Positions of pluripotency genes and hepatocyte-specific genes in the nucleus before and after mouse ES cell differentiation

K. Udagawa¹ and T. Ohyama¹,²

¹Major in Integrative Bioscience and Biomedical Engineering, Graduate School of Science and Engineering, Waseda University, Tokyo, Japan
²Department of Biology, Faculty of Education and Integrated Arts and Sciences, Waseda University, Tokyo, Japan

Corresponding author: T. Ohyama
E-mail: ohyama@waseda.jp

Received October 31, 2012
Accepted March 19, 2013
Published March 24, 2014
DOI http://dx.doi.org/10.4238/2014.March.24.2

ABSTRACT. Spatial positioning of genes in the cell nucleus plays an important role in the regulation of genomic functions. Evidence for changes in gene positioning associated with transcriptional activity has been reported. However, our understanding of this phenomenon is still quite limited. We examined how pluripotency genes and hepatocyte-specific genes behave during the differentiation of mouse embryonic stem (ES) cells into hepatocytes, by targeting the loci of the Klf4, Nanog, Oct4, Sox2, Cyp7a1, Pck1, Tat, and Tdo2 genes, and using three-dimensional fluorescence in situ hybridization analyses. We found that each gene has a distinctly inherent localization profile in the ES cell nucleus. During differentiation, the Klf4, Nanog, Oct4, Cyp7a1, Pck1, and Tat loci shifted toward the nuclear center, while the Sox2 and Tdo2 loci shifted toward the periphery. The Klf4, Nanog, Oct4, and Tdo2 seem to prefer the outer regions, rather than the inner regions, when they are active. We also found that the radial positioning of the focused genes in the hepatocyte cell nucleus was highly correlated with
the local GC content and the gene density of the surrounding region, but not with gene activity.

**Key words:** Gene localization; ES cell; Hepatocyte; Cell differentiation; Pluripotency gene; Hepatocyte-specific gene

**INTRODUCTION**

Gene expression in eukaryotes is controlled by various mechanisms, including regulation by modifications of genomic DNA and histones (epigenetic regulation), nucleosome positioning, or higher-order chromatin structures (Fuks, 2005; Ohyama, 2005; Kouzarides, 2007; Cairns, 2009; Bai and Morozov, 2010; Tanase et al., 2010). Regarding this issue, attention is now being paid to the spatial positions of genes in the nucleus (Takizawa et al., 2008a; Joffé et al., 2010; Meister et al., 2011). A number of correlations were found between the nuclear localization of genes and their expression in mouse cells (Kosak et al., 2002; Hewitt et al., 2004; Ragoczy et al., 2006; Williams et al., 2006; Takizawa et al., 2008b). The mouse β-globin gene tends to move from a nuclear peripheral region to an inner region in response to gene activation during the differentiation of erythroid cells (Ragoczy et al., 2006). The IgH and Igk genes also show similar phenomena during B lymphocyte development (Kosak et al., 2002). On the other hand, many genes maintain the same radial positioning in the nucleus regardless of their gene activities (Hewitt et al., 2004; Sadoni et al., 2008; Jost et al., 2011).

Although numerous experiments have been performed to understand the biological significance of gene positioning, various issues have been left unsolved, e.g., where an active/inactive gene localizes in the interphase cell nucleus, whether gene localization influences its expression during differentiation, and whether some correlation exists between the direction of gene movement and cell differentiation.

The current study focused on how pluripotency genes and hepatocyte-specific genes behave during cell differentiation into hepatocytes. We employed a system that can differentiate mouse embryonic stem (ES) cells into hepatocytes (Teratani et al., 2005; Tanase et al., 2011; Udagawa et al., 2012), and examined the localizations of the Klfl, Nanog, Oct4, Sox2, Cyp7a1, Pck1, Tat and Tdo2 genes. Oct4, Sox2 and Klfl encode three of the four so-called Yamanaka factors (Takahashi and Yamanaka, 2006). Nanog is also known as a gene encoding a core pluripotency factor (Mitsui et al., 2003; Silva et al., 2006). The remaining genes, Cyp7a1, Pck1, Tat and Tdo2, are hepatocyte-specific genes (Teratani et al., 2005; Banas et al., 2007; Tanase et al., 2011). To our knowledge, the behavior of these gene loci in the nucleus, during the differentiation of mouse ES cells into hepatocytes, has not been analyzed.

