Epigenetic Characterization of the *FMR1* Promoter in Induced Pluripotent Stem Cells from Human Fibroblasts Carrying an Unmethylated Full Mutation

C.E.F. de Esch[1], M. Ghazvini[2,3], F. Loos[3], N. Schelling-Kazaryan[4], W. Widagdo[1], S.T. Munshi[5], E. van der Wal[6], H. Douben[1], N. Günhanlar[5], S.A. Kushner[5], W.W.M.P. Pijnappel[6], F.M.S. de Vrij[5], N. Geijsen[4,7], J. Gribnau[3,8], and R. Willemsen[1]

[1]*Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, the Netherlands*

[2]*iPS Cell Facility, Erasmus Medical Center, Rotterdam, the Netherlands*

[3]*Department of Reproduction and Development, Erasmus Medical Center, Rotterdam, the Netherlands*

[4]*KNAW Hubrecht Institute and UMC Utrecht, Utrecht, the Netherlands*

[5]*Department of Psychiatry, Erasmus Medical Center, Rotterdam, the Netherlands*

[6]*Molecular Stem Cell Biology, Department of Clinical Genetics and Department of Pediatrics, Division of Metabolic Diseases and Genetics, Center for Lysosomal and Metabolic Diseases, Erasmus Medical Center, Rotterdam, the Netherlands*

[7]*Department Companion Animals, Utrecht University School for Veterinary Medicine, Utrecht, the Netherlands*

Published in Stem Cells Reports (2014), 3: 548–555

ABSTRACT

Silencing of the *FMR1* gene leads to fragile X syndrome, the most common cause of inherited intellectual disability. To study the epigenetic modifications of the *FMR1* gene during silencing in time, we used fibroblasts and induced pluripotent stem cells (iPSCs) of an unmethylated full mutation (uFM) individual with normal intelligence. The uFM fibroblast line carried an unmethylated *FMR1* promoter region and expressed normal to slightly increased *FMR1* mRNA levels. The *FMR1* expression in the uFM line corresponds with the increased H3 acetylation and H3K4 methylation in combination with a reduced H3K9 methylation. After reprogramming, the *FMR1* promoter region was methylated in all uFM iPSC clones. Two clones were analyzed further and showed a lack of *FMR1* expression, whereas the presence of specific histone modifications also indicated a repressed *FMR1* promoter. In conclusion, these findings demonstrate that the standard reprogramming procedure leads to epigenetic silencing of the fully mutated *FMR1* gene.

INTRODUCTION

The most common inherited form of intellectual disability, fragile X syndrome (FXS), is caused by the absence of the *FMR1* gene product, the fragile X mental retardation protein (FMRP). In the majority of FXS patients, the transcriptional silencing of the *FMR1* gene is initiated by an expansion of a naturally occurring CGG repeat in the 5' UTR of the *FMR1* gene, to more than 200 units[1,2]. This so-called full mutation results in hypermethylation of the cytosines in the repeat region and the *FMR1* promoter region during early human embryonic development[3,4]. This results in a lack of *FMR1* transcription and consequently an absence of FMRP. Along with hypermethylation, the *FMR1* promoter in FXS is characterized by additional epigenetic marks specific for transcriptionally repressed chromatin including reduced histone H3 and H4 acetylation, reduced histone H3K4 methylation, and increased histone H3K9 methylation[5-8]. However, the timing and molecular mechanisms involved in the CGG expansion, the concomitant DNA methylation, and the additional epigenetic changes that occur during embryonic development are not yet fully understood.

Insights into these processes may lead to a more complete understanding of the developmental processes underlying fragile X syndrome, which, in turn, could lead to new therapeutic strategies. Because murine fragile X models cannot be used to investigate epigenetic *FMR1* inactivation as methylation of the full mutations does not occur, human FXS embryonic stem cells have been studied. These studies showed that FMRP is expressed during early embryonic development, but that epigenetic silencing of *FMR1* occurs upon differentiation[9,10]. A further attempt to study the epigenetic changes over time made use of induced pluripotent stem cells (iPSCs) generated from human FXS fibroblasts. In contrast to human embryonic FX stem cells, these pluripotent cells were shown to already carry a fully methylated *FMR1* promoter and additional heterochromatin marks, so the epigenetic silencing mechanisms in time could not be studied[11-13].

In 1991, a familial case was reported in which two brothers with normal intelligence were shown to have a full *FMR1* mutation without the concomitant hypermethylation of the CGG repeat and the promoter region[14]. In order to unravel the molecular mechanisms behind the epigenetic silencing in fragile X syndrome, we derived iPSCs from these human fibroblasts, to analyze the epigenetic characteristics of the *FMR1* promoter after reprogramming and during differentiation. Here, we report the characterization of these iPSCs and show, unexpectedly, that the *FMR1* promoter of the unmethylated full mutation cell line becomes methylated during reprogramming and stays methylated after differentiation into neural progenitor cells.

RESULTS

Fibroblast Characterization

Fibroblasts from a normal male carrying an unmethylated full mutation first described by Smeets et al. (1995)[14] (uFM) and fibroblasts from a clinically diagnosed male fragile X syndrome patient (14 years old, FXS) and an unrelated unaffected male control line (3 years old, control) were analyzed for *FMR1* 5' UTR CGG repeat length, methylation status, *FMR1* expression, and the histone marks associated with the *FMR1* promoter. As expected, the control line showed a CGG repeat length within the normal range (<55), whereas the uFM and the FXS line showed CGG repeat lengths in the full mutation range (approximately 233 and 380 repeats, respectively) (**Figure S1**). Also, as expected, the part of the *FMR1* promoter analyzed after bisulfite conversion was not methylated in the control and the uFM cell lines, whereas in the FXS cell line the *FMR1* promoter was methylated (**Figures 1A** and **S2** for location of the primers). Because the methylation status is predictive of *FMR1* expression, indeed the control line showed normal expression levels and the uFM line showed normal to slightly increased *FMR1* expression, whereas the FXS cell line did not express *FMR1* transcripts (**Figure 1B**). Additionally, bisulfite Sanger sequencing of a region of the *FMR1* promoter containing 22 CpGs was carried out, which confirmed the absence of methylation of the *FMR1* promoter in the uFM fibroblast line (**Figure 1C**).

Fibroblast Reprogramming and iPSC Characterization

The fibroblasts were reprogrammed to iPSC lines according to established protocols[15,16]. First, four iPSC clones were generated that showed typical characteristics of pluripotent stem cells: morphology similar to that of embryonic stem cells (data not shown), expression of alkaline phosphatase (data not shown), silencing of the multicistronic lentiviral transgene (data not shown), reactivation of genes indicative of pluripotency (data not shown), immunoreactivity for OCT4, NANOG, TRA-1-60, TRA-1-81, and SSEA4 (**Figure S3**), propagation for a long time in culture (up to passage 30), and maintenance of a normal diploid karyotype (data not shown). All four cell lines generated embryonic bodies that, after differentiation *in vitro*, expressed markers of endoderm, mesoderm and ectoderm (**Figure S3**). These four lines were extensively characterized and the results are described below. Second, we generated eight additional iPSC clones from the uFM fibroblast line solely in order to confirm the methylation status of the *FMR1* promoter by quantitative PCR (**Figure 2D**). These additional iPSC clones were generated from the uFM fibroblast line by the same methods as described, except this time we used naive human stem cell medium (WIS-NHSM) as defined by Gafni et al. (2013)[17]. This medium facilitates the derivation of naive pluripotent iPSCs with properties highly similar to mouse naive ES cells.

