183, BM293391–BM293823, BU576966–BU576966, CA530650–CA580325, CA870176–CA882792, CA882932–CA896558, CD538085–CD544029, CD544034–CD555913, CD559752–CD565790, CF153424–CF161651, and CF161657–CF175178.

Acknowledgments

We thank M. A. Espiritu, A. Ebrahimi, J. J. Evans, S. J. Olson, M. Roque-Briewer, and N. Caffo at Applied Biosystems for contract-based sequencing and S. Chacko for setting up the mouse genome database on Biowulf. This study utilized the high-performance computational capabilities of the Biowulf/LoBoS3 cluster at the National Institutes of Health (NIH), Bethesda, Maryland, United States of America. Sequencing of cDNA clones was solely supported by the research and development funds of the National Institute on Aging (NIA). The project was mainly supported by the Intramural Research Program of the NIA. The collection of HS cells has been funded in part with federal funds from the National Cancer Institute, under contract number NO1-CO-5600.

Conflicts of interest. The authors have declared that no conflicts of interest exist.

Author contributions. MSHK conceived and designed the experiments. YP, RM, GF, PRM, CAS, UCB, YW, MGC, TH, KA, HA, LS, TST, WLK, TY, SAJ, SP, and MSHK performed the experiments. AAS, YP, RM, DBD, YQ, VV, GF, J. Kelso, WH, and MSHK analyzed the data. RN, KRB, DDT, RJH, DLL, DS, J. Keller, EK, GHK, AU, AV, JR, TK, BLMH, AC, MD, J. Kelso, and WH contributed reagents/materials/analysis tools. AAS, YP, RM, VV, GF, and MSHK wrote the paper. ▪

References

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, et al. (1991) Complementary DNA sequencing: Expressed seqeunce tags and human genome project. Science 252: 1651–1656.

Anisimov SV, Tarasov KV, Tweedie D, Stern MD, Wobus AM, et al. (2002) SAGE identification of gene transcripts with profiles unique to pluripotent mouse R1 embryonic stem cells. Genomics 79: 169–176.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: Tool for the unification of biology—the Gene Ontology Consortium. Nat Genet. 25: 25–29.

Audic S, Claverie JM (1997) The significance of digital gene expression profiles. Genome Res 7: 986–995.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc B Met 57: 289–300.

Boguski MS, Lowe TMJ, Tolstoshev CM (1993) dbEST: Database for "expressed sequence tags." Nat Genet 4: 332–333.

Carter MG, Hamatani T, Sharov AA, Carmack CE, Qian Y, et al. (2003) *In situ*-synthesized novel microarray optimized for mouse stem cell and early developmental expression profiling. Genome Res 13: 1011–1021.

Chapman S, Schenk P, Kazan K, Manners J (2002) Using biplots to interpret gene expression patterns in plants. Bioinformatics 18: 202–204.

Galli R, Fiocco R, De Filippis L, Muzio L, Gritti A, et al. (2002) *Emx2* regulates the proliferation of stem cells of the adult mammalian central nervous system. Development 129: 1633–1644.

Hochedlinger K, Jaenisch R (2002) Nuclear transplantation: Lessons from frogs and mice. Curr Opin Cell Biol 14: 741–748.

Hubbard T, Barker D, Birney E, Cameron G, Chen Y, et al. (2002) The Ensembl genome database project. Nucleic Acids Res 30: 38–41.

Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, et al. (2002) A stem cell molecular signature. Science 298: 601–604.

Kargul GJ, Dudekula DB, Qian Y, Lim MK, Jaradat SA, et al. (2001) Verification and initial annotation of the NIA mouse 15K cDNA clone set. Nat Genet 28: 17–18.

Kent WJ (2002) BLAT: The BLAST-like alignment tool. Genome Res 12: 656–664.

Ko MSH (2001) Embryogenomics: Developmental biology meets genomics. Trends Biotechnol 19: 511–518.

Ko MSH, Kitchen JR, Wang X, Threat TA, Hasegawa A, et al. (2000) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development. Development 127: 1737–1749.

Latham KE, Schultz RM (2001) Embryonic genome activation. Front Biosci 6: D748–D759.

Makino S, Fukuda K, Miyoshi S, Konishi F, Kodama H, et al. (1999) Cardiomyocytes can be generated from marrow stromal cells *in vitro*. J Clin Invest 103: 697–705.

Marra M, Hillier L, Kucaba T, Allen M, Barstead R, et al. (1999) An encyclopedia of mouse genes. Nat Genet 21: 191–194.

