Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Impact of DNA methylation on 3D genome structure

Diana Buitrago
Institute for Research in Biomedicine - The Barcelona Institute of Science and Technology
Mireia Labrador
Institute for Research in Biomedicine - The Barcelona Institute of Science and Technology
Juan Pablo Arcon
Institute for Research in Biomedicine (IRB Barcelona) - The Barcelona Institute of Science and Technology https://orcid.org/0000-0003-3350-1576
Rafael Lema
IRB
Oscar Flores
Institute for Research in Biomedicine - The Barcelona Institute of Science and Technology
Anna Esteve-Codina
CNAG-CRG, Centre for Genomic Regulation (CRG), Universitat Pompeu Fabra (UPF) https://orcid.org/0000-0003-0361-2873
Julie Blanc
CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST) https://orcid.org/0000-0002-5367-0628
Nuria Villegas
Institute for Research in Biomedicine - Barcelona Institute of Science and Technology (BIST)
David Bellido
Centres Científics i Tecnològics de la Universitat de Barcelona - CCiTUB
Marta Gut
CNAG-CRG
Pablo Dans
Institute for Research in Biomedicine - Barcelona Institute of Science and Technology (BIST)
Simon Heath
Barcelona Institute of Science and Technology https://orcid.org/0000-0002-9550-0897
Ivo Gut
CNAG-CRG, Centre for Genomic Regulation (CRG) - CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST)
Isabelle Brun Heath
Parc Científic de Barcelona
Modesto Orozco (modesto.orozco@irbbarcelona.org)
Institute for Research in Biomedicine

Article

Keywords: DNA methylation, Saccharomyces cerevisiae, epigenetic regulation, chromatin structure and function

Posted Date: July 7th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-36311/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

---

Version of Record: A version of this preprint was published at Nature Communications on May 28th, 2021. See the published version at https://doi.org/10.1038/s41467-021-23142-8.

Abstract

The extreme complexity of epigenetic regulation in higher organisms makes the determination of the intrinsic effect of DNA methylation in chromatin structure and function challenging. We investigated the role of DNA methylation in a simpler model system, budding yeast (Saccharomyces cerevisiae), an organism in which methylation-related machinery is normally absent thus making it a perfect model system to study the intrinsic role of methylation in chromatin structure and function. With this aim, we expressed the murine DNA Methyl Transferases in S. cerevisiae and analyzed the correlation between DNA methylation, nucleosome positioning, gene expression and 3D genome organization. We showed that despite the lack of machinery for positioning and reading of methylation marks, the methylation pattern follows a conserved pattern, the level of DNA methylation being very low at the 5' end of the gene, and then increasing gradually toward the 3' end, mimicking mammalian behavior. DNA methylation and gene expression correlate as DNA methylation is lower at the TSS and higher at the TTS in highly expressed genes compared to lowly expressed ones, mimicking again mammalian behavior. We found that methylated DNA is unlikely to be wrapped around nucleosomes, but is concentrated in linkers and nucleosome free regions. DNA methylation increases chromatin condensation in the peri-centromeric region, decreases overall DNA flexibility and favors the heterochromatin state. Taken together, these results demonstrate that methylation intrinsically modulates chromatin structure and function even in the absence of cellular machinery evolved to recognize and process the methylation signal.

Introduction

DNA methylation is one of the most important epigenetic marks which introduce major changes in cellular functionality, some of them with systemic impact and coupled to important pathologies in mammals (for a review see 1). For example, mutations in DNA Methyltransferase 3b (DNMT3b) are implicated in Immunodeficiency, Centromere instability and Facial anomalies (ICF) syndrome 2 mutations in DNA Methyltransferase 3a (DNMT3a) are found in acute myeloid leukemia (AML) patients 3 while those in DNA Methyltransferase 1 (DNMT1) cause autosomal dominant cerebellar ataxia, deafness and narcolepsy 4. DNA methylation plays a key role in development (reviewed in 5) and cell differentiation 6, and a correct methylation level is crucial for the regulation of parental imprinting and X chromosome inactivation mechanisms 7. Not surprisingly, changes in DNA methylation patterns have been associated with many different types of cancer in humans 8, 9, 10, 11 and reviewed in 12, 13.

It is commonly stated that DNA methylation in the promoter region of a gene is a hallmark of repression 1, 14. However, several studies have shown that DNA methylation in the gene body could also affect gene expression, and that the increase in methylation in promoter regions is not always correlated with gene repression, making the effect of DNA methylation on gene expression far more complicated than a simple on/off signal 15, 16. Two possible mechanisms might link DNA methylation with gene regulation: i) a direct effect involving proteins with methylated DNA binding domains, which act as anchoring point for other effector proteins regulating gene activity and ii) an indirect effect related to changes in DNA properties. In the second case, at least two possibilities emerge: i) DNA methylation changing the binding affinity of transcription factors to DNA and ii) DNA methylation affecting chromatin structure, which in turn modulates gene expression 17. Accurate *in silico* and *in vitro* studies have demonstrated that DNA methylation makes DNA less flexible and less likely to form nucleosomes 18, 19 , but other studies relying on *in vitro* nucleosome reconstitution claim the opposite: *i.e.* that DNA methylation increases the affinity of histones for DNA 20 and DNA methylation promote compaction on pre-assembled nucleosomes 21. Experiments using *in vitro* reconstituted nucleosomes and *in vivo* studies on mammals and plants are also confusing, with some suggesting that methylation occurs preferentially on nucleosomal DNA 22, 23, 24 and others concluding the opposite 25, 26, 27, 28. This controversy can probably be explained by the high complexity of mammalian genomes, with a myriad of factors controlling directly or indirectly nucleosome positioning 29, and by the existence of higher orders of chromatin structure beyond the nucleosome fiber.