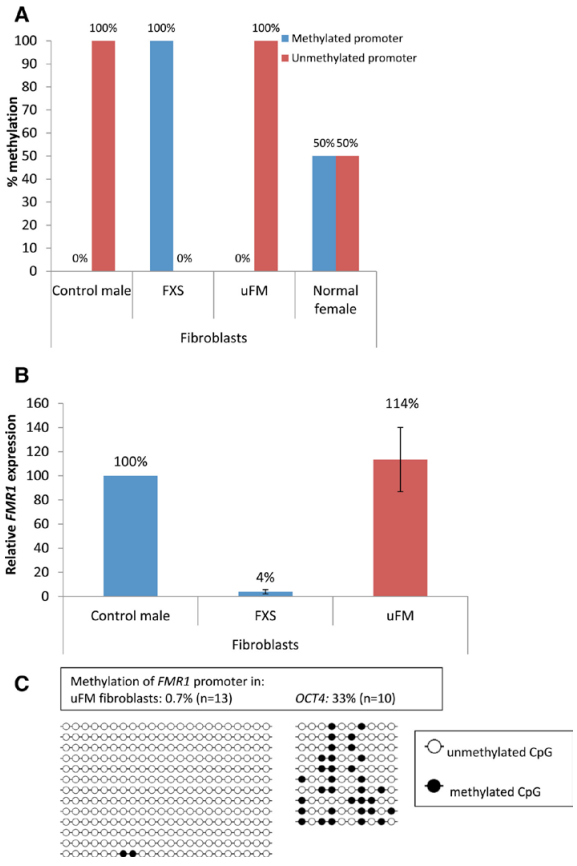

Figure 1. Methylation Status and FMR1 Expression Levels in the Fibroblast Cell Lines (A) Methylation status of a region of the FMR1 promoter in fibroblasts of the male control line, fragile X line (FXS), and the unmethylated full mutation line (uFM). Values were normalized to CLK2 promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n = 2–3 separate measurements). (B) Real-time quantitative PCR data showing FMR1 transcript levels in fibroblasts of the male control line, fragile X line (FXS), and the unmethylated full mutation line (uFM) normalized to CLK2 expression. Values are means ± SEM relative to appropriate male control line (n = 2–3 separate measurements). (C) The percentage of methylated CpGs in the FMR1 promoter and as a control the OCT4 promoter, in 13 and ten clones, respectively, after Sanger sequencing of bisulfite converted DNA of the uFM fibroblast line. Each line represents a clone, and each circle represents a CpG site, which is methylated (closed circle) or unmethylated (open circle). See also Figures S1 and S2.

Reprogramming Effects on CGG Repeat Length, FMR1 Expression, and Methylation

Analysis of the CGG repeat in the 5' UTR of the *FMR1* promoter indicated that the repeat length in the cell lines carrying a full mutation did not contract to levels below 200 CGGs during reprogramming (**Figure S1**). The iPSC clone of the control cell line contained a CGG length under 55 repeats. Nonetheless, the CGG repeat length contracted slightly in the FXS iPSC line after reprogramming, from 380 repeats to approximately 290 repeats. In contrast, the repeat was expanded in the two uFM iPSC clones to approximately 330 and 380 repeats (**Figure S1**). As expected, the iPSC clone of the control cell line showed *FMR1* expression, in contrast to the FXS iPSC clone that did not show *FMR1* expression. Unexpectedly, the two uFM iPSC clones did not express *FMR1* either (**Figure 2A**). Further analysis showed that the bisulfite converted *FMR1* promoter region was methylated in the FXS iPSC clone as well as in both uFM iPSC clones, whereas the control iPSC cell line did not show any methylation (**Figure 2B**). Bisulfite Sanger sequencing confirmed the methylation status of the two uFM iPS clones (**Figure 2C**). The additional eight iPS clones generated from the uFM fibroblast line

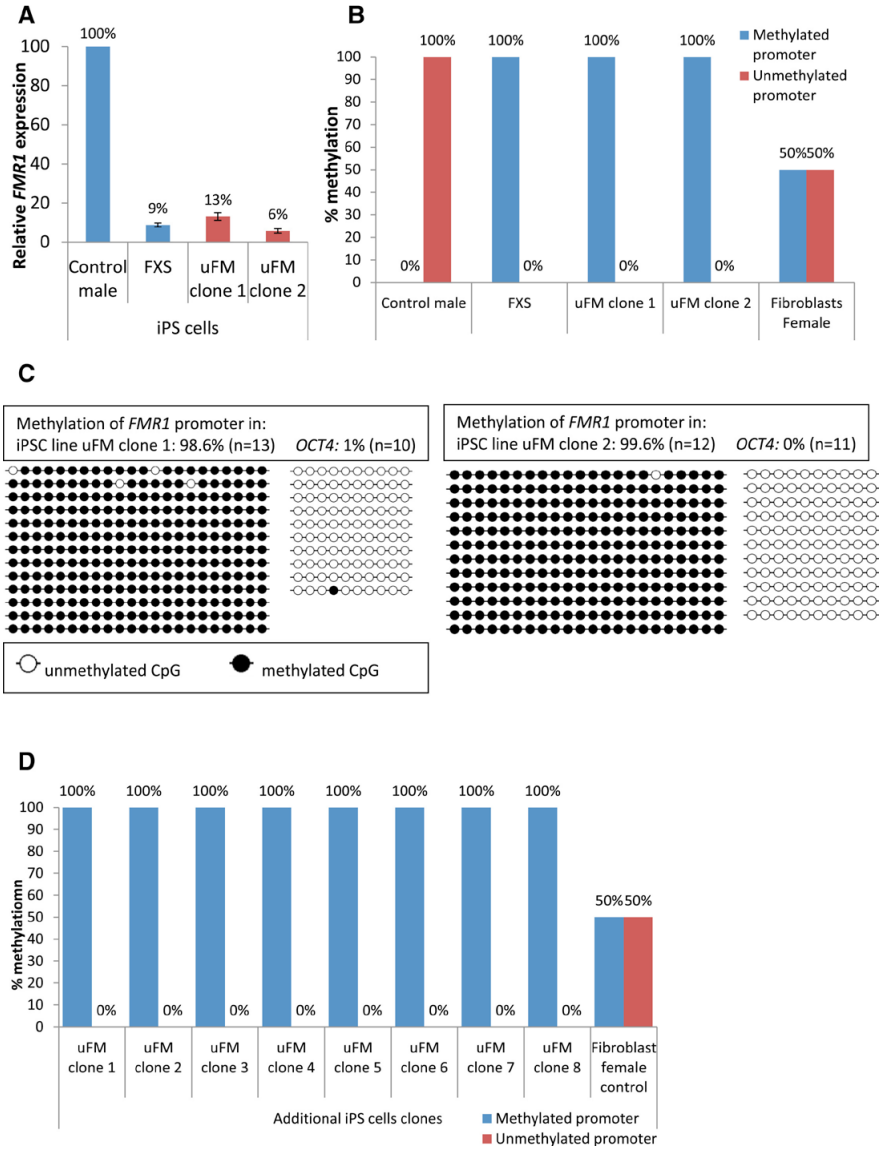

Figure 2 Methylation Status and *FMR1*Expression Levels in the Induced Pluripotent Stem Cells

(A) Real-time quantitative PCR data showing *FMR1* transcript levels in induced pluripotent stem cells (iPSCs) of the male control line, fragile X line (FXS), and the unmethylated full mutation clones (uFM clone 1 and clone 2) normalized to *CLK2* expression. Values are mean ± SEM relative to appropriate male control line (n = 2–3 separate measurements).

(B) Methylation status of a region of the *FMR1*promoter in iPSCs of the male control line, fragile X line (FXS), and the unmethylated full mutation clones (uFM clone 1 and clone 2). Values were normalized to *CLK2* promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n = 2–3 separate measurements).

