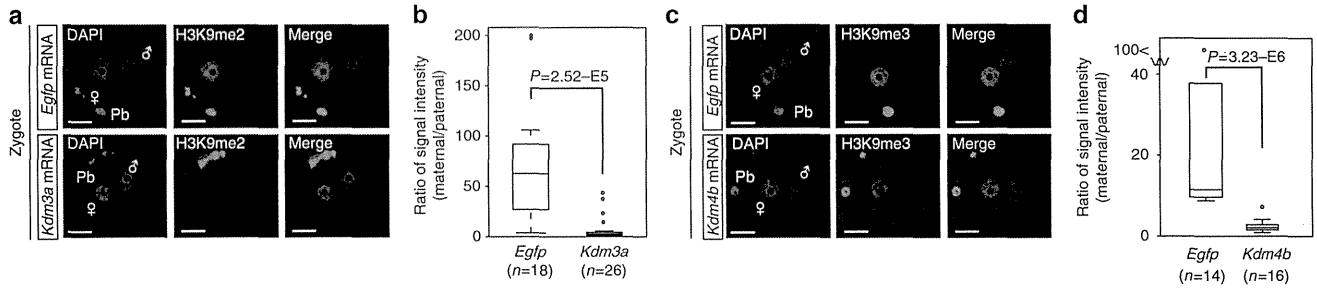
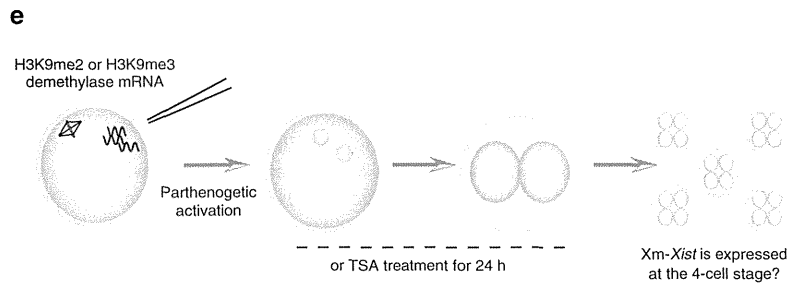
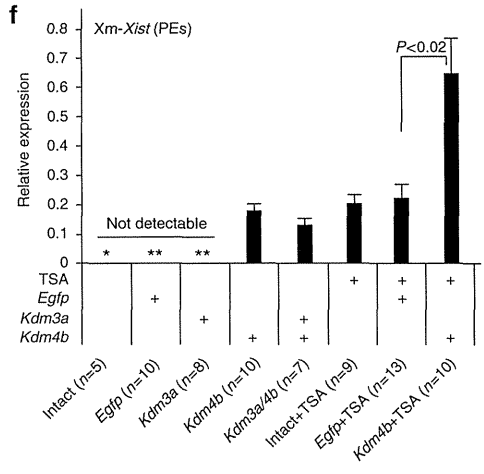
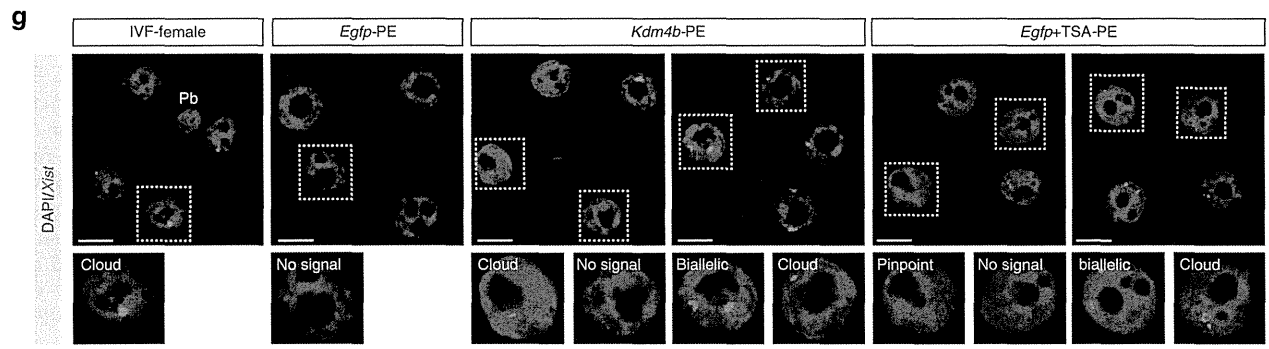
Fig. 1f), suggesting that H3K9me3 demethylation caused Xm-*Xist* derepression.

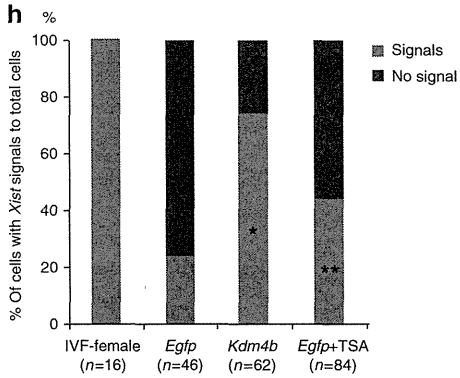
We next assessed the effects of a histone deacetylase inhibitor, trichostatin A (TSA), on Xm-*Xist* expression. TSA-treated PEs (Intact + TSA-PEs and *Egfp* + TSA-PEs) also activated Xm-*Xist* (Fig. 1f). No significant changes were detected in Xm-*Xist* expression levels between *Kdm4b*-PEs and *Egfp* + TSA-PEs. However, although co-injection with *Kdm4b* and *Kdm3a* mRNAs did not increase Xm-*Xist* expression levels as compared with *Kdm4b*-PEs, a combination of TSA and *Kdm4b*-mRNA significantly increased Xm-*Xist* expression as compared with *Egfp* + TSA-PEs (2.9-fold, $P<0.04$, Student's *t*-test; Fig. 1f). Moreover, derepression of Xm-*Xist* transcription occurred in the absence of *Rnf12* overexpression (Supplementary Fig. 3), and *Jpx* and *Ftx*,

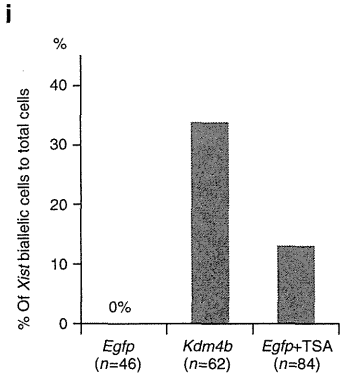












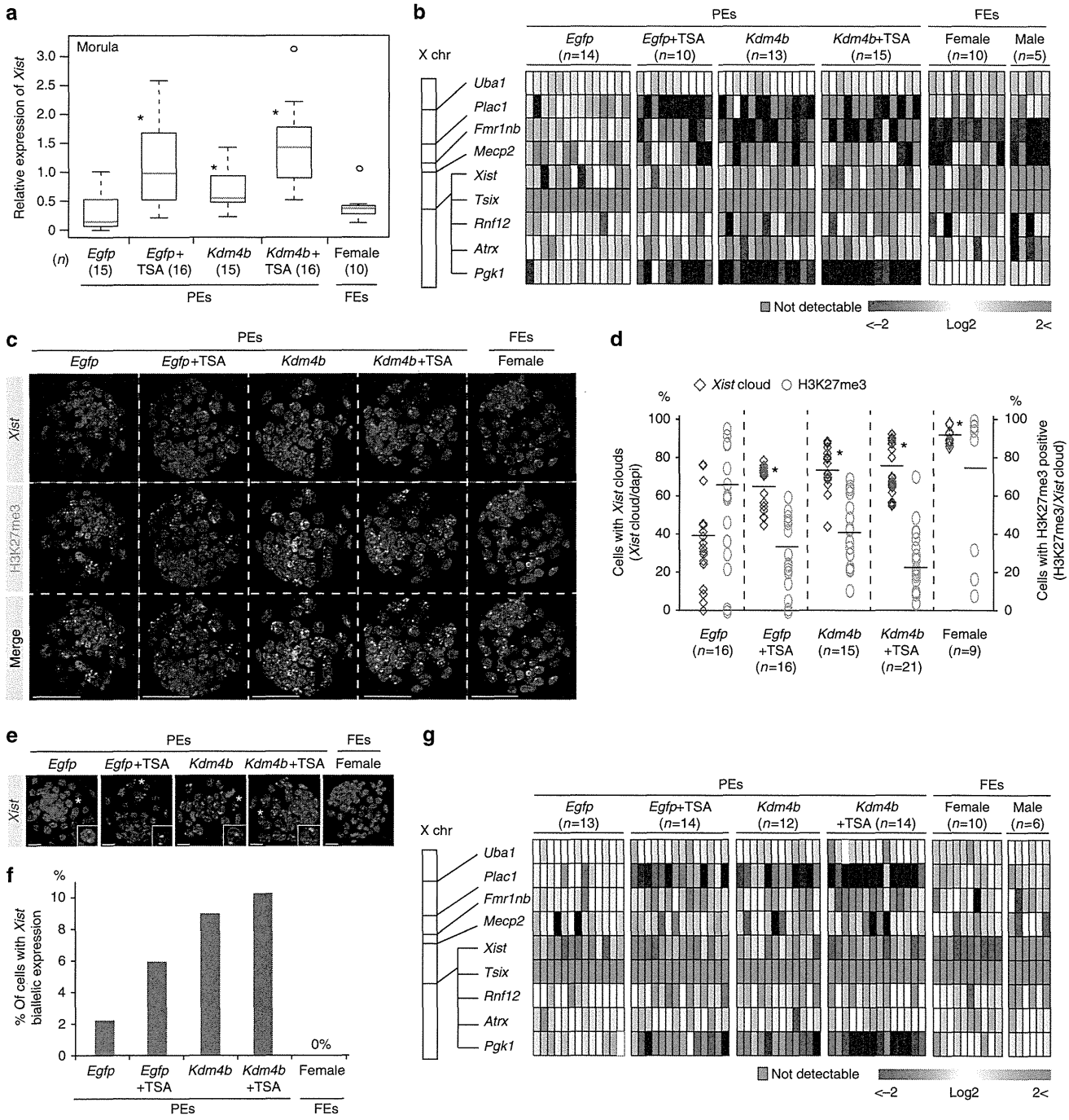


Figure 2 | Global XCI and *Xist* expression states of Xm at late preimplantation stages. (**a**) Analysis of *Xist* expression using qPCR of individual embryos in the morula. An asterisk indicates $P<0.05$ (Student's *t*-test) compared with *Egfp*-PEs. FEs, fertilized embryos. (**b**) Large-scale qPCR analysis of *Xist* and eight X-linked genes in individual blastocysts after culturing for 96 h. Coloured bars indicate expression levels. (**c**,**d**) IF (H3K27me3, green) combined with RNA FISH (*Xist*, red) analysis in 96-h blastocyst stage. 4′,6-diamidino-2-phenylindole (DAPI)-stained nuclei are shown in blue. (**c**) Representative confocal projection. Scale bars, 50 μm. (**d**) The graph shows *Xist* expression and H3K27me3 modification states in individual embryos. The horizontal axis indicates the average percentage in the group. $*P<3.1\times10^{-28}$ (Fisher's exact test). *n*, number of embryos analysed. (**e**,**f**) Xm-*Xist* biallelic expression states in PEs at 96 h. The asterisk indicates cells with biallelic expression. Scale bars, 20 μm. (**f**) Summary of the ratio of biallelic cells to *Xist*-positive cells in 96-h blastocyst stage in each group. The number of cells is shown in Supplementary Table 3. (**g**) qPCR analysis of *Xist* and eight X-linked genes in individual blastocysts after culturing for 120 h. (**h**,**i**) IF (H3K27me3, green) combined with RNA FISH (*Xist*, red) analysis in 120 h blastocysts. $*P<5.4\times10^{-23}$ (Fisher's exact test). Scale bars, 50 μm. (**j**) The ratio of biallelic cells to *Xist*-positive cells in 120 h blastocysts. In qPCR analysis, the average expression level of Xm-*Xist* in *Egfp*-PEs was set as 1 (also see the Methods section). *Gapdh* and *β-actin* were used as internal controls. Data are summarized in Supplementary Table 4.

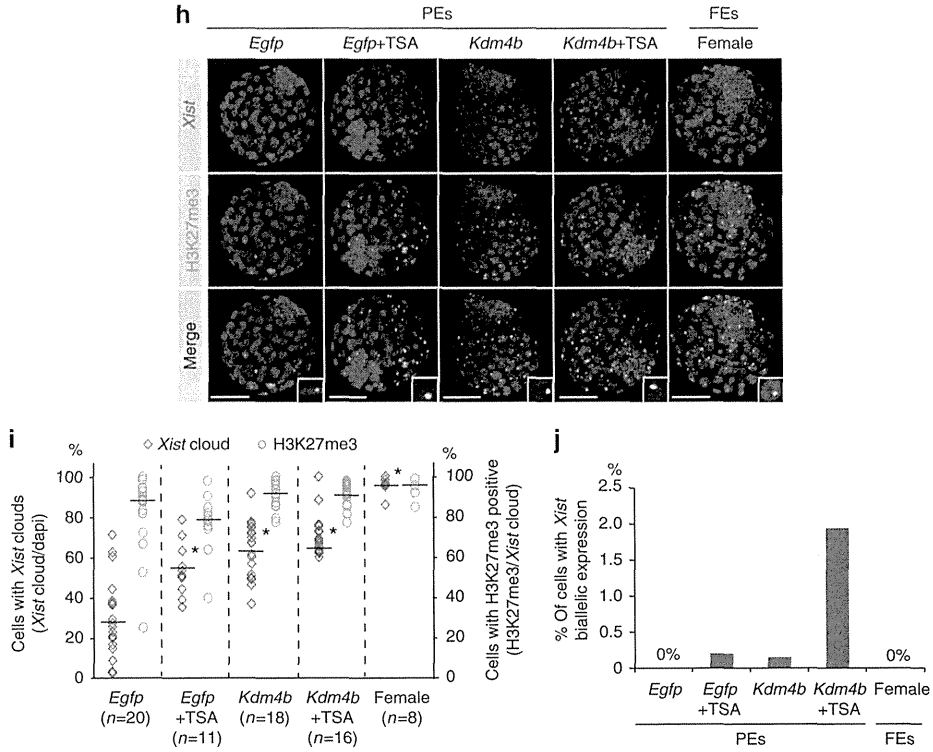


Figure 2 | Continued.

which have been identified as *Xist* activators[2], were not expressed at the four-cell stage. These results showed that KDM4B- and TSA-mediated Xm-*Xist* derepression was not involved in the abnormal expression of known *Xist* activators.

Next, we examined Xm-*Xist* derepression states at the single-cell level by fluorescence *in situ* hybridisation (FISH) of *Xist* RNA. Consistent with the qPCR results, Xm-*Xist* signals were significantly increased in *Egfp* + TSA- and *Kdm4b*-PEs (Fig. 1g,h and Supplementary Table 1). However, neither TSA treatment nor *Kdm4b* overexpression consistently activated Xm-*Xist* in all cells (Fig. 1g,h). FISH analysis also revealed *Xist* biallelic cells in *Kdm4b*- and *Egfp* + TSA-PEs (*Kdm4b*-PEs: 34% and *Egfp* + TSA-PEs: 13%; Fig. 1g,i). Taken together, these results indicated that *Kdm4b* overexpression and TSA treatment induced Xm-*Xist* derepression at the same developmental stage as Xp-*Xist* activation.

Xm-*Xist* transcripts establish XCI at the blastocyst stage. To investigate whether Xm-*Xist* transcripts from the four-cell stage induced XCI in late preimplantation stages, we cultured *Kdm4b*-, *Egfp* + TSA- and *Kdm4b* + TSA-PEs until the blastocyst stage. Development into blastocysts occurred in >80% of PEs in all groups (Supplementary Table 2). At the morula stage (72 h), although *Egfp*-PEs initiated Xm-*Xist* expression, the levels of *Xist* expression were significantly lower compared with those in *Kdm4b*-injected and/or TSA-treated PEs (Fig. 2a). At the 96-h blastocyst stage, we examined the expression levels of *Xist* and X-linked genes (*Tsix*, *Rnf12*, *Pgk1*, *Fmr1nb*, *Atrx*, *Uba1*, *Mecp2* and *Plac1*) in individual embryos. The significant upregulation of *Xist* observed in PEs that had been injected with *Kdm4b* mRNA and/or TSA continued in 96-h blastocyst stage (Fig. 2b). In PEs exhibiting Xm-*Xist* upregulation, *Tsix* expression was not detectable and *Rnf12* was not overexpressed as compared with *Egfp*-PEs (Fig. 2b). The average expression levels of *Pgk1*, *Plac1* and *Fmr1nb* in *Kdm4b*-overexpressing or TSA-treated PE groups were significantly reduced (Supplementary Fig. 4a). However, single embryos in the same group exhibited heterogeneity in the expression levels of these targets (Fig. 2b). Moreover, although Xm-*Xist* was not overexpressed in *Kdm4b*-overexpressing or TSA-treated PE groups compared with levels seen in female fertilized embryos (FEs), the levels of *Plac1* and *Pgk1* were strongly downregulated in *Kdm4b*-, *Egfp* + TSA- and *Kdm4b* + TSA groups (Fig. 2b). These results suggested that Xm-*Xist* expression states differed at the single-cell level in individual embryos.