Matsui Y, Zsebo K, Hogan BL (1992) Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell 70: 841–847.

Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Barrell D, et al. (2003) The InterPro database 2003 brings increased coverage and new features. Nucleic Acids Res 31: 315–318.

Ochi K, Chen G, Ushida T, Gojo S, Segawa K, et al. (2003) Use of isolated

mature osteoblasts in abundance acts as desired-shaped bone regeneration in combination with a modified poly-DL-lactic-co-glycolic acid (PLGA)-collagen sponge. J Cell Physiol 194: 45–53.

Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, et al. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420: 563–573.

Ortiz M, Wine JW, Lohrey N, Ruscetti FW, Spence SE, et al. (1999) Functional characterization of a novel hematopoietic stem cell and its place in the c-Kit maturation pathway in bone marrow cell development. Immunity 10: 173–182.

Pertea G, Huang X, Liang F, Antonescu V, Sultana R, et al. (2003) TIGR gene indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. Bioinformatics 19: 651–652.

Piao Y, Ko NT, Lim MK, Ko MSH (2001) Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method. Genome Res 11: 1553–1558.

Pruitt KD, Maglott DR (2001) RefSeq and LocusLink: NCBI gene-centered resources. Nucleic Acids Res 29: 137–140.

Ramalho-Santos M, Yoon S, Matsuzaki Y, Mulligan RC, Melton DA (2002) "Stemness": Transcriptional profiling of embryonic and adult stem cells. Science 298: 597–600.

Sasaki N, Nagaoka S, Itoh M, Izawa M, Konno H, et al. (1998) Characterization of gene expression in mouse blastocyst using single-pass sequencing of 3995 clones. Genomics 49: 167–179.

Savatier P, Lapillonne H, van Grunsven LA, Rudkin BB, Samarut J (1996) Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. Oncogene 12: 309–322.

Shamblott MJ, Axelman J, Wang S, Bugg EM, Littlefield JW, et al. (1998) Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc Natl Acad Sci U S A 95: 13726–13731.

Slack JMW (1991) From egg to embryo: Regional specifications in early development. Cambridge, United Kingdom: Cambridge University Press. 348 p.

Solter D, de Vries WN, Evsikov AV, Peaston AE, Chen FH, et al. (2002) Fertilization and activation of the embryonic genome. In: Rossant J, Tam PPL, editors. Mouse development: Patterning, morphogenesis, and organogenesis. San Diego, California: Academic Press. pp. 5–19.

Tanaka S, Kunath T, Hadjantonakis AK, Nagy A, Rossant J (1998) Promotion of trophoblast stem cell proliferation by FGF4. Science 282: 2072–2075.

Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, Wang X, et al. (2000) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. Proc Natl Acad Sci U S A 97: 9127–9132.

Tanaka TS, Kunath T, Kimber WL, Jaradat SA, Stagg CA, et al. (2002) Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. Genome Res 12: 1921–1928.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, et al. (1998) Embryonic stem cell lines derived from human blastocysts. Science 282: 1145–1147.

VanBuren V, Piao Y, Dudekula DB, Qian Y, Carter MG, et al. (2002) Assembly, verification, and initial annotation of the NIA mouse 7.4K cDNA clone set. Genome Res 12: 1999–2003.

Waddington CH (1957) The strategy of the genes: A discussion of some aspects of theortical biology. London: Allen and Unwin. 262 p.

Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520–562.

Yanagimachi R (2002) Cloning: Experience from the mouse and other animals. Mol Cell Endocrinol 187: 241–248.

ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Experimental Cell Research 288 (2003) 35–50

Experimental Cell Research

www.elsevier.com/locate/yexcr

Redifferentiation of dedifferentiated chondrocytes and chondrogenesis of human bone marrow stromal cells via chondrosphere formation with expression profiling by large-scale cDNA analysis

Hideaki Imabayashi,[a,b] Taisuke Mori,[a,f] Satoshi Gojo,[c] Tohru Kiyono,[d] Tomoyasu Sugiyama,[e] Ryotaro Irie,[e] Takao Isogai,[e] Jun-ichi Hata,[a,f,1] Yoshiaki Toyama,[b] and Akihiro Umezawa[a,f,1,*]

[a] *National Research Institute for Child Health and Development, Tokyo, Japan*
[b] *Department of Orthopedic Surgery, Keio University School of Medicine, Tokyo, Japan*
[c] *Department of Cardiovascular Surgery, Saitama Medical Center, Kawagoe, Japan*
[d] *Virology Division, National Cancer Center Research Institute, Tokyo, Japan*
[e] *Helix Research Institute, Kisarazu, Japan*
[f] *Department of Pathology, Keio University School of Medicine, Tokyo, Japan*