Figure 2 Methylation Status and *FMR1*Expression Levels in the Induced Pluripotent Stem Cells
(continued)
(C) The percentage of methylated CpGs in the *FMR1* promoter and as a control the *OCT4*promoter, after Sanger sequencing of bisulfite converted DNA of the uFM iPSC clones. Each line represents a clone, and each circle represents a CpG site, which is methylated (closed circle) or unmethylated (open circle).
(D) Methylation status of a region of the *FMR1*promoter in additionally generated iPSC clones of the unmethylated full mutation fibroblast line in naive human stem cell medium. Values were normalized to *CLK2*promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n = 2 separate measurements). See also Figures S1–S3.

in WIS-NSHM medium also showed complete methylation of the bisulfite converted *FMR1* region (**Figure 2D**). Thus, the originally unmethylated extended CGG repeat found in the uFM fibroblasts became methylated at some point during the reprogramming process.

Chromatin immunoprecipitation (ChIP) experiments with the fibroblast lines showed that the *FMR1* promoter of the control line carried active histone marks, H3 acetylation and H3K4 dimethylation with values similar to the positive control, namely, the active gene APRT, and values much higher than the negative control CRYAA (crystalline), which only serves as a positive control for repressed genes. The inactive mark H3K9 trimethylation was not enriched in the control fibroblasts (**Figures 3A–3C**). The uFM fibroblast line carried histone marks representative of an actively transcribed gene, namely, H3 acetylation and H3K4 methylation at similar levels as the control line. The inactive mark H3K9 methylation could not be detected in the uFM fibroblast line (**Figures 3A–3C**). The *FMR1* promoter of the FXS cell line only showed enrichment of the repressive mark H3K9 methylation (**Figures 3A–3C**). ChIP analysis of the *FMR1* promoter in iPSCs showed enrichment of the active marks H3 acetylation and H3K4 methylation in the control iPSC clone, to levels higher than the positive control APRT. The FXS iPSCs and clone 1 of the uFM iPSCs showed an increase of the repressive mark H3K9 methylation to values above the repressive control CRYAA, whereas enrichment of the active marks could not be detected in FXS iPSCs and uFM iPSC clones 1 and 2 (**Figures 3D–3F**).

Next, we investigated the effects of differentiation into neural progenitor cells (NPCs) on *FMR1* expression and methylation (see **Figure S4** for staining with marker SOX2). NPCs derived from the FXS and uFM iPSCs lacked *FMR1* expression and carried a methylated *FMR1* promoter. The NPCs derived from the control iPSC clone showed clear *FMR1* expression and an unmethylated promoter region (**Figures 4A and 4B**). These findings indicate that the reprogramming process leads to methylation of the expanded *FMR1* CGG repeat sequence, which results in a stable shut down of *FMR1* gene expression.

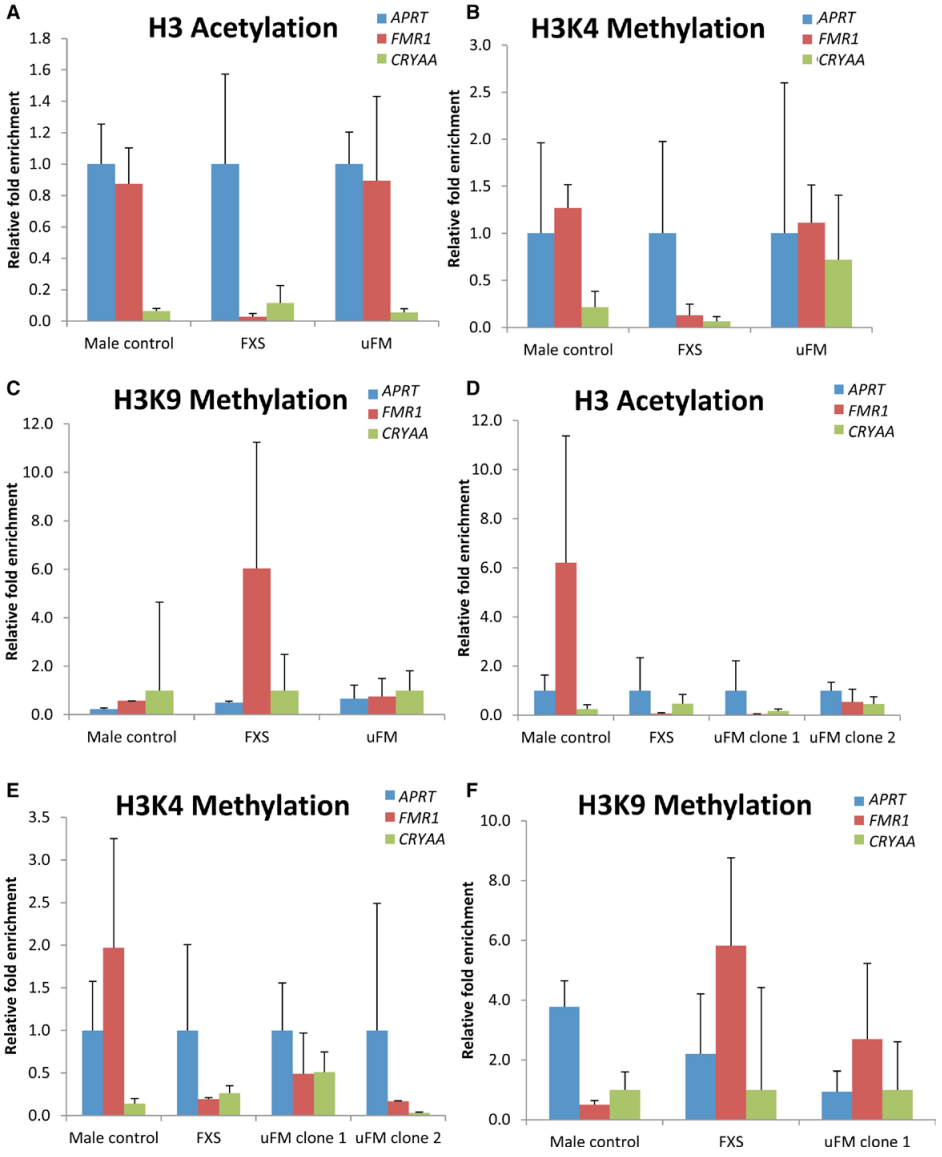

Figure 3 Chromatin Immunoprecipitation Analysis of H3 Acetylation, H3K4 Methylation, and H3K9 Methylation in the *FMR1* Promoter of Fibroblasts and iPSCs

Chromatin immunoprecipitation analysis of H3 acetylation, H3K4 methylation, and H3K9 methylation in the *FMR1* promoter of fibroblasts (A–C) and iPSCs (D–F), respectively. Results were normalized to the appropriate positive control (*APRT* or *CRYAA*), averaged from at least two different experiments and shown with their respective SEs.

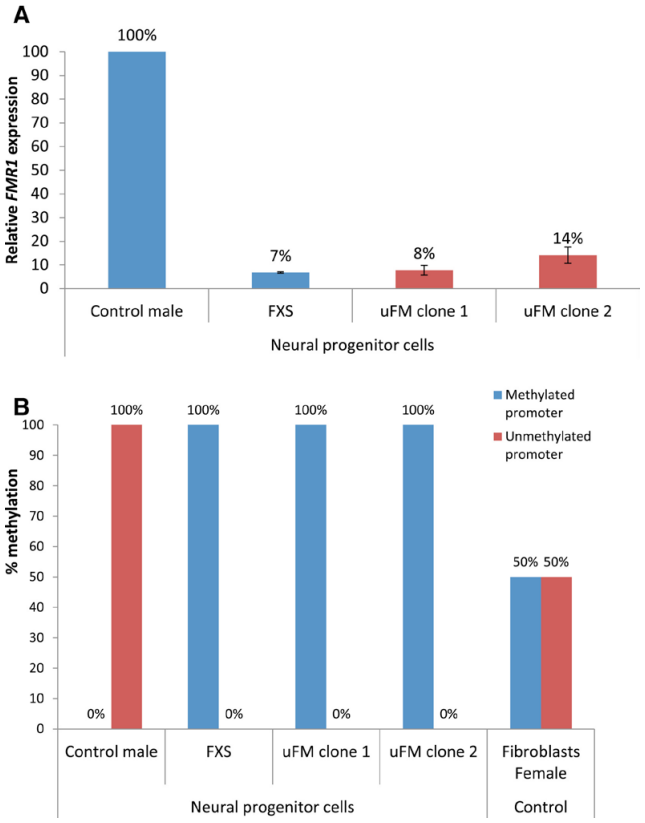

Figure 4 Methylation Status and *FMR1*Expression Levels in Neural Progenitor Cells

(A) Real-time quantitative PCR data showing *FMR1* transcript levels in neural progenitor cells (NPCs) of the male control line, fragile X line (FXS), and the unmethylated full mutation clones (uFM clone 1 and clone 2) normalized to *CLK2* expression. Values are mean ± SEM relative to appropriate male control line (n = 2 separate measurements).