To gain further insights into *Xist* expression states and repression of X-linked genes on Xm alleles, we conducted IF to detect the H3K27me3 state, which is a hallmark of XCI[20], and performed *Xist* FISH analysis. The *Xist* RNA FISH probe recognizes *Xist* and *Tsix*. Therefore, the cloud state of the FISH signal defined *Xist* expression. Consistent with the qPCR results, the number of cells with *Xist* signals increased significantly in all *Kdm4b*-overexpressing or TSA-treated PEs (Fig. 2c and Supplementary Table 3). However, in all *Kdm4b*-overexpressing or TSA-treated PEs, <45% of *Xist* cloud-containing cells had an H3K27me3 signal (Fig. 2c,d and Supplementary Table 3). We also found that *Xist* biallelic cells were present in all PE groups (*Egfp*-PEs: 2.2%, *Egfp* + TSA-PEs: 5.9%, *Kdm4b*-PEs: 9.0% and *Kdm4b* + TSA-PEs: 10.3%; Fig. 2e,f). These results suggested that there were various Xm-*Xist* expression states present at the single-cell level, affecting the heterogeneity of X-linked genes, and that strong suppression of some X-linked genes in *Kdm4b*-overexpressing or TSA-treated PEs could be attributed to biallelic expression of Xm-*Xist*.

As *Kdm4b* overexpression and TSA treatment did not affect the extent of H3K27me3 modification (Supplementary Fig. 5), its acquisition in PEs may have been slightly delayed compared with that in FEs. In support of this notion, the developmental timing of PEs lags behind that of FEs[9]. Therefore, we extended the culture period to 120 h and again performed qPCR, IF and FISH analyses.

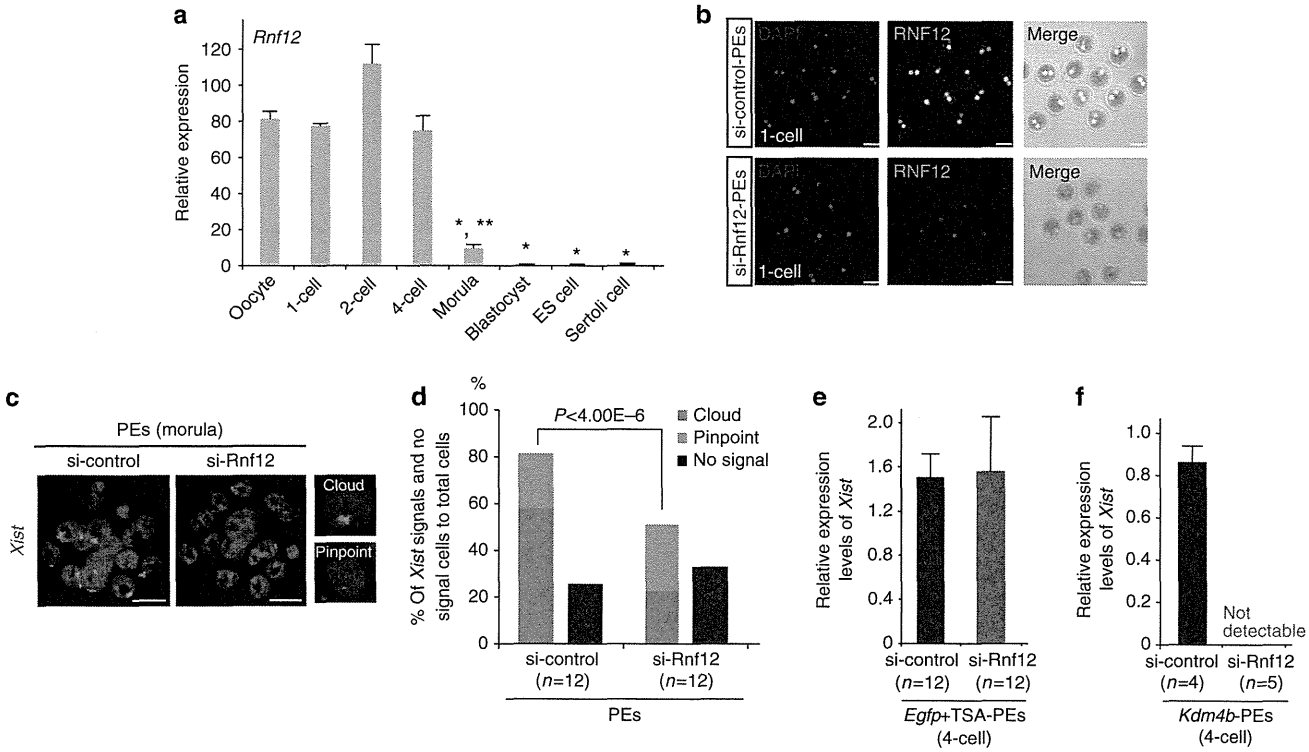


Figure 3 | RNF12 is required for *Xist* expression in various types of preimplantation embryos. (**a**) *Rnf12* expression profiles in preimplantation stages, ES cells and somatic cells. Ten oocytes ($n=3$), ten IVF-1 cells ($n=3$), ten IVF-2 cells ($n=3$) and five IVF-4 cells ($n=3$) were pooled. The expression level of *Rnf12* in the morula stage, as determined using qPCR, represents the average of 16 individual embryos. The numbers of ES and Sertoli cells represent the averages of three independent cell lines and male pups, respectively. The error bars indicate the mean ± s.e.m. $*P<0.004$ compared with oocytes; $**P<0.0001$ compared with ES cells (Student's *t*-test). (**b**) IF analysis of RNF12 at the one-cell stage (green). Samples were fixed 11 h after parthenogenetic activation (18–19 h after siRNA injection). The same laser beam intensities were used to excite the *Rnf12*-knockdown and control samples. 4′,6-diamidino-2-phenylindole (DAPI)-stained nuclei are shown in blue. Two independent experiments were conducted. Scale bars, 50 μm. (**c**,**d**) *Xist* FISH analysis of si-Rnf12 PEs at the morula stage. At 72 h after activation, PEs injected with siRNA were analysed. Representative images of siRNA-treated embryos (**c**). Scale bars, 20 μm. The percentage of total *Xist*-positive signals and -negative cells to total cells in si-Rnf12 and si-control PEs. Biallelic expression was counted as two signals. *n*, number of embryos analysed (**d**). (**e**,**f**) qPCR analysis of Xm-*Xist* expression at the four-cell stage of embryos treated with TSA (**e**) or injected with *Kdm4b* mRNA (**f**). PEs derived from maternal si-Rnf12-treated oocytes. A detailed experimental scheme is shown in Supplementary Fig. 6d. A pool of eight to ten four-cell embryos represents one biological replicate.

As seen in 96-h blastocyst stage, qPCR analysis revealed that *Plac1* and *Pgk1* were significantly repressed in *Kdm4b*-overexpressing or TSA-treated PEs, although heterogeneity was observed (Fig. 2g and Supplementary Fig. 4b). However, *Fmr1nb*, which was significantly repressed in 96-h blastocyst stage of both *Kdm4b*- and TSA-treated groups, did not show marked suppression (Fig. 2g and Supplementary Fig. 4b), suggesting that *Xist* expression states were altered in 120 h blastocysts.

In FISH and IF analysis, a significant increase in the number of *Xist*-positive cells was detected in *Kdm4b*-overexpressing and/or TSA-treated PEs, as compared with *Egfp*-PEs, after culture for 120 h (Fig. 2h,i and Supplementary Table 4). There was a marked increase in the ratio of H3K27me3 spots to cloud *Xist* signals in blastocysts in all groups compared with that observed after 96 h of culture as follows: 88.1% in *Egfp*-PEs, 78.7% in *Egfp* + TSA-PEs, 91.8% in *Kdm4b*-PEs and 91.2% in *Kdm4b* + TSA-PEs (Fig. 2h,i and Supplementary Table 4). Interestingly, >98% of the cells in the *Xist* cloud state exhibited monoallelic *Xist* expression in all groups at 120 h (Fig. 2j).

Taken together, these results indicated that *Kdm4b* overexpression and TSA treatment induced global XCI of Xm in blastocysts and that the counting mechanism automatically functioned in late blastocysts, as has also been observed for Xp alleles[21].

KDM4B-mediated Xm-*Xist* expression depends on RNF12. During preimplantation phases, Xp-*Xist* expression is induced by maternal RNF12 (ref. 5). Thus, we investigated whether Xm-*Xist* expression also depended on RNF12. High *Rnf12* expression levels were maintained until the four-cell stage (around 80-fold higher than in ES cells; Fig. 3a). At the morula stage, although the expression level was significantly decreased compared with that in oocytes, *Rnf12* expression was still more than ninefold higher than that in ES cells (Fig. 3a), indicating that maternal and early zygotic RNF12 may be critical for Xm-*Xist* activation. To examine the dependency of RNF12 on Xm-*Xist* expression, we inhibited maternal and zygotic RNF12 expression using *Rnf12*-siRNA (Supplementary Fig. 6a). IF analysis at the one-cell stage showed a marked decline in RNF12 signal intensity in si-*Rnf12* embryos compared with that in the si-control embryos (Fig. 3b). Significant repression by si-*Rnf12* was maintained in the four-cell stage (Supplementary Fig. 6b,c). Using this knockdown system, we examined whether Xm-*Xist* activation was induced by RNF12 at the morula stage. *Xist* FISH analysis revealed that *Xist* signals (cloud and pinpoint) in PEs treated with si-Rnf12 were significantly reduced at the morula stage as compared with those observed in the controls (si-Rnf12: 50.0% versus si-control: 80.0%; Fig. 3c,d and Supplementary Table 5).

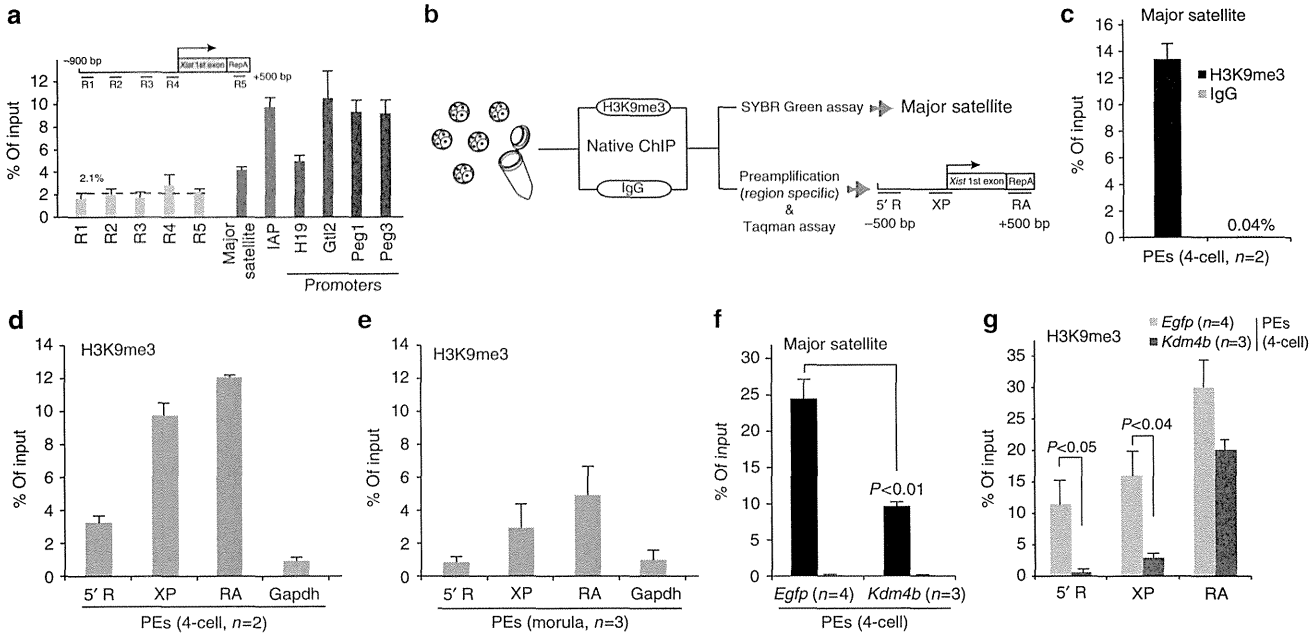


Figure 4 | H3K9me3 states at the Xm-*Xist* promoter region in preimplantation embryos. (**a**) ChIP-qPCR analysis of Xp (sperm) at *Xist* 5′-regions, the *H19*, *Gtl2*, *Peg1* and *Peg3* promoter region, and regions in repetitive elements. $n=2$–4. Rabbit IgG was used as a negative control. The percentage of input for negative control DNA was >1% for all genes tested. The data were not normalized for local nucleosome occupancies. (**b**) Schematic representation of eChIP-qPCR analysis. H3K9me3 states at major satellite regions (**c**) and at *Xist* regions and the *Gapdh* promoter region (**d**) in PEs at the four-cell stage. Two independent experiments were performed. In each experiment, 250 embryos were used. (**e**) H3K9me3 states at the *Xist* and *Gapdh* promoter regions in morula-stage embryos. Three independent experiments were conducted and 40 embryos were used for each assay. H3K9me3 states in major satellite (**f**) and *Xist* regions (**g**) in *Egfp*- and *Kdm4b*-PEs at the four-cell stage. Three (*Kdm4b*-PEs) and four (*Egfp*-PEs) independent experiments were conducted. In each experiment, 170–250 embryos were prepared. The percentages of input for negative controls (IgG) were <0.2% (**f**) and 1.9% (**g**), respectively. Error bars indicate the mean ± s.e.m. The *P*-values were determined using Student's *t*-tests.

Next, we examined whether Xm-*Xist* derepression from the four-cell stage by ectopic *Kdm4b* expression or TSA treatment was regulated by RNF12 (Supplementary Fig. 6d). Four-cell embryos of *Kdm4b*- or *Egfp* + TSA-PEs treated with si-Rnf12 were analysed by qPCR. Depletion of RNF12 did not affect Xm-*Xist* expression in *Egfp* + TSA-PEs compared with the si-control PEs (Fig. 3e), suggesting that factors other than RNF12 may contribute to histone acetylation-mediated Xm-*Xist* activation. This observation is consistent with results obtained using *Rnf12*$^{-/-}$ mice, which demonstrated that *Xist* is activated by RNF12 and other unidentified factors[5,22]. In contrast, *Xist* expression in *Kdm4b*-PEs derived from oocytes treated with si-*Rnf12* did not induce detectable expression of Xm-*Xist* (Fig. 3f). These results demonstrated that KDM4B-mediated Xm-*Xist* expression depended on RNF12 and suggested that H3K9me3 prevented the expression.

Promoter demethylation of H3K9me3 causes Xm-*Xist* derepression. As activation of *Xist* by RNF12 is essential for establishing iXCI[5], we attempted to determine the mechanism responsible for the transcriptional derepression of Xm-*Xist* by KDM4B-mediated demethylation of H3K9me3. We first examined whether H3K9me3 levels were enriched at the Xp-*Xist* promoter region. Nucleosomes were extracted from the sperm genome (Supplementary Fig. 7). ChIP-qPCR revealed the low H3K9me3 levels of Xp-*Xist* in the 5′-regions containing the major promoter for *Xist* expression (average: 2.1%; Fig. 4a), as compared with those of the *H19*, *Gtl2*, *Peg1* and *Peg3* promoter loci (average: 8.5%) and regions in repetitive elements (intracisternal A-particles and major satellite DNAs; average: 7.0%), which are known to be associated with H3K9me3 (ref. 23). These results indicated that the *Xist* promoter region was hypomethylated in sperm, in agreement with Xp-*Xist* being expressed in early embryogenesis.

