Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 325 (2004) 265–275
BBRC
www.elsevier.com/locate/ybbrc

Global analysis of the expression patterns of transcriptional regulatory factors in formation of embryoid bodies using sensitive oligonucleotide microarray systems

Wataru Gunji[a,b], Takahito Kai[a,b], Eiko Sameshima[c], Naomi Iizuka[c], Hiroaki Katagi[c], Takahiko Utsugi[b], Fumihiro Fujimori[a], Yasufumi Murakami[a,b,c,*]

[a] *Department of Biological Science and Technology, Faculty of Industrial Science and Technology, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan*

[b] *Department of Structural Genomics, Genome and Drug Discovery Research Center, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-0022, Japan*

[c] *Bio Matrix Research Inc. No603 room, Toukatsu techno plaza, 5-4-6 Kashiwanoha, Kashiwa, Chiba 278-0882, Japan*

Received 1 October 2004
Available online 22 October 2004

Abstract

We manufactured a highly sensitive oligonucleotide microarray system comprised entirely of transcription regulatory factors (a TF oligo microarray) in order to comprehensively analyze the expression profiles of transcription factors in mice. We compared the expression profiles of transcription regulatory factors in mouse embryonic stem (ES) cells and ES-differentiated cells by using this TF oligo microarray, a cDNA microarray, a GeneChip system, and quantitative RT-PCR. The TF oligo microarray was able to comprehensively analyze the expression profile of transcription regulatory factors. In addition, we used the manufactured TF oligo microarray to analyze the expression patterns of transcriptional regulatory factors during the formation of embryoid bodies. The TF array was able to reveal the chronologic expression profile of transcription regulatory factors involved in embryogenesis or the maintenance of pluripotency in ES cells.

Keywords: DNA microarray; Expression profile; Transcription factor; Oligonucleotide; Embryonic stem cells; Embryoid body; Differentiation

The processes of development and differentiation of mammalian cells are determined by the site/stage-specific regulatory mechanism of gene expression mediated by multiple transcription factors. Embryonic stem (ES) cells are pluripotent cells derived from the inner cell mass of developing blastocysts [1,2]. Regulation of the differentiation and development in mouse ES cells is controlled by critical transcription factors, such as STAT3 [3,4], Oct-3/4 [5,6], and Nanog [7,8]. In turn, these factors hierarchically regulate the expression of other transcription factors that regulate the differentiation of various cell lineages in mouse embryogenesis. Therefore, global analysis of the hierarchical structures of the transcription regulation involved in embryogenesis is important for elucidating the mechanism of the development and differentiation of mammalian cells.

The DNA microarray is a very useful tool for comprehensive determination of the expression levels and other information on novel genes identified through genome analysis. Thus, DNA microarray investigations are essential for a comprehensive analysis of the complex network of gene expression [9–11]. On the other hand, it is very difficult to detect low expression levels and slight changes of gene expression by DNA

* Corresponding author. Fax: +81 4 7122 1360.
E-mail address: yasufumi@rs.noda.tus.ac.jp (Y. Murakami).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.10.025

microarray, due to the limitations of fluorescence detection systems. Conventional DNA microarray systems usually monitor the expressions of tens of thousands of genes at a stretch. However, the dynamic ranges of detection systems are narrower than the ranges of expression levels between genes with low-level expression and those with high-level expression, such as housekeeping genes. Transcription factor's expressions and its changes seem to be low, so a higher sensitive DNA microarray system is required. Because both the expression and the changes in expression of transcription factors are generally low, a more sensitive DNA microarray system is required.

Probes are an important factor for the sensitivity of a DNA microarray and may be categorized into two types, cDNA probes and the oligonucleotide probes. The oligonucleotide probes may be further categorized into two types, short oligonucleotide probes of 20–30 mer, such as Affymetrix [12,13], and long oligonucleotide probes with more than 50 mer nucleotides [14,15]. cDNA probes are in need of manipulation for the amplification and purification of PCR product. A cDNA microarray has high hybridization efficiency, but cDNA causes cross-hybridization due to low sequence specificity. Short oligonucleotide probes have the advantage that they can be designed with high specificity sequences [16]. However, except for the probes directly synthesized on the array surface, amino-group modification is required for immobilization in the spotting of short oligonucleotide probes to slide glass. Long oligonucleotide probes are able to achieve higher sequence specificity than short oligonucleotide probes, and can hybridize to target DNA with high efficiency. Moreover, long oligonucleotide probes do not require amino-group modification.

In this study, we manufactured a highly sensitive oligonucleotide microarray system comprised of transcription regulatory factors of mice (TF oligo microarray) in order to comprehensively analyze the expression profiles of transcription factors having low expression levels. To evaluate this oligonucleotide microarray, we compared the sensitivity or accuracy among this manufactured TF oligo microarray, a cDNA microarray, and a GeneChip system, by using the expression profiles of transcription regulatory factors in mouse ES cells and ES-differentiated cells. Moreover, we analyzed the expression patterns of transcriptional regulatory factors during embryoid bodies (EBs) formation as an in vitro early embryogenesis model using a TF oligo microarray. We identified not only the changes of expression level in transcription regulatory factors that are known to be involved in the embryogenesis or in the maintenance of pluripotency of ES cells, but also the changes in many novel transcription factors that showed no previous indication of involvement in these processes. Accumulation of the expression profiles of transcription factors during embryogenesis will lead to elucidation of the molecular mechanism of the development and differentiation of mammalian cells.

Materials and methods

Cell culture and RNA isolation

D3 ES cells (CRL-1934; ATCC, Manassas, VA) were cultured on gelatin-coated tissue culture dishes, and murine embryonic fibroblast STO cells (CRL-1503; ATCC) were cultured in KnockOut Dulbecco's modified Eagle's medium (Invitrogen, Carlsbad, CA) supplemented with 20% KnockOut Serum Replacement (Invitrogen), 2 mM L-glutamine (Invitrogen), 0.1 mM MEM non-essential amino acid (Invitrogen), 50 U/mL penicillin–50 μg/mL streptomycin (Invitrogen), 1000 U/mL murine leukemia inhibitory factor (LIF; Invitrogen), and 0.1 mM 2-mercaptoethanol (Sigma, St. Louis, MO) at 37 °C in a 5% CO_2 atmosphere. For embryoid bodies (EBs) formation, undifferentiated ES cells were cultured in suspension in low-cell-binding dishes (Nunc, Roskilde, Denmark) for 4 days in medium without both LIF and 2-mercaptoethanol. Differentiation of EBs was induced by addition of 1 μM *all-trans* retinoic acid (RA; Sigma) for 7 days [17]. Using an RNeasy Mini Kit (Qiagen, Hilden, Germany), total RNA was extracted from ES cells that were induced to differentiate by addition of RA at day 0 or day 7. RNA quality was assessed with an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA).