(B) Methylation status of a region of the *FMR1*promoter in NPCs of the male control line, fragile X line (FXS), and the unmethylated full mutation clones (uFM clone 1 and clone 2). Values were normalized to *CLK2* promoter activity first. The normalized exponential values were then presented as a percentage relative to the female fibroblast control line, for which the normalized exponential values were set to 50% for each primer set (n = 2–3 separate measurements). See also Figures S2 and S4.

DISCUSSION

We undertook this study in an attempt to unravel the epigenetic mechanisms involved in the silencing of the *FMR1* gene in fragile X syndrome by the use of a fibroblast line carrying an unmethylated full mutation. There have been several attempts to study epigenetic silencing in fragile X syndrome. Eiges et al. (2007)[9] have shown that FXS human embryonic stem cells (hESCs) still express FMRP at a level similar to that in unaffected hESCs, whereas the FMRP level decreases as the hESCs were differentiated. Based on these results, it was expected that

by reprogramming FXS fibroblasts into pluripotent stem cells, the hypermethylated state of the *FMR1* promoter region would be reversed. However, by now several research groups have shown that iPSCs derived from FXS patients show epigenetic marks characteristic for heterochromatin similar to the full mutation fibroblasts they originated from[11-13]. These observations could be explained by the fact that the FXS iPSCs may not have all the characteristics of early pluripotency, but that they represent a later stage of human development[11-13,17].

Another approach was used in studies with human fragile X lymphoblastic cells; here, a fully mutated and hypermethylated *FMR1* gene was reactivated by treatment with 5-azadeoxycytidine, a hypomethylating agent. Although such treatment significantly reduced DNA methylation in some cells, it could not restore all remaining epigenetic marks to control levels[5,6,18,19]. Drugs such as 4-phenylbutyrate, sodium butyrate or trichostatin A, which block the activity of histone deacetylases, did not restore *FMR1* expression to normal levels[5,6,8,19]. In addition, treatment with a compound that reduces the *in vitro* expression of the FRAXA fragile site, acetyl-l-carnitine, did not restore the *FMR1* expression either[8]. Recently, 5-azadeoxycytidine treatment was also tested on fragile X iPSCs, and it appeared to restore *FMR1* expression in both iPSCs and differentiated neurons, which offers possibilities to use these cells as an epigenetic model[13].

The availability of a fibroblast cell line carrying an unmethylated full mutation (uFM) provided a new opportunity to study the epigenetic silencing mechanisms in time. We first characterized the uFM fibroblast cell line together with a normal male fibroblast control line and a FXS fibroblast cell line carrying a fully methylated *FMR1* promoter. Although increased *FMR1* mRNA levels (up to five times) were reported in lymphoblastoid cells of premutation carriers (55~200 unmethylated CGGs), our findings of normal to slightly increased *FMR1* mRNA levels in the uFM fibroblasts are similar to the findings of Pietrobono et al. (2005)[7], who examined a lymphoblastic cell line from the same individual. The lack of DNA methylation ensures that the chromatin is less densely packed and more accessible for transcription, which explains the *FMR1* expression in this cell line. Our ChIP results differ from the original ChIP analysis of the uFM lymphoblastoid cell line[8]. We found a similar increase in H3K4 methylation; however, we did not find decreased H3 acetylation levels or intermediate H3K9 levels in the uFM fibroblasts. These differences could be explained by the fact that we have analyzed a distinct cell type (fibroblasts versus lymphoblastoid cells), and by differences in the ChIP protocol (e.g., quantification methods and reference genes used). Because the uFM fibroblast line lacked methylation of the *FMR1* promoter site despite the high number of CGG repeats, we expected to find an unmethylated *FMR1* promoter and normal levels of *FMR1* mRNA after reprogramming into iPSCs. Surprisingly, we found the promoter region of *FMR1* to be hypermethylated in all iPSC clones. Other epigenetic chromatin marks also indicated a repressed *FMR1* promoter similar to the marks observed in the fragile X iPSC line. After differentiation of these iPSCs into neural progenitor cells, the *FMR1* promoter remained methylated and thus silenced.

There are three possible explanations for our findings. First, it is possible that the reprogramming process resulted in iPSCs that were solely derived from methylated FM fibroblasts and not of the unmethylated cells. This assumes that methylated FM fibroblasts were present in our culture, which according to our bisulfite sequencing results seems highly unlikely. Second, there may be an unknown genetic factor present in this individual that was protective against DNA methylation during embryonic development but which was absent in his fibroblasts or which was altered or blocked during the reprogramming process. In our case, the brother of this individual was also carrier of an unmethylated full mutation. Being a carrier of an unmethylated full mutation is already a very rare phenomenon, but the fact that two children escaped methylation in one family clearly points towards the involvement of a maternal-paternal genetic component or environmental factors. Finally, the reprogramming process might activate genes that induce de novo methylation of the *FMR1* promoter. Although the *FMR1* gene in this individual escaped methylation during embryonic development, the full mutation in his fibroblasts might be recognized by epigenetic remodelers, e.g., by histone and/or DNA methyltransferases (DNMTs) that are not recruited in embryonic development. This would also explain the unmethylated full mutation observed in human embryonic FXS stem cells because these cells never went through this reprogramming process. A strategy to test this hypothesis would be, for example, to perform the reprogramming of the uFM fibroblasts as well as FXS fibroblast lines under conditions that inhibit the functioning of DNMT 3a and 3b.

In conclusion, standard reprogramming of somatic uFM fibroblasts into pluripotent stem cells by the use of four transcription factors did not lead to demethylation of the expanded CGG repeat and even induced methylation of an unmethylated template. Very recently, Gafni et al. (2013)[17] suggested that a more naive ground state pluripotent stem cell in which epigenetic memory is completely erased could be obtained by a unique combination of cytokines and small molecule inhibitors (WIS-NHS medium). Their study also demonstrated the reactivation of the *FMR1* gene in FXS iPSCs after the reprogramming of FXS fibroblast under naive conditions. However, in contrast to these findings, the use of this WIS-NHS medium did not prevent the occurrence of the de novo methylation of the extended *FMR1* repeat in our uFM iPS clones. In conclusion, our results show that although this fibroblast line may offer a unique system to study the de novo methylation of an extended *FMR1* repeat during reprogramming, the mechanism behind the silencing of the *FMR1* gene in fragile X syndrome remains elusive.

EXPERIMENTAL PROCEDURES

Cell Culture

The rare fibroblast cell line established from a normal male carrying an unmethylated full mutation first described by Smeets et al. (1995)[14] (uFM) was used. This line has been subcloned,

so that a homogenous population of cells that carry a fully extended repeat was obtained. Fibroblasts from a clinically diagnosed male fragile X syndrome patient (14 years, FXS), and an unrelated unaffected male (3 years, control) and female control fibroblast line (9 years) were all obtained from the cell repository of the department of Clinical Genetics, Erasmus MC, Rotterdam. For culture conditions, see the Supplemental Experimental Procedures.

iPSC Generation

Reprogramming of human primary skin fibroblasts was performed as described previously[16]. Briefly, fibroblasts were infected with a single, multicistronic lentiviral vector encoding OCT4, SOX2, KLF4, and MYC and cultured on γ-irradiated mouse embryonic feeder (MEF) cells until iPSC colonies could be picked[16]. The second round of reprogramming of the uFM fibroblast line was done in naive ES medium (WIS-NHSM medium) according to Gafni et al. (2013)[17] (see the Supplemental Experimental Procedures). These cells were used to affirm the methylation status of the *FMR1* promoter after reprogramming by methylation specific quantitative PCR. For further details, see the Supplemental Experimental Procedures.