Preparing sufficient numbers of embryos or oocytes for ChIP combined with deep sequencing (ChIP-seq) analysis is problematic. Some ChIP-qPCR methods requiring small samples have been reported[24]; however, most of these studies are based on a cross-linking ChIP method, in which the ChIP efficiency is lower than that of native ChIP methods[25]. Thus, we developed a new native ChIP method combined with a Taqman qPCR system for quantification of transcription in single cells (termed eChIP-qPCR) and focused on three loci at *Xist* 5′-regions containing the major promoter and repeat A, which is essential for establishment of iXCI[26] (Fig. 4b). We first tested whether the quantification system was biased by using diluted DNA from bulk ES cells. The results of ChIP-qPCR from pre-amplified DNA were comparable to those obtained without pre-amplification (Supplementary Fig. 8).

Using this system, we examined H3K9me3 states at three *Xist* regions and at *Gapdh* promoter regions (as a negative control) in PEs at the four-cell stage. We first investigated whether our eChIP method was efficient by examining major satellite repeats that have been identified as H3K9me3-rich regions in preimplantation embryos[23]. Consistent with a previous report, H3K9me3 was highly enriched at major satellite regions (Fig. 4c). The three *Xist* regions were also highly methylated compared with the *Gapdh* promoter region, as follows: 5′-R, 3.7-fold upregulated; XP, 9.8-fold upregulated; RA, 12.1-fold upregulated; (Fig. 4d).

As Fig. 2a showed Xm-*Xist* spontaneous derepression at the morula stage, we next investigated whether H3K9me3 levels at the *Xist* promoter region were low at this stage. eChIP-aPCR

analysis revealed that enrichment of H3K9me3 was markedly reduced compared with that at the *Gapdh* promoter region in the four-cell stage (5′-R, 0.83-fold upregulated; XP, 2.85-fold upregulated; RA, 4.8-fold upregulated; Fig. 4e). These results suggested that demethylation at the promoter region was essential for Xm-*Xist* derepression.

We then asked whether *Xist* promoter demethylation was involved in the Xm-*Xist* derepression observed in *Kdm4b*-PEs at the four-cell stage. The H3K9me3 levels at major satellite regions in *Kdm4b*-PEs were significantly reduced compared with those in *Egfp*-PEs (*Kdm4b*-PEs: 9.74% versus *Egfp*-PEs: 24.63%, $P<0.01$, Student's *t*-tests; Fig. 4f). At three *Xist* regions, the H3K9me3

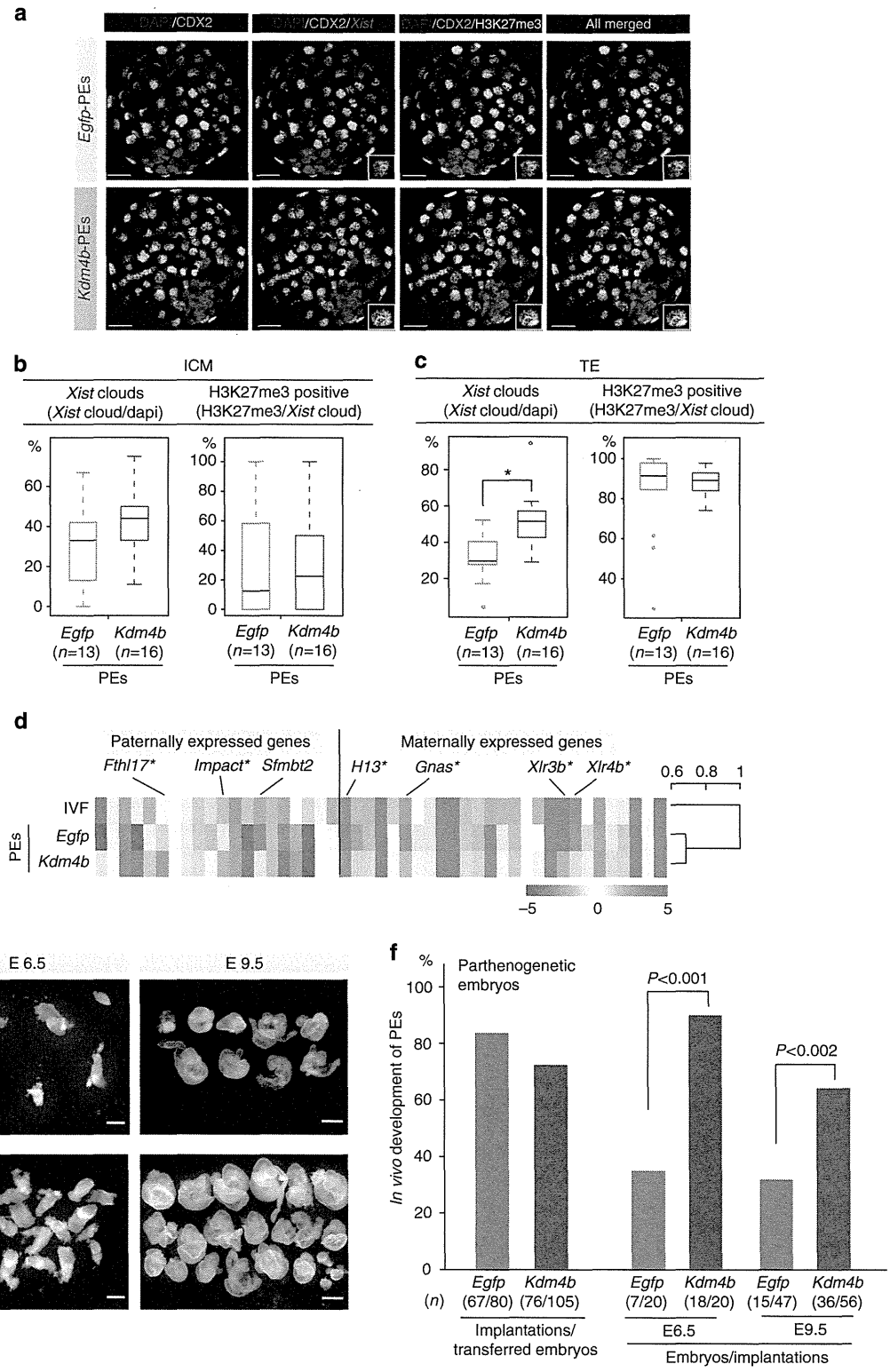


Figure 5

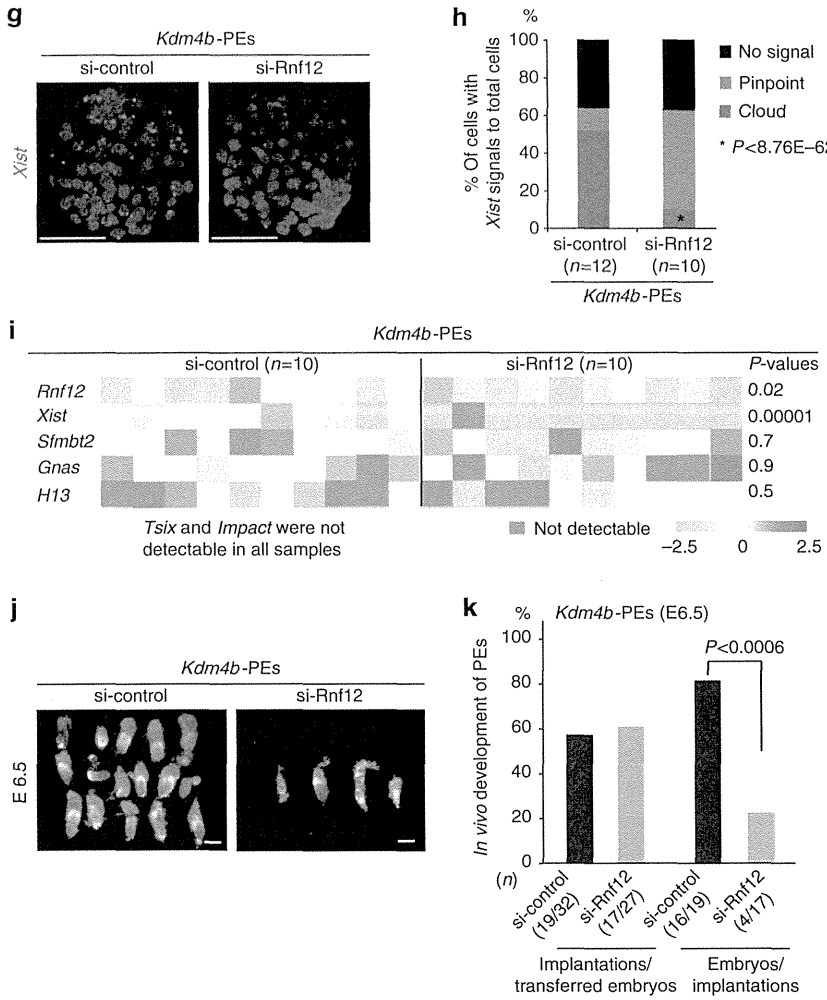


Figure 5 | Loss of Xm-XCI is the primary cause of developmental failure immediately after implantation in most PEs. (**a**) IF combined with FISH analysis of blastocysts in *Egfp*-PEs (upper panel) and *Kdm4b*-PEs (lower panel). CDX2-positive cells were identified as belonging to the trophectoderm (TE). Representative pictures of Z-sections. 4′,6-diamidino-2-phenylindole (DAPI) (blue), CDX2 (green), *Xist* (red) and H3K27me3 (white). Scale bars, 20 μm. The rates of cells with *Xist* (left) or H3K27me3 (right) in the inner cell mass (ICM) (**b**) and TE (**c**), respectively. *n*, number of embryos analysed. The number of cells analysed is shown in Supplementary Table 7. $*P<4.3\times10^{-23}$ (Fisher's exact test). (**d**) Expression states and clustering analysis of imprinted genes. *Sfmbt2* important for placentation and differentially expressed genes (asterisk) are shown. The scale bar indicates normalized values of $\log_2$. (**e**) Embryos with extra-embryonic tissues at E6.5 and E9.5 for *Kdm4b*- and *Egfp*-PEs, respectively. Upper and lower images indicate *Egfp*- and *Kdm4b*-PEs, respectively. Left and right column sides show E6.5 and E9.5, respectively. Scale bars, 200 μm (E6.5) and 500 μm (E9.5). (**f**) Summary of the developmental abilities of *Kdm4b*-PEs and *Egfp*-PEs at postimplantation stages (E6.5 and E9.5). Five and 12 independent recipients were analysed at E6.5 and E9.5, respectively. (**g**,**h**) *Xist* analysis in *Rnf12*-knockdown and control *Kdm4b*-PEs. Representative images of FISH analysis. Scale bars, 50 μm (**g**) and *Xist* expression states (**h**). (**i**) Expression of imprinted and X-linked genes in *Rnf12*-knockdown and control *Kdm4b*-PEs. *P*-values were determined using Student's *t*-tests. (**j**) Embryos with extra-embryonic tissues at E6.5 in *Rnf12*-knockdown and control *Kdm4b*-PEs. Scale bars, 200 μm. (**k**) Summary of the developmental ability of *Rnf12*-knockdown and control *Kdm4b*-PEs at E6.5. Five independent recipients were analysed. The *P*-values were determined using Fisher's exact test.

levels of the promoter region in *Kdm4b*-PEs were significantly reduced, as follows: 5′-R, *Kdm4b*-PEs: 0.84% versus *Egfp*-PEs: 11.48%, $P<0.05$, Student's *t*-tests; XP, *Kdm4b*-PEs: 3.13% versus *Egfp*-PEs: 16.08%; $P<0.04$, Student's *t*-tests; and RA, *Kdm4b*-PEs: 20.16% versus *Egfp*-PEs: 29.99%; Fig. 4g. Taken together, these results demonstrated that H3K9me3 at the promoter region protected Xm-*Xist*, preventing RNF12-mediated activation from the four-cell stage. We concluded that silencing of Xm-*Xist* by imprinting to establish iXCI involved H3K9me3.

Maternal repressive H3K9me3 mark is absent in ES cells. Previous studies have shown that *Xist* is ectopically expressed in embryos cloned from somatic and ES cells[27,28]. However, the cause of aberrant *Xist* expression in cloned embryos remains unknown. Given that high H3K9me3 levels at the promoter region in PEs are lost during development (Fig. 4d,e), we investigated whether the maternal repressive H3K9me3 mark was lost in ES cells and whether *Xist* in ES cells was permissive against RNF12 during oocyte-mediated reprogramming.

To test this possibility, we first examined H3K9me3 states at *Xist* regions in various types of male ES cells, using published data[29,30]. The levels of H3K9me3 at *Xist* regions containing promoters in TT2 and E14 ES cell lines were low compared with those in positive control regions (Supplementary Fig. 9). ChIP-seq analysis revealed that although ectopic expression of KDM4B

in male ES cells induced global H3K9me3 demethylation (Supplementary Fig. 10a–d), it did not alter H3K9me3 levels at *Xist* regions. Moreover, these levels were low compared with those of a known H3K9me3-rich region (Supplementary Fig. 10e)[16]. Furthermore, the expression of *Xist* in cloned embryos was also not affected by ectopic expression of KDM4B (Supplementary Fig. 10f). These results indicated that the maternal repressive H3K9me3 mark was lost.

To establish whether RNF12 is involved in *Xist* activation during oocyte-mediated reprogramming, oocytes treated with si-*Rnf12* were used as recipients for nuclear transfer (Supplementary Fig. 10g). At the four-cell stage, derepression of *Xist* transcription in ES-cloned embryos depended on RNF12 (>50% repression in si-*Rnf12* group; Supplementary Fig. 10h). Taken together, these data indicated that the intrinsic H3K9me3 mark, which was essential for repression of *Xist* by RNF12, was lost during embryo development. This indicated that the primary cause of aberrant *Xist* expression in cloned embryos was involved in loss of intrinsic H3K9me3 at *Xist* regions.

Effects of iXCI disruption on FEs. The effects of Xm-*Xist* derepression on postimplantation development remain unclear. First, we asked whether ectopic *Kdm4b* expression caused Xm-*Xist* derepression in FEs. *Kdm4b*-FEs developed into blastocysts with high efficiency (>80% of two-cell embryos; Supplementary Table 6). At the 96-h blastocyst stage, *Xist* transcription was derepressed in male *Kdm4b*-FEs, while *Pgk1* and *Plac1* levels were reduced to <13% of those observed in controls (Supplementary Fig. 11a). In female *Kdm4b*-FEs, FISH analysis revealed that there were cells with *Xist* biallelic expression (Supplementary Fig. 11b). Although the expression level was slightly elevated (1.3-fold), X-linked genes were also significantly repressed in female *Kdm4b*-FEs (Supplementary Fig. 11c). These results showed that ectopic Xm-*Xist* derepression caused X-linked gene silencing and elimination of iXCI.

To test the effects of iXCI disruption on postimplantation development, we conducted *in vivo* transplantation experiments. Interestingly, our results demonstrated that XCI on Xm during preimplantation did not affect developmental competence to term (*Kdm4b*-FEs: 63.2% versus *Egfp*-FEs: 53.1%; Supplementary Fig. 11d,e), suggesting that aberrant XCI in preimplantation embryos was restored during postimplantation development, probably through an automatic counting function. These results were consistent with the observation that the developmental competency of embryonic cloned embryos was high, despite the ectopic expression of *Xist* and the occurrence of global XCI (Supplementary Fig. 12)[28,31].

Loss of XCI impairs the postimplantation development of PEs. It is still unknown whether the embryonic lethality observed immediately after implantation in the majority of PEs (around 70–80%) can be attributed to the loss of dysregulation of X-linked genes or to loss of expression of autosomal paternally imprinted genes. Figure 2 showed that the percentage of *Xist*-positive cells in *Kdm4b*-PEs was significantly higher than that in *Egfp*-PEs. Thus, we reasoned that *Kdm4b*-PEs would be suitable for studying this long-standing question.

We first performed a detailed analysis of Xm-XCI states in *Kdm4b*-PEs using IF against H3K27me3 and CDX2, a marker of the trophectoderm, in combination with *Xist* FISH, at the blastocyst stage. This analysis revealed that *Xist*-positive cells were significantly increased in the trophectoderm of *Kdm4b*-PEs, although the ratio of H3K27me3-positive cells was comparable to that of *Egfp*-PEs (Fig. 5a–c and Supplementary Table 7). However, no significant difference was observed in the inner cell mass between groups (Fig. 5a–c and Supplementary Table 7), indicating that loss of H3K9me3 in the maternal genome led to establishment of Xm-XCI as an imprinted Xp-XCI.