To determine the intrinsic impact of DNA methylation (i.e. that independent of specific methylation recognition machinery) on genome organization, we use budding yeast as a model system. *S. cerevisiae*'s genome is deprived of any cytosine methylation 30 30 and its structure is very well characterized 31, 32, 33, 34, 35. It consists of a Rabl-like structure where the 16 centromeres are attached to the Spindle Pole Body (SPB) at one pole of the nucleus and the chromosome telomeres extend outward toward the nuclear membrane. *S. cerevisiae* does not have methylation/demethylation machinery and no methylated DNA binding domain has been characterized, precluding the impact of any evolution refined biological mechanism for regulating methylation imprinting and reading. Addition of methylation machinery to yeast should then yield to a perfect system to check for the intrinsic impact of DNA methylation in chromatin structure.

Different authors 36, 37 have shown that ectopic expression of either DNMT3a and DNMT3L or DNMT3a and DNMT1, could induce DNA methylation in yeast, but the level of methylation achieved in both cases were too low to perform genome wide analysis of the effects of DNA methylation on chromatin structure or gene expression. More recently, Morselli *et al.* achieved a higher level of methylation by

expressing DNMT3b at high level and collecting the cells at saturation 27. Their results demonstrate an intrinsic anticorrelation between nucleosome and methylation and a link between DNA methylation and H3K4 and H3K36 methylation. A comparable methylation pattern was just reproduced by Finnegan et al 38 expressing human DNMT3s in the yeast *Komagataella phaffii.*

Using a similar approach but expressing murine DNMT1, DNMT3L, DNMT3a and DNMT3b simultaneously in our *S. cerevisiae* cells, we achieve an even higher methylation rates than the one reported by Morselli *et al.* 27, creating a perfect model for the study of the intrinsic impact of methylation in chromatin structure and function. We demonstrated that even in the absence of any directing machinery, methylations occurred in a reproducible pattern reminiscent of that seen in mammals, with methylation concentrated at linkers and depleted at nucleosomes. High levels of methylation affect gene expression in a complex manner, altering specific pathways in the cell. Quite interestingly, Hi-C experiments revealed significant changes in global chromatin structure with a noticeable increase in gene condensation. Overall, our results demonstrate that DNA methylation affects chromatin structure and function even in the absence of specific methylation-recognition machinery, suggesting that the methylation imprinting has an intrinsic impact in chromatin structure and function.

Results

Description of the system

In order to reach a high level of DNA methylation, we expressed the 4 DNMTs simultaneously: yeast cells were transformed with a combination of 4 plasmids each of them expressing one murine DNA Methyl Transferase (DNMT). The *de novo* DNMTs, DNMT3a and 3b, and the maintenance DNMT, DNMT1, were expressed under the control of the *tet*O promoter while DNMT3L, the expression of regulatory DNMT was controlled by the *Gal1* promoter 39. The expression and stability of the DNMTs were assessed by Western blotting and no sign of protein degradation could be detected, even after 48 hours of induction (Supplementary Fig. S1A-D). Expression of the 4 DNMTs slightly affected cell growth and cell viability (Supplementary Fig. S2A and S2B). Flow cytometry analysis performed on non-synchronized cultures after 24 hours of induction showed a clear increase in the percentage of cells in G2/M in the samples expressing the DNMTs compared to the two control populations transformed with the empty plasmids suggesting that the slower growth could be due to a longer G2 phase (Supplementary Fig S2C).

Differential gene expression analysis (Supplementary Table S1) showed that DNMTs ectopic overexpression did not induce expression of genes normally activated by stress, suggesting that the effects that we observed are not caused by stress but are a direct consequence of DNA methylation.

DNA methylation was first assessed by HPLC/MS and then analysed at single base pair resolution in several independent transformants, using Illumina whole genome bisulfite sequencing (WGBS). This was done both for cells in exponential growth phase synchronized in G1, and for cells at saturation. Methylated cytosines were, as expected, almost exclusively found in CpG context (Table 1) and showed a very reproducible pattern from one sample to another (Fig. 1A, G), a result verified using longer reads (<10 kbp) obtained by Nanopore Technology (ONT), see Supplementary Fig. S3A-C), which allowed us to confirm the high level of methylation in long repetitive sequences in heterochromatin (Supplementary Fig. S4A,C). The global distribution of methylation across all CpG sites is shown in Supplementary Fig. S5A,B. For the cells synchronized in G1 we can see that ~50% of CpG sites have <5% methylation, but there are almost 20% of sites with methylation over 20%, and ~5% of sites with methylation over 50%. For the cells at saturation the methylation levels are generally higher, with 50% of sites having >15% methylation and 25% of sites having methylation over 50%.