In Vitro Differentiation of Embryonic Bodies

To form embryonic bodies (EBs), iPSC colonies from two wells per line were broken up by collagenase IV treatment and transferred to ultralow attachment 6-well plates (Corning). For the germ layer differentiation conditions, see the Supplemental Experimental Procedures. After 2 weeks in culture, the cells were fixed with formalin and immunostainings were performed (see the Supplemental Experimental Procedures).

Neural Differentiation

Human iPSCs were differentiated according to Brennand et al. (2011)[20], with modifications (see the Supplemental Experimental Procedures). After 1 week, NPCs were dissociated with collagenase (100 U/ml), replated, and used for staining and methylation analysis after three to five passages.

Karyotype Analysis and Immunocytochemistry

Standard staining procedures were followed; for details, see Supplemental Experimental Procedures.

CGG Repeat Length, FMR1 Expression, and Methylation Analysis

CGG repeat size was determined in a PCR using the primers 5'- CGGAGGCGCCGCTGCCAGG-3' and 5'-TGCGGGCGCTCGAGGCCCAG-3' with the Expand high fidelity PCR kit (Roche) supplemented with 2.5 M betaine (see the Supplemental Experimental Procedures). For details of the *FMR1* expression analysis, see the Supplemental Experimental Procedures. Genomic DNA was modified by bisulfite treatment according to the EpiTect Bisulfite Kit.

The diluted converted DNA was then measured using quantitative PCR with two different primer set designed specifically for a region of the *FMR1* promoter (see **Figure S1** for the locations). One primer set contained the methylated DNA sequence and the other contained the unmethylated DNA sequence of a region of the *FMR1* promoter after bisulfite conversion (see the Supplemental Experimental Procedures).

Bisulfite Sanger Sequencing

Genomic DNA (1,000 ng) was modified by bisulfite treatment according to the EpiTect Bisulfite Kit. Then a region of the *FMR1* promoter containing 22 CpGs was amplified using PlatinumTaq (Invitrogen) (see **Figure S1** for location of the primers). PCR products were cloned into pGEM-T Easy (Promega), and single clones were sequenced by Sanger sequencing (see Supplemental Experimental Procedures).

Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed according to the Upstate ChIP protocol with some small modifications (see Supplemental Experimental Procedures). Eluted DNA fragments were used for quantitative PCR analysis (see **Figure S1** for location of the primers). The Ct values of the histone modifications were first normalized for the nonspecific immunoglobulin G antibody treatment and then for the amount of input DNA. Data were then presented in relative fold enrichment after further normalization to the APRT gene for H3 acetylation and H3K4 methylation and CRYAA for H3K9 methylation. Data from at least two separate experiments were averaged, and both reference genes were previously used by Urbach et al. (2010)[11] and Bar-Nur et al. (2012)[13].

ACKNOWLEDGMENTS

The authors thank Dr. A. Schambach for kindly sharing the iPSC reprogramming vector. This work was supported by the Netherlands Organisation for Health Research and Development (R.W.; ZonMw; 912-07-022), E-Rare program entitled “Cure FXS” (R.W.; no. EU/FISPS09102673), the Prinses Beatrix Spierfonds (W.W.M.P.P. and E.v.d.W., project number: W.OR13-21), a fellowship of the Hersenstichting Nederland (F.M.S.d.V.) and the NeuroBasic-PharmaPhenomics consortium (S.A.K.), and NWO VICI and ERC grants (J.G.).

REFERENCES

1. Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P., et al. (1991). Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65, 905–914.
2. Pearson, C.E., Nichol Edamura, K., and Cleary, J.D. (2005). Repeat instability: mechanisms of dynamic mutations. Nat. Rev. Genet. 6, 729–742.
3. Sutcliffe, J.S., Nelson, D.L., Zhang, F., Pieretti, M., Caskey, C.T., Saxe, D., and Warren, S.T. (1992). DNA methylation represses FMR-1 transcription in fragile X syndrome. Hum. Mol. Genet. 1, 397–400.
4. Willemsen, R., Bontekoe, C.J., Severijnen, L.A., and Oostra, B.A. (2002). Timing of the absence of FMR1 expression in full mutation chorionic villi. Hum. Genet. 110, 601–605.
5. Coffee, B., Zhang, F., Warren, S.T., and Reines, D. (1999). Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. Nat. Genet. 22, 98–101.
6. Coffee, B., Zhang, F., Ceman, S., Warren, S.T., and Reines, D. (2002). Histone modifications depict an aberrantly heterochromatinized FMR1 gene in fragile x syndrome. Am. J. Hum. Genet. 71, 923–932.
7. Pietrobono, R., Tabolacci, E., Zalfa, F., Zito, I., Terracciano, A., Moscato, U., Bagni, C., Oostra, B., Chiurazzi, P., and Neri, G. (2005). Molecular dissection of the events leading to inactivation of the FMR1 gene. Hum. Mol. Genet. 14, 267–277.
8. Tabolacci, E., Pietrobono, R., Moscato, U., Oostra, B.A., Chiurazzi, P., and Neri, G. (2005). Differential epigenetic modifications in the FMR1 gene of the fragile X syndrome after reactivating pharmacological treatments. Eur. J. Hum. Genet. 13, 641–648.
9. Eiges, R., Urbach, A., Malcov, M., Frumkin, T., Schwartz, T., Amit, A., Yaron, Y., Eden, A., Yanuka, O., Benvenisty, N., and Ben-Yosef, D. (2007). Developmental study of fragile X syndrome using human embryonic stem cells derived from preimplantation genetically diagnosed embryos. Cell Stem Cell 1, 568–577.
10. Gerhardt, J.I., Tomishima, M.J., Zaninovic, N., Colak, D., Yan, Z., Zhan, Q., Rosenwaks, Z., Jaffrey, S.R., and Schildkraut, C.L. (2013). The DNA replication program is altered at the FMR1 locus in fragile X embryonic stem cells. Mol, Cell 53, 19–31.
11. Urbach, A., Bar-Nur, O., Daley, G.Q., and Benvenisty, N. (2010). Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. Cell Stem Cell 6, 407–411.
12. Sheridan, S.D., Theriault, K.M., Reis, S.A., Zhou, F., Madison, J.M., Daheron, L., Loring, J.F., and Haggarty, S.J. (2011). Epigenetic characterization of the FMR1 gene and aberrant neurodevelopment in human induced pluripotent stem cell models of fragile X syndrome. PLoS ONE 6, e26203.
13. Bar-Nur, O., Caspi, I., and Benvenisty, N. (2012). Molecular analysis of FMR1 reactivation in fragile-X induced pluripotent stem cells and their neuronal derivatives. J. Mol. Cell Biol. 4, 180–183.
14. Smeets, H.J., Smits, A.P., Verheij, C.E., Theelen, J.P., Willemsen, R., van de Burgt, I., Hoogeveen, A.T., Oosterwijk, J.C., and Oostra, B.A. (1995). Normal phenotype in two brothers with a full FMR1 mutation. Hum. Mol. Genet. 4, 2103–2108.
15. Takahashi, K., Okita, K., Nakagawa, M., and Yamanaka, S. (2007). Induction of pluripotent stem cells from fibroblast cultures. Nat. Protoc. 2, 3081–3089.
16. Warlich, E., Kuehle, J., Cantz, T., Brugman, M.H., Maetzig, T., Galla, M., Filipczyk, A.A., Halle, S., Klump, H., Scholer, H.R., et al. (2011). Lentiviral vector design and imaging approaches to visualize the early stages of cellular reprogramming. Mol. Ther. 19, 782–789.
17. Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature 504, 282–286.