**MATERIAL AND METHODS**

**Cell and culture conditions**

The mouse ES cell line E14TG2a (Hooper et al., 1987) was grown in Glasgow minimum essential medium, supplemented with 10% fetal bovine serum, 0.1 mM 2-mercaptoethanol, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, and 1000 U/mL leukemia inhibitory factor (Chemicon, Temecula, CA, USA), on gelatin-coated dishes without feeder cells at 37°C in 7.5% CO₂. To perform the three-dimensional fluorescence in situ hybridization...
(3D-FISH) analyses, the cells were transferred onto gelatin-coated coverslips (thickness 0.17 ± 0.02 mm) and grown for 6 h before fixation.

**Cell differentiation**

Mouse ES cells were differentiated into hepatocytes according to a procedure previously described (Udagawa et al., 2012). The cells were grown on collagen-coated coverslips for 5-7 days in step 4 of the procedure, and then subjected to FISH analyses.

**DNA probes**

Bacterial artificial clones (BACs) were purchased from Invitrogen (Carlsbad, CA, USA), and labeled with digoxigenin-11-dUTP (Roche Diagnostics GmbH, Mannheim, Germany) by nick translation. The following BACs were used: RP23-357M17 (for Klf4 detection), RP23-406B15 (Nanog), RP23-75C13 (Oct4), RP23-425G5 (Sox2), RP23-247G5 (Cyp7a1), RP23-68J18 (Pck1), RP23-134P22 (Tat), and RP23-36B19 (Tdo2). The DNA probes were dissolved at a final concentration of 20 ng/µL in the hybridization mixture (50% formamide, 10% dextran sulfate in 1X SSC).

**Three-dimensional FISH**

The cells were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) (without calcium and magnesium) for 10 min, and washed three times in PBS for 3 min. Permeabilization was performed as previously reported (Hepperger et al., 2007, 2008). Briefly, the cells were immersed in 0.5% Triton X-100/PBS for 15 min, and then incubated in 20% glycerol/PBS for at least 1 h. The cells were then subjected to five rounds of freeze/thaw treatment in liquid nitrogen. The cells were incubated in 0.1 N HCl for 10 min, washed three times in 2X SSC for 3 min, and stored in 50% formamide/2X SSC at 4°C overnight. The DNA probes and nuclear DNA were denatured simultaneously at 75°C for 2 min. Hybridization was performed in a humid chamber at 37°C for 3 days. The cells were washed twice with 2X SSC at 37°C for 5 min, and then stringently washed three times with 0.1X SSC at 60°C for 5 min. After blocking in 4X SSC solution containing 4% bovine serum albumin (BSA) and 0.2% Tween-20, the probe was detected with rhodamine-conjugated anti-digoxigenin Fab fragments (Roche). The nuclear DNA was stained with 4',6-diamidino-2-phenylindole (DAPI) (100 ng/mL) for 10 min, and the samples were mounted in VECTASHIELD Mounting Medium (Vector Laboratories, Burlingame, CA, USA). Air drying was carefully avoided during all steps.

**Immunostaining**

The differentiation of mouse ES cells into hepatocytes was confirmed at the single cell level by immunostaining for cytokeratin-18 (CK18), a hepatocyte-specific gene product. This was performed simultaneously with the detection of the hybridization probe. The primary and secondary antibodies used to detect CK18 were goat anti-mouse CK18 antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA) and FITC-conjugated donkey anti-goat IgG (Santa Cruz Biotechnology), respectively. Only CK18-positive cells were used for data acquisition for hepatocytes.
Image acquisition, processing and analysis

Images were obtained using an Olympus FluoView FV1000-D confocal laser scanning microscope, equipped with a PlanApoN 60X OSC (NA = 1.4) objective lens. Three-dimensional stacks were obtained with 51 (for hepatocytes) or 61 (for ES cells) steps of 0.3 µm. Only the cell nuclei with two FISH signals were subjected to the following analysis. The relative distances between the nuclear center and the FISH signals were determined, using a slight modification of a previous method (Udagawa et al., 2012). Briefly, the images were processed using median filtering (3 x 3) and “fill holes” filtering in the MetaMorph software (Molecular Devices, Inc.). Then, to define the borders of the nucleus and the BAC signals, the thresholds were visually determined. Subsequently, the coordinates of the nuclear center and the BAC signals were calculated. Image data were converted into the Text Image format using the ImageJ software (http://rsbweb.nih.gov/ij), and imported into R software (http://www.R-project.org). The relative radial distances between the nuclear center and the BAC signals were measured using a custom-written R script (Udagawa et al., 2012).