Homogeneity of the samples

To check the homogeneity of the samples, we examined the long nanopore derived reads, which allow assaying the methylation status at a large number of CpG sites on the same DNA molecule. We can see (Supplementary Fig. S6) that the histograms of methylation level for the 2 control samples show an almost identical distribution with a peak at around 0.02, whereas for the methylated samples from cells at saturation the peak is around 0.33. The histogram for the methylated samples from replicating cells has a peak close to the non-methylated controls, but with a long right tail.

For the methylated samples, most CpGs have an intermediate level of methylation. To investigate whether the intermediate methylation was due to heterogeneity in the population of reads (implying that the sequencing library contains a mixture of highly methylated and lowly methylated DNA molecules for a given genomic region), we modelled the distribution of methylation within reads using a mixture model with three components, one non-methylated, one lowly methylated and one highly methylated. The fit of the observed data to the

mixture model was compared with that to a model with a homogenous read population using likelihood ratio tests (Supplementary methods).

The results in Supplementary Table S2 show that in the samples at saturation, almost 95% of the reads are highly methylated (with roughly 30% of CpGs on the read being methylated), indicating a fairly homogenous population. On the contrary, in the samples collected in exponential phase, only 20% of the reads are highly methylated with the remaining reads having a lower methylation rate (with roughly 8% of methylated CpGs on a read being methylated). The likelihood ratio tests show high significant support for that the presence of multiple components for both samples ($p<1.0e-16$ in both cases). The presence of two distinct populations of reads in the samples in exponential phase suggests that the DNA methylation maintenance machinery may be not fully functional, leading to a difference in average methylation between the original and daughter DNA strands

Effect on methylation of the different DNMTs

We tested the functionality of each DNMT in our system by comparing the DNA methylation pattern and levels obtained when only 3 of the 4 DNMTs were expressed (Fig. 1A,B). This was performed always with replicating cells, synchronized in G1 to make the data fully comparable. For all combinations of 3 DNMTs the methylation pattern was broadly the same (Fig. 1A). The overall methylation level, however, was 2 to 3 times lower with 3 DNMTs compared to when all 4 DNMTs were expressed, demonstrating that each DNMT is functional (Fig. 1B).

While the different DNMT combinations displayed broadly the same pattern of methylated CpGs, when examined in detail some differences emerged. To check whether these differences are due to intrinsic sequence specificity of the different DNMTs, logistic regression was used to assess the effect of local sequence context on the methylation rate at CpG sites. This analysis was applied to 6 samples (2 samples with all 4 DNMTs to assess the reproducibility of the results, and the 4 samples each lacking one of the DNMTs). Figure 1C shows that the two replicates with all 4 DNMTs being expressed provide very similar results, proving the robustness of the experiments. Removing DNMT1 or DNMT3L does not significantly alter the sequence preferences found in the experiments with the 4 DNMTs active. In contrast, removing either DNMT3a or DNMT3b has a large effect on the sequence context. In general, cells lacking DNMT3b show a strong bias for CpGs in the 5’ ATCGAG 3’ motif (with the percentage of methylated CpGs in an ATCGAG motif being more than six times lower when DNMT3b is removed, compared to the samples with one of the other DNMTs missing or to the sample with all 4 DNMTs induced). Cells lacking DNMT3a also show a sequence bias, in this case towards sequences that are more C rich. Note that the lack of methyl DNA binding protein makes these sequence preferences intrinsic of DNMTs and not the result of positioning of the methylases in certain sequences due to auxiliary proteins.

Pattern of methylation

Despite the lack of specific proteins directing epigenetic signaling of DNA, the pattern of methylation obtained in our model organism is very similar to that observed in higher eukaryotes, with DNA methylation being low at the Transcription Start Site (TSS) increasing toward the end of the genes and reaching a maximum at the Transcription Termination Site (TTS) (Fig. 1D). Previous studies have linked DNA methylation with histone post translational modifications, DNMT3L being recruited by unmethylated H3K4 and DNMT3A and 3B recruited by H3K36me2 and H3K36me3 respectively27, 40, 41, 42, 43. We confirmed that the H3K4 methylation pattern in the stationary cells both in the control and in the methylated sample was the same as the one described in exponential cells, and that DNA methylation was anti-correlated with H3K4 trimethylation (Fig. 1E), suggesting that H3K4 and H3K36 methylation are capable of tightly controlling DNA methylation even upon DNMTs overexpression by a direct mechanism.

DNA methylation and Nucleosome positioning.

To investigate the correlation between DNA methylation and nucleosome positioning, we obtained MNase-Seq data of the samples in stationary phase incubated with all 4 DNMTs or with empty plasmids. We observed that DNA methylation was anti-correlated with nucleosomes and tended to be accumulated in the linker regions (Fig. 2A,C). These results agree with a bulk of *in vivo* 25, 26, 27, 28, *in vitro* and *in silico* data18, 19 and ruled out the suggestion that methylation is favored at well positioned nucleosomes 22, 23.

We observed a small (4%), but very significant increase (pvalue 2.2e−16) in the number of fuzzy nucleosomes and a mirror decrease in the number of well positioned nucleosomes in the methylated samples (Supplementary Table S3). In well positioned nucleosomes, DNA

methylation was almost absent at the dyad and increased toward the entry and exit points of the nucleosome, while fuzzy nucleosomes had higher methylation levels with a more constant level across the nucleosome (compare Fig. 2C with Fig. 2D). Detailed analysis (Fig. 2E), failed to detect any periodicity in the methylation within the nucleosome bound sequence, arguing against a preferential methylation of DNA at accessible sites of the nucleosome. Considering that all methylases were active, this suggests that (at least in the absence of directing machinery) all DNA segments covered by histones are equally inaccessible to methylation imprinting.