Next, we carried out transcriptome analysis in *Egfp*-PEs, *Kdm4b*-PEs and FEs using microarray. Clustering analysis based on gene expression patterns showed that all three groups could be distinguished clearly from each other (Supplementary Fig. 13a). Comparison of transcripts between *Egfp*-PEs and *Kdm4b*-PEs identified transcripts that were significantly differentially expressed: 671 transcripts were upregulated and 711 transcripts were downregulated ($P<0.05$, Student's *t*-test and >1.5-fold changes in *Kdm4b*-PEs). Chromosome distribution analysis showed that upregulated transcripts in *Kdm4b*-PEs were distributed across various chromosomes (2.33–7.23%; Supplementary Fig. 13b). However, downregulated transcripts in *Kdm4b*-PEs were mostly concentrated on the X chromosome, which particularly involved the *Xlr* and *Magea* families (10.26%; Supplementary Fig. 13c–e).

Comparison of the imprinted genes between *Egfp*- and *Kdm4b*-PEs revealed that only six genes were significantly differentially expressed (paternally expressed genes: *Impact* and *Fthl17*; maternally expressed genes: *Gnas*, *H13*, *Xlr3b* and *Xlr4b*; Fig. 5d). Clustering analysis based on the expression of imprinted genes showed that the expression levels in *Kdm4b*-PEs were similar to those in *Egfp*-PEs rather than to those of biparental embryos (Fig. 5d). Thus, H3K9me3 demethylation does not result in restoration of expression states in paternally expressed genes.

We conducted *in vivo* transplantation experiments using *Kdm4b*-PEs. Surprisingly, at E6.5, the developmental rates of *Kdm4b*-PEs were markedly increased compared with those of *Egfp*-PEs (*Kdm4b*-PEs: 90% versus *Egfp*-PEs: 35%; $P<0.001$, Fisher's exact test; Fig. 5e,f). At E9.5, although the stages of the recovered embryos varied, *Kdm4b*-PEs retained a significantly higher developmental ability compared with controls (*Kdm4b*-PEs: 64.3% versus *Egfp*-PEs: 31.9%; $P<0.002$, Fisher's exact test; Fig. 5e,f).

However, we did not rule out the possibility that the significant improvement in *Kdm4b*-PE development resulted from restoration of the expression levels of some imprinted genes. To determine whether the improvement in developmental competency could be attributed to the gain of XCI, we constructed *Kdm4b* + si-*Rnf12*-PEs. In *Kdm4b*-PEs with si-*Rnf12* at the blastocyst stage, *Xist* expression analysis by FISH revealed that *Xist* cloud signals in control *Kdm4b*-PEs were present in 51.9% of cells, while those in *Rnf12*-knockdown *Kdm4b*-PEs were present in only 10% of cells, and most of the signals were pinpoint rather than cloud (Fig. 5g,h and Supplementary Table 8).

qPCR analysis showed that although *Xist* signals were significantly reduced in *Kdm4b*-PEs with si-*Rnf12* (12.5% of the control on average), the expression levels of *Impact*, *H13* and *Gnas*, which are expressed in response to ectopic *Kdm4b* expression (Fig. 5d), did not change when compared with those of controls (Fig. 5i). We further demonstrated that RNF12 depletion did not affect *Tsix* and *Sfmbt2* expression levels in *Kdm4b*-PEs (Fig. 5i). These results clearly indicated that RNF12 depletion led to *Xist* downregulation in *Kdm4b*-PEs, without altering the features of PEs.

Finally, *in vivo* transplantation experiments demonstrated that *Xist* repression by RNF12 depletion significantly inhibited developmental competency at E6.5 in *Kdm4b*-PEs (*Kdm4b* + si-control: 84.2% versus *Kdm4b* + si-*Rnf12*: 23.5%; $P<0.0006$, Fisher's exact test; Fig. 5j,k). Taken together, these results demonstrated that the developmental defects seen in PEs immediately after implantation could be attributed to the lack of XCI, but not to loss of expression of paternally expressed genes.

Discussion

In this study, we demonstrated that maternal imprinting of Xm, which protected against *Xist* activation by RNF12 in the preimplantation stages, was mediated by H3K9me3.

Xm-*Xist* imprints are established during oogenesis and autosomal imprinting also occur in the phases[6,10]. In many imprinted genes, DNA methylation at the promoter regions is the primary regulator and H3K9me3 modifications overlap with these regions[32]. However, it is not clear why Xm-*Xist* regions are targeted by H3K9me3, but not by DNA methylation. One of the possibilities is that during primordial germ cell development, *Xist* must be silenced to activate the inactivated allele before inducing the expression of *Dnmt3a/3l*, which encodes a *de novo* DNA methyltransferase that is activated during oogenesis[33]. Consistent with this concept, *Xist* repression begins in primordial germ cells at E10.5 (ref. 34). Thus, comparison of H3K9me3 states at promoter regions in non-growing oocytes with those in growing oocytes will greatly facilitate understanding of the molecular mechanisms of iXCI.

We found that *Kdm4b*-, *Egfp* + TSA- and *Kdm4b* + TSA-PEs did not show complete XCI at the blastocyst stage as compared with female FEs. These results suggested that other repressive marks were imposed on Xm to silence Xm-*Xist* expression. Alternatively, removal of H3K9me3 may be incomplete because demethylation at RA regions was mild (Fig. 4g). However, it is not clear why RA regions show resistance against demethylation by KDM4B. As suggested in a previous study, this mechanism may be related to the three-dimensional structure of the A-repeat, which has been reported to constitute stable regions in the *Xist* transcript[35]. Further studies using ChIP-seq and/or chromatin-conformation capture sequencing technologies in preimplantation embryos are required for comprehensive understanding of *Xist* regulation.

In ES cells, RNF12 induces *Xist* expression through degradation of REX1, which is required for suppression of *Tsix*[36]. Interestingly, we did not detect *Tsix* expression from the morula to the blastocyst stages in *in vitro*-fertilized (IVF) embryos, implying that the molecular mechanism of RNF12-mediated *Xist* activation differs between imprinted and randomly induced XCI. It is not known whether the role of RNF12 in *Xist* activation during the preimplantation stages was direct or indirect. Recent RNF12 studies reported the specific binding of RNF12 to *Smad7* in mouse ES cells[37], suggesting that signalling via transforming growth factor-β family members may be associated with imprinted *Xist* activation.

In this study, we revealed the molecular mechanisms underlying imprinting of XCI and demonstrated the role of XCI in various types of embryo development in mice. Recent studies using somatic- and ES-cloned embryos revealed that aberrant *Xist* reprogramming is a major cause of developmental failure in cloned embryos[27,28]. We found that RNF12 was highly expressed in oocytes compared with somatic and ES cells (>80-fold). Moreover, we showed that H3K9me3 levels at *Xist* promoter regions were low in ES cells and that *Xist* expression in ES-cloned embryos depended on RNF12. These data provided the first evidence that RNF12 inhibited developmental reprogramming. Therefore, the use of RNF12-depleted oocytes as recipient cells would improve cloning efficiency. However, *Xist* activation in cloned embryos may be induced by factors other than RNF12, as supported by the observation that *Xist* was still expressed at ~40% of control levels, even after marked depletion of RNF12. Consistent with this notion, a recent study has demonstrated that RNF12 is dispensable for random XCI *in vivo*[38].

Xm-*Xist* derepression from the four-cell stage could rescue developmental defects in PEs. This finding demonstrated that the primary cause of developmental failure immediately after implantation was a lack of XCI, but not a lack of expression of paternally imprinted genes. We also tested whether *Kdm4b*-PEs could extend development; however, we did not observe extended *Kdm4b*-PE development after E9.5, implying that expression of paternally imprinted genes is required for subsequent development in PEs[11,12].

Our data resolved several long-standing unanswered questions about XCI during preimplantation in various types of embryos (Supplementary Fig. 14). Moreover, given that injection of *Kdm4b* mRNA into PEs improved their developmental ability, genetic mutation leading to embryonic lethality could be rescued by transient expression of epigenomic modifiers during preimplantation phases.

Methods

Embryo manipulations. All mice were maintained and used in accordance with the Guidelines for the Care and Use of Laboratory Animals of the Japanese Association for Laboratory Animal Science and the National Research Institute for Child Health and Development (NRICHD) of Japan. All animal experiments were performed according to protocols approved by the Institutional Animal Care and Use Committee of the NRICHD (Permit Number: A2006-009).

Adult female B6D2F1 mice were purchased from Clea Japan (Tokyo, Japan) and oocytes were collected following standard methods[27]. PEs were generated using Ca-free M16 medium containing 8 mM $SrCl_2$ and Cytochalasin B (5 μg ml^{-1}) (Sigma-Aldrich, St Louis, MO, USA), and cultured KSOM (EMD Millipore, Darmstadt, Germany). Injection experiments (mRNA, short interfering RNA (siRNA) and nuclear transfer) were conducted using a Prime Tech Piezo drive (Sutter Instrument Company, Novato, CA, USA). To produce cloned embryos, nuclear-transferred oocytes were parthenogentically activated. Manipulated embryos were cultured to the developmental stages, as follows: 4-cell, 48 h; morula, 72 h; and blastocyst, 96 and 120 h after parthenogenetic activation or ICSI, respectively. All embryos were cultured at 37 °C in KSOM in an atmosphere containing 5% CO_2. In the TSA experiment, the embryos were cultured for 24 h in activation and culture media containing 50 nM TSA (Sigma-Aldrich). IVF fertilization and nuclear transfer were performed following published procedures[27]. To determine the effects of ectopic KDM4B expression on *Xist* expression in cloned embryos, doxycycline was added to ES cell culture and KSOM medium to a final concentration of 2 μg ml^{-1}. Pseudopregnant ICR mice (Clea Japan) were used as embryo recipients. At E6.5, E9.5 and E18.5, the embryos were recovered from the uterus.

***In vitro* mRNA synthesis.** The coding region of *Kdm3a* was amplified from mouse testis complementary DNA using PCR with KOD-Plus-Neo DNA polymerase (Toyobo, Osaka, Japan). Forward and reverse primers contained T7 promoter and poly(T)$_{120}$ sequences, respectively. A step-down PCR amplification method was used, following the manufacturer's instructions (Toyobo). Poly(A)-containing PCR products were subjected to *in vitro* transcription using a mMESSAGE mMACHINE T7 ULTRA Kit (Life Technologies, Carlsbad, CA, USA), following the manufacturer's instructions. To generate a *Kdm4b* DNA template for *in vitro* transcription, pCMV-SPORT6 containing the full-length *Kdm4b* mRNA was used as the PCR template (DNAFORM, Kanagawa, Japan, Clone ID 3490671). *Egfp* cDNA was cloned using the pGEM-T Easy Vector System (Promega, Madison, WI, USA) and transcribed *in vitro* using the mMESSAGE mMACHINE T7 ULTRA Kit (Life Technologies) following the manufacturer's instructions. The concentrations of the mRNAs were adjusted to 150 ng ml^{-1} (*Egfp*), 550 ng ml^{-1} (*Kdm3a*), or 450 ng ml^{-1} (*Kdm4b*) to maintain a constant number of injected mRNA molecules. The primer sequences used for generating the templates for *in vitro* transcription are shown in Supplementary Table 9.

***Rnf12* knockdown.** siRNA targeting *Rnf12* (si-Rnf12 sense 5′-GAAGUCAAAUG GAUCGCUUTT-3′ A and antisense 5′-AAAGCGAUCCAUUUGACUUCTG-3′ GC, and the negative control siRNA (si-control: 4390846) were purchased from Life Technologies. The final concentration of each siRNA was 50 ng ml^{-1}. The siRNA was injected into MII oocytes using the Piezo drive and then incubated for 6–7 h in KSOM medium at 37 °C in an atmosphere containing 5% CO_2 before mRNA injection. For the NT experiment using *Rnf12*-knockdown oocytes, oocytes were incubated for 5–6 h after siRNA injection, and NT was then conducted and activated as described above.

Immunofluorescence. Oocytes injected with mRNAs were subjected to ICSI. After 10–11 h, the zygotes were fixed with 2% paraformaldehyde (PFA) in PBS containing 0.1% polyvinyl alcohol (PBS-PVA) for 15 min at room temperature (RT). Zygotes were then permeabilized using 0.2% Triton X-100 in PBS-PVA for 15 min at RT and blocked in 1% BSA in PBS-PVA for 1 h at RT. The primary antibodies used in the assay were as follows: anti-H3K9me3 (ab8898, 1:500 dilution, Abcam, Cambridge, UK), anti-H3K9me2 (ab1220, 1:500, Abcam) and anti-H3K27me3

(07-449, 1:150, EMD Millipore). The primary antibodies were diluted with blocking solution (PBS-PVA containing 1% BSA) and the embryos were incubated overnight at 4 °C. After washing in blocking solution, the embryos were incubated for 1 h at RT with Alexa Flour 634- or 546-conjugated anti-mouse or anti-rabbit IgG secondary antibodies (1:500, Life Technologies). After the embryos were washed, the nuclei were stained with 1 μg ml^{-1} 4′,6-diamidino-2-phenylindole and the embryos were placed on a glass slide and observed with a LSM510 laser scanning confocal microscope (Carl Zeiss, Oberkochen, Germany). Signal intensities of maternal and paternal pronuclei were calculated using NIH ImageJ software (http://rsb.info.nih.gov/ij/).

In *Rnf12*-knockdown experiments, one-cell and four-cell PEs were fixed at 10–11 h (18–19 h after siRNA injection) and 48 h after activation, respectively. Anti-RNF12 (1:500, Abnova, Taipei, Taiwan) and Alexa Fluor 488-conjugated anti-mouse IgG antibodies (1:500, Life Technologies) were used as the primary and secondary antibodies, respectively. *Rnf12*-knockdown and negative-control PEs were observed under the same conditions, to assess knockdown efficiency. Signal intensities were calculated using ImageJ software.

Fluorescent *in situ* hybridization. The zona pellucida of embryos was removed using acid Tyrode solution (Sigma-Aldrich) and then fixed and permeabilized with 2% PFA-PVA containing 0.25% Triton X-100 for 10 min on ice. The samples were placed on glass slides, evaporated to dryness, dehydrated sequentially in 70 and 100% ethanol and then air-dried. Hybridization buffer containing an *Xist* probe (provided by T. Sado) was prepared using a Nick Translation Kit (Abbott, Abbott Park, IL, USA) and Cy3-dUTP (GE Healthcare Life Sciences, Fairfield, CT, USA) and was then applied to the slides. The slides were then incubated and washed as previously described[26]. Fluorescence was visualized using the LSM510.

IF combined with FISH. The zona pellucida of embryos was removed using acid Tyrode solution (Sigma-Aldrich) and fixed with 2% PFA-PVA for 15 min at RT in four-well dishes. The fixed samples were permeabilized with 0.5% Triton X-100 in PBS-PVA for 20 min on ice. After washing with PBS-PVA, the samples were blocked in 1% BSA-PBS-PVA containing 1.3 U ml^{-1} RNaseOUT (Life Technologies) for 30 min at RT. After washing, the embryos were incubated with primary antibodies (anti-CDX2 (BioGenex, San Ramon, CA, USA), diluted 1:30 and anti-H3K27me3 diluted 1:150 in blocking buffer containing 1.3 U ml^{-1} RNaseOUT) for 1 h at RT. Secondary antibody reactions were performed as described above. The samples were placed on glass slides, evaporated to dryness, dehydrated sequentially in 70 and 100% ethanol and air dried. The samples were then analysed by FISH according to the procedures described above.

Analysis of IF combined with FISH data. *Xist* cloud signals detected in three-dimensional images using *Z*-sections of the LSM Image Browser (Carl Zeiss) were judged as positive. Only cells that did not overlap at interphase were used in the analysis. Biallelic expression was defined as cells with two *Xist* cloud spots. Statistical analysis was performed using Fisher's exact test.

Gene expression analysis. Total RNA was extracted using an RNeasy Micro Kit (Qiagen, Venlo, The Netherlands) and treated with DNase following the manufacturer's instructions. mRNAs were reverse transcribed using an oligo(dT) primer and SuperScriptIII Reverse Transcriptase (Life Technologies). For quantitative gene expression analysis with high specificity, TaqMan probes (Life Technologies) were used in all assays. In four-cell stage embryos, *Xist* was assayed in triplicate and only the samples that were detected in two or three replicates were judged as positive. In morulae and blastocysts, expression of target genes was assayed in duplicate. *Gapdh* was used as the internal control in the four-cell-stage assays and *Rnf12* was used in the time-lapse assays. *Gapdh* and *Actb* (encoding β-actin) were used as internal controls at the morula and blastocyst stages. For normalization of qPCR analysis (Fig. 2b,g), the expression levels of all embryos were normalized to the average expression levels of *Egfp*-PEs. The TaqMan probes and primer sets used in this study are shown in Supplementary Table 8.