18. Chiurazzi, P., Pomponi, M.G., Willemsen, R., Oostra, B.A., and Neri, G. (1998). In vitro reactivation of the FMR1 gene involved in fragile X syndrome. Hum. Mol. Genet. 7, 109–113.
19. Chiurazzi, P., Pomponi, M.G., Pietrobono, R., Bakker, C.E., Neri, G., and Oostra, B.A. (1999). Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene. Hum. Mol. Genet. 8, 2317–2323.
20. Brennand, K.J., Simone, A., Jou, J., Gelboin-Burkhart, C., Tran, N., Sangar, S., Li, Y., Mu, Y., Chen, G., Yu, D., et al. (2011). Modelling schizophrenia using human induced pluripotent stem cells. Nature 473, 221–225.

SUPPLEMENTAL DATA SUPPLEMENTAL FIGURES

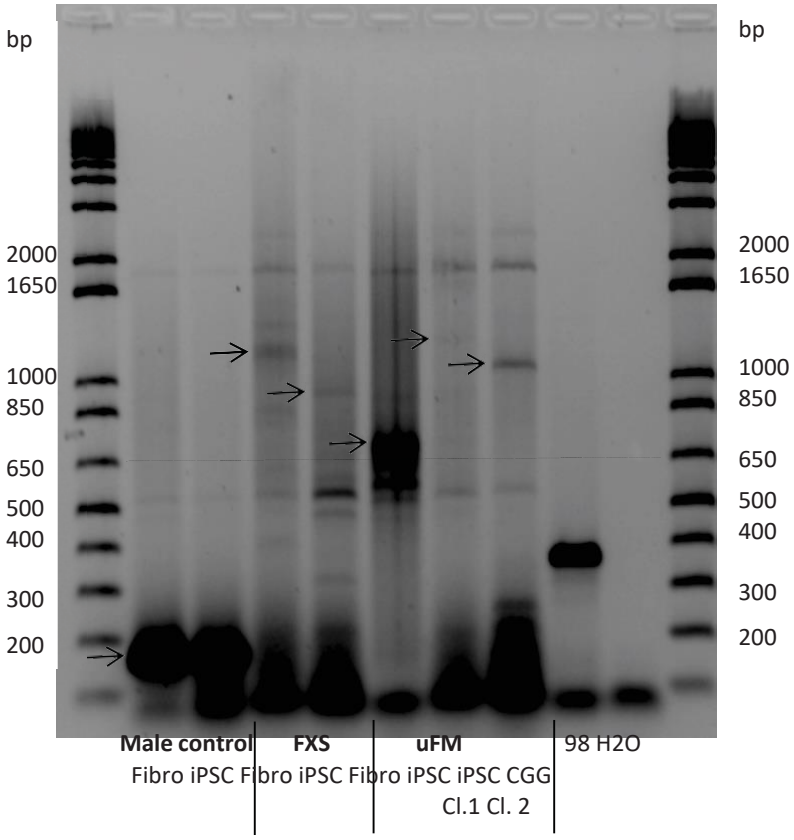

Figure S1: Repeat length analysis before and after reprogramming into iPS cells

The male control fibroblasts and the derived control iPS cells show a repeat length below 55. The fragile X (FXS) fibroblast line shows a repeat length of approximately 380 repeats. The derived FXS iPSCs show a small contraction of the repeat length to approximately 290 repeats. The unmethylated full mutation (uFM) fibroblast line has a repeat length of approximately 233 repeats (~750bp). The two uFM iPS cell clones show an expanded repeat length of approximately 380 (clone 1) and 330 repeats (clone 2). A DNA sample containing 98 CGG repeats was run on the same gel, in addition to a water control. All products of the CGG repeat PCR were run on one agarose gel with a ladder on both sides. A band at the marker level of 650 base pairs corresponds with a CGG repeat length of 200 repeats. Related to Figures 1 and 2.

13261 ccactccacc tcccgctcag tcagactgcg ctactttgaa ccggaccaaa ccaaaccaaa

ChIP F ----------> F1 ----------> <----------

13381 gggataaccg gatgcatttg atttcccacg ccactgagtg cacctctgca gaaatgggcg

<-ChIP R BSM F---- >

13441 ttctggccct cgcgaggcag tgcgacctgt caccgccctt cagccttccc gccctccacc

< ---BSM R

13501 aagcccgcgc acgcccggcc cgcgcgtctg tctttcgacc cggcacccgg gccggttccc

< ---------- R1

13561 agcagcgcgc atgcgcgcgc tcccaggcca cttgaagaga gagggcgggg ccgaggggct

BSNM F--- >

13621 gagcccgcgg ggggagggaa cagcgttgat cacgtgacgt ggtttcagtg tttacacccg

Start of transcription

13681 cagcgggccg ggggttcggc ctcagtcagg cgctcagctc cgtttcggtt tcacttccgg

<--- BSNM R

13741 tggagggccg cctctgagcg ggcggcgggc cgacggcgag cgcgggcggc ggcggtgacg

CGG-repeat

13801 gaggcgccgc tgccaggggg cgtgcggcag cgcggcggcg gcggcggcgg cggcggcggc

13861 ggaggcggcg gcggcggcgg cggcggcggc ggctgggcct cgagcgcccg cagcccacct

Start of translation

13921 ctcgggggcg ggctcccggc gctagcaggg ctgaagagaa gatggaggag ctggtggtgg

start of intron 1

13981 aagtgcgggg ctccaatggc gctttctaca aggtacttgg ctctagggca ggccccatct

Figure S2: The promoter region of the *FMR1* gene with the location of the primers used in this study
Sequence numbering from GenBank L29074. Primers used in this study are indicated by the underlined sequence in combination with the name and arrow above the sequence. Individual cytosines belonging to methylation sites are indicated in bold as well as the CGG repeat. In addition, the start of transcription, translation and intron 1 are indicated as well. For the methylation-specific qPCR we analysed 11 methylation sites using the methylation specific primers (BSM F and R) with the unmethylated primers set (BSNM F and R). Bisulfite sequencing by using the F1 and R1 primers was based on Pietrobono *et al*, 2002, and contained 22 methylation sites. Related to Figures 1, 2, and 4.