Figure 3A shows some representative nucleosome fiber structures for the DUG2 gene (chromosome II) as a test case, modeled from Mnase-seq signals for control and methylated cells. Both ensembles of structures reproduce the experimental results to a great extent (Figures 3B, C and Supplementary Fig. S7A, B) and confirm the presence of fuzzier nucleosomes in methylated cells. Due to the presence of fuzzier nucleosomes, a wide range of nucleosome configurations are possible leading to the sampling of very elongated conformations with large radius of gyration (see Fig. 3A, lower panel, and Supplementary Fig. S7C). The diversity of nucleosome arrangements was further analyzed by comparing the 3D distances between N and $N+x$ nucleosomes (Supplementary Fig. S7D), which in methylated cells show a more dispersed profile. Finally, it is noticed that regions with high degree of methylation probability are mainly linker DNA (Fig. 3B, 3C and Supplementary Fig. S7B), though the probability of finding some nucleosomal DNA methylated is not insignificant, as shown in some fiber configurations (Fig. 3A, lower panel). This is because those nucleosome positions are retained despite having low coverage in the Mnase-seq signal.

Analysis by Nucleosome Dynamics 44 of the crucial region around the promoter demonstrates that high level of methylation induces more dynamic and fuzzy nucleosomes (Fig. 4A), as well as a statistically significant ($p < 2.2e-16$) narrowing of Nucleosome Free Regions (NFRs; Fig, 4B,C). These changes, which are typically considered to be signals of gene inactivation 45, 46 cannot be explained here by the coordinated effect of methyl-DNA binding proteins coupled to chromatin remodelers and must be then considered intrinsic to the changes on physical properties of DNA induced by methylation and its direct impact in protein-DNA interaction. The fact that high methylation is correlated also with narrower NFR also at the 3' end of the genes (Supplementary Fig S8) agrees with this hypothesis.

DNA methylation and Gene expression

Despite the lack of any directing mechanism, the methylation pattern is quite homogenous throughout the gene only in lowly expressed genes while in the highly expressed ones, DNA methylation is low at the promoter and increases toward the end of the gene, suggesting a direct link between DNA methylation and gene expression (Fig. 5A) in the absence of any specific methylation-recognition mechanism. A differential expression analysis (Fig. 5B) shows that genes which are very lowly methylated do not change their expression level between the control and transformed cells, while high methylation levels lead to important changes in gene activity in both directions: towards greater and lower expression (Fig. 5C and Supplementary table S1). Particularly, we see a very strong correlation between gene expression and methylation level for a subset of genes involved in meiosis and that appear to share a common sequence in their regulatory region (Fig. 5C, D), corresponding to the binding site of Ume6p, a subunit of the histone deacetylase complex Rpd3p known to repress early meiotic gene expression. It is tempting to hypothesize that methylation of a Ume6p binding site known as URS1, could affect intrinsically Ume6p binding directly (through changes in direct protein-DNA interactions or indirectly through changes in chromatin structure), leading to a deregulation of its target genes. Supporting this hypothesis, we observed that the level of expression of the target genes increases proportionally with the level of methylation of the Ume6p binding site (Supplementary Fig. S9 and Supplementary Table S4). Also, for most of these genes, we observe a 5–10bp shift of the –1 or +1 nucleosome (Supplementary Fig. S10), consistent with changes in expression. In summary, it seems that two physically driven mechanisms: changes in protein-DNA binding interactions due to the presence of a methyl group and methylation-induced nucleosome rearrangements work in a coordinated way to induce a change in gene activity which would be typically assigned to the effect of specific methyl DNA binding proteins, which are absent in *S. cerevisiae.*

DNA methylation and genome 3D structure

We performed Hi-C experiments in control and methylated populations at saturation to explore the intrinsic effect of DNA methylation in the global chromatin structure. As shown in figures 6A and 6B (and Supplementary Fig. S11A, B), DNA methylation leads globally to an increase in *cis (*especially close to the centromeres (< 100 kb); 6C and 6D) and a very significant decrease in *trans* contacts (Fig. 6C and 6E). To obtain further insights into the effect of DNA methylation on chromatin structure, we modeled the spatial organization of each chromosome using a restraint-based model derived from the interaction counts (see Methods) from which we obtained ensembles reproducing with an astonishing quality the HiC maps (Supplementary Fig S12). Clearly, in all chromosomes (except the short chrI), the 100kb region centered around their centromeres is more condensed upon DNA methylation (Fig. 6F), decreasing the overall chromosome

flexibility with the effect being very clear for chromosomes V, IX, XI and XV (Fig. 6G and Supplementary Figs. S12). Very interestingly, the nucleosome depletion that is observed at the centromeres in stationary phase in the control sample (compared to cells in exponential growth) is not detected in the methylated one (Supplementary Fig. S13), suggesting that the profound reorganization of the centromeric region occurring in quiescent cells is reduced in the methylated samples.