Genechip hybridization

Isolated total RNAs were amplified and labeled as described in the GeneChip Expression Analysis Technical Manual, Rev.3 (Affymetrix, Santa Clara, CA). First, total RNA (7 μg) was converted into double-stranded cDNA using the Superscript Choice system for cDNA synthesis (Invitrogen) and T7-oligo(dT) primer (Affymetrix). Second-strand synthesis was accomplished using a 0.2 mM dNTP mix, 10 U *Escherichia coli* DNA ligase, 40 U *E. coli* DNA polymerase, and 2 U *E. coli* RNase H (Invitrogen). Double-stranded cDNA was purified by using a GeneChip sample Cleanup Module (Affymetrix). In vitro transcription reactions were performed using an Enzo BioArray High Yield RNA Transcript Labeling Kit (Affymetrix), which includes T7 RNA polymerase and Biotin-labeled ribonucleotides. Biotin-labeled cRNA was purified using a GeneChip Sample Cleanup Module. The concentration of cRNA was calculated from light absorbance at 260 nm using a UV spectrophotometer. Next, cRNA (15 μg) was fragmented at 94 °C in the presence of a fragmentation buffer (Affymetrix). Fifteen micrograms of the cRNA was hybridized to Affymetrix GeneChip Mouse Expression Array 430A (Affymetrix), which contains 22,690 probe sets representing 14484 full-length genes and EST clusters plus controls. The array was incubated for 16 h at 45 °C, then automatically washed and stained with streptavidin–phycoerythrin (Invitrogen). The Probe Array was scanned using a GeneArray Scanner (Affymetrix).

cDNA labeling

For the cDNA microarray and the TF oligo microarray, total RNA was labeled with aminoallyl-dUTP (Amersham Biosciences, Piscataway, NJ) using a CyScribe Post-Labeling Kit (Amersham Biosciences) as follows. Total RNA (20 μg) was mixed with 3 μL Anchored oligo(dT), heat-denatured at 70 °C for 5 min, and cooled at room temperature. To this mixture were added 4 μL of 5× CyScript buffer, 2 μL of 0.1 M DTT, 1 μL nucleotide mix, 1 μL aminoallyl dUTP, and 1 μL SyScript reverse transcriptase, and then the mixture was incubated at 42 °C for 90 min. After reverse transcription, 1 μL of 2.5 M NaOH was added and the mixture was incubated at 37 °C for 20 min. Then, 10 μL of 2 M Hepes-free acid was added. Aminoallyl modified cDNA was purified with a CyScribe GFX Purification Kit (Amersham

Biosciences). The purified cDNA was coupled with Cy3 reactive dye (Amersham Biosciences) or Cy5 reactive dye (Amersham Biosciences) in 15 μL of 0.1 M $NaHCO_3$ at room temperature for 60 min. For inactivating dye, 15 μL of 4 M hydroxylamine HCl was added, and the mixture was incubated at room temperature for 15 min. CyDye-labeled cDNA was purified with a CyScribe GFX Purification Kit.

cDNA microarray hybridization

We manufactured a cDNA microarray derived from an NIA mouse 15 and 7.4K cDNA clone set [18,19]. Pre-hybridization was done in a pre-hybridization buffer at 50 °C for 1 h. Hybridization solution was prepared from a mixture of Cy3-labeled and Cy5-labeled cDNA, 5× SSC, 0.3% SDS, and 20 μg/μL acetylated BSA. The cDNA microarray was covered with a Gap cover glass (MATSUNAMI, Osaka, Japan), hybridization solution was injected from the edge onto the chip, and the chip was placed in a hybridization cassette (TeleChem International, Sunnyvale, CA). The hybridization cassette was placed in a water bath and incubated at 50 °C for 14 h. After hybridization, the cDNA microarray was washed once in 2× SSC/0.1% SDS solution at 42 °C for 5 min, once in 0.1× SSC/0.1% SDS solution at room temperature for 10 min, and finally twice in 0.1× SSC solution at room temperature for 2 min using a Wash Station (TeleChem International). The cDNA microarray was dried by centrifugation. Hybridization images were scanned using a ScanArray Express (Perkin–Elmer, Norwalk, CT).

TF oligo microarray production

We identified 916 mouse homologs of human transcriptional regulators by amino acid homology searches of the Online Mendelian Inheritance in Man (OMIM) database (http://www.ncbi.nlm.nih.gov/Omim/). Moreover, we added 842 known mouse transcriptional regulators from the Mouse Genome Informatics (MGI) database (http://www.informatics.jax.org/), and removed 196 genes that overlap between mouse homologs of human transcriptional regulators and those of MGI (May, 2002). Identified transcriptional regulatory factors totaled 1562 genes. For each of the 1562 transcription regulators, we designed sequence-specific 75 mer probes.

Oligonucleotide probes were dissolved at 25 μM in 3× SSPE solution (Invitrogen), and printed in duplicate on γ-amino propyl silane-coated UltraGAPS Coated Slides (Corning, New York, NY) using a MicroGrid II (Genomic Solutions, Ann Arbor, MI). After printing, oligonucleotide probes were cross-linked to the slide with 600 mJ of ultraviolet light by using a UV Stratalinker 1800 (Stratagene, La Jolla, CA).

TF oligo microarray hybridization

For pre-hybridization, an oligo microarray was incubated in a pre-hybridization buffer (5× SSC, 0.1% SDS, and 0.1% BSA) at 42 °C for 1 h. The hybridization solution consisted of a mixture of Cy3-labeled and Cy5-labeled cDNAs, 15% formamide, 5× SSC, 0.5% SDS, and 20 μg/μL acetylated BSA (Invitrogen). Hybridization was performed with a hybridization chamber (Agilent Technologies) and hybridization Incubator (Apogent Technologies, Portsmouth, NH) at 42 °C for 17 h. Following the hybridization, the TF oligo microarray was washed once in 2× SSC/0.1% SDS solution at 42 °C for 5 min, once in 0.1× SSC/0.1% SDS solution at room temperature for 10 min, and twice in 0.1× SSC solution at room temperature for 2 min using a Wash Station. After washing, the oligo microarray was dried by centrifugation. Hybridization images were scanned using a ScanArray Express fluorescence laser scanner.

Microarray analysis

cDNA microarray and TF oligo microarray. Fluorescence intensity in the cDNA microarray and TF oligo microarray were quantified using QuantArray 3.0 software (Perkin–Elmer). In each array, genes whose expression levels were lower than those of the negative control spots under both conditions were removed as cutoffs. The signal intensity data were imported into the program GeneSpring 6.0 (Silicon Genetics, Redwood City, CA). The ratios of the signal intensities between the two cell types were calculated through Lowess normalization.

GeneChip system. Signal intensities data were analyzed using Microarray Suite 5.0 software (Affymetrix). All data were normalized using global normalization (i.e., using all probe sets). Subsequently, the ratios of the expression levels between the two cell types were calculated. Of the genes in the calculated ratio, genes classified as "absent" by the GeneChip software under both conditions were removed as cutoffs.

Quantitative RT-PCR

First-strand cDNA was synthesized from 10 μg of total RNA using the Superscript Choice system for cDNA synthesis (Invitrogen) as follows. Total RNA (10 μg) was mixed with 2 μL oligo(dT) primer (0.5 μg/μL), heat-denatured at 70 °C for 10 min, and cooled on ice. To this mixture were added 4 μL 5× first-strand buffer, 2 μL of 0.1 M DTT, 10 μL of 10 mM dNTP mixture, and 1 μL of 200 U/μL SuperScript II, followed by incubation at 42 °C for 60 min. After reverse transcription, the reaction mixture was incubated at 70 °C for 15 min and cooled on ice. For RNA degradation, 1 μL RNaseH was added, followed by incubation at 37 °C for 20 min. Synthesized cDNA was diluted with nuclease free water to 100 ng/μL and stored at −20 °C before use.