Generation of *Kdm4b*-inducible ES cell lines and ES cell culture. The XhoI- and ClaI-linearized pGEM-IRES-EGFP plasmids were inserted into the cognate sites of pPB-CAG-EBNX (provided by A. Bradley) to generate pPB-CAG-IRES-EGFP. A Tet3G fragment with BglII and XhoI cleavage sites was amplified from a pEF1a-Tet3G template (Clontech Laboratories, Mountain View, CA, USA) using PCR and inserted into pPB-CAG-IRES-EGFP, generating the vector pPB-CAG-Tet3G-IRES-EGFP. The XhoI and BamHI cleavage sites in pPB-UbC (provided by A. Bradley) were replaced with the p-TRE3G multiple cloning sites (Clontech). The *Kdm4b* coding sequence, with terminal ClaI and BamHI cleavage sites, was amplified by PCR and inserted into the corresponding sites of pPB-TRE3G, yielding pPB-TRE-Kdm4b.

The NCH.4.6 male mouse ES cell line (C57B6/N × C57B6/N), which had a normal karyotype, was electroporated with pPB-TRE-Kdm4b, pPB-CAG-Tet3G-IRES-EGFP and pCMV-hyPBase (provided by A. Bradley). All ES cells used in this study were cultured in knockout DMEM (Life Technologies) containing recombinant human leukemia inhibitory factor culture supernatant for mouse ES cell culture (Wako Pure Chemical Industries, Ltd, Osaka, Japan), as well as GlutaMAX, 2-mercaptoethanol, non-essential amino acids and 15% KSR (all from Life Technologies). Doxycycline (2 μg ml^{-1}; Sigma-Aldrich) was added to ES cell culture medium to induce ectopic KDM4B expression.

Western blotting. Cells were extracted using a stock lysis buffer containing 1 M Tris–HCl, 5 M NaCl, 10% Triton-X and protease inhibitors, and were subjected to e-PAGEL (ATTO, Amherst, NY, USA) electrophoresis. The membranes were washed in TBS containing 0.1% Tween 20 (TBS-T) and blocked in 5% skim milk in TBS-T for 1 h. The membranes were incubated with anti-KDM4B antibodies (1:500 dilution; Active Motif, Carlsbad, CA, USA) overnight at 4 °C, washed and incubated with a rabbit horseradish peroxidase-conjugated secondary antibody (1:5,000 dilution; Sigma-Aldrich) for 1 h at RT. Immunoblottings were visualized using SuperSignal chemiluminescent substrate (Thermo Scientific, Waltham, MA, USA) and an ImageQuant LAS4000 system (GE Healthcare). After capturing the images, the membranes were washed with WB Stripping Buffer (Nacalai Tesque, Kyoto, Japan) for 10 min, washed with TBS-T and incubated with an anti-β-actin antibody conjugated to fluorescein isothiocyanate (1:2,000 dilution; Sigma-Aldrich) for 1 h at RT.

ChIP analysis of K4B-ES cells. Trypsinized feeder-free ES cells (2×10^7) were collected and fixed with 1% formaldehyde. The cells were resuspended in SDS lysis buffer (ChIP Reagent, Nippon Gene Co., Ltd.) and the lysate was sonicated to fragment chromatin using a S220 Focused-ultrasonicator (Covaris, Woburn, MA, USA). The chromatin was purified by centrifugation and immunoprecipitated with Protein A-beads (Veritas Life Sciences, Ribeirão Preto, Brazil) conjugated to anti-H3K9me3 antibodies (Abcam: ab8898) or rabbit IgG (Abcam: ab37415) in Buffer A with protease inhibitor (LowCell ChIP kit, Diagenode, Denville, NJ, USA) overnight at 4 °C. The chromatin beads were washed with Buffers A and C (LowCell ChIP kit). After washing, the chromatin beads were incubated in ChIP direct elution buffer (ChIP Reagent) for 6 h at 65 °C, following incubation with 2 μl proteinase K (20 mg0 ml^{-1}) for 2 h at 55 °C. The DNA immunoprecipitated from the supernatant was purified using Agencourt AMPure XP beads (Beckman Coulter, Inc., Pasadena, CA, USA) according to the manufacturer's instructions.

ChIP combined with deep sequencing. ChIP-Seq libraries were prepared using the NEBNext ChIP-Seq Library Prep Master Mix Set and Multiplex Oligos from Illumina (New England BioLabs Inc., Ipswich, MA, USA) according to the manufacturer's instructions. Ten nanograms of ChIP or input DNA was subjected to end repair, dA-tailing and adaptor ligation, and amplified using nine cycles of PCR. The final library size was checked using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). After the concentration of each library was determined using qPCR with a KAPA Library Quantification Kit-Illumina/Universal system (KK4824, Kapa Biosystems, Wilmington, MA, USA), the libraries were sequenced using the HiSeq 1000 sequencing system (Illumina, San Diego, CA, USA) to generate 100 bp × 2 paired-end data.

ChIP-seq data analysis. Reads from each sample were first trimmed by removing adapter sequences and low-quality bases at ends using Trimmomatic 0.22 (http://www.usadellab.org/cms/index.php?page=trimmomatic). Approximately 115 million reads for each of the ChIP and input libraries were aligned to the mouse reference genome (mm10:http://genome.ucsc.edu/cgi-bin/hgGateway) using the Burrows-Wheeler Aligner 0.6.2. Uniquely mapped reads were selected using a custom script, converted from SAM to BAM format using SAMtools 0.1.18 and processed using Picard 1.83 to mark PCR duplicates. Reads with a mapping quality of <20 were removed using SAMtools 0.1.18. The resulting BAM files (a pair of files for ChIP and input libraries) were visualized using the Integrative Genomics Viewer (http://www.broadinstitute.org/igv/) and subjected to peak detection using the MACS algorithm implemented in Avadis NGS software (Agilent). In scatter plot analysis using 1 and 15 K4B-ES cell lines, the numbers of mapped reads were counted for 10,000-bp windows (with a sliding size of 5,000 bp). To adjust for differences in total amount of reads, the number of mapped reads in each window was transformed into reads per million format. Calculation methods are available on request.

ChIP-qPCR analysis of sperm. Sperm were obtained from BDF1 mice aged 9–12 weeks. Preparation of sperm chromatin was performed according to published protocols with modifications[39,40]. For each native ChIP experiment, 5×10^7 sperm were used. Sperm were washed twice with PBS. The pellet was suspended in PBS containing 0.5% Triton-X, 10 mM dithiothreitol (DTT) and protease inhibitor (Diagenode), and incubated on ice for 1.5 h. After washing with PBS, pelleted sperm nuclei were suspended in 400 μl PBS containing 1 mM $CaCl_2$ and 1 mM DTT, and incubated for 5 min at 37 °C. After incubation, 1 μl (2×10^6 gel units per ml) micrococcal nuclease (New England BioLabs) was added to the nuclei, which were then incubated for 5 min at 37 °C. EDTA was added to a concentration of 0.5 mM and solubilized chromatin was clarified by centrifugation for 15 min at 15,000 r.p.m. at 4 °C. The pellets were suspended in PBS containing $CaCl_2$ and DTT (at the same concentrations as used above), and treated again with micrococcal nuclease. To examine whether H3K9me3-modified nucleosomes were

present in sperm chromatin, soluble (chromatin) and insoluble (pellet) fractions were subjected to western blotting using anti-H3K9me3 antibodies (ab8898; 1:1,000), as described above.

Chromatin was incubated with Protein A beads conjugated to anti-H3K9me3 antibodies (ab8898) or rabbit IgG (ab37415) overnight at 4 °C in ChIP buffer (40 mM Tris–HCl, pH 7.5, 1 M NaCl and 10 mM EDTA). Pelleted beads were washed twice with Buffer 1 (50 mM Tris–HCl, pH 7.5, 500 mM NaCl and 10 mM EDTA) and Buffer 2 (50 mM Tris-HCL, pH 7.5, 300 mM NaCl and 10 mM EDTA). The pelleted beads were suspended in ChIP direct elution buffer and incubated with proteinase K for 2 h at 37 °C. The immunoprecipitated DNA was then purified using Agencourt AMpure XP beads.

ChIP-qPCR analysis was performed according to published methods using SYBR Green[39]. The sequences of each primer set are listed in Supplementary Table 9.

eChIP-quantitative qPCR. The zona pellucidae of the embryos were removed by acid Tyrode's solution and washed in PBS containing 0.1% PVA. The embryos were suspended in PBS containing 0.5% Triton-X, 0.5 mM DTT and protease inhibitor, and incubated on ice for 30 min. After incubation, 1 mM $CaCl_2$ was added to the buffer and samples were incubated for 5 min at 37 °C. After incubation, 0.5 µl (2×10^6 gel units per ml) micrococcal nuclease (New England BioLabs) was added to the nuclei, which were then incubated for 5 min at 37 °C. EDTA was added to a concentration of 0.5 mM and solubilized chromatin was clarified by centrifugation at 15,000 r.p.m. for 15 min at 4 °C. The same procedures were repeated one more time. Chromatin was incubated with Protein A beads conjugated to anti-H3K9me3 antibodies (ab8898) or rabbit IgG (ab37415), prepared as described above, overnight at 4 °C in ChIP buffer (40 mM Tris–HCL, pH 7.5, 1 M NaCl and 10 mM EDTA). Pelleted beads were washed twice with Buffer 1 (50 mM Tris–HCL, pH 7.5, 500 mM NaCl and 10 mM EDTA) and then with Buffer 2 (50 mM Tris–HCL, pH 7.5, 300 mM NaCl and 10 mM EDTA). The pelleted beads were then suspended in ChIP direct elution buffer and incubated with proteinase K for 2 h at 55 °C. The immunoprecipitated DNA was then purified using Agencourt AMpure XP beads.

Eluted DNA (20 µl) was divided into two aliquots; one (4 µl) was used for a SYBR Green assay targeting a major satellite and the other (16 µl) was subjected to pre-amplification using a Single Cell-to-CT kit (Ambion, Austin, TX, USA) according to the manufacturer's instructions. The number of PCR cycles at the pre-amplification step was 20. The primer and probe sequences used are shown in Supplementary Table 9.

Microarray analysis. Five *Egfp*-PE, *Kdm4b*-PE and IVF blastocysts (120 h) were lysed using ISOGEN (Nippongene) and RNA was extracted by phenol–chloroform and isopropanol precipitation. cDNA was synthesized using the Ovation RNA Amplification System V2 kit (NuGEN, West Cumbria, UK) and hybridized with SurePrint G3 Mouse GE 8x60K Microarray (Agilent Technologies). Analysis was conducted using GeneSpringV12.5 (Agilent Technologies). Transcripts were considered to be expressed if raw values were >100 and a flag was present in at least one of the groups.

References

1. Augui, S., Nora, E. P. & Heard, E. Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.* **12,** 429–442 (2011).
2. Lee, J. T. Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nat. Rev. Mol. Cell Biol.* **12,** 815–826 (2011).
3. Wutz, A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat. Rev. Genet.* **12,** 542–553 (2011).
4. Takagi, N. & Sasaki, M. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature* **256,** 640–642 (1975).
5. Shin, J. *et al.* Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. *Nature* **467,** 977–981 (2010).
6. Tada, T. *et al.* Imprint switching for non-random X-chromosome inactivation during mouse oocyte growth. *Development* **127,** 3101–3105 (2000).
7. Chiba, H. *et al. De novo* DNA methylation independent establishment of maternal imprint on X chromosome in mouse oocytes. *Genesis* **46,** 768–774 (2008).
8. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. & Jaenisch, R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* **11,** 156–166 (1997).
9. Liu, N. *et al.* Genome-wide gene expression profiling reveals aberrant MAPK and Wnt signaling pathways associated with early parthenogenesis. *J. Mol. Cell Biol.* **2,** 333–344 (2010).
10. Obata, Y. & Kono, T. Maternal primary imprinting is established at a specific time for each gene throughout oocyte growth. *J. Biol. Chem.* **277,** 5285–5289 (2002).
11. Kono, T. *et al.* Birth of parthenogenetic mice that can develop to adulthood. *Nature* **428,** 860–864 (2004).
12. Kawahara, M. *et al.* High-frequency generation of viable mice from engineered bi-maternal embryos. *Nat. Biotechnol.* **25,** 1045–1050 (2007).
13. Santos, F., Peters, A. H., Otte, A. P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev. Biol.* **280,** 225–236 (2005).
14. Cantone, I. & Fisher, A. G. Epigenetic programming and reprogramming during development. *Nat. Struct. Mol. Biol.* **20,** 282–289 (2013).
15. Lewis, A. *et al.* Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nat. Genet.* **36,** 1291–1295 (2004).
16. Yuan, P. *et al.* Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev.* **23,** 2507–2520 (2009).
17. Fodor, B. D. *et al.* Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes Dev.* **20,** 1557–1562 (2006).
18. Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486,** 415–419 (2012).
19. Nesterova, T. B., Barton, S. C., Surani, M. A. & Brockdorff, N. Loss of Xist imprinting in diploid parthenogenetic preimplantation embryos. *Dev. Biol.* **235,** 343–350 (2001).
20. Plath, K. *et al.* Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300,** 131–135 (2003).
21. Okamoto, I., Tan, S. & Takagi, N. X-chromosome inactivation in XX androgenetic mouse embryos surviving implantation. *Development* **127,** 4137–4145 (2000).
22. Barakat, T. S. *et al.* RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS Genet.* **7,** e1002001 (2011).
23. Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat. Genet.* **40,** 411–420 (2008).
24. Dahl, J. A. & Collas, P. A rapid micro chromatin immunoprecipitation assay (microChIP). *Nat. Protoc.* **3,** 1032–1045 (2008).
25. Turner, B. in *Mapping Protein/DNA Interactions by Cross-Linking* (2001).
26. Hoki, Y. *et al.* A proximal conserved repeat in the Xist gene is essential as a genomic element for X-inactivation in mouse. *Development* **136,** 139–146 (2009).
27. Fukuda, A. *et al.* Identification of inappropriately reprogrammed genes by large-scale transcriptome analysis of individual cloned mouse blastocysts. *PLoS ONE* **5,** e11274 (2010).
28. Inoue, K. *et al.* Impeding Xist expression from the active X chromosome improves mouse somatic cell nuclear transfer. *Science* **330,** 496–499 (2010).
29. Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149,** 590–604 (2012).
30. Karimi, M. M. *et al.* DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8,** 676–687 (2011).
31. Ono, Y. & Kono, T. Irreversible barrier to the reprogramming of donor cells in cloning with mouse embryos and embryonic stem cells. *Biol. Reprod.* **75,** 210–216 (2006).
32. Dindot, S. V., Person, R., Strivens, M., Garcia, R. & Beaudet, A. L. Epigenetic profiling at mouse imprinted gene clusters reveals novel epigenetic and genetic features at differentially methylated regions. *Genome Res.* **19,** 1374–1383 (2009).
33. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.* **9,** 129–140 (2008).
34. Sugimoto, M. & Abe, K. X chromosome reactivation initiates in nascent primordial germ cells in mice. *PLoS Genet.* **3,** e116 (2007).
35. Duszczyk, M. M., Wutz, A., Rybin, V. & Sattler, M. The Xist RNA A-repeat comprises a novel AUCG tetraloop fold and a platform for multimerization. *RNA* **17,** 1973–1982 (2011).
36. Gontan, C. *et al.* RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* **485,** 386–390 (2012).
37. Zhang, L. *et al.* RNF12 controls embryonic stem cell fate and morphogenesis in zebrafish embryos by targeting Smad7 for degradation. *Mol. Cell* **46,** 650–661 (2012).
38. Shin, J. *et al.* RLIM is dispensable for X-chromosome inactivation in the mouse embryonic epiblast. *Nature* **511,** 86–89 (2014).
39. Brykczynska, U. *et al.* Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat. Struct. Mol. Biol.* **17,** 679–687 (2010).
40. Hammoud, S. S. *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460,** 473–478 (2009).

Acknowledgements

We are grateful to T. Sado for critical reading of this manuscript and discussions. We thank T. Sugawara and H. Kobayashi for helpful comments; A. Bradley for providing the PiggyBac vector; K. Kusakabe, T. Kikuchi, Y. Harada, S. Kanai and Y. Takahashi for technical assistance with some experiments and analysis; and T. Kawasaki for preparation of figures. This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan; a grant from the Ministry of Health, Labour and Welfare (MHLW) to H.A. and A.U.; a Grant-in-aid for Scientific

Research (21390456); a grant from JST-CREST to H.A.; and a JSPS KAKENHI Grant-in-Aid for Research Activity Start-up to A.F. (24890300).