The three-dimensional arrangement of telomeres is also largely modified by methylation, as we observed a significant decrease in the number of interactions (Supplementary Fig. S14), resulting in a large dispersion of telomeres in methylated cells, something that is visible in the HiC-derived ensembles (Fig. 6H, I). It is worth noting that telomeres tend to cluster in quiescent cells to form hyperclusters 47 but it seems that this reorganization is not happening when DNA is methylated. Once again, intrinsic changes linked to methylation lead to a 3D organization that is closer to the one described in exponential growth (which is when the expression of the DNMTs started to be induced), than the one we and others observed in quiescent cells.

Comparing the changes in interactions between the control and methylated samples for each individual chromosome (Supplementary Fig. S15), we observed the strongest effects of methylation for chrIII (largest increase in intra-chromosomic-contacts, Fig. 7A-E) and for chrXII (largest decrease in intra-chromosomic contacts, Fig. 7 F-H). Looking closer at chrIII, we noticed an important gain of interactions in the methylated sample between the left telomeric region containing the silenced HMLα and the peri- centromeric region (Fig. 7A-C). This gain of interaction correlates with a significant decrease of the distance between the HMLα and the MATa loci (Fig. 7D), which leads to very significant changes in the ensemble of chromosome conformations compatible with the HiC restrains (Fig. 7E).

ChrXII is a very peculiar chromosome, as it carries the rDNA locus that consists of 150- 200 repeats of the rDNA genes and localizes at the nucleolus where ribosomal RNAs are transcribed. In the control sample, interactions between the upstream and downstream regions are found, which can be explained by the global decrease of transcription occurring in stationary phase leading to a more compact nucleolus compared with exponential growth 48. However, in methylation conditions (Fig 7F-H), the segregation between the upstream and downstream regions of the chromosome is very strong, reproducing again the situation expected for a cell in exponential growth. The absence of interactions between the two domains probably allows a general increase in the relative flexibility that could explain why the general reduction in flexibility related to methylation seen in other chromosomes is inverted for this chromosome (Figure 6G).

Globally, the control samples show a quite spherical nucleus (Fig. 8A) with the chromatin concentrated in the exterior (Fig. 8B), while in the methylated samples the nucleus tends to elongate with a more dense packing of chromatin in the interior (Fig. 8D and E). In both cases, centromeres are all localized to one pole of the structure, organized as a rosette (Fig. 8C and F), but they do not appear as clustered as reported for the interphasic nucleus 34, 35, 49.

Discussion

Yeast expressing murine methyl transferases show, *in vivo,* a specific, reproducible, pattern of DNA methylation that is very similar to that of genes-containing regions in mammals. This is a striking result as yeast is an organism deprived of any DNA methylation machinery and is not ready to place or recognize methylation marks. Methylation leads to a slight decrease in the viability and in the doubling time maybe caused by a longer than normal G2/M phase. However, those changes are moderate, and again it is quite surprising that an organism not prepared to have methylated DNA tolerates well a large amount of methylation in its genome.

Our synthetic model system helped us to highlight some previously unknown intrinsic sequence specificity for the methyl-transferases, i.e. those that cannot be explained by accessory proteins. For example, little sequence specificity is found for DNMT1 and DNMT3L, while significant sequence specificities are found for DNMT3a and DNMT3b. Thus, cells lacking DNMT3b show a strong bias for methylation of CpG in 5’ ATCGAG 3’ motif, while cells lacking DNMT3a seems to have a bias toward CpG sequences embedded in C-rich environments. Again, these differences are intrinsic and not coupled to any specific directing mechanism, which demonstrate the intrinsic ability of methylation to alter cellular phenotype.

Methylation in our engineered yeast model is preferentially located between nucleosomes, and when it occurs in nucleosome-occupied regions, it is associated with significant alterations in nucleosome positioning reflected in an increase in nucleosome fuzziness. The fact that methylated DNA is less frequent at nucleosomes confirms our previous *in silico* and *in vitro* models18, 19 but does not rule out the possibility that in higher eukaryotes methylated-DNA binding domains might stabilize the presence of methylated CpG in mammal nucleosomes, leading to a situation of “loading-spring” which might facilitate fast nucleosome reorganization upon release of the stabilizing protein. There is, however, no question that methylation and nucleosome position are intrinsically anti-correlated. Furthermore, our results demonstrate that there is not any short-distance periodicity pattern which might indicate methylation at periodically exposed

regions of nucleosome DNA. Very interestingly, when methylation occurs in the promoter region it tends to narrow the NFR region, a fingerprint of lowly expressed genes in mammals, which is found here in absence of any methylation-specific chromatin remodeler.

In mammalian cells, the relationship between methylation and gene expression is complex, with high levels of gene expression often associated with low promoter methylation but elevated gene body methylation, and the causality relationships are not clear. Our simple synthetic system allows a more direct interrogation to the relationship between methylation and gene expression. We observed that despite the lack of specific methylated DNA binding domains, highly and lowly expressed genes have quite different methylation profiles, with much higher levels of methylation near (±850 bp) the TSS of silent genes while very actively expressed genes have much higher methylation levels at the TTS. In the absence of specific proteins modulating this profound difference, we can speculate that nucleosome positioning is one of the main factors responsible for this differential behavior, which suggests that methylation and nucleosome positioning might act in concert in the regulation of gene function in mammals, adding an extra layer of control of gene expression.