Real-time quantitative PCR was performed using an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA). The PCR mixture included TaqMan Universal PCR master mix, 900 nM forward and reverse primers, 200 nM TaqMan probe, and 100 ng template cDNA. The PCR conditions consisted of 50 °C for 2 min and 95 °C for 10 min, followed by 40 cycles of amplification at 95 °C for 15 s and 60 °C for 1 min. Primers and probes for amplification of Tbp1 were used as endogenous controls to normalize the amount of total RNA per sample.

Results and discussion

Microarray production and experimental design

To manufacture a highly sensitive microarray, we adopted long oligonucleotide probes for the following reasons. We can select sequences with high specificity to achieve the advance of superior accuracy. First, the use of these probes allowed us to select sequences with high specificity and thereby improve the accuracy. Second, the probes have high hybridization efficiency and do not require amino-group modification. We designed specific 75 mer probes for 1562 mouse transcription regulatory factors which were identified as homologs of human transcriptional regulators and known mouse transcriptional regulators. In addition to the probe design, we optimized various experimental equipment and processes (for example, the slide glasses, spotting solution, probe concentration, hybridization buffer, target labeling, temperature, and washing conditions) to create a high-sensitivity microarray system. Moreover, in order to avoid the saturation of the fluorescence

signals due to the limitation of the dynamic range of fluorescence detection, we diluted oligonucleotide DNA probe solution of some transcription factors prior to spotting. Because they were highly expressed constitutively on many mouse tissue samples and cultured cells in our examination of their expression profiles using TF oligo microarray (data not shown).

Total RNA from ES cells induced to differentiate by incubation with retinoic acid (RA) for 1 week and total RNA from ES cells not induced to differentiate were isolated for evaluation of the TF oligo microarray, with regard to reproducibility, sensitivity, accuracy, and comparison with other microarray systems. In the GeneChip system, two kinds of total RNA were amplified and labeled with biotin by in vitro transcription. The biotin-labeled cRNAs derived from the two kinds of cells were divided into four aliquots. The ratios of the expression levels between the two cell types were calculated from the 16 combination in quadruplicated hybridization experiments. Of the genes in the calculated ratio, genes classified as ''absent'' by Microarray Suite in both two conditions were removed as cutoff. A total of 6005 genes (26.5% of the total examined; Table 1) were considered absent in all 16 combinations of replicated hybridization. In the cDNA microarray system and the TF oligo microarray system, we labeled two kinds of total RNA with both Cy3 and Cy5 fluorescent dye (dye-swap). Labeled cDNA mixtures of two kinds of color combination were used for competitive hybridizations. The hybridization of each combination was done in duplicate for a total of four microarrays. In the cDNA microarray, genes whose expression levels were lower than that of a negative control (buffer) in both control and differentiation samples were removed as cutoffs. The total number of cutoff genes was 818 (3.7% of the total examined) for all replicated hybridizations. Among the genes in each TF oligo microarray, genes whose expression levels were lower than 1.5 times the value of the negative control probes in both samples were removed as cutoffs. A negative control probe was designed with a 75 mer non-redundant sequence for all known mouse genes. The total number of cutoff genes was 125 (8.0% of the total examined) for all replicated hybridizations. A considerably larger number of genes were removed using the GeneChip system than when using the other methods. This may be attributed to the low efficiency of hybridization due to the use of short oligonucleotide probes. On the other hand, the percentage of genes detected with the TF oligo microarray was similar to that detected with the cDNA microarray. This result indicated that the efficiency of hybridization of the 75 mer oligonucleotide is similar to that of cDNA probe. Fig. 1A compares the expression profiles of undifferentiated ES cells with those of differentiated ES cells induced by RA. The intensity value of each gene was averaged with the signals in four replicated hybridizations. The ratio-intensity plot shows the distribution of intensities and ratios for each gene (Fig. 1B). Ratio-intensity plots have been shown to be useful for revealing intensity-specific artifacts [20]. In the GeneChip system, the ratios lying on the low intensity signal appear widely dispersed. On the other hand, the ratios lying on the low intensity signal appear converged in the cDNA microarray. The distributions of spots obtained from the cDNA microarray and TF oligo microarray are similar, but that of the GeneChip system is different. Then, to examine whether the distribution of genes at the low expression level was reproducible, we calculated the coefficient of variation (CV) value of the expression ratio for all genes between replicated hybridization experiments in each microarray system. CV is the standard deviation divided by the mean. The average CV value for all genes in the GeneChip system was 17.3%, that in the cDNA microarray was 16.8%, and that in the TF oligo microarray was 15.9%. Fig. 2 shows the changes of CV value in the difference of expression levels in each microarray system. The reproducibility on the low expression level in the GeneChip system was poor, but the CV value decreased in proportion to the rise in expression level. In contrast with the GeneChip system, both the cDNA microarray and TF oligo microarray had high reproducibility, regardless of the expression levels. Since the fluorescent-labeled DNAs used for replicated hybridization in this experiment were divided from a homogeneous pool, the reproducibility of microarray data was almost dependent upon the microarray system itself. These results in low expression level may have been due to the differences in the characteristics of the probes in each microarray system.

Analysis of transcription factors in each microarray systems

In order to examine the changes in the expression of transcription factors with each microarray system, the

Table 1
Comparison of gene numbers detected in each microarray system

System	All genes		Transcription factors	
	All probes	Detected	All probes	Detected
GeneChip	22,690	16,685 (73.5%)	1096	759 (48.6%)
cDNA microarray	22,194	21,376 (96.3%)	828	801 (51.3%)
TF oligo microarray	1562	1437 (92.0%)	1562	1437 (92.0%)