Received 28 August 2002, revised version received 13 December 2002

Abstract

Characterization of dedifferentiated chondrocytes (DECs) and mesenchymal stem cells capable of differentiating into chondrocytes is of biological and clinical interest. We isolated DECs and bone marrow stromal cells (BMSCs), H4-1 and H3-4, and demonstrated that the cells started to produce extracellular matrices, such as type II collagen and aggrecan, at an early stage of chondrosphere formation. Furthermore, cDNA sequencing of cDNA libraries constricted by the oligocapping method was performed to analyze difference in mRNA expression profiling between DECs and marrow stromal cells. Upon redifferentiation of DECs, cartilage-related extracellular matrix genes, such as those encoding leucine-rich small proteoglycans, cartilage oligomeric matrix protein, and chitinase 3-like 1 (cartilage glycoprotein-39), were highly expressed. Growth factors such as FGF7 and CTGF were detected at a high frequency in the growth stage of monolayer stromal cultures. By combining the expression profile and flow cytometry, we demonstrated that isolated stromal cells, defined by $CD34^-$, c-kit^-, and $CD140\alpha^{-\text{ or low}}$, have chondrogenic potential. The newly established human mesenchymal cells with expression profiling provide a powerful model for a study of chondrogenic differentiation and further understanding of cartilage regeneration in the means of redifferentiated DECs and BMSCs.

Keywords: Dedifferentiation; Redifferentiation; Chondrocytes; Marrow stromal cells; Regeneration; Stem cells; Transplantation

Introduction

Chondrocytes differentiate from mesenchymal cells during embryonic development [1], and the phenotype of the differentiated chondrocyte is characterized by the synthesis, deposition, and maintenance of cartilage-specific extracellular matrix (ECM) molecules, including type II collagen and aggrecan [2–4]. It is rapidly lost during serial monolayer culture since the phenotype of differentiated chondrocytes is unstable in culture [5–8]. This process is referred to as "dedifferentiation" and is a major impediment to use of in mass cell populations for cell 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

* Corresponding author. Department of Pathology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan. Fax: +81-3-3353-3290.
E-mail address: umezawa@1985.jukuin.keio.ac.jp (A. Umezawa).
[1] Present address: National Research Institute for Child Health and Development, Tokyo, 157-8535, Japan.

0014-4827/03/$ – see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0014-4827(03)00130-7

Table 1
Monoclonal antibodies[a]

Antigen	Clones	Application	Source
CD14	MY4	F	BC
CD29, integrin β1	4B4	F	BC
CD31, PECAM-1	5.6E	F, H	IT
CD34	QBEnd-10	F	IT
CD44, Pgp-1	J-173	F, H	IT
CD45, LCA	DW124-5-2	F	BC
CD50, ICAM-3	HP2/19	F, H	IT
CDw90, Thy.1	F15-42-1	F, H	CT
CD105, endoglin	1G2	F	IT
CD106, VCAM-1	51-10C9	F	P
CD117, c-kit	Nu-ckit	F	CT
CD140a, PDGFR α	αR1	F	P
CD144, VE-cadherin	55-7H1	F	P
CD166, ALCAM	3A6	F, H	RDI

[a] BC, Beckman Coulter; CT, Cytotech; F, fluorescence-activated cell sorting; H, immunohistochemistry; ICAM-3, intercellular adhesion molecule-3; IT, Immunotech; LCA, leukocyte common antigen; P, Pharmingen; PECAM-1, platelet endothelial cell adhesion molecule-1; PDGFR, platelet-derived growth factor receptor; Pgp-1, phagocytic glycoprotein-1/hyaluronate receptor; RDI, Research Diagnostics, Inc.

type II collagen and cartilage proteins [9]. When cultured three-dimensionally in a scaffold such as agarose, collagen, and alginate, redifferentiated chondrocytes reexpressed the chondrocytic differentiation phenotype. We have successfully induced differentiation of chondrocytes and marrow stromal cells by chondrosphere formation.