Our results suggest that methylation can also directly affect the binding of a transcription factor highlighting another intrinsic role of methylation in gene activity, which is often ignored, but that can be important in modulating DNA-protein binding free energy. Interestingly, we found that the methylation-induced change in protein- DNA binding is responsible for a dramatic increase of expression of early meiotic genes. In that case, our results suggest that the methylation affects the intrinsic binding of the histone deacetylase complex Rpd3p, thus hindering the placement of repressive marks on these genes.

Analysis of HiC data shows that the characteristic Rabl configuration previously described 31, 34, 35 is maintained in stationary phase both in the control and in the methylated sample. However, methylation induces an increase of intra-chromosomic contacts and a significant decrease of the inter-chromosomic ones, which leads to significant changes in the chromosome conformation. We observed that a significant (Fisher's exact test, $p = 1.76e-7$) proportion of the interactions that are gained or lost upon methylation involves regions containing one or several tRNA genes, supporting the involvement of RNA polymerase III in the overall organization of chromatin structure 50.

Changes in heterochromatin regions of *S. cerevisiae* (telomeres, the mating type locus and the rDNA locus) are also clear upon methylation, mimicking the situation found in mammals. First, we observed a general loss of interactions between telomeres which leads to a significant change in chromosome structure and that can be explained by a generally higher rigidity of the chromatin upon methylation. Second, we observed that chromosome III underwent some conformational changes in response to methylation, with the silenced HMLα locus getting closer to the MAT locus, a conformation expected in exponentially growing MATa cells 51, 52, but not in stationary cells. Finally, the separation commonly observed in yeast between the two regions of chrXII delimited by the rDNA locus is clearly weaker in our cells in stationary phase when compared to the structure published for cells in exponential phase 31, but the situation is reverted in methylated samples (in stationary phase), suggesting that methylation of the rDNA can freeze the heterochromatin in exponential phase-like conformation.

Our previous results 18, 19 strongly suggest that in general, methylation increases the stiffness of DNA and this should impact the structure of the chromatin both locally and globally. This is confirmed here by Hi-C experiments coupled to 3D modeling that show a condensation of the centromeric region for each chromosome individually, resulting in more rigid and condensed chromosomes unable to transiently interact with other chromosomes. Locally, at the gene level, the situation is more variable, as the higher rigidity of methylated DNA leads to fuzzier nucleosome architectures and longer linkers (as shown by MNase-seq experiments), which in turn can lead to a local increase in the structural variability of the nucleosome fiber, dominated by the length of the linkers. Overall these results demonstrate that the local increase in rigidity induced by methylation have direct consequences on the global structure of chromatin, even in the absence of protein machinery ready to recognize methylation signals.

Taken together, our results suggest that in the absence of any mechanism directing methylation and reading the methylation signals, methylation leads to significant changes in chromatin structure and function. Very interestingly and quite surprisingly, many of these methylation effects resemble those annotated to methylation in mammals, where an exquisite machinery for imprinting and reading of methylation signals is present. It can be concluded then that changes in physical properties of DNA induced by methylation (reflected in alterations in DNA deformability or in DNA-protein interactions) can be responsible for a significant part of the phenotypic effects triggered by DNA methylation in mammals and that the cellular machinery involved in DNA methylation just amplifies an intrinsic signal coded in the physical properties of DNA.

Changes in methylation across the genome are key features of developmental pathways in both normal differentiation 53, 54 and many cancers 2. We observed that in our synthetic model, methylation makes the 3D genomic structure in the stationary phase closer to that of

G1 cells than in the unmethylated samples, which suggest that one of the intrinsic roles of DNA methylation could be to freeze chromatin conformation to maintain the state of cell.

Materials And Methods

Plasmid construction

pYADE4 yeast plasmids encoding full length DNMT1 and DNMT3a with modified sequences around the translation start sites were kindly provided by Dr Jan Fronck, pYES3/CT encoding DNMT3b was provided by Dr Shen Li 36, 55. DNMT3L cloned into pYES3/CT to produce a Nterminal FLAG tagged DNMT3L was provided by Dr Jia-Lei Hu 37.

pCM188 (marker cgURA3) and pCM185 (marker cgHIS or cgLEU), centromeric vectors 39 which differ for the number of Tet operators (respectively 2 and 7) were kindly provided by Dr Jessie Colin.

*Sma*I restriction fragment (from pYADE4-DNMT1) containing full length DNMT1 cDNA was inserted at *Pme*I site of pCM185 (LEU) to give pCM185(LEU)-DNMT1.

*Bam*HI-*Mlu*I restriction fragment (from pYADE4-DNMT3a) containing full length cDNA from DNMT3a was ligated to pCM185 (HIS) linearized with *Bam*HI and *Mlu*I to give pCM185(HIS)-DNMT3a.

*Bam*HI-*Not*I restriction fragment from pYES3/CT-DNMT3b containing full length DNMT3b was ligated to pCM188 (URA) linearized with *Bam*HI and *Not*I to obtain pCM188 (URA)-DNMT3b.

Yeast strains and culture conditions

Strain YPH499 (*Mata ura3–52 lys2–801 ade2–101 trp1-Δ63 his3-Δ200 leu2-Δ1)* was transformed with 2, 3 or 4 expression plasmids by the standard lithium acetate procedure. Transformants were selected on plates of appropriate selective medium with 2% Raffinose and 10µg/ml doxycycline to repress any expression.