Author contributions

A.F., K.E. and H.A. conceived and designed the study. A.F. performed the experiments and analysis of embryo manipulation, IF, FISH, qPCR, ChIP, cultured ES cells, vector construction and microarray analysis, and developed the eChIP-qPCR technique. J.T., K.H. and K.N. conducted ChIP-seq experiments and analyses. T.M. and H.A. constructed vectors and cultured ES cells. H.A. and A.U. supervised the study. A.F., K.N. and H.A. wrote the manuscript.

Additional information

Accession codes: The original data for the microarray have been deposited in the GEO at http://www.ncbi.nlm.nih.gov/geo/ (accession number: GSE53662). The original data of ChIP-seq have been deposited in DDBJ at http://cibex.nig.ac.jp/index.jsp with accession number: DRA001041.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Fukuda, A. *et al.* The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice. *Nat. Commun.* 5:5464 doi: 10.1038/ncomms6464 (2014).

Journal of Human Genetics (2014) 59, 326–331

www.nature.com/jhg

ORIGINAL ARTICLE

Compilation of copy number variants identified in phenotypically normal and parous Japanese women

Ohsuke Migita[1,11], Kayoko Maehara[1,11], Hiromi Kamura[1], Kei Miyakoshi[2], Mamoru Tanaka[3], Seiichi Morokuma[4], Kotaro Fukushima[4], Tomihiro Shimamoto[5], Shigeru Saito[6], Haruhiko Sago[7], Keiichiro Nishihama[8], Kosei Abe[1], Kazuhiko Nakabayashi[1], Akihiro Umezawa[9], Kohji Okamura[10] and Kenichiro Hata[1]

With increasing public concern about infertility and the frequent involvement of chromosomal anomalies in miscarriage, analyses of copy number variations (CNVs) have been used to identify the genomic regions responsible for each process of childbearing. Although associations between CNVs and diseases have been reported, many CNVs have also been identified in healthy individuals. Like other types of mutations, phenotypically indefinite CNVs may have been retained and accumulated during anthropogenesis. Therefore to distinguish causative variants from other variants is a formidable task. Furthermore, because previous studies have predominantly focused on European and African populations, comprehensive detection of common Asian CNVs is eagerly awaited. Here, using a high-resolution genotyping array and samples from 411 Japanese women with normal parity without significant complications, we have compiled 1043 copy number variable regions. In total, the collected regions cover 164 Mb, or up to 0.5% of the genome. The copy number differences in these regions may be irrelevant not only to infertility but also to a wide range of diseases. The utility of this resource in reducing the candidate pathogenetic variants, especially in Japanese subjects, is also demonstrated.
Journal of Human Genetics (2014) **59**, 326–331; doi:10.1038/jhg.2014.27; published online 1 May 2014

INTRODUCTION

The advent of new technologies has allowed the identification of structural variants that have a more significant impact on human diversity than does the entire set of single-nucleotide polymorphisms (SNPs). Copy number variations (CNVs) are one such type of structural variant and constitute the largest proportion of genomic variations.[1–3] CNVs result from the duplication or deletion of a DNA segment and are commonly observed in human genomes.[4–7] When a genomic event results in a CNV, not only the copy number of a gene can be altered but also its genic sequences. Therefore, CNVs can cause disease or contribute to disease susceptibility,[8–10] and they have been compiled in several databases for public use.[9–11]

Although a number of deleterious changes may have been negatively selected during human evolution, it is likely that phenotypically neutral changes have been retained, transmitted and accumulated over generations. Increasing numbers of CNVs are found in phenotypically normal human individuals.[1] Accordingly, each ethnic group tends to have distinct features in terms of the positions, copy numbers and frequencies of their CNVs, and it is possible that fixed CNVs have contributed to ethnic differences in phenotypic variations and disease susceptibility.[12–15] Therefore, it is important to have a list of CNVs for each ethnic group, especially for medical purposes. However, the number of reported CNVs from Asian populations is small compared with those of Europeans and Africans. Extensive examination of Asian CNVs is eagerly awaited by Asian researchers.

The compilation of nonpathogenic variations, in addition to disease-related variations, is also important for a better understanding of the genetic landscape of the human genome. Data sets including both these sorts of variations should be helpful in pinpointing causative mutations. Even when we search for variations using patient samples, most of the variations identified would be normal polymorphisms, together with a few pathogenic mutations. Although we can consider most of the available variation data nonpathogenic, it is difficult to know which variations are pathogenic. Therefore, the collection of data from normal controls is essential. To investigate

[1]Department of Maternal–Fetal Biology, National Research Institute for Child Health and Development, Tokyo, Japan; [2]Department of Obstetrics and Gynecology, School of Medicine, Keio University, Tokyo, Japan; [3]Department of Obstetrics and Gynecology, St Marianna University School of Medicine, Kanagawa, Japan; [4]Department of Obstetrics and Gynecology, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; [5]Department of Obstetrics and Gynecology, Miyazaki Prefectural Miyazaki Hospital, Miyazaki, Japan; [6]Department of Obstetrics and Gynecology, University of Toyama, Toyama, Japan; [7]Department of Maternal–Fetal and Neonatal Medicine, National Center for Child Health and Development, Tokyo, Japan; [8]Illumina KK, Tokyo, Japan; [9]Department of Reproductive Biology and Pathology, National Research Institute for Child Health and Development, Tokyo, Japan and [10]Department of Systems BioMedicine, National Research Institute for Child Health and Development, Tokyo, Japan
[11]These authors contributed equally to this work.
Correspondence: Dr K Okamura or Dr K Hata, Department of Maternal–Fetal Biology, National Research Institute for Child Health and Development, 2-10-1 Okura, Setagaya Ward, Tokyo 157-8535, Japan.
E-mail: okamura-k@ncchd.go.jp or hata-k@ncchd.go.jp
Received 27 September 2013; revised 21 March 2014; accepted 26 March 2014; published online 1 May 2014

phenotypically 'normal' samples in this study, we considered reproduction and child development, and chose parous Japanese women, who had experienced normal pregnancies and deliveries.

Although the origin of the Japanese population remains controversial, the last major migration to the Japanese Archipelago is thought to have occurred approximately 2000 years ago.[16,17] The population has been mixed well with various Asian ethnic groups during previous migrations, but has remained relatively isolated for 2000 years. However, although the current population of Japan is 127 million, far fewer CNVs have been documented in Japanese samples than in Europeans. Compiling a list of Japanese CNVs is also important from the perspective of medical science in Japan.

MATERIALS AND METHODS

Subject recruitment and SNP genotyping with a high-resolution microarray

We examined 411 unrelated Japanese women who had had one or more normal parities, with no significant abnormalities in any of their pregnancies, deliveries or neonates. Ethical approval was also obtained from each review board of the hospitals that participated in the study. The informed consent of all the subjects was obtained. To avoid cell-culture-induced chromosomal rearrangements, genomic DNAs were extracted directly from blood using the QIAsymphony DNA Midi Kit (Qiagen, Venlo, The Netherlands) with the QIAsymphony SP instrument and analyzed with a high-resolution SNP-based genotyping microarray, HumanOmni2.5-8 BeadChip (Illumina, San Diego, CA, USA). Only data that met the quality control guidelines of the manufacturer were used for further analyses.

Identification of CNVs and CNVRs

Two distinct algorithms were used to maximize the specificity of our CNV calling: a likelihood-based method with CNVPartition version 3.2.0 (http://www.illumina.com/software/illumina_connect.ilmn) and a hidden Malkov method with PennCNV version (27 August 2009).[18] The parameters applied with these tools were referred to those typically used by many research groups (at least three consecutive probes to define a CNV, using the GC wave adjustment option, etc.). These programs computed confidence scores that can be used to filter out CNV regions that are likely to be false positives. However, we should note that the two programs calculated the scores in different ways, with different scales. To minimize false positives, we first chose only CNVs with high confidence scores; that is, more than 100 with CNVPartition, and selected copy number variable regions (CNVRs) that overlapped those called by PennCNV for at least 80% of their lengths. For PennCNV, we generated a list of B allele frequencies using a collection of signal intensities for 47 samples from HapMap Japanese in Tokyo with the compile_pfb script (Figure 1a).

Multiplex PCR assay

Multiplex polymerase chain reaction (PCR) assay was used to confirm regions that had been called homozygously deleted. The reactions were performed with both a control primer pair that generated a 296-bp fragment and a test primer pair that amplified a target region. The thermal cycling conditions were initial denaturation at 95 °C for 2 min, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at 60 °C for 30 s and extension at 72 °C for 30 s, and a final extension at 72 °C for 3 min. Detailed information on these primers is given in Supplementary Table S3.

RESULTS

Genetic ancestry of the subjects

First, the population structure was inferred with the Structure software (http://pritchardlab.stanford.edu/structure.html) to confirm the Japanese ancestry of the subjects.[19] A cluster analysis of our samples together with the sequences of 499 HapMap individuals from three ancestral populations (European, African and Asian) was performed using 1959 unlinked tag SNPs on chromosome 21. The expected ancestry of all the subjects was confirmed with a minimum coefficient of 0.85. We also performed a principal components analysis with the pca.jar program (Biobank Japan project; http://genome-analysis.src.riken.jp/PCP/). The results indicated that all but one subject were derived from the main islands of Japan and that the remaining singleton was Ryukyuan.[20]

Characterization of CNVs and CNVRs

The CNVPartition software (Illumina) identified 26 150 candidate regions as CNVs. We then used another program, PennCNV,[18] which is based on an integrated hidden Markov algorithm, to maximize the specificity of the analysis. If a candidate CNV was also supported by PennCNV for at least 80% of its length, it was retained. In this way,

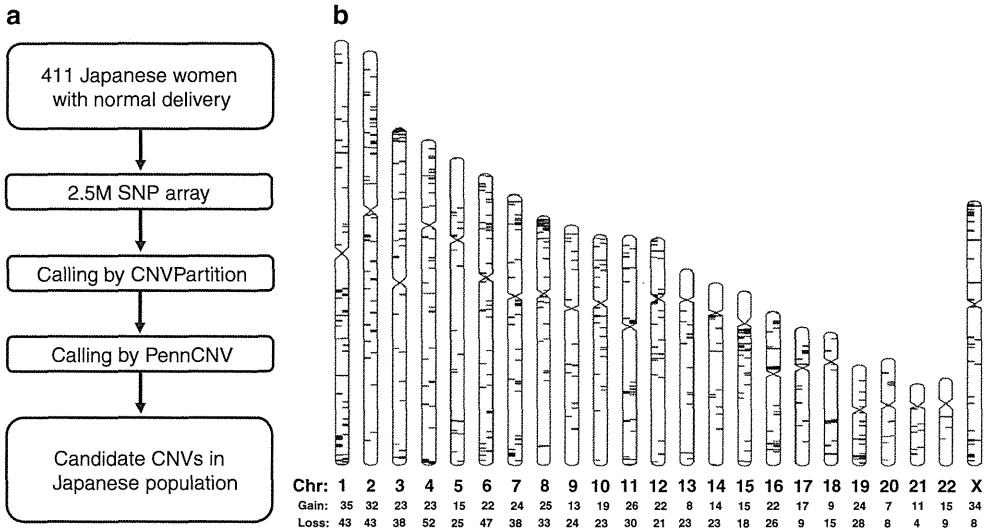


Figure 1 (**a**) Data processing flow. The initial 26 150 regions identified with CNVPartition were validated with PennCNV. (**b**) Chromosomal distribution of the CNVRs. Each CNVR is shown by a horizontal bar. Gain- and loss-type CNVs are distinguished by bars on the left and right, respectively. The numbers of each type of CNV are also shown, and are drawn with Idiographica (http://www.ncrna.org/idiographica/). CNV, copy number variation; CNVR, copy number variable regions.

Table 1 Comparison of the CNVRs with those reported in other studies and in the DGVs

	Present study	*McCarroll et al.*[21]	*Conrad et al.*[22]	*Koike et al.*[23]	*DGV Jul. 2013*
CNVs reported	1043	592	1768	169	202 430
CNVs spanning our data	1043	88 (71/1043)[a]	156 (112/1043)[a]	37 (45/1043)[a]	30 322 (1033/1043)[a]
Number of samples	411 Japanese females	45 HapMap JPT	45 HapMap JPT	57 Japanese females and 123 Japanese males	Collective (including non-Japanese samples)
Experimental method	SNP array (Illumina HumanOmni2.5-8 BeadChip)	SNP array (Affymetrics Genome-Wide Human SNP Array 6.0)	Custom CGH array (NimbleGen and Agilent)	SNP array (Affymetrics Genome-Wide Human SNP Array 6.0)	N/A
CNV calling	CNVPartition and then PennCNV	Birdseye and custom program	Custom program	PennCNV	N/A

Abbreviations: CNV, copy number variations; CNVR, copy number variable region; DGV, database of genomic variants; JPT, Japanese in Tokyo; N/A, not available; SNP, single-nucleotide polymorphism.
[a]The number of CNVRs overlapped with those in the present study is indicated within parentheses.

Table 2 CNVRs overlapping between the Japanese and other populations

Population	*Sample size*	*Reported CNVRs*	*Frequency of overlapping regions among studies*[a]
Japanese (present study)	411	1043	—
Korean[24]	100	576	10% (106/1043)
Tibetan[14]	29	139	4.9% (51/1043)
Chinese[13] (Han, Tibetan and five other ethnic group)	155	1440	17% (173/1043)
Han Chinese[13]	80	1407	17% (175/1043)
Swiss[25]	717	917	16% (163/1043)
Rwandan[25] (sub-Saharan African)	450	1185	14% (141/1043)
HapMap[26] (mixed)	112	3262	13% (134/1043)

Abbreviation: CNVR, copy number variable regions.
[a]Number of overlapped CNVRs is indicated within parentheses.

Table 3 List of genes lying within a homozygously deleted region

No.	*Coordination*	*Frequency*	*Suffered gene*		
1	Chr 1: 161 570 803–161 644 281*	2/411	*FCGR3B*	*FCGR2B*	
2	Chr 2: 111 884 593–111 886 246*	3/411	*BCL2L11*		
3	Chr 4: 69 367 146–69 489 473*	302/411	*UGT2B17*		
4	Chr 5: 180 377 470–180 424 820*	32/411	*BTNL3*		
5	Chr 6: 32 551 892–32 555 728	2/411	*HLA-DRB1*		
6	Chr 7: 115 584 568–115 593 688*	1/411	*TFEC*		
7	Chr 7: 141 761 027–141 795 404*	6/411	*MGAM*		
8	Chr 11: 18 949 220–18 961 743	1/411	*MRGPRX1*		
9	Chr 19: 41 350 895–41 379 321*	10/411	*CYP2A6*		
10	Chr 19: 43 590 229–43 772 302	86/411	*PSG5*	*PSG4*	*PSG9*
11	Chr 19: 46 622 776–46 636 139*	3/411	*IGFL3*		
12	Chr 19: 52 132 392–52 150 601*	114/411	*SIGLEC5*		

Abbreviation: PCR, polymerase chain reaction.
Asterisk indicates a homozygously deleted region validated by PCR.

we identified 6871 CNVs and 1043 regions with variable copy numbers from 411 Japanese individuals, with an average of 16.7 CNVs per diploid genome (Supplementary Table S1). Detailed information on all the SNP probes used for the CNV calls is tabulated (Supplementary Table S2). The mean length of the CNVs was 79.9 kb, ranging from 169 bases to 2.27 Mb. These 6871 CNVs corresponded to 1043 CNVRs (588 losses and 455 gains). Figure 1 shows the chromosomal distribution of the observed CNVRs. The total length of all of these CNVRs was 163 720 kb, which is equivalent to 0.5% of the whole human genome. The CNVRs can be divided into gain regions and loss regions, depending on whether their copy numbers have increased or decreased. Of the 1043 regions identified, 1033 overlap the latest database of genomic variants (DGVs) (released on 23 July 2013) reported at the DGV. More than half the CNVRs, including 72% of the gain CNVRs and 36% of the loss CNVRs, intersect RefSeq gene loci.