Databases and genomic properties

Evaluations of gene density and GC content in Mus musculus were performed according to Hepperger et al. (2008). The genes involved in the 2 Mbp region centered at each target gene were counted using the NCBI Map viewer, developed at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/mapview, Build 38.1 from January 2012), and the sequence of the region was downloaded. The GC content was calculated using GENETYX-MAC (version 14.0.0, GENETYX Corporation).

Statistical analyses

Statistical significance was determined using the two-sample Kolmogorov-Smirnov test, for a pairwise comparison of the relative radial distributions. The P values were calculated with R software, and those less than 0.05 were considered significant. For correlation analyses, the median values of the relative distances from all nuclei were used. Pearson correlation coefficients and P values were also calculated with R software.

RESULTS

The effects of cell differentiation on gene localization were investigated for the following eight genes: Klf4, Nanog, Oct4 and Sox2; and Cyp7α1, Pck1, Tat and Tdo2 (Table 1). The former four are pluripotency genes (Mitsui et al., 2003; Silva et al., 2006; Takahashi and Yamanaka, 2006; Yamanaka, 2008), and the latter four are hepatocyte-specific genes (Teratani et al., 2005; Banas et al., 2007; Tanase et al., 2011). Oct4, Sox2 and Klf4 are three of the four genes encoding Yamanaka factors (Takahashi and Yamanaka, 2006). The product of Nanog is also known as a core pluripotency factor (Mitsui et al., 2003; Silva et al., 2006). These genes are active in ES cells, but inactive in hepatocytes (Panigada et al., 1999; Mitsui et al., 2003; Amaral et al., 2009; Tanase et al., 2011). The Cyp7α1, Pck1, Tat and Tdo2 genes are strongly expressed in hepatocytes differentiated from mouse ES cells, and thus, they are used as differentiation markers (Teratani et al., 2005; Banas et al., 2007; Tanase et al., 2011).
Gene movement during ES cell differentiation

<table>
<thead>
<tr>
<th>Gene Localization in ES Cells</th>
</tr>
</thead>
</table>
| To determine the spatial positions of the genes described above in the nuclei of mouse ES cells, we first performed 3D-FISH analyses (Figure 1). The radial distribution of the locus of a target gene was analyzed by measuring the distances between the FISH signals and the nuclear center. The measured distances were normalized and sorted into 10 space fractions of a nucleus that was equally divided according to the radial distance from the center (Figure 2). The Nanog, Cyp7α1 and Pck1 loci showed distinct localizations: nearly 40% of the total BAC signals were localized in the second- (Nanog and Cyp7α1) or third- (Pck1) most outer fraction. In contrast, the localization profile of the Tdo2 locus was broad and very close to the random distribution profile, except a small peak at 0.55. The gene localization profiles of the pluripotency genes differed from each other. Regarding the hepatocyte-specific genes, the distribution profiles of Cyp7α1, Pck1 and Tat were obviously different from that of Tdo2. In addition, the former three also showed slight differences between each other. Although the Klf4 and Cyp7α1 genes and the Sox2 and Tdo2 genes reside on the same chromosome, respectively (Table 1: the former pair, chromosome 4; the latter pair, chromosome 3), the two genes in each pair showed different distribution profiles from each other in both cases. In conclusion, each gene has its inherent localization profile in the ES cells. No group-specific profile was observed.

Figure 1. 3D-FISH on nuclei of mouse embryonic stem (ES) cells and their differentiated hepatocytes. The pictures are representative z-stack projections. The loci of each gene are shown in red. The nuclei were stained with DAPI and are colored blue. Differentiation into hepatocytes was confirmed by CK18 immunostaining (green). Scale bar = 5 µm.