Selected transformants (2 to 4 transformants per combination of plasmids) were grown on selective liquid medium with 2% Raffinose and 10µg/ml doxycycline up to OD600 = 0.5. Then, yeast cells were spun 10min at 1000x g, washed twice with sterilized water, and resuspended into selective media with 1% Raffinose and 2% Galactose without doxycycline to allow expression of all four DNMTs. For experiments on synchronized cells, cells were treated with alpha-factor (3µM final) for 4 hours to synchronise cells in G1 or with Nocodazole to synchronize cells in G2.

After different times of induction, cells were collected and treated for subsequent experiments: protein extraction for western blotting, gDNA extraction for whole genome bisulfite sequencing, RNA extraction for RNA-sequencing or Semi-intact cell preparation for Mnase digestion and nucleosome mapping.

Flow cytometry analysis

0.5 ml of culture (OD_{600} = 0.6–0.8) were collected and centrifuged for 5 mn at 1000 g at RT. Pellets were washed twice with 1x ice-cold PBS and resuspended in 50 µl of 1x ice- cold PBS. 20 µl of cells were fixed with 1 ml of 70 % EtOH overnight at 4 °C. Samples were washed with 1x Saline Sodium Citrate buffer (SSC; 150 mM NaCl, 15 mM Na citrate, pH 7.8 for 20x SSC). The pellet was resuspended in 0.5 ml of 1x SSC, treated with 0.5 mg/ml RNase A (Roche) for 1.5 h and then with 0.5 mg/ml Proteinase K (Roche) for another 1.5 h at 50 °C. After incubation, cells were briefly sonicated for 10 mn, medium potency, by using the *Bioruptor* system (intervals of 10 s on–20 s off). 250 µl of the cells were added in 0,5 ml of 1x SSC containing 1 µM Sytox Green (Sigma) and were incubated 10–20mn in the dark (room temperature) before analyzing the DNA content using a Beckam Coulter GalliosTM flow cytometer.

HPLC/MS

HPLC/MS/MS analysis was based on the protocol described 56 54. A Kinetez 2.6 µm HILIC 100A column (150 mm × 4.6 mm) (Phenomenex) and a Acquity UPLC system (Waters Corp., Milford, MA, USA) coupled to a mass spectrometer API 3000™ (AB Sciex, Foster City, CA, USA) triple quadrupole working in MRM(multiple reaction monitoring) method in positive mode. Two eluents were used: eluent A2

(Acetonitrile) and eluent B1 (0.1 M ammonium formiate adjusted at pH 3.2) with a isocratic gradient 8 min of total running time at 90 % A and 10 % B for the nucleosides elution. The separation was performed in a flow of 1400 µl min−1, with 10 µl injection volume and two replicates each, totaling two biological replicates and two technical replicates of each sample. The standard nucleosides cytosine and methyl-cytosine (Sigma) were diluted in HCl 0.01N and stored at −20 °C. The m/z transitions from 112 to 95 (cytosine) and from 126 to 81 (methyl cytosine) were chosen for MRM experiments. The peak area obtained was analyzed by Analyst 1.4.2 (AB Sciex). Quantification (%) was performed according to 5mdC concentration divided by 5mdC concentration plus dC concentration multiplied by 100.

Western Blot

Proteins were extracted by resuspending the pellet of cells from a 20ml cultures at OD600 = 1 in 400µl of RIPA buffer (50mM Tris pH7.5, 150mM NaCl, 1% NP40, 0.5% NaDeoxycholate, 0.1% SDS) containing 1mM PMSF and protease inhibitors (cOmplete ULTRA Tablets, Mini, EASYpack, Roche). 400µl of glass beads were added and samples were processed using FastPrep (MP) for 3 times for 20sec pulses @4.5m/s. After centrifugation 5min at 5000rpm, supernatant were recovered and quantified by bradford. 20 µg of protein were loaded on 6 or 8 % acrylamide gel and subjected to PAGE, proteins were then transferred onto an immobilon membrane (millipore) for subsequent hybridization with anti-DNMT1 (ref ab87654, Abcam), anti-DNMT3a (ab2850, Abcam), anti-DNMT3b (ab122932, Abcam) or anti-Flag (F7425, Sigma) antibody overnight followed by secondary antibody anti Rabbit (Goat)-HRP conjugated (65–6120, Invitrogen). The signal was revealed using ECLTM prime WB detection reagent (Amercham, GE Heathlcare).

DNA METHYLATION PATTERN

Whole-genome bisulfite sequencing (WGBS)

WGBS was performed following the procedure outlined in 9. Briefly, genomic DNA (1- 2µg) was spiked with unmethylated λ DNA (5 ng of λ DNA per µg of genomic DNA) (Promega). The DNA was sheared by sonication to 50–500 bp using a Covaris E220 and fragments of size 150–300 bp were selected using AMPure XP beads (Agencourt Bioscience Corp.). Genomic DNA libraries were constructed using the Illumina TruSeq Sample Preparation kit (Illumina Inc.) following the Illumina standard protocol: end repair was performed on the DNA fragments, an adenine was added to the 3’ extremities of the fragments and Illumina TruSeq adapters were ligated at each extremity. After adaptor ligation, the DNA was treated with sodium bisulfite using the EpiTexy Bisulfite kit (Qiagen) following the manufacturer’s instructions for formalin-fixed and paraffin-embedded (FFPE) tissue samples. Two rounds of bisulfite conversion were performed to assure a conversion rate of over 99%. Enrichment for adaptor- ligated DNA was carried out through 7 PCR cycles using the PfuTurboCx Hotstart DNA polymerase (Stratagene). Library quality was monitored using the Agilent 2100 BioAnalyzer (Agilent), and the concentration of viable sequencing fragments (molecules carrying adaptors at both extremities) estimated using quantitative PCR with the library quantification kit from KAPA Biosystem. Paired-end DNA sequencing (2x100bp) was then performed using the Illumina Hi-Seq 2000.