As far as we know, three studies have examined the Japanese population with array-based methods: two of them used samples from HapMap and the other used healthy individuals.[21–23] These results are summarized with our data set (Table 1). Although those three studies had together already reported 82 regions, more than half the regions reported in the present study were not detected by them. It is probable that the higher resolution of our analysis and our larger sample size allowed us to detect additional CNVRs. Depending on the types of platform used, array-based CNV studies occasionally show discrepancies in the regions of CNVs.[21,22] Differences in the array architectures, scanning machines and calling algorithms could affect the final data sets. Using reported CNV data from SNP arrays, we counted the overlapping regions among studies that focused on other populations or HapMap data[13,14,24–26] (Table 2). The similarities among these studies are comparable, but our results suggest a greater similarity between the Japanese and Chinese populations.

Homozygous deletions found in parous Japanese women

In our study, 1628 homozygous deletions that could affect 112 RefSeq gene loci were called in a total of 822 chromosomes. Although the CNV analysis was unable to determine the precise breakpoints, our data indicate that some exonic sequences are disrupted by homozygous deletions (Table 3). Using multiplex PCR with both control and test primer pairs, we confirmed the null genotypes caused by deletions (Supplementary Figure S1 and Supplementary Table S3). Five genes, *FCGR3B*, *FCGR2B*, *UGT2B17*, *HLA-DRB1* and *CYP2A6*, are described as disease related in the OMIM database. The *FCGR3*, *FCGR2B* and *HLA-DRB1* genes have roles in the immune system. *FCGR3B* and *FCGR2B* encode the crystallizable region of immunoglobulin G. Several studies have shown that a low copy number at the *FCGR3B–FCGR2B* locus is associated with a susceptibility to systemic lupus erythematosus in the Caucasian population,[27–29] but not in the Chinese population.[28] *UGT2B17* encodes a protein that belongs to the family of UDP-glucuronosyltransferases enzymes, which catalyzes the glucuronidation of steroid hormones. A case–control study of

osteoporosis-related fracture suggested that a CNV at the *UGT2B17* locus contributes to osteoporosis.[30] Jakobsson *et al.*[31] found that its null genotype was more common in Koreans (67%) than in Swedish (9%). Our array results also showed a high frequency (74%) of the null genotype. The CYP2A6 protein metabolizes nicotine and coumarin in the liver. The lack of a *CYP2A6* gene may affect nicotine levels in individuals and probably has a protective effect against tobacco dependence.[32] Another study reported that the frequency of homozygotes for the *CYP2A6* gene deletion was lower in Japanese lung cancer patients than in control samples.[33] Except for *HLA-DRB1*, these disease-related genes have been reported to be frequently deleted in Asian populations.[25,34–36] Because we limited our samples to parous women only, it is unlikely that the CNVRs identified in the present study are related to human reproduction.

DISCUSSION

In the present study, we compiled a catalog of copy number variable regions identified in phenotypically normal Japanese samples, especially those with a history of full-term pregnancy and deliveries without major complications. The data set will be useful in the search for novel or rare CNVs that increase the individual's susceptibility to congenital diseases and complications during pregnancy. It is unlikely that the newly identified CNVs are related to infertility or miscarriage. CNVs in parous women without complications have never before

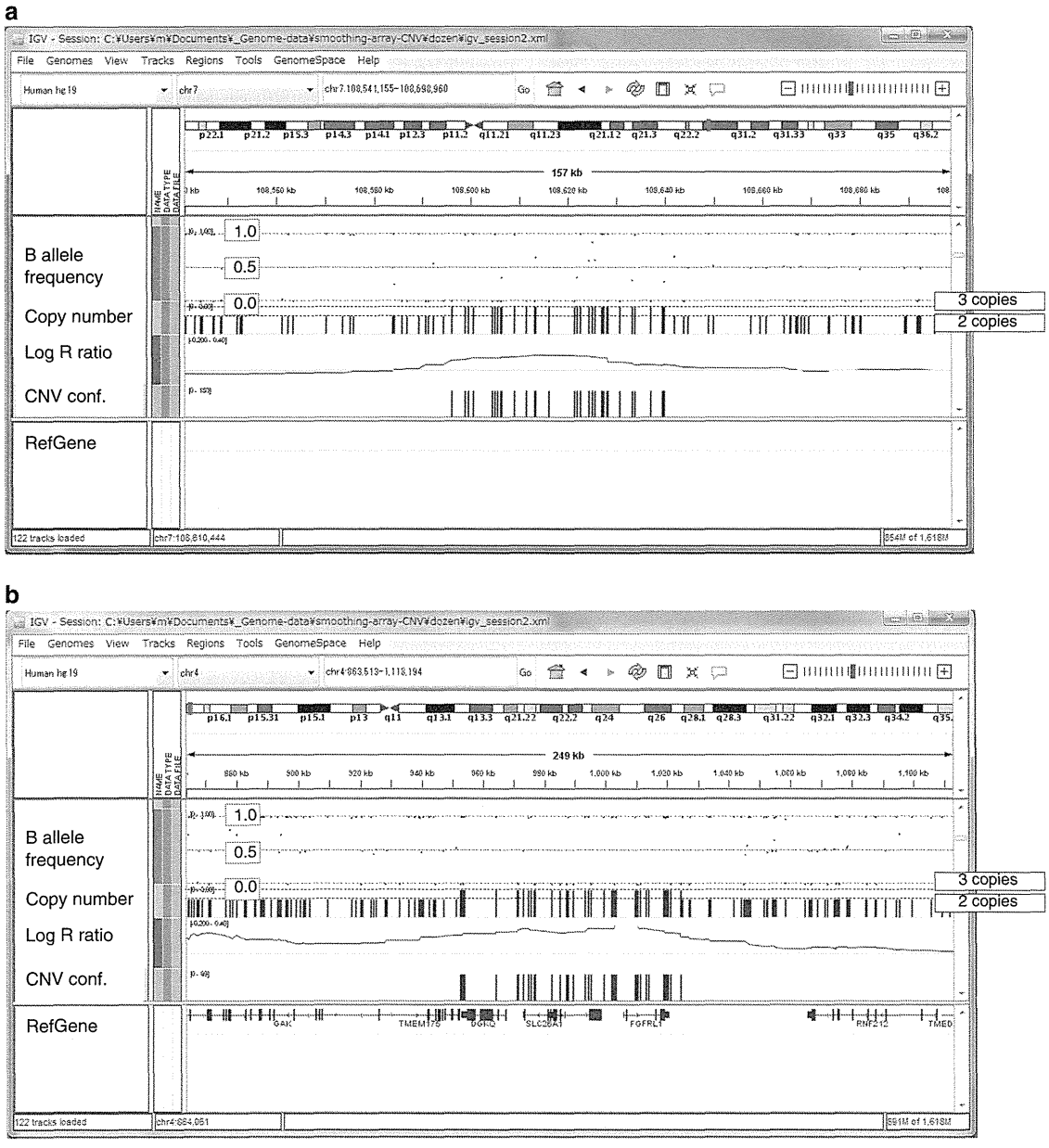


Figure 2 (**a**) A copy number variation (CNV) located on chromosome 7. The panel shows the region at nucleotides 108 541 155–108 698 960 in hg19. This CNV was called with high-intensity probes. The B allele frequencies (BAFs) were separated into four levels, which corresponded to AAA, AAB, ABB and BBB, respectively. (**b**) Another CNV located in the subtelomeric region on chromosome 4. The panel shows the region at nucleotides 863 513–1 113 194. Despite high-intensity probes used, as in the example shown above, the four levels of BAFs were not observed, suggesting that the call might be implausible. Such CNVs tended to be called in G + C-rich regions; for example, 58% G + C content in this case. The snapshot was made with the IGV program (http://www.broadinstitute.org/igv/). A full color version of this figure is available at *Journal of Human Genetics* online.

been investigated. Although the copy numbers of these regions were not thoroughly validated with other methods; such as, quantitative PCR, according to DGV, most of the CNVRs identified here have been reported in previous studies, indicating that they should be observed by other methods or techniques. Because our identification strategy was based on a microarray technique, it is inevitable that errors would have occurred. Besides routine data processing, we also carefully curated the data by examining the B allele frequencies and signal intensities ($\log R$ ratio) for each CNVR using the GenomeStudio software (Illumina) (Figure 2). We found that many implausible calls were situated in regions with high G + C contents; for example, in subtelomeric regions. All of them were copy number gain-type CNVs rather than copy number loss-type CNVs. Although further research is required, it is important to note that CNVRs tend to be detected in those regions by SNP microarrays. Even if such CNVRs are false positives, our data set is still useful for screening large numbers of candidate CNVs.

It is unclear whether CNVs are selectively neutral on the basis of genetic drift, but they are certainly distributed throughout all human populations. Using the genotypes of mitochondrial DNA and Y chromosome, geneticists and anthropologists have surmised various intriguing scenarios about the history of humans.[37–40] However, these genetic materials have been transmitted exclusively through maternal and paternal lineages, respectively. In contrast, the CNVs reported here occur in the more extensive remaining genome regions; that is, on autosomes or the X chromosome. Therefore, they have acted some times as maternal alleles at and at other times as paternal alleles. They might also have been subjected to crossingover. CNV data from various parts of the world are essential to substantiate these hypothetical scenarios.

Chromosomal anomalies are found with conventional cytogenetic techniques in approximately half of all early sporadic miscarriages.[41] It is possible that miscarriages and pregnancy losses are also caused by submicroscopic chromosomal changes, including CNVs. Twenty-eight CNVs have been reported as candidate miscarriage-related variations when instances of recurrent pregnancy loss were examined by Rajcan-Separovic *et al.*[42] When 17 Caucasian and three African-American couples with recurrent pregnancy losses and their miscarriage samples were examined, CNVs that may have been related to miscarriages were reported.[42] They reported 11 novel CNVs in miscarriage samples and three in the parent samples and suggested that these CNVs were probably mutations causing susceptibility to miscarriage. Of the 11 CNVs in the miscarriage samples, one on chromosome 12 (130 060 706–130 430 847 in hg18) and another one on chromosome X (6 498 521–8 091 951) overlapped with our data set. Whereas the first one on chromosome 12 was up to 370 kb in length and encompassed the *GPR133* gene, the corresponding variable region in our data set is much shorter and includes no known genes. The *GPR133* gene encodes one of the orphan G-protein-coupled receptors, but its function is unknown.[43] It is possible that this receptor protein has a role in several signal-transduction pathways via classical receptor/G-protein interactions. Therefore, the CNV mentioned above may be a variant that causes miscarriage. However, one of the CNVs on chromosome X is consistent with our data set, suggesting that it is a commonly observed variant. In fact, Rajcan-Separovic *et al.*[42] tried to define the common CNVs using a collective repository in the DGV, but insufficient phenotypic information was available to refine the data. Taking these observations together, it seems that to define a set of common CNVs, it will be necessary to collect a large number of control data that focus on a specific phenotype; such as, normal parity in this case.

The Japanese are an admixture of ancient Asian populations that inhabited regions outside the Japanese Archipelago. We investigated the similarities among the CNVRs detected in various populations and noted that around 15% of Japanese CNVRs overlap those of other populations (Table 2). It has been suggested that the number of overlapping CNVs is influenced by the number of subjects. For instance, Japanese and Tibetan data showed dissimilarity because of the limited number of Tibetan subjects. Although the sample sizes of the Korean and Chinese populations are smaller than those of the European and African populations, similarities between the Japanese and other East Asian populations were similar to those of the European and African populations. This probably suggests strong similarities between the Japanese and other East Asian populations.

Previous studies have predominantly targeted European and African populations, but CNVs have been observed at different frequencies or copy numbers in different populations; for example, variations in the salivary amylase gene.[44] Many CNVs; such as, those at the *AMY1* locus, may be associated with diabetes, asthma, hypertension, allergy and other diseases of affluence in each ethnic group. Although CNVRs may result from the accumulation of tolerable structural mutations in the course of an ethnic history, they could start to influence the population's susceptibility to disease once its lifestyle is altered. The allelic frequencies of SNPs and short indels in each population have recently been documented.[45] The complete documentation of the CNVRs in each ethnic group is similarly important. The development of an innovative method to achieve this; such as, one involving next-generation sequencing and informatics, is another challenge.

CONFLICT OF INTEREST

The authors received no financial support from Illumina KK and the company had no role in the study design. The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to all the participants in the present study, including the 411 women. Computation time was partly provided by the supercomputer system, Shirokane, at the Human Genome Centre, Institute of Medical Science, University of Tokyo. This work was supported by CREST Program 'Epigenomic analysis of the human placenta and endometrium constituting the fetal-maternal interface' of Japan Science and Technology Agency (JST) and Health and Labor Sciences Research Grants for Research into Rare and Intractable Diseases (H23 Jitsuyoka (Nanbyo)-Ippan-003 and H25 Jisedai-Ippan-001), and was also partly supported by KAKENHI 23770273, 24657151, 24390251, 24592494, 24659742, 25293345 and 25860258.

1 Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D. *et al.* Global variation in copy number in the human genome. *Nature* **444,** 444–454 (2006).
2 Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315,** 848–853 (2007).
3 Pang, A. W., Migita, O., Macdonald, J. R., Feuk, L. & Scherer, S. W. Mechanisms of formation of structural variation in a fully sequenced human genome. *Hum. Mut.* **34,** 345–354 (2013).
4 Frazer, K. A., Chen, X., Hinds, D. A., Pant, P. V., Patil, N. & Cox, D. R. Genomic DNA insertions and deletions occur frequently between humans and nonhuman primates. *Genome Res.* **13,** 341–346 (2003).
5 Locke, D. P., Segraves, R., Carbone, L., Archidiacono, N., Albertson, D. G., Pinkel, D. *et al.* Large-scale variation among human and great ape genomes determined by array comparative genomic hybridization. *Genome Res.* **13,** 347–357 (2003).
6 Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P. *et al.* Large-scale copy number polymorphism in the human genome. *Science* **305,** 525–528 (2004).
7 Sharp, A. J., Locke, D. P., McGrath, S. D., Cheng, Z., Bailey, J. A., Vallente, R. U. *et al.* Segmental duplications and copy-number variation in the human genome. *Am. J. Hum. Genet.* **77,** 78–88 (2005).