Table 1. Target genes in the current study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pluripotency gene</td>
<td></td>
</tr>
<tr>
<td>Klf4 [Kruppel-like factor 4 (gut)]</td>
<td>4</td>
</tr>
<tr>
<td>Nanog [Nanog homeobox]</td>
<td>6</td>
</tr>
<tr>
<td>Oct4 (POU domain, class 5, transcription factor 1)</td>
<td>17</td>
</tr>
<tr>
<td>Sox2 (SRV-box containing gene 2)</td>
<td>3</td>
</tr>
<tr>
<td>Hepatocyte-specific gene</td>
<td></td>
</tr>
<tr>
<td>Cyp7α1 (Cytochrome P450, family 7, subfamily a, polypeptide 1)</td>
<td>4</td>
</tr>
<tr>
<td>Pck1 (Phosphoenolpyruvate carboxykinase 1, cytosolic)</td>
<td>2</td>
</tr>
<tr>
<td>Tat (Tyrosine aminotransferase)</td>
<td>8</td>
</tr>
<tr>
<td>Tdo2 (Tryptophan 2,3-dioxygenase)</td>
<td>3</td>
</tr>
</tbody>
</table>
Figure 2. Localization of pluripotency genes and hepatocyte-specific genes in the nuclei of mouse embryonic stem (ES) cells and hepatocytes. The relative distance was calculated as the ratio of the measured distance from the nuclear center to the length of the nuclear radius containing the signal. Each single cell generated two distance data points. At first, the entire distance data from the cell population analyzed were sorted into 10 fractions (x-axis). The number of data points in each fraction was then divided by the entire number of distance data points. The resulting value (%) was plotted at the center of the fraction. N = total number of cell nuclei analyzed. Distributions in the mouse ES cells and the hepatocytes are shown by solid lines and dashed lines, respectively. Gray lines show the random distribution. The P values were calculated between ES cells and hepatocytes by the two-sample Kolmogorov-Smirnov test.

Change in gene localization during differentiation into hepatocytes

We next studied the influence of differentiation on the spatial positioning of the genes. Mouse ES cells were differentiated into hepatocytes, using the method previously described (Udagawa et al., 2012), and they were then subjected to 3D-FISH analysis (Figure 1). The completion of differentiation into hepatocytes was confirmed at the single cell level by immunostaining of cytokeratin-18 (CK18), which is a hepatocyte-specific gene product, and only the CK18-positive cells were used for further analyses. The localization profile of Nanog was close to a random distribution profile after differentiation (Figure 2; dashed lines). In this case, the profile
slightly shifted towards the nuclear center. Regarding the Sox2 locus, the change in the localization profile was very slight before and after differentiation. On the other hand, the other six gene loci showed considerable changes in their profiles. The profiles of Klf4, Oct4, Cyp7α1, Pck1 and Tat revealed clear shifts towards the nuclear center. In the cases of Cyp7α1 and Pck1, broadening of the profiles was also observed. In contrast, the profile of Tdo2 shifted toward the periphery and formed a clear “peak”. The profile of Sox2 also shifted toward the periphery but the extent was very slight. In summary, the Klf4, Nanog, Oct4, Cyp7α1, Pck1 and Tat loci shifted toward the nuclear center, while the Sox2 and Tdo2 loci shifted toward the periphery.

**Relationship between gene position, gene density and GC content**

Gene localization is reportedly correlated with some genomic properties (Murmann et al., 2005; Hepperger et al., 2008; Jost et al., 2011). Therefore, we examined whether such a correlation exists for the genes under analysis. For each gene, we obtained the median value of the entire distance data (based on Figure 2), and plotted it against the gene density of the surrounding region (Figure 3A, B) and the local GC content of the surrounding region (Figure 3C, D). In linear regression analysis, the Pearson correlation coefficient $R$ and the $P$ value were calculated for each case. In panel A, the $P$ value was quite large, and thus, it seemed that the variable sets in this panel did not correlate or showed very low correlation. On the other hand, in panels B and D, the $P$ values were less than 0.05, indicating that the relative distance from the nuclear center correlated very well with both the local gene density and local GC content in hepatocytes. These two cases each showed a clear negative correlation between the sets of variables. In panel C, the variable set might have a slight correlation, judging from the $P$ value of ~0.1. In Figure 3, the highest correlation coefficient was found between the gene position and the local GC content in hepatocytes ($R = -0.92, P = 0.001$; Figure 3D). In conclusion, in hepatocytes, both the local GC content and gene density of the surrounding region highly correlated with gene localization.