Read mapping and estimation of cytosine methylation levels

The WGBS reads were processed using the gemBS pipeline v3.0 57 using as reference *S. cerevisiae* S288c. Reads with MAPQ scores < 20 and read pairs mapping to the same start and end points on the genome were filtered out after the alignment step. The first 5 bases from each read were trimmed before the variant and methylation calling step to avoid artifacts due to end repair. For each sample, CpG sites were selected where both bases were called with a Phred score of at least 20, corresponding to an estimated genotype error level of < = 1%. Sites with >500x coverage depth were excluded to avoid centromeric/telomeric repetitive regions. CpGs were considered methylated when the number of mapped reads was larger than 10 and the estimated methylation percentage was above 0.1.

DNMT specificity analysis

We extracted two bases downstream and upstream from each CpG (having at least ten WGBS reads mapped) and trained a logistic regression model (using R) for the number of converted and non-converted Cs, using the extracted motifs as predictors for each WGBS sample (samples removing one of the DNMTs, T859, T860, T861 and T869; and two samples with the four DNMTs, T862 and T863). We computed for each sample the effect of each motif and its standard deviation, and used it to determine those with a significant effect on methylation level (estimated effect above two standard deviations). We found motifs specific for each sample lacking one of the DNMTs (motifs with significant effect in the sample removing one DNMT but not significant in the sample with all DNMTs) and compared their relative frequencies in all samples.

Nanopore sequencing

Suspensions of spheroplasts from methylated and control *S. cerevisiae* strains were loaded on Sage Science gel cassettes to perform lysis under electrophoretic conditions. DNA content in each sample was estimated by the cell count. A number of spheroplasts equivalent to 10µg of genomic DNA were resuspended in 70 µl of HLS Suspension buffer (Sage Science, Mammalian white Blood cell suspension kit, #CEL-MWB1) and loaded on the gel cassettes (Sage Science, SageHLS HMW DNA extraction kit #HEX–0012).

The custom Sage HLS (Sage Science) protocol used (Extraction Collection DC55V 1h15m) was accommodated for the yeast small chromosome sizes. This custom protocol did not include a DNA fragmentation step. In brief, during the extraction step, the High Molecular Weight (HMW) yeast gDNA was bound in agarose while the solubilised and degraded proteins and other contaminants were kept in solution. The Sage Science Buffer A was used as a lysis buffer for this step. In the last step of the protocol, the HMW DNA was retrieved from the gel through an automated elution process that was optimized to elute all the yeast chromosomes in the elution module number 2 of the cassette.

Elution modules 1, 2, 3 & 4 were selected for the library preparation of the control and methylated *S. cerevisiae* samples. For each condition, the selected elution modules were pooled, purified with 1-fold excess of Agencourt AMPure XP beads (Beckman Coulter, A63882) and eluted in water. Two barcoded libraries containing both type of samples were prepared using the Oxford Nanopore Ligation sequencing kit (ONT, SQK-LSK109) combined with the Oxford Nanopore Native Barcoding Expansion kit (EXP-NBD103 1D) following manufacturer's instructions.

After connecting the flows cells to the MinION Mk1b device, the MinKNOW interface QC (Oxford Nanopore Technologies) was run in order to assess the flow cell quality. Once the priming of the flow cell was finished, from 200ng to 600ng of the final barcoded library was loaded into R9.4.1 FLO-MIN106 or FLO-MIN106D flow cells and the sequencing data were collected during 48 hours. The quality parameters of the sequencing runs were further monitored by the MinKNOW platform in real time. The MinKNOW versions used was 1.15.4. The basecalling was performed using Guppy 2.3.7.

Reads were mapped using minimap2 2.9-r720, and CpG methylation was called using nanopolish 0.11.0.

NUCLEOSOME MAPPING

Semi-intact Yeast cell preparation

Semi-intact cells were prepared as previously described 58. Briefly, cells were grown at 30°C in 300 ml YPD to = 1 x 107 cells/ml. For each 250 ml of cells (107cells/ml), semi- intact cells were prepared as follows. Cells were collected by centrifugation (700 g, 7 min, RT), resuspended in 25 ml 100 mM Pipes, pH 9.4, 10 mM DTT, incubated with gentle agitation at 30°C for 10 min, and collected by centrifugation (1,000 g, 5 min, RT). Cells were resuspended in 6 ml YP, 0.2% glucose, 50 mM KPO4, pH 7.5, 0.6 M sorbitol. 10u zymolase was added, and the suspension was incubated with gentle shaking 30°C for 30 min. Spheroplasting was monitored by light microscopy. Great care was taken not to overdigest cells to avoid lysis. Spheroplasts were collected by centrifugation at 1,000 g for 5 min at RT, re-suspended with a plastic pipette in 40 ml YE 1% glucose, 