8 Marshall, C. R. & Scherer, S. W. Detection and characterization of copy number variation in autism spectrum disorder. *Methods Mol. Biol.* **838,** 115–135 (2012).
9 Swaminathan, G. J., Bragin, E., Chatzimichali, E. A., Corpas, M., Bevan, A. P., Wright, C. F. *et al.* DECIPHER: web-based, community resource for clinical interpretation of rare variants in developmental disorders. *Hum. Mol. Genet.* **21,** R37–R44 (2012).
10 Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D. *et al.* DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* **84,** 524–533 (2009).
11 Macdonald, J. R., Ziman, R., Yuen, R. K., Feuk, L. & Scherer, S. W. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* **42,** D986–D992 (2014).
12 Li, J., Yang, T., Wang, L., Yan, H., Zhang, Y., Guo, Y. *et al.* Whole genome distribution and ethnic differentiation of copy number variation in Caucasian and Asian populations. *PLoS ONE* **4,** e7958 (2009).
13 Lou, H., Li, S., Yang, Y., Kang, L., Zhang, X., Jin, W. *et al.* A map of copy number variations in Chinese populations. *PLoS ONE* **6,** e27341 (2011).
14 Zhang, Y. B., Li, X., Zhang, F., Wang, D. M. & Yu, J. A preliminary study of copy number variation in Tibetans. *PLoS ONE* **7,** e41768 (2012).
15 Kanduri, C., Ukkola-Vuoti, L., Oikkonen, J., Buck, G., Blancher, C., Raijas, P. *et al.* The genome-wide landscape of copy number variations in the MUSGEN study provides evidence for a founder effect in the isolated Finnish population. *Eur. J. Hum. Genet.* **21,** 1411–1516 (2013).
16 Hanihara, K. Dual structure model for the population history of the Japanese. *Jpn Rev.* **2,** 1–33 (1991).
17 Japanese Archipelago Human Population Genetics C., Jinam, T., Nishida, N., Hirai, M., Kawamura, S., Oota, H. *et al.* The history of human populations in the Japanese Archipelago inferred from genome-wide SNP data with a special reference to the Ainu and the Ryukyuan populations. *J. Hum. Genet.* **57,** 787–795 (2012).
18 Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S. F. *et al.* PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17,** 1665–1674 (2007).
19 Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155,** 945–959 (2000).
20 Kumasaka, N., Yamaguchi-Kabata, Y., Takahashi, A., Kubo, M., Nakamura, Y. & Kamatani, N. Establishment of a standardized system to perform population structure analyses with limited sample size or with different sets of SNP genotypes. *J. Hum. Genet.* **55,** 525–533 (2010).
21 McCarroll, S. A., Kuruvilla, F. G., Korn, J. M., Cawley, S., Nemesh, J., Wysoker, A. *et al.* Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat. Genet.* **40,** 1166–1174 (2008).
22 Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y. *et al.* Origins and functional impact of copy number variation in the human genome. *Nature* **464,** 704–712 (2010).
23 Koike, A., Nishida, N., Yamashita, D. & Tokunaga, K. Comparative analysis of copy number variation detection methods and database construction. *BMC Genet.* **12,** 29 (2011).
24 Moon, S., Kim, Y. J., Hong, C. B., Kim, D. J., Lee, J. Y. & Kim, B. J. Data-driven approach to detect common copy-number variations and frequency profiles in a population-based Korean cohort. *Eur. J. Hum. Genet.* **19,** 1167–1172 (2011).
25 Vogler, C., Gschwind, L., Rothlisberger, B., Huber, A., Filges, I., Miny, P. *et al.* Microarray-based maps of copy-number variant regions in European and sub-Saharan populations. *PLoS ONE* **5,** e15246 (2010).
26 Shaikh, T. H., Gai, X., Perin, J. C., Glessner, J. T., Xie, H., Murphy, K. *et al.* High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res.* **19,** 1682–1690 (2009).
27 Fanciulli, M., Norsworthy, P. J., Petretto, E., Dong, R., Harper, L., Kamesh, L. *et al.* FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. *Nat. Genet.* **39,** 721–723 (2007).
28 Willcocks, L. C., Lyons, P. A., Clatworthy, M. R., Robinson, J. I., Yang, W., Newland, S. A. *et al.* Copy number of FCGR3B, which is associated with systemic lupus erythematosus, correlates with protein expression and immune complex uptake. *J. Exp. Med.* **205,** 1573–1582 (2008).
29 McKinney, C. & Merriman, T. R. Meta-analysis confirms a role for deletion in FCGR3B in autoimmune phenotypes. *Hum. Mol. Genet.* **21,** 2370–2376 (2012).
30 Yang, T. L., Chen, X. D., Guo, Y., Lei, S. F., Wang, J. T., Zhou, Q. *et al.* Genome-wide copy-number-variation study identified a susceptibility gene, UGT2B17, for osteoporosis. *Am. J. Hum. Genet.* **83,** 663–674 (2008).
31 Jakobsson, J., Ekström, L., Inotsume, N., Garle, M., Lorentzon, M., Ohlsson, C. *et al.* Large differences in testosterone excretion in Korean and Swedish men are strongly associated with a UDP-glucuronosyl transferase 2B17 polymorphism. *J. Clin. Endocrinol. Metab.* **91,** 687–693 (2006).
32 Pianezza, M. L., Sellers, E. M. & Tyndale, R. F. Nicotine metabolism defect reduces smoking. *Nature* **393,** 750 (1998).
33 Miyamoto, M., Umetsu, Y., Dosaka-Akita, H., Sawamura, Y., Yokota, J., Kunitoh, H. *et al.* CYP2A6 gene deletion reduces susceptibility to lung cancer. *Biochem. Biophys. Res. Commun.* **261,** 658–660 (1999).
34 Lv, J., Yang, Y., Zhou, X., Yu, L., Li, R., Hou, P. *et al.* FCGR3B copy number variation is not associated with lupus nephritis in a Chinese population. *Lupus* **19,** 158–161 (2010).
35 Xue, Y., Sun, D., Daly, A., Yang, F., Zhou, X., Zhao, M. *et al.* Adaptive evolution of UGT2B17 copy-number variation. *Am. J. Hum. Genet.* **83,** 337–346 (2008).
36 Oscarsona, M., McLellana, R. A., Gullsteènb, H., Yuec, Q. Y., Langd, M. A., Bernale, M. L. *et al.* Characterisation and PCR-based detection of a CYP2A6 gene deletion found at a high frequency in a Chinese population. *FEBS Lett.* **448,** 105–110 (1999).
37 Hammer, M. F. & Horai, S. Y. chromosomal DNA variation and the peopling of Japan. *Am. J. Hum. Genet.* **56,** 951–962 (1995).
38 Shinka, T., Tomita, K., Toda, T., Kotliarova, S. E., Lee, J., Kuroki, Y. *et al.* Genetic variations on the Y chromosome in the Japanese population and implications for modern human Y chromosome lineage. *J. Hum. Genet.* **44,** 240–245 (1999).
39 Dulik, M. C., Zhadanov, S. I., Osipova, L. P., Askapuli, A., Gau, L., Gokcumen, O. *et al.* Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between Native Americans and Indigenous Altaians. *Am. J. Hum. Genet.* **90,** 229–246 (2012).
40 Oppenheimer, S. Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. *Philos. Trans. R. Soc. Lond. Ser. B* **367,** 770–784 (2012).
41 van den Berg, M., van Maarle, M., van Wely, M. & Goddijn, M. Genetics of early miscarriage. *Biochim. Biophys. Acta* **1822,** 1951–1959 (2012).
42 Rajcan-Separovic, E., Diego-Alvarez, D., Robinson, W. P., Tyson, C., Qiao, Y., Harvard, C. *et al.* Identification of copy number variants in miscarriages from couples with idiopathic recurrent pregnancy loss. *Hum. Reprod.* **25,** 2913–2922 (2010).
43 Bohnekamp, J. & Schoneberg, T. Cell adhesion receptor GPR133 couples to Gs protein. *J. Biol. Chem.* **286,** 41912–41916 (2011).
44 Perry, G. H., Dominy, N. J., Claw, K. G., Lee, A. S., Fiegler, H., Redon, R. *et al.* Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* **39,** 1256–1260 (2007).
45 1000 Genomes Project Consortium, Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M. *et al.* An integrated map of genetic variation from 1092 human genomes. *Nature* **491,** 56–65 (2012).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

Regenerative Therapy 1 (2015) 30–37

Contents lists available at ScienceDirect

JSRM

Regenerative Therapy

journal homepage: http://www.elsevier.com/locate/reth

Original article

Characterization of *in vivo* tumorigenicity tests using severe immunodeficient NOD/Shi-scid IL2Rγ^{null} mice for detection of tumorigenic cellular impurities in human cell-processed therapeutic products

Shinji Kusakawa [a, b], Kazuhiko Machida [c], Satoshi Yasuda [a, b], Nozomi Takada [a, d, 1], Takuya Kuroda [a, b], Rumi Sawada [a], Hanayuki Okura [d, 1], Hideki Tsutsumi [c], Shin Kawamata [b], Yoji Sato [a, b, e, f, g, *]

[a] *Division of Cell-Based Therapeutic Products, National Institute of Health Sciences, Tokyo, Japan*
[b] *Foundation for Biomedical Research and Innovation, Kobe, Japan*
[c] *Testing Department, Central Institute for Experimental Animals, Kawasaki, Japan*
[d] *Platform for Realization of Regenerative Medicine, Foundation for Biomedical Research and Innovation, Kobe, Japan*
[e] *Department of Quality Assurance Science for Pharmaceuticals, Graduate School of Pharmaceutical Sciences, Nagoya City University, Nagoya, Japan*
[f] *Department of Cellular and Gene Therapy Products, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, Japan*
[g] *Department of Translational Pharmaceutical Sciences, Graduate School of Pharmaceutical Sciences, Kyushu University, Fukuoka, Japan*

ARTICLE INFO

Article history:
Received 11 September 2014
Received in revised form
14 December 2014
Accepted 28 December 2014

Keywords:
Tumorigenicity test
NOG mice
Cellular therapy
Regenerative medicine
Quality control

ABSTRACT

The contamination of human cell-processed therapeutic products (hCTPs) with tumorigenic cells is one of the major concerns in the manufacturing and quality control of hCTPs. However, no quantitative method for detecting the tumorigenic cellular impurities is currently standardized. NOD/Shi-scid IL2Rγ^{null} (NOG) mice have shown high xeno-engraftment potential compared with other well-known immunodeficient strains, e.g. nude mice. Hypothesizing that tumorigenicity test using NOG mice could be a sensitive and quantitative method to detect a small amount of tumorigenic cells in hCTPs, we examined tumor formation after subcutaneous transplantation of HeLa cells, as a model of tumorigenic cells, in NOG mice and nude mice. Sixteen weeks after inoculation, the 50% tumor-producing dose (TPD_{50}) values of HeLa cells were stable at 1.3×10^4 and 4.0×10^5 cells in NOG and nude mice, respectively, indicating a 30-fold higher sensitivity of NOG mice compared to that of nude mice. Transplanting HeLa cells embedded with Matrigel in NOG mice further decreased the TPD_{50} value to 7.9×10 cells, leading to a 5000-fold higher sensitivity, compared with that of nude mice. Additionally, when HeLa cells were mixed with 10^6 or 10^7 human mesenchymal stem cells as well as Matrigel, the TPD_{50} values in NOG mice were comparable to those of HeLa cells alone with Matrigel. These results suggest that the *in vivo* tumorigenicity test using NOG mice with Matrigel is a highly sensitive and quantitative method to detect a trace amount of tumorigenic cellular impurities in human somatic cells, which can be useful in the quality assessment of hCTPs.

1. Introduction

Cell-processed therapeutic products (CTPs) derived from human somatic/stem cells are eagerly expected to treat patients with severe diseases involving functional damage of organs and tissues. To transplant hCTPs into patients, however, tumorigenicity is raised as one of the issues of these products. Tumorigenicity is defined as the capacity of a cell population transplanted into an animal model to

* Corresponding author. 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Tel./fax: +81 3 3700 9373.
E-mail address: yoji@nihs.go.jp (Y. Sato).

Peer review under responsibility of the Japanese Society for Regenerative Medicine.

[1] Present address: Research on Disease Bioresources, Platform of Therapeutics for Rare Disease and Health Policy, National Institute of Biomedical Innovation, Kobe, Japan.

http://dx.doi.org/10.1016/j.reth.2014.12.001

produce a tumor by proliferation at the site of transplantation and/or at a distant site by metastasis [1]. Assessment of tumorigenicity is quite important to manufacture products with consistent quality. Currently, the World Health Organization (WHO) Technical Report Series (TRS) No. 878 Annex 1 is the only international guideline that addresses tumorigenicity tests of animal cells for the production of biologicals. However, the tumorigenicity test described in WHO TRS 878, which involves the administration of 10^7 cells to ten nude mice, would not be sensitive enough to detect a trace amount of tumorigenic cellular impurities in hCTPs [2]. In addition, the *in vivo* tumorigenicity test proposed in WHO TRS 878 covers only viable animal cells used as cell substrates for manufacturing biological products but not cells used directly for therapy by transplantation into patients. Thus, to date, no suitable tumorigenicity test has been established for hCTPs.

To establish methods to detect a trace amount of tumorigenic cellular impurities in hCTPs, the usage of several new generations of highly immunodeficient animal models are proposed. Rag2-γC double-knockout mice [3], NOD/Shi-scid IL2Rγ^{null} (NOG) mice [4], and NOD/SCID/IL-2rγKO (NSG) mice [5] indicate multiple immunodeficiencies, including defects in T, B, and natural killer (NK) cells, and a reduction in the function of macrophages and dendritic cells. NOG mice exhibit extremely high engraftment rates of human HeLa S3 cells compared with T-cell-deficient nude mice and T and B-cell-deficient SCID mice [6]. NSG mice are reported to show efficient tumor formation by single human melanoma cells in combination with Matrigel, a basement membrane-like extracellular matrix extract [7]. However, for the use of these highly immunodeficient mouse strains to detect tumorigenic cellular impurities in hCTPs as a part of the quality assessment/control, the performance of the tumorigenicity tests using these strains shall be validated using well known tumor cell lines.

In the present study, we examined the tumor formation potential of HeLa cells transplanted in NOG mice with Matrigel and compared their tumorigenicity with that in nude mice. To determine the sensitivity for the detection of tumor cells contamination in non-tumorigenic human somatic cells, we mixed various dose of HeLa cells in human mesenchymal stem cells and conducted tumorigenicity tests using NOG mice and Matrigel. We also performed soft agar colony formation assay, which is commonly used to detect anchorage-independent cell growth *in vitro*, and compared tumor cell detection level by soft agar with the *in vivo* tumorigenicity test.

2. Materials and methods

2.1. Cells

Human cervical cancer HeLa cells were obtained from the Health Science Research Resources Bank (HSRRB, Osaka, Japan). The cells were maintained in Eagle's minimum essential medium (Sigma), supplemented with 10% fetal bovine serum (FBS; Sigma), 0.1 mM non-essential amino acids (Life Technologies), 50 U/ml penicillin, and 50 μg/ml streptomycin (Life Technologies). Human mesenchymal stem cells (hMSCs) were purchased from Lonza and cultured in MSCGM™ medium (Lonza). Cells were cultured in a humidified atmosphere of 5% CO_2 and 95% air at 37 °C, and were passaged upon reaching 80% confluence. hMSCs were used at passage 6 and passages 6–8 for *in vivo* tumorigenicity tests and soft agar colony formation assay, respectively.

2.2. Preparation of cell suspensions for transplantation

Upon reaching approximately 80% confluence, cells were washed twice with phosphate buffered saline (PBS) and treated with 0.25% trypsin-EDTA solution (Life Technologies) for detachment from culture dishes. HeLa cells and/or hMSCs were counted and prepared in 100 μl of ice-cold HeLa cell culture medium or a 1:1 (v/v) mixture of HeLa cell culture medium and Matrigel (product #354234, BD Biosciences, San Jose, CA) for transplantation.

2.3. Tumorigenicity test with immunodeficient mice

Male BALB/cA nu/nu mice (nude; CLEA Japan, Inc., Tokyo) and male NOG mice maintained in the Central Institute for Experimental Animals (CIEA, Kanagawa, Japan) were used for *in vivo* tumorigenicity studies. Prepared cell suspensions were injected using 1 ml syringes with a 25 G needle (Terumo) into 8-week-old mice (n = 6 or 10). The mice were palpated weekly for 16 weeks to observe nodule formation at the injection site. Tumor size was assessed by external measurement of the length and width of the tumors in two dimensions using a caliper as soon as tumors reached measurable size. The tumor volume (TV) was calculated using the formula volume = 1/2 × length (mm) × (width [mm])^2. The successive engraftment was determined according to progressive nodule growth at the injection site. Mice were euthanized and necropsied when tumors reached approximately 20 mm in any dimension or when a sign of deconditioning was noted. The tumorigenicity of HeLa cells was evaluated by measuring tumor-forming capacity, which indicates the tumorigenic phenotype [8,9]. Tumor-forming capacity is defined as 50% tumor-producing dose (TPD_{50}), which represents the threshold dose of cells forming tumors in 50% of the animals. TPD_{50} values were calculated using the Spearman-Kärber method [10–12] at each time point. Not all animals transplanted with the highest dose formed tumors, in which case it was assumed that the tumor incidence of animals at 10 or 100 times the uppermost dose step (a dummy set of data) would have been 100% for the Spearman-Kärber method to be applicable [12].

The protocol of the present study was reviewed beforehand and approved by the Animal Ethics Committees of CIEA (Permit Number: 13041A) and the National Institute of Health Sciences (NIHS, Tokyo) (Permit Number: 359, 359-1, 359-2, 359-3). All animal experiments were performed according to the Ethical Guidelines for Animal Experimentation from the CIEA and the NIHS. All animals were sacrificed under isoflurane inhalation anaesthesia, and all efforts were made to minimize suffering.

2.4. Histology and immunohistochemistry

The engrafted tumors were isolated and fixed with 10% neutral buffered formalin (Wako). The paraffin-embedded sections were investigated by hematoxylin and eosin (H&E) stain and immunohistochemical studies using Bond-max stainers (Leica Biosystems). Some sections were incubated at 100 °C for 10 min in a target retrieval solution consisting of 10 mM citrate buffer (ER1; Leica Microsystems), and then placed at room temperature for 20 min. Mouse anti-human HLA class I-A, B, C monoclonal antibody (EMR8-5; Hokudo, Sapporo, Japan), and rabbit anti-vimentin monoclonal antibody (SP20; Nichirei Bioscience, Tokyo) were used as the primary antibodies. The antibodies for mouse immunoglobulin were visualized using Bond polymer refine detection kits (Leica Microsystems). Sections were counterstained with hematoxylin.

2.5. Soft agar colony formation assay

A soft agar colony formation assay was performed using a CytoSelect™ 96-well Cell Transformation Assay kit (CellBio labs, San Diego, CA) as previously described [13]. Prewarmed 25 μl of 2 × DMEM/20% FBS and 25 μl of 1.2% agar solution were mixed and