Mesenchymal stem cells persist in adult bone marrow as a population of cells [10] with the capacity to differentiate to osteogenic [11], chondrogenic [12–14], or adipogenic [15] lineages both in vitro and when implanted subcutaneously in SCID mice [14]. These cells can be isolated from the marrow by standardized techniques and expanded in culture through many generations, while retaining their capacity to differentiate along these pathways when exposed to appropriate culture conditions. This property of mesenchymal stem cells opens up therapeutic opportunities for the treatment of lesions in mesenchymal tissues, and protocols have been devised for the treatment of defects in articular cartilage [16], bone [17], tendon [18], and meniscus [19] and for bone marrow stromal recovery [20] and osteogenesis imperfecta [21]. It also provides an opportunity to study the differentiation of mesenchymal cells in chondrocytes in the developing limb in detail.

We have generated cDNA libraries and carried out large-scale sequencing of cDNA constructed with the oligocapping methods [22] with the objective of conducting a global survey of gene expression and differentiation of genes that are specific for redifferentiated chondrocytes and marrow stromal-derived chondrogenic progenitor. We have successfully established human diploid chondrocyte strains and marrow stromal cell strains, and although these strains are mortal, they serve a useful tool as reproducible systems of human chondrogenesis with cell-type-specific profiling patterns.

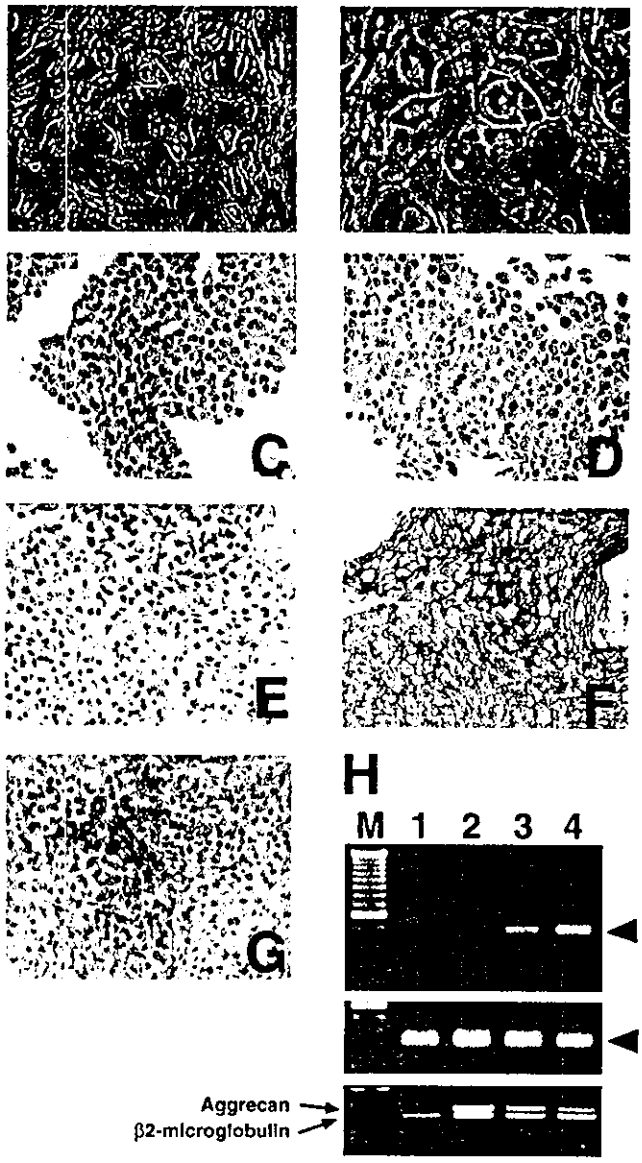

Fig. 1. Redifferentiation of dedifferentiated chondrocytes. Morphology of dedifferentiated chondrocytes in monolayer culture is fibroblast-like (original magnification: A, × 100; B, × 200). Aggregated cells started to accumulate alcian blue-positive substances in their cytoplasms, but not deposit matrix in an extracellular space [C, hematoxylin and eosin (HE) stain; D, Alcian blue stain]. Redifferentiation of the dedifferentiated chondrocytes on 5th day of the chondrosphere culture (E, HE stain; F, Alcian blue stain; G, toluidine blue stain). The differentiated cells produced an extracellular matrix that showed metachromasie with toluidine blue stain. Reverse transcription-polymer chain reaction analysis of gene expression for type II collagen (upper column, arrowhead) and aggrecan (lower column) during redifferentiation (H). β2-Microglobulin was used for an internal control of each RNA (arrowhead in the middle column; lower band indicated by arrow in the lower column). M, molecular markers; lane 1, monolayer culture; lane 2, 1 week of redifferentiation; lane 3, 2 weeks; and lane 4, 3 weeks.