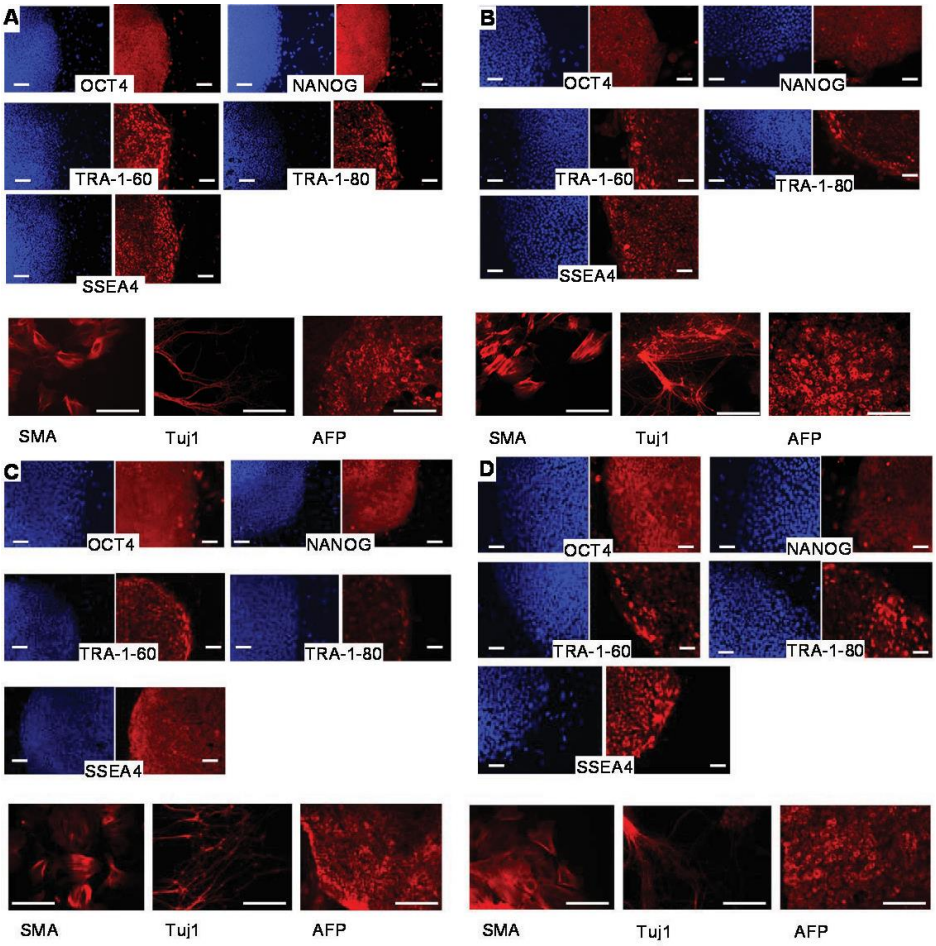

Figure S3: Expression of pluripotency markers by iPSCs and germ layer marker expression after *in vitro* differentiation

From left to right and top to bottom you can see images showing OCT4, NANOG, Tra-1-60, Tra-1-80 and SSEA4 expression as well as expression of the mesodermal marker smooth muscle actin (SMA), the ectodermal marker Tuj1 (or B-tubulin III) and endodermal marker alpha fetoprotein (AFP) (all in red) in the control line (A), fragile X cell line (B) and the uFM clones (C and D, clone 1 and clone 2 respectively). For each pluripotency marker a nuclear Hoechst staining is displayed in blue. Scale bars: 50 µm for the pluripotency markers and 100 µm for the germ layer markers. Related to Figure 2.

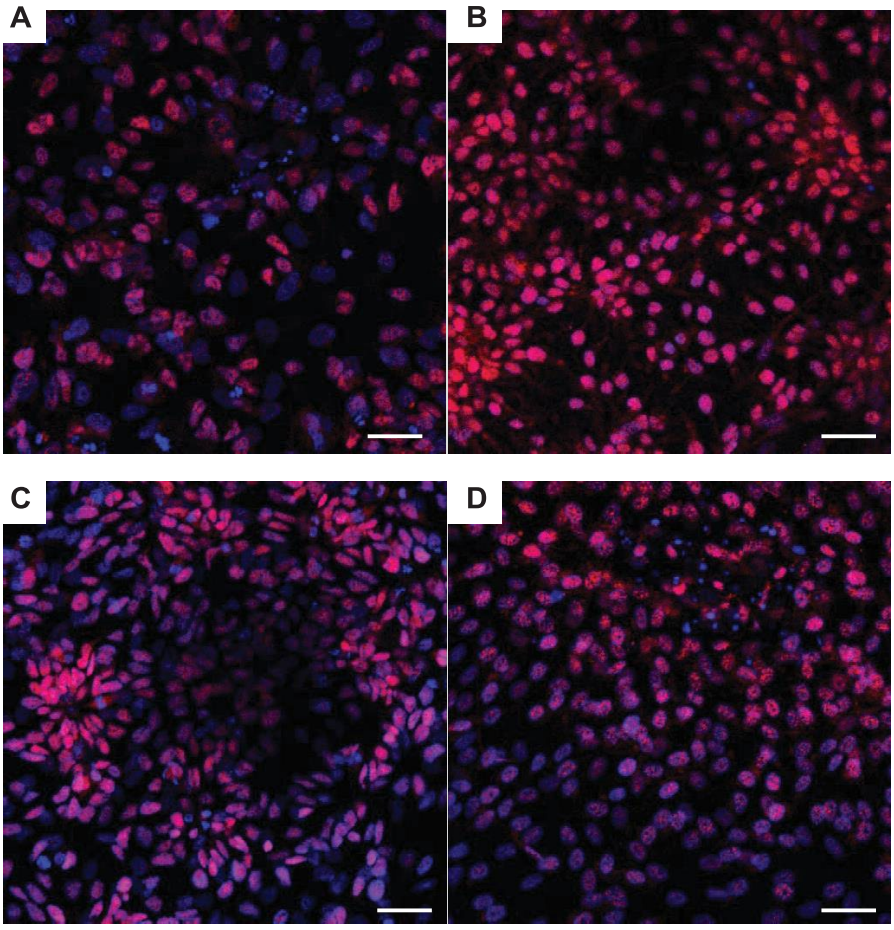

Figure S4: **SOX2 expression by neural progenitor cells**
Expression of SOX2 in red by neural progenitor cells of the control line (**A**), fragile X cell line (**B**) and the uFM clones (**C** and **D,** clone 1 and clone 2 respectively). In each image a nuclear DAPI staining (in blue) is displayed as well. Scale bar 100 μm. Related to Figure 4.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES CELL CULTURE CONDITIONS

The fibroblasts were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco-Invitrogen) containing 10% fetal calf serum and 1% penicillin/streptomycin.

The first set of generated iPS cell lines, namely the male control line, the FXS line and the uFM iPS clone 1 and 2 were cultured in conventional ES cell culture medium containing DMEM/F12 (Gibco-Invitrogen) supplemented with 20% knock-out serum replacement (Gibco-Invitrogen), 2 mM L-glutamine, 50 units of penicillin/streptomycin/glutamine, 0.1 mM MEM-non-essential aminoacids (PAA Laboratories GmbH), 0.1 mM B-mercaptoethanol, and 10 ng/ml bFGF (Invitrogen) filtered through a 0.22 µm filter (Corning). Human iPS lines growing on conventional medium were passaged weekly using collagenase IV (1 mg/ml, Invitrogen) on γ-irradiated MEFs.

The second round of iPS cells were grown in WIS-NHSM medium containing 475 ml knockout DMEM (Invitrogen), 20% knockout serum (invitrogen), human insulin (Sigma, 12.5 µg/ml), 10 µg recombinant human Lif (Peprotech), 8 ng/ml recombinant bFGF (Peprotech) and 1 ng/ml recombinant TGF-β1 (Peprotech), 1 mM glutamine (Invitrogen), 1% MEM-non-essential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma), Penicillin-Streptomycin (Invitrogen) and small molecule inhibitors: PD0325901 (1 µM, ERK1/2i, Axon Medchem); CHIR99021 (3 µM, GSK3i, Axon Medchem); SP600125 (10 µM, JNKi, TORCIS) and SB203580 (10 µM, p38i, Axon Medchem) Y-27632 (5 µM, Axon Medchem) and protein kinase C inhibitor G06983 (5 µM, TOCRIS). Naive human iPS clones were grown on y-irradiated MEFs on gelatin-coated plates and passaged by single-cell trypsinization (0.05% EDTA) every 4 days. These naive cells also showed the main characteristics of induced pluripotent stem cells including morphology similar to that of embryonic stem cells, silencing of retroviral transgenes and reactivation of pluripotency genes (data not shown).

In vitro differentiation of embryonic bodies

Floating EBs were cultured in iPSC medium without bFGF for a minimum of 6 days with supplemented SB431542 for ectoderm conditions only. The embryonic bodies (EBs) designated for endoderm were then transferred to gelatin coated 12-wells plates containing the following medium: RPMI 1640 (Gibco-Invitrogen), supplemented with 20% FBS, 1: 100 dilution of penicillin/streptomycin/glutamine and alpha-thioglycerol (0.4mM). Mesoderm differentiation from the EBs was inducted in gelatin-coated 12-wells plates with DMEM low glucose medium supplemented with 15% FBS, 1:100 dilution of penicillin/streptomycin/glutamine and 1:100 dilution of MEM-non-essential amino acids. The formation of ectoderm was induced in matrigel-coated plates with the following medium: neurobasal medium (Gibco) and DMEM/F12 (v/v 50/50) supplemented with 1:100 dilution of penicillin/streptomycin/

glutamine and 1:100 dilution of MEM-non-essential aminoacids, 0.02% BSA (Gibco), 1:200 N2 (Gibco) and 1:100 B27 (Gibco).