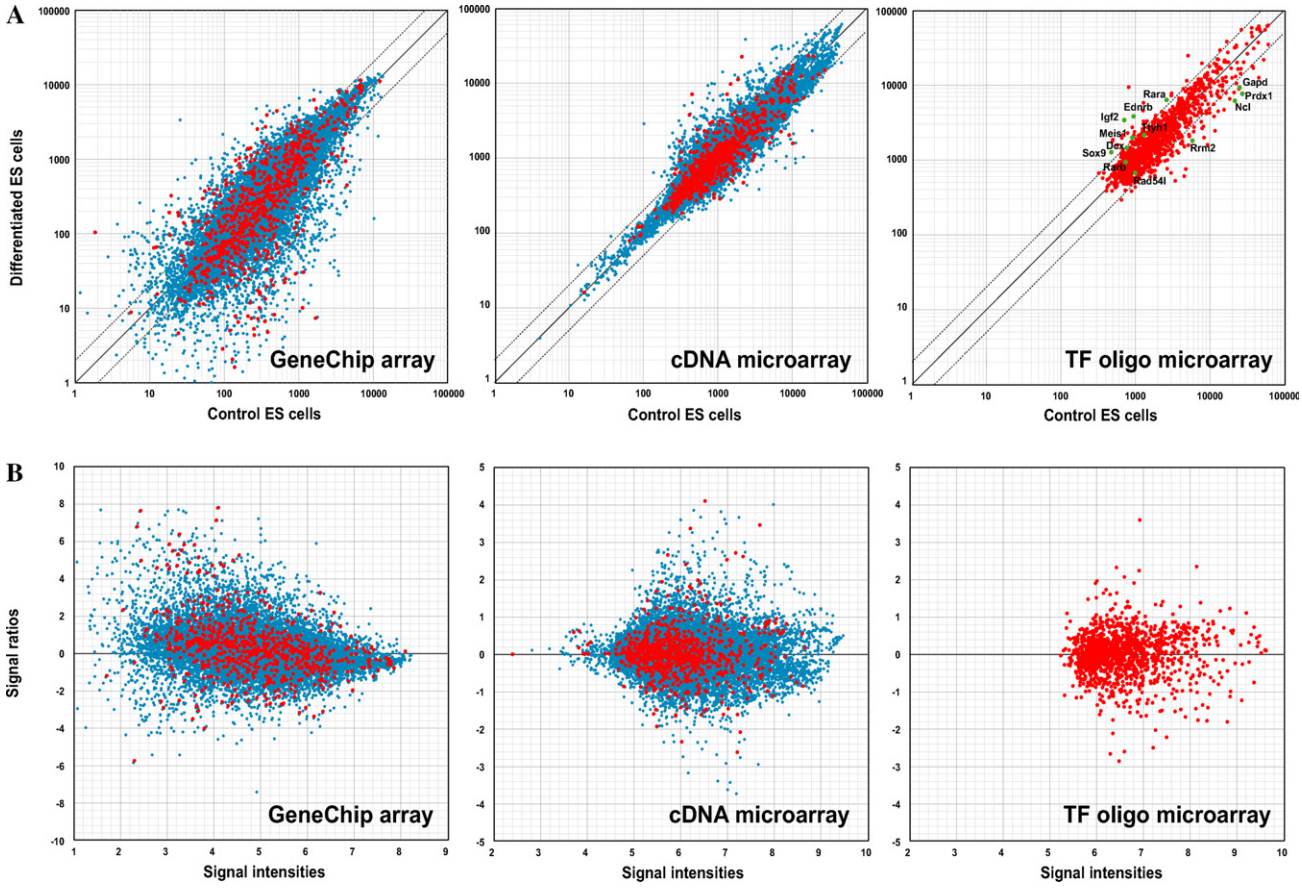

Fig. 1. (A) The scatter plots show a comparison of the expression profiles of ES cells with those of differentiated ES cells induced by RA in a GeneChip system (left panel), a cDNA microarray (center panel), and a highly sensitive TF oligonucleotide microarray (right panel). For each gene, the intensity values were averaged with the signals in four replicated hybridizations. A broken line shows twofold expression changes. To examine the expression changes of transcription factors, the expression profiles of 1562 genes identified as transcription factors were extracted from the expression profiles of all genes in the GeneChip system and cDNA microarray. Blue plots are all genes, and red are transcription factors. (B) The ratio-intensity plot shows the distribution of the average log ratio (differentiation/control) versus the average log intensity (differentiation × control) for each gene in replicated experiments.

expression profiles of 1562 genes identified as transcription factors were extracted from the expression profiles of all genes in the GeneChip system and cDNA microarray. The GeneChip Mouse Expression Array 430A contains the probes of 1096 transcription factors. Among them, the signals for 759 genes (48.6% of all transcription factors; Table 1) were detected at significant levels under both conditions in each of the 4 hybridizations. The 22K cDNA microarray contains the probes of 828 transcription factors. Among the 828 transcription factors, signals for 801 genes (51.3%) were detected at a significant level under both conditions in each of the 4 hybridizations. In TF oligo microarray, signals for 1437 genes (92.0%) were detected at a significant level under both conditions in each of the 4 hybridizations. The TF oligo microarray was able to detect the expression signals of most of the transcription factors. However, the other two systems were able to detect no more than about half of them, because many of the genes were removed from the statistical analysis, and none of them were used as probes originally. Table 2 shows the genes whose expression levels changed more than 2-fold in the TF oligo microarray. Several of the genes in the cDNA microarray and GeneChip system were not detected or not present. These results suggested that the TF oligo microarray has high sensitivity and can detect low expression levels that are not detectable using conventional systems.

Because both the expressions of the transcription factors and the changes in these expressions occur at low levels, there is need of a microarray system that can detect low signal levels. We were able to comprehensively analyze the expression profile of transcription regulatory factors using a TF oligo microarray, since this array has a high efficiency of hybridization due to long oligonucleotide probes. Moreover, the fluorescence signals of all oligonucleotide probes were admitted into the dynamic range of fluorescence detection by diluting the probe concentration for highly expressed transcription factors.

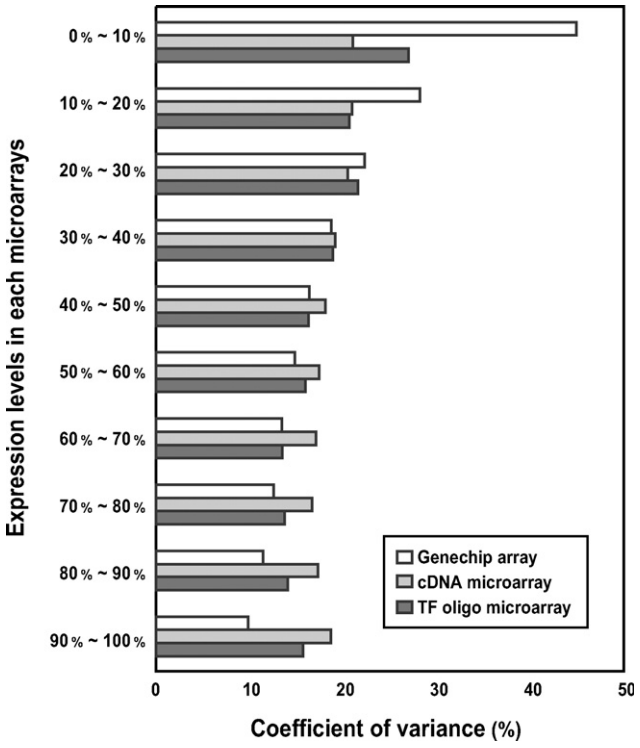

Fig. 2. The CV value in the difference of expression levels in each microarray system. In each microarray, all genes which were not removed on the cutoff step were sorted into 10 ranges according to the difference in their expression levels (differentiation × control).

Validation of microarray data by quantitative RT-PCR

Expression ratios in the TF oligo microarray and other microarray systems were confirmed using quantitative RT-PCR. We attempted to manufacture an oligo microarray having sufficiently high accuracy to enable identification of a little expression change such as 1.5-fold. Such data are disregarded in conventional DNA microarray. So, we included transcription factors whose expression levels changed as little as 1.5-fold in the TF oligo microarray as candidate quantitative RT-PCR. Quantitative RT-PCR was performed using probes of transcription factors: Dcx, Ednrb, Igf2, Meis1, Ncl, Prdx1, Rad54l, Rara, Rarb, Rrm2, Sox9, and Ttyh1, and Gapd as a control. Same lots of total RNA were used in the quantitative RT-PCR and microarray experiments. The ratio of target expression levels was normalized by using those of the Tbp (*Mus musculus* TATA box-binding protein) gene, and the fold change for each gene in differentiated ES cells compared to control ES cells was calculated (Fig. 3). The gene expression data provided by the TF oligo microarray were most similar to those of quantitative RT-PCR, and TF oligo microarray could also detect low expression levels which other microarrays could not. The result of quantitative RT-PCR suggested that the change of signal levels between different samples greater than 1.5-fold found by TF oligo microarray turned to significant changes of gene expression. The TF oligo microarray has high sensitivity and high accuracy due to long oligonucleotide primers with specific sequence. Although the sensitivity of the cDNA microarray had excellent reproducibility, the data on the cDNA microarray were not in very good agreement with those of quantitative RT-PCR in the validation of microarray data. cDNA probes have high efficiency of hybridization, but their accuracy may decline since they cause cross-hybridization due to low sequence specificity. In summary, we were able to comprehensively analyze the expression profiles of transcription regulatory factors with high sensitivity and high accuracy by using a TF oligo microarray. It is expected that the use of the TF oligo microarray will allow comprehensive analysis and highly reliable expression profiles of transcription regulatory factors.