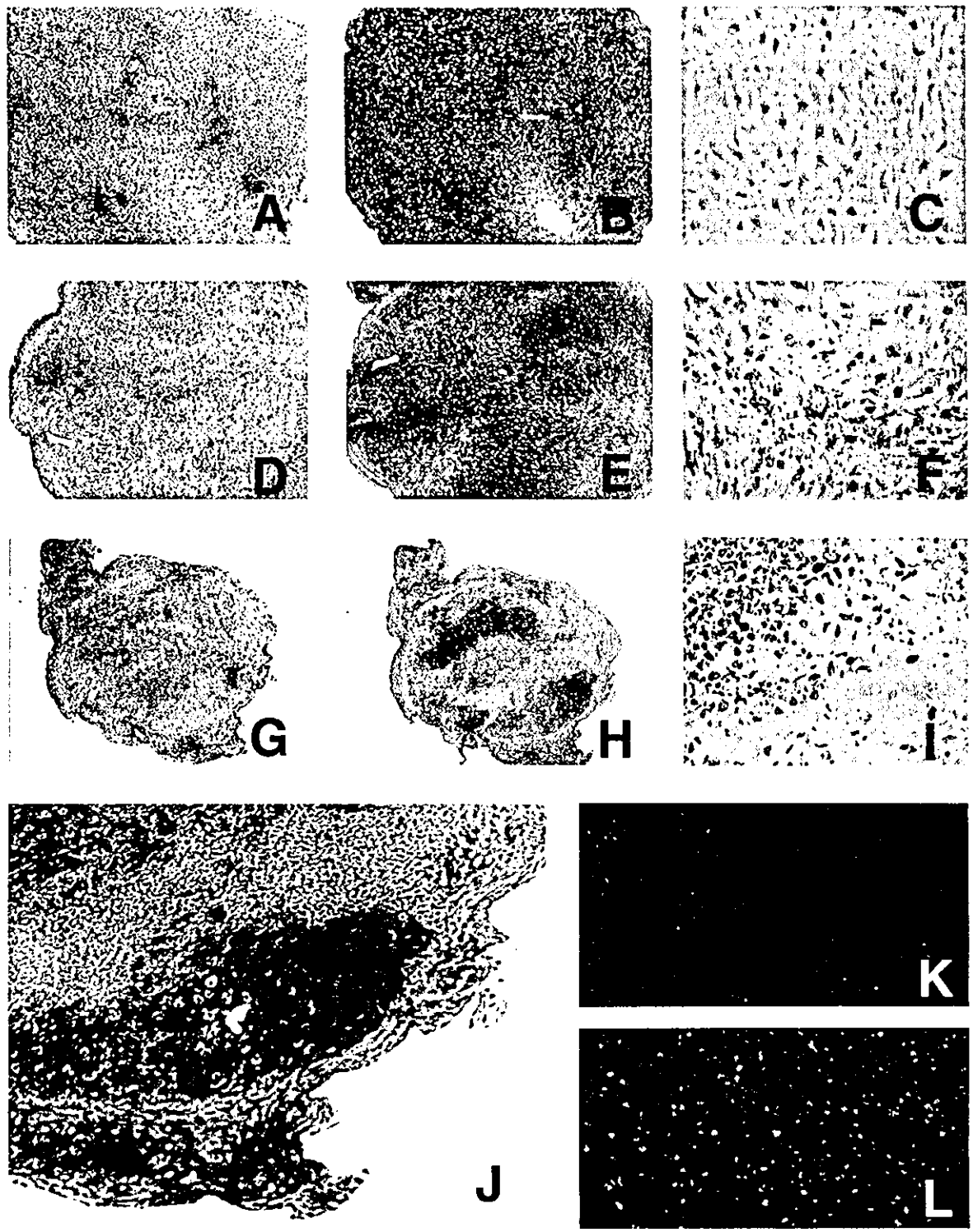

Fig. 2. Chondrosphere culture for a long period. Chondrosphere culture of dedifferentiated chondrocytes at 3 weeks [A, hematoxylin and eosin (HE) stain; B, Alcian blue stain; C, aggrecan], 2 months (D, HE stain; E, Alcian blue stain; F, aggrecan), 11 months (G, HE stain; H, Alcian blue stain; I, aggrecan). (J) High-power view of H. Immunocytochemical analysis was performed for the expression of aggrecan (C, F, and I). Immunocytochemical analysis for the expression of type II collagen in chondrosphere culture of dedifferentiated chondrocytes at just after 2 h of incubation (K) and at 3 weeks (L). Original magnification: A, B, D, E, G, and H, ×40; C, F, and I, ×200; J, ×100.