Neural differentiation

Briefly, iPS colonies were dissociated from MEFs with collagenase (100 U/ml) and transferred to non-adherent plates in hES cell medium on a shaker in an incubator at 37ºC/5% CO_2. After two days, embryonic bodies (EBs) were placed in neural induction medium (DMEM/F12, 1x N2, 2 µg/ml heparin, penicillin/streptomycin) and cultured for another four days in suspension. EBs were gently dissociated and plated onto laminin-coated dishes in NPC medium (DMEM/F12, 1x N2, 1x B27-RA, 1 µg/ml laminin and 20 ng/ml FGF2, penicillin/streptomycin). All cell culture reagents were obtained from Invitrogen.

Karyotype analysis

For karyotype analysis, cells in a well of a 6-wells plate were treated with colcemid (100 ng/ml) for 1 hour. Then cells were harvested with trypsin, treated with hypotonic solution and fixed. Metaphases were spread onto glass slides and stained with DAPI (Dako). Chromosomes were classified according to the International System for Human Cytogenetic Nomenclature. At least 10 metaphases were analyzed per cell line.

Immunocytochemistry and antibodies used in this study

iPS cells or NPCs were washed with PBS once, fixed with 4% formalin solution for 5 min and washed again with PBS. Cells were then incubated with 50 mM glycine for 5 min, washed with PBS and permeabilized with 0.5% Triton X-100 for 5 min (only for OCT4 and NANOG). After blocking for 45 min at room temperature with 0.1% PBS-Tween containing 2% fetal bovine serum (Invitrogen), primary antibody staining was performed for 1 hour in room temperature with antibodies diluted in blocking solution. Cells were then washed and incubated with the appropriate secondary Cy3 or Alexa Fluor A555 antibody (1:200, Jackson Immunoresearch Laboratories or Invitrogen) for 45 min. Afterwards, cells were washed with twice 0.1% PBS-Tween, with a nuclear staining step in between (Hoechst or DAPI). Cells were covered with Mowiol and a glass slide. Staining for alkaline phosphatase was carried out using the Alkaline Phosphatase kit (Sigma-Aldrich) according to the manufacturer's instructions. The antibodies used for pluripotency stainings or neural marker stainings were goat anti-human OCT3/4 (1:100, Santa Cruz Biotechnology), goat anti-human NANOG (1:50, R&D Systems), mouse anti-human TRA-1-60, TRA-1-80, and SSEA4 (1:100 Santa Cruz Biotechnology) and rabbit anti-SOX2 (1:1000 Millipore). Antibodies used for *in vitro* differentiation stainings were anti-human smooth muscle actin (SMA, 1:50, DAKO), rabbit anti-human alpha-fetoprotein (AFP, 1:200, Dako), mouse anti-human β-tubulin III (TujI) (1:200, Sigma-Aldrich).

CGG repeat length, *FMR1* expression and methylation analysis

CGG length PCR

In order to isolate total genomic DNA, cell were treated with lysis buffer containing 100 mM NaCl, 10 mM Tris, 15 mM EDTA, 0,5% SDS, and 5% Proteinase K. After overnight incubation at 55°C, DNA was extracted and precipitated using a standard protocol containing saturated salt solution and ethanol. PCR was performed with 35 cycles of 35 seconds denaturing at 98°C, 35 seconds of annealing at 55°C, and 5 minutes elongation at 72°C. PCR products were analyzed with standard agarose gel electrophoresis.

FMR1 expression

RNA was isolated using the RNAeasy kit (Qiagen), and 1 µg of RNA was reverse transcribed using iScript (BioRad). Real-time PCR was carried out in triplicate using Kappa mix and a 7300 Real-time PCR system (Applied Biosystems). A forward primer located in exon 4 was used in combination with a reverse primer located in exon 5 to measure *FMR1* expression: 5'-GGTGGTTAGCTAAAGTGAGGA-3' and 5'-GTGGCAGGTTTGTTGGGATTA-3'.

CLK2 was used as reference gene with forward primer 5'-CCTACAACCTAGAGA-AGAAGCGAG-3' and reverse primer 5'-CACTGCCAAAGTCTACCACC-3'(de Brouwer *et al.* 2006). *FMR1* expression was normalized to *CLK2* expression and data was presented as an average value from 2 to 3 independent measurements. The expression values of the male control and the female control cell lines were combined and their average relative fold enrichment was set to 100%.

FMR1 methylation analysis

The primers for the methylated sequence are F 5'-GGTCGAAAGATAGACGCGC-3', R 5'-AAACAATGCGACCTATCACCG-3'; and for the unmethylated sequence are F 5'- TGTT-GGTTTGTTGTTTGTTTAGA-3', R 5'-AACATAATTTCAATATTTACACCC-3' and for the promoter of the unmethylated bisulfite converted reference gene *CLK2*: F 5'-CGGTT-GATTTTGGGTGAAGT-3' and R 5'-TCCCGACTAAAATCCCACAA-3'. All reactions were carried out in triplicate using SYBR Green ROX mix and a 7300 Real-time PCR system. Experiments were only analyzed when the Ct values of the female control sample were under 30 for both primer sets, as an indication for an efficient bisulfite conversion and DNA recovery. For each sample, the values for the methylated and the unmethylated sequences were normalized to *CLK2* promoter activity first to obtain delta Cts. The normalized exponential values from the measurements of both primers sets were then set to 50% for the female control cell line. These values represent the random X-inactivation in female control cells. The normalized exponential data of the remaining samples was then presented as a percentage relative to the female control data. Average ratios from at least two independent measurements were used for each sample.

Bisulfite Sanger sequencing

The following primers were used *FMR1* F1 5'-GAGTGTATTTTTGTAGAAATGGG-3' and R1 5'-TCTCTCTTCAAATAACCTAAAAAC-3' (see supplemental figure 1 for location of primers), while the *OCT4* promoter containing 10 CpG sites was amplified using the forward primer 5'-GAGGGAGAGAGGGGTTGAGTAG-3' and the reverse primer 5'-CCTCCAAAAAAACCTTAAAAACTTAAC-3' (based on Al-khtib *et al*. 2012).

Chromatin immunoprecipitation assay

In short, approximately 2.5 million cells were crosslinked with 1% formaldehyde for 5 minutes at room temperature. After quenching the reaction with 125 mM glycine, cells were subsequently suspended in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris pH 8.1) containing proteinase inhibitor (Roche, Complete). Chromatin was then sonicated using the Bioruptor (Diagonide) to create 200bp-1000bp DNA fragments. All chromatin was pre-cleared by treatment with salmon sperm agarose beads (Millipore) for 0.5 hour at room temperature. Immunoprecipitation was performed overnight using 7.5 µg anti-acetylated histone H3 (Millipore), anti-dimethyl histone H3K4 (Millipore), anti-trimethyl histone H3K9 (Millipore), or anti-IgG antibody (Millipore) in dilution buffer. Next, crosslinking was reversed by incubation with 0.2M NaCl at 65ºC and DNA was purified using a PCR clean-up kit (Mobio). Quantitative PCR analysis was carried out using primers for the *FMR1* promoter region F 5'-AACTGGGATAACCGGATGCAT-3' and R 5'- GGCCAGAACGCCCATTTC-3' (see supplemental figure 1 for location) as well as appropriate positive and negative controls namely *APRT* F 5'-GCCTTGACTCGCACTTTT-3', and R 5'- TAGGCGCCATCGATTTTA-3' and *CRYAA* F 5'-CCGTGGTACCAAAGCTGA-3', and R 5'-AGCCGGCTGGGGTAGAA-3'.