**Figure 3.** Relationship between gene localization and genomic properties. The median values of the relative distances are plotted against the local gene density (A, B) and the local GC content (C, D). The regression lines for the mouse embryonic stem (ES) cells and the hepatocytes are solid and dashed, respectively. $R = $ Pearson’s correlation coefficient. The local gene density (A, B) and the local GC content (C, D) were calculated in the 2 Mbp window with the focused gene set at the center. Since the region surrounding Pck1 has undetermined sequences of ~3 kb in total in the mouse genome database, the values for the Pck1 locus were calculated while excluding the gaps.
DISCUSSION

The current study examined the nuclear localization of pluripotency genes and hepatocyte-specific genes, using mouse ES cells and hepatocytes differentiated from them. No group-specific profile was observed in the ES cells, and each gene exhibited an inherent localization profile. During differentiation, the Klf4, Nanog, Oct4, Cyp7α1, Pck1 and Tat loci shifted toward the nuclear center, while the Sox2 and Tdo2 loci shifted toward the periphery. Furthermore, the results showed that both the local GC content and gene density of the surrounding region highly correlated with gene localization in hepatocytes.

*Klf4, Nanog and Oct4 move towards the inner region of the nucleus with silencing, and* Cyp7α1, Pck1 and Tat *show the same behaviors as the former three upon activation*

Of the eight genes studied, Klf4, Nanog, Oct4, Cyp7α1, Pck1 and Tat tended to re-localize towards the inner regions during differentiation, although the extent of the shift was small for Nanog (Figure 2). Klf4, Nanog and Oct4 are active in ES cells and inactive in hepatocytes, while Cyp7α1, Pck1 and Tat are active in hepatocytes and inactive in ES cells. Thus, regarding Klf4, Nanog, Oct4 and Tdo2, our analysis indicated that they seem to prefer the outer regions, rather than the inner regions, when they are active. Some genes favor inner regions of the nucleus when they are active (Kosak et al., 2002; Hewitt et al., 2004; Ragoczy et al., 2006; Williams et al., 2006; Finlan et al., 2008; Reddy et al., 2008; Takizawa et al., 2008b), while others become active at the nuclear periphery (Hewitt et al., 2004). The behaviors of Sox2, Cyp7α1, Pck1 and Tat were in the former category, on the other hand, Klf4, Nanog, Oct4 and Tdo2 were in the latter. Thus, both types were involved in the genes used in the current study.

In the active state, the major populations of Klf4, Nanog (in the state of ES cells) and Tdo2 (in hepatocytes) were located in the second-most outer space of the nucleus. Thus, the loci of these genes were not likely to be at the very edge of the nucleus or the nuclear membrane. Accordingly, in the active state, they can physically escape from the direct association with the nuclear envelope or from the repressive microenvironment juxtaposed with the nuclear envelope that may be required for gene silencing (Takizawa et al., 2008a; Deniaud and Bickmore, 2009; Joffe et al., 2010; Peric-Hupkes, 2010).

**Gene localization is highly dependent on GC content in hepatocytes**

Using the genes encoding β-actin, albumin, β-globin, CD11b, lysozyme, MYB, MPO and OCT4 and the DNA region homologous to the human Beckwith-Wiedeman syndrome region, Hepperger et al. (2008) showed that gene localization correlates with GC content but not gene density, in both mouse ES cells and the macrophages differentiated from the ES cells. Our results for the pluripotency genes and the hepatocyte-specific genes in the mouse ES cells agreed with this report, although the P value for the correlation was ~0.1 (Figure 3C). The correlation coefficient ($R$) value in the figure was calculated to be -0.62. This value is comparable to those reported in the above cited study, in which the $R$ values were almost the same between the ES cells and macrophages, -0.61 and -0.63, respectively (Hepperger et al., 2008). Our study showed that in hepatocytes, gene localization correlated very well with not only...
the local GC content but also gene density, and the $R$ value of the former was -0.92 (Figure 3D) and that of the latter -0.71 (Figure 3B). In conclusion, at least for pluripotency genes and hepatocyte-specific genes, the nuclear localization of the genes correlated with the local GC content and gene density of the surrounding region in hepatocytes.

ACKNOWLEDGMENTS

Research supported by the Ministry of Education, Culture, Sports, Science & Technology - Japan (MEXT).

REFERENCES