Expression analysis during embryoid bodies formation

We used a manufactured TF oligo microarray to analyze the expression patterns of transcriptional regulatory factors during the formation of embryoid bodies. The embryogenesis and differentiation progressed with site/stage-specific regulation of transcription for many genes by several transcription factors. In the post-implantation stage of embryogenesis, the expressions of transcriptional regulatory genes and signal transducers, and the genes regulated by these are changed drastically. Therefore, examining these expression changes chronologically is important for understanding the mechanism of development and differentiation. ES cells can form three-dimensional aggregates called embryoid bodies (EBs) spontaneously by removal of the feeder cells and LIF from the suspension culture. EBs can differentiate to various cell lineages, so EBs are a powerful tool for characterizing the function of the early stage cells of embryogenesis, and they mimic post-implantation embryonic tissues [21]. Furthermore, the distribution of developmental marker genes in EBs strictly correlates with their expressions observed during specific stages of embryogenesis [21]. Moreover, EBs have the advantage that the function of a specific gene can easily be inhibited by targeted mutagenesis or RNA interference. Although the expressions of several marker genes of gastrulation have been apparent during the formation of EBs, the molecular mechanism by which EBs mimic post-implantation embryonic tissues remains to be clarified. We considered that the accumulation of chronological data on the expression of transcription factors during formation of EBs would serve to identify the transcription factors that play important roles in the embryogenesis or maintenance of pluripotent ES cells. The TF oligo microarray is more suitable than the microarray comprised of tens of thousands of genes

Table 2
The genes whose expression levels changed more than 2-fold in the TF oligo microarray

Gene symbol	TF oligo microarray	cDNA microarray	GeneChip
2410129E14Rik	12.06	10.97	No probe
Vim	5.08	2.34	No probe
Igf2	5.00	17.19	9.59
Ndn	4.70	6.55	4.58
Ednrb	4.19	No probe	No probe
Gata2	3.90	3.14	7.17
App	3.75	6.10	4.92
2900026H06Rik	3.74	No probe	No probe
Snap91	3.31	3.37	9.58
Crim1	3.11	2.95	No probe
Rgs2	3.10	1.95	No probe
Gata3	3.03	3.82	58.11
Idb4	2.84	3.93	No probe
Sox9	2.77	No probe	No probe
Elavl2	2.65	2.85	No probe
Ttc3	2.65	2.07	2.75
Fxr2h	2.60	0.85	0.92
Il9r	2.54	No probe	Not detect
Rara	2.51	0.97	1.03
Acas1	2.48	No probe	1.34
Evx1	2.43	No probe	Not detect
Runx3	2.42	No probe	Not detect
Prkar1a	2.39	1.38	No probe
Mrg1	2.39	3.50	1.22
Rorc	2.37	No probe	No probe
Mllt1	2.33	No probe	2.20
Cyln2	2.32	No probe	4.18
Meis1	2.28	1.84	No probe
Nab1	2.23	1.42	1.45
Hoxb5	2.17	No probe	144.17
Mapk8ip	2.16	2.49	No probe
Rom1	2.15	No probe	Not detect
Shox2	2.13	No probe	31.98
Gata4	2.10	1.04	1.83
Ntn1	2.08	No probe	7.69
Bmi1	2.07	2.23	No probe
Rcn2	2.02	2.96	2.31
Il11ra1	2.00	No probe	No probe
Otx2	0.14	0.26	No probe
Nr3c1	0.16	0.79	1.66
Zic3	0.16	No probe	No probe
Dnmt3b	0.18	0.32	0.09
Hmga1	0.21	0.23	0.25
Spink3	0.23	No probe	0.19
Hells	0.24	0.31	0.23
Rfx3	0.28	No probe	No probe
Prdx1	0.29	0.59	0.42
Ncl	0.29	0.53	No probe
Ccne1	0.29	0.35	0.20
Pcna	0.31	0.44	No probe
Cdh1	0.31	0.48	0.37
Rrm2	0.32	0.60	0.40
Utf1	0.33	No probe	0.13
Mcmd6	0.34	0.36	0.30
Hprt	0.34	0.40	0.36
Csda	0.34	0.39	No probe
Apoe	0.35	0.19	0.11
Pitx2	0.35	No probe	0.26
Top2a	0.35	0.52	No probe
Zfp98	0.36	No probe	Not detect
Nme2	0.36	No probe	0.53
Pabpc1	0.37	0.55	0.65
Tieg	0.37	Not detect	0.40
Akr1b3	0.38	0.32	0.34
2310005B10Rik	0.38	0.48	0.29
Txn1	0.39	0.50	0.61
Pmaip1	0.39	0.73	0.15
Lbr	0.40	0.48	No probe
D19Wsu57e	0.40	0.51	0.37
Apex1	0.40	0.63	No probe
Trp53	0.41	0.36	No probe
Tgif	0.41	0.46	0.15
Foxh1	0.41	No probe	No probe
Orc1	0.41	No probe	No probe
Apc2	0.42	No probe	40.44
Rpa1	0.42	0.56	No probe
Rps4x	0.42	0.74	0.72
Sox13	0.43	0.64	0.59
Cdc2a	0.43	0.49	0.38
Eif2a	0.43	0.56	0.79
Mcmd	0.43	0.35	No probe
Pou5f1	0.44	0.16	0.11
2310042L19Rik	0.45	No probe	No probe
Nmyc1	0.45	0.55	No probe
Tcf7	0.45	No probe	0.16
Rps3	0.46	0.46	No probe
Xrcc1	0.46	0.68	0.54
Hmga2	0.46	0.61	0.15
4931417M11Rik	0.46	0.50	0.32
Trim28	0.46	0.57	0.58
Rad53	0.47	0.75	0.36
Ccna2	0.47	0.51	0.49
Recc1	0.47	1.77	0.96
Trip13	0.47	0.48	No probe
Klf9	0.48	0.50	0.13
Hmgb2	0.48	Not detect	No probe
Gpiap1	0.48	0.94	0.49
Sdh1	0.48	0.54	No probe
Tfdp1	0.49	0.60	No probe
Cnbp	0.49	0.81	0.58
Cebpa-rs1	0.49	0.59	0.34
Mcmd4	0.49	0.68	No probe
Sgk	0.49	0.56	0.33
Rnac-pending	0.50	0.58	No probe

for analysis of the expression profiles of transcription factors whose expression seems to be low, because the TF oligo microarray can monitor the comprehensive expression profile of transcription regulatory factors with high sensitivity and high accuracy. We performed a time course analysis of the expression profiles for transcription factors during formation of EBs as an in vitro model of early embryogenesis by using the TF oligo microarray.

Leahy et al. [22] reported that the expression pattern of EBs prior to day 3 in suspension culture is equivalent to that of pregastrulation stage embryos, and EBs between days 3 and 5 contain cell types present in embryos during gastrulation. Therefore, the transcription factors whose expression levels changed over the first 96 h of

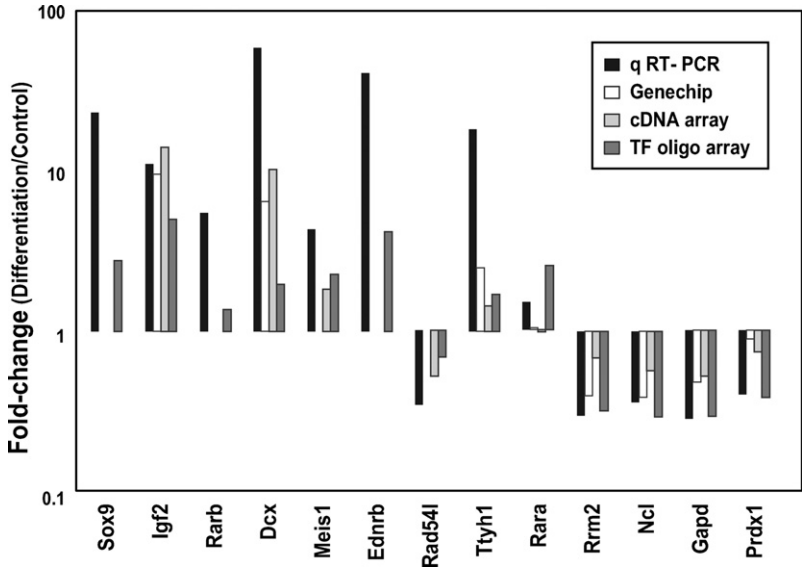

Fig. 3. Validation of the microarray data by quantitative RT-PCR. Expressed genes at various levels in TF oligo microarray and other microarray systems were confirmed using quantitative RT-PCR. Quantitative RT-PCR was performed using probes of transcription factors: Dcx, Ednrb, Igf2, Meis1, Ncl, Prdx1, Rad54l, Rara, Rarb, Rrm2, Sox9, and Ttyh1, and Gapd as a control. Quantitative RT-PCR and microarray experiments used same lots of total-RNA. The ratio of RNA levels was normalized by using that of the Tbp (*Mus musculus* TATA box-binding protein) gene. For each gene, the fold change between differentiated ES cells and control ES cells was calculated. The genes on the graph are arranged in order of their low spot intensities.

the formation of EBs may be involved in the regulation of the pregastrulation and gastrulation stages. We harvested EBs at 24, 48, 72, and 96 h in low-cell-binding dishes in medium without both LIF (leukemia inhibitory factor) and 2-mercaptoethanol. Undifferentiated ES cells were harvested just before induction of EBs formation as a 0-h control. The total RNAs of EBs in suspension culture at four time points—24, 48, 72, and 96 h—were compared with those of EBs at 0 h by using four TF oligo microarrays (Fig. 4). The number of transcription factors that showed a change in expression levels was increased with the time shift. A total of 53 transcription factors whose expression increased or decreased more than 2-fold at least at one point are shown in Fig. 5. Some of these genes—i.e., Ctsl, Cyr61, Dazl, Klf5, and Pim1—were verified by quantitative RT-PCR. We were able to confirm that the changes in expression detected on TF oligo microarray were also apparent on quantitative RT-PCR.

The genes whose expression levels decreased compared with the 0-h controls contained a specific undifferentiated marker; Zfp42 (Rex-1) [23]. Remarkable down-regulation of STAT3, which plays an important role in the maintenance of pluripotent ES cells, was observed. STAT3 activation is essential for maintenance of an LIF-mediated undifferentiated state in ES cells [3,4]. However, the genes targeted by STAT3 in the maintenance of pluripotent ES cells remain to be elucidated. We observed that the inactivation of STAT3 due to LIF deficiency had an effect on the expression

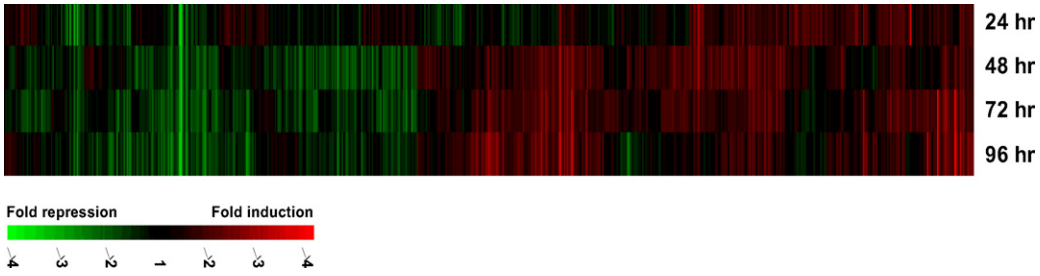

Fig. 4. Cluster images of all transcriptional regulatory factors showing the time-course effects on the RNA levels during formation of EBs. ES cells differentiated into EBs were harvested at 24, 48, 72, and 96 h in suspension culture without both LIF and 2-mercaptoethanol. The total RNAs of EBs at four time points were compared with that of EBs at 0 h by using four TF oligo microarrays. The signal intensity data from the TF oligo microarray were imported into the program GeneSpring 6.0, and the ratios of the signal intensities between cells at each time point and that of the cells at 0 h were calculated. The color scale represents the ratio of hybridization signals between cells at each time point and those at 0 h, and ranges from −4 (brightest green) to 4 (brightest red).

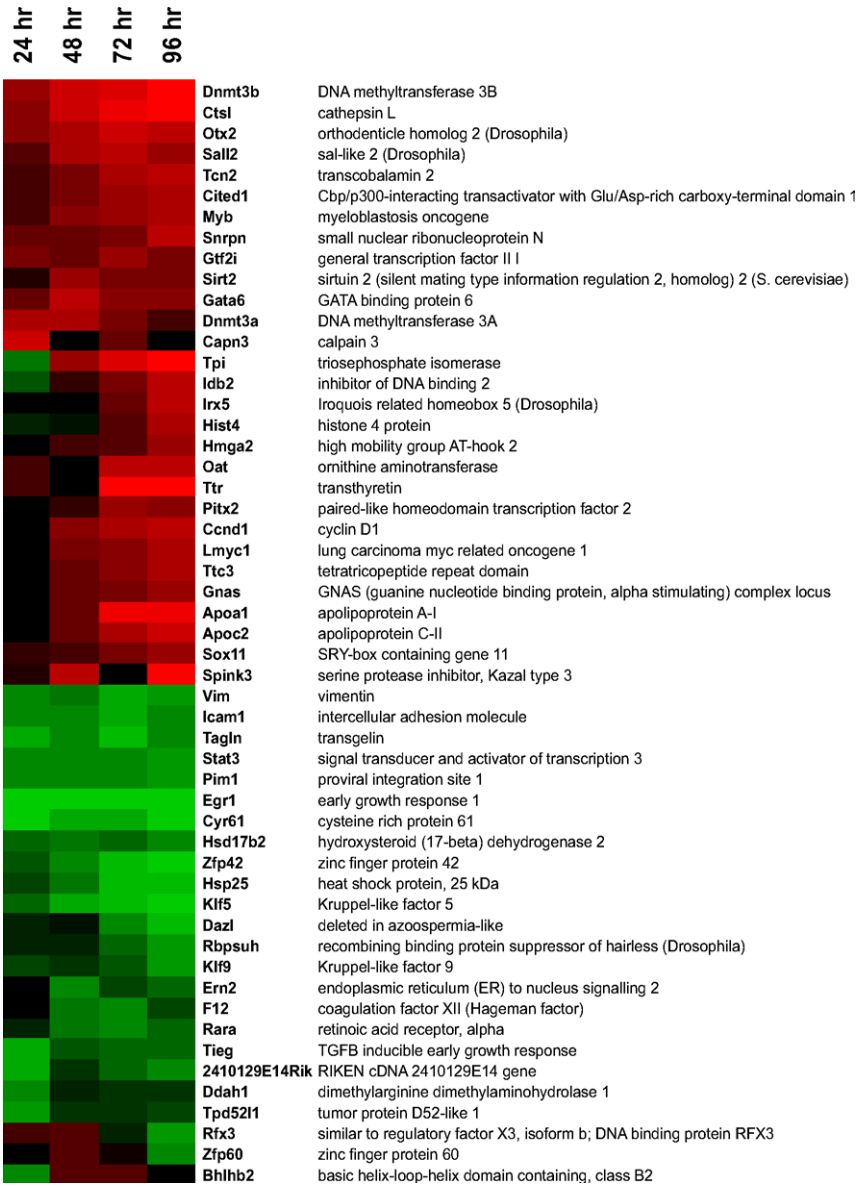

Fig. 5. Cluster images of transcription factors whose expression changed more than 2-fold at at least one time point. Fifty-three genes were selected. The names of the genes are given along with a brief description.

profiles of many transcription factors. The expression level of STAT3 at 4 h after removal of LIF was already reduced to approximately half that in the undifferentiated state. Owing to this, the genes whose expression levels decreased in the early stage of the formation of EBs were expected to be involved in the maintenance of the pluripotency of cells with regulation by STAT3.

Rbpsuh plays a central role for the signal transduction of Notch receptors [24]. The Notch intracellular domain in the nucleus interacts with the Rbpsuh and activates transcription of various target genes [25]. Notch-Rbpsuh signaling mediates the cell fate in development, and the gene expression of Rbpsuh has been reported in pre-implantation development from unfertilized eggs until late blastocyst stage [26]. In our experiment, Rbpsuh expression decreased after induction of differentiation, so this result suggested that the Notch signaling pathway through Rbpsuh is involved in the development of not only preimplantation in the undifferentiated state but also post-implantation. The mimicking of expression patterns of post-implantation embryonic tissues in the formation of EBs requires not only removal of LIF but also aggregation of cells. The aggregations of ES cells are induced by using a low-binding culture dish. We observed the down-regulation of intercellular adhesion molecule Icam1 at 24 h during the formation of EBs [27]. In the signals of cell differentiation, the cell-to-cell interaction due to cell aggregation will play essential roles in the determination of axis and other cell fates.

On the other hand, several genes (Gata6, Ttr, Otx2, Irx5, Dnmt3a, and Dnmt3b) that have previously been reported to be expressed in embryogenesis were induced during the formation of EBs. Members of the GATA family of zinc finger transcription factor, Gata6 is visceral endoderm marker [28–30]. And Ttr (transthyretin) is a visceral endoderm marker [31,32]. In addition to the known differentiation or undifferentiation markers, the expression levels of some genes likely to participate in embryogenesis were up-regulated. The homeobox gene Otx2 is expressed in the visceral endoderm and plays important roles in the initial patterning of the anterior–posterior axis in the mouse embryo [33]. Irx5, a member of the *Iroquois* homeobox gene family, has been shown to be expressed during early mouse embryogenesis [34,35]. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation of the mouse genome during early post-implantation development, and their proteins are expressed at different stages of embryogenesis [36,37]. The genes of both Dnmt3a and Dnmt3b were expressed at 24 h in the formation of EBs, so the genes which regulate these genes may be expressed just after abolition of the undifferentiated state. In this way, changes in the expression of many transcription factors during the formation of EBs were correlated with expression patterns reported in the stages of embryogenesis.

Furthermore, the changes in expression of many transcription factors not known to participate in the embryogenesis or the maintaining of pluripotency in ES cells were also identified. The expression of Sox11 (SRY-box containing gene 11) was up-regulated at 96 h during formation of EBs. The sox gene family plays an important role in the regulation of embryonic development and cell fate determination [38]. Sox11 has been markedly detected in the central nervous system of mouse embryos or a wide range of tissues [39], so it may play an important role for the regulation of the gene expressions in these tissues. These results suggest that the transcription factors whose expression was induced during the early stage of the formation of EBs were more involved in early embryogenesis than in maintenance of ES cells. At only 4 days after induction of the formation of EBs, many transcription factors related to embryogenesis were up-regulated or down-regulated, so key factors which control the expression of transcription factors are expected to show changes in expression in the early stage of the formation of EBs.

In this way, the TF aligo microarray could be used to analyze the chronological expression of transcription regulatory factors that are known to be involved in embryogenesis or the maintenance of pluripotency in ES cells. In the present study, the changes in the expression of many transcription factors that are not known to participate in the embryogenesis or the maintenance of pluripotency in ES cells were also identified. The participation of these transcription factors into embryogenesis and maintaining of pluripotency in ES cells remains to be identified. If we inhibit the functions of transcription factors with knock-down methods such as RNA interference, we may be able to clarify their involvement in embryogenesis by examining the changes in phenotype and expression of differentiation markers. We are currently performing functional analyses using RNA interference and forced expression of the transcription factors whose expression changes were observed in this experiment. Additional analysis of the observations made using the TF oligo microarray system will lead to identification of novel transcription factors that play important roles in the embryogenesis or maintenance of pluripotent ES cells. Accumulation of the expression profiles of transcription factors for the embryogenesis or differentiation into various lineages from stem cells may lead to elucidation of the molecular mechanism of the development and differentiation of mammalian cells.

Acknowledgments

We thank Masanori Togashi and Dr. Kohji Itoh (Science University of Tokyo) for supporting sequence design of microarray probes. This work was supported by the grants supplied by the Ministry of Economy, Trade and Industry (METI), the Organization for Pharmaceutical Safety and Research, and the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

References

[1] M.J. Evans, M.H. Kaufman, Establishment in culture of pluripotential cells from mouse embryos, Nature 292 (1981) 154–156.
[2] G.R. Martin, Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells, Proc. Natl. Acad. Sci. USA 12 (1981) 7634–7638.
[3] H. Niwa, T. Burdon, I. Chambers, A. Smith, Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3, Genes Dev. 12 (1998) 2048–2060.
[4] T. Matsuda, T. Nakamura, K. Nakao, T. Arai, M. Katsuki, T. Heike, T. Yokota, STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells, EMBO J. 18 (1999) 4261–4269.
[5] J. Nichols, B. Zevnik, K. Anastassiadis, H. Niwa, D. Klewe-Nebenius, I. Chambers, H. Scholer, A. Smith, Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4, Cell 95 (1998) 379–391.
[6] H. Niwa, J. Miyazaki, A.G. Smith, Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells, Nat. Genet. 24 (2000) 372–376.
[7] K. Mitsui, Y. Tokuzawa, H. Itoh, K. Segawa, M. Murakami, K. Takahashi, M. Maruyama, M. Maeda, S. Yamanaka, The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells, Cell 113 (2003) 632–642.

[8] I. Chambers, D. Colby, M. Roberson, J. Nichols, S. Lee, S. Tweedle, A. Smith, Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells, Cell 113 (2003) 643–655.
[9] W. Gunji, T. Kai, Y. Takahashi, Y. Maki, W. Kurihara, T. Utsugi, F. Fujimori, Y. Murakami, Global analysis of the regulatory network structure of gene expression in *Saccharomyces cerevisiae*, DNA Res. 11 (2004) 163–177.
[10] C.L. Tucker, J.F. Gera, P. Uetz, Towards an understanding of complex protein networks, Trends Cell Biol. 11 (2001) 102–106.
[11] T. Ideker, V. Thorsson, J.A. Ranish, R. Christmas, J. Buhler, J.K. Eng, R. Bumgarner, D.R. Goodlett, R. Aebersold, L. Hood, Integrated genomic and proteomic analyses of a systematically perturbed metabolic network, Science 291 (2001) 929–934.
[12] R.J. Lipshutz, S.P. Fodor, T.R. Gingeras, D.J. Lockhart, High density synthetic oligonucleotide arrays, Nat. Genet. 21 (1999) 20–24.
[13] D.J. Lockhart, H. Dong, M.C. Byrne, M.T. Follettie, M.V. Gallo, M.S. Chee, M. Mittmann, C. Wang, M. Kobayashi, H. Horton, E.L. Brown, Expression monitoring by hybridization to high-density oligonucleotide arrays, Nat. Biochem. 14 (1996) 1675–1680.
[14] M.D. Kane, T.A. Jatkoe, C.R. Stumpf, J. Lu, J.D. Thomas, S.J. Madore, Assessment of the sensitivity and specificity of oligonucleotide (50 mer) microarrays, Nucleic Acids Res. 28 (2000) 4552–4557.
[15] F. Li, G.D. Stormo, Selection of optimal DNA oligos for gene expression arrays, Bioinformatics 17 (2001) 1067–1076.
[16] A. Relogio, C. Schwager, A. Richter, W. Ansorge, J. Valcarcel, Optimization of oligonucleotide-based DNA microarrays, Nucleic Acids Res. 30 (2002) e51.
[17] W.G. van Inzen, M.P. Peppelenbosch, M.W. van den Brand, L.G. Tertoolen, S.W. de Laat, Neuronal differentiation of embryonic stem cells, Biochim. Biophys. Acta 1312 (1996) 21–26.
[18] G.J. Kargul, D.B. Dudekula, Y. Qian, M.K. Lim, S.A. Jaradat, T.S. Tanaka, M.G. Carter, M.S. Ko, Verification and initial annotation of the NIA mouse 15K cDNA clone set, Nat. Genet. 28 (2001) 17–18.
[19] V. VanBuren, Y. Piao, D.B. Dudekula, Y. Qian, M.G. Carter, P.R. Martin, C.A. Stagg, U.C. Bassey, K. Aiba, T. Hamatani, G.J. Kargul, A.G. Luo, J. Kelso, W. Hide, M.S. Ko, Assembly, verification, and initial annotation of the NIA mouse 7.4K cDNA clone set, Genome Res. 12 (2002) 1999–2003.
[20] J. Quackenbush, Microarray data normalization and transformation, Nat Genet. Suppl. (2002) 496–501.
[21] I. Desbaillets, U. Ziegler, P. Groscurth, M. Gassmann, Embryoid bodies: an in vitro model of mouse embryogenesis, Exp. Physiol. 85 (2000) 645–651.
[22] A. Leahy, J.W. Xiong, F. Kuhnert, H. Stuhlmann, Use of developmental marker genes to define temporal and spatial patterns of differentiation during embryoid body formation, J. Exp. Zool. 284 (1999) 67–81.
[23] E. Ben-Shushan, J.R. Thompson, L.J. Gudas, Y. Bergman, Rex-1, a gene encoding a transcription factor expressed in the early embryo, is regulated via Oct-3/4 and Oct-6 binding to an octamer site and a novel protein, Rox-1, binding to an adjacent site, Mol. Cell. Biol. 18 (1998) 1866–1878.
[24] T. Honjo, The shortest path from the surface to the nucleus: RBP-J kappa/Su(H) transcription factor, Genes Cells 1 (1996) 1–9.
[25] S. Artavanis-Tsakonas, M.D. Rand, R.J. Lake, Notch signaling: cell fate control and signal integration in development, Science 284 (1999) 770–776.
[26] S. Cormier, S. Vandormael-Pournin, C. Babinet, M. Cohen-Tannoudji, Developmental expression of the Notch signaling pathway genes during mouse preimplantation development, Gene Expr. Patterns 4 (2004) 713–717.
[27] D.E. Staunton, S.D. Marlin, C. Stratowa, M.L. Dustin, T.A. Springer, Primary structure of ICAM-1 demonstrates interaction between members of the immunoglobulin and integrin supergene families, Cell 52 (1988) 925–933.
[28] E.E. Morrisey, H.S. Ip, M.M. Lu, M.S. Parmacek, GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm, Dev. Biol. 177 (1996) 309–322.
[29] E.E. Morrisey, Z. Tang, K. Sigrist, M.M. Lu, F. Jiang, H.S. Ip, M.S. Parmacek, GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo, Genes Dev. 12 (1998) 3579–3590.
[30] M. Koutsourakis, A. Langeveld, R. Patient, R. Beddington, F. Grosveld, The transcription factor GATA6 is essential for early extraembryonic development, Development 126 (1999) 723–732.
[31] K. Abe, H. Niwa, K. Iwase, M. Takiguchi, M. Mori, S.I. Abe, K. Abe, K. Yamamura, Endoderm-specific gene expression in embryonic stem cells differentiated to embryoid bodies, Exp. Cell Res. 229 (1996) 27–34.
[32] D.R. Soprano, K.J. Soprano, M.L. Wyatt, D.S. Goodman, Induction of the expression of retinol-binding protein and transthyretin in F9 embryonal carcinoma cells differentiated to embryoid bodies, J. Biol. Chem. 263 (1988) 17897–17900.
[33] C. Kimura, M.M. Shen, N. Takeda, S. Aizawa, I. Matsuo, Complementary functions of Otx2 and Cripto in initial patterning of mouse epiblast, Dev. Biol. 235 (2001) 12–32.
[34] A. Bosse, A. Stoykova, K. Nieselt-Struwe, K. Chowdhury, N.G. Copeland, N.A. Jenkins, P. Gruss, Identification of a novel mouse Iroquois homeobox gene, Irx5, and chromosomal localisation of all members of the mouse Iroquois gene family, Dev. Dyn. 218 (2000) 160–174.
[35] D.R. Cohen, C.W. Cheng, S.H. Cheng, C. Hui, Expression of two novel mouse Iroquois homeobox genes during neurogenesis, Mech. Dev. 91 (2000) 317–321.
[36] M. Okano, D.W. Bell, D.A. Haber, E. Li, DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development, Cell 99 (1999) 247–257.
[37] D. Watanabe, I. Suetake, T. Tada, S. Tajima, Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis, Mech. Dev. 118 (2002) 187–190.
[38] O. Lioubinski, M. Muller, M. Wegner, M. Sander, Expression of Sox transcription factors in the developing mouse pancreas, Dev. Dyn. 227 (2003) 402–408.
[39] M. Hargrave, E. Wright, J. Kun, J. Emery, L. Cooper, P. Koopman, Expression of the Sox11 gene in mouse embryos suggests roles in neuronal maturation and epithelio-mesenchymal induction, Dev. Dyn. 210 (1997) 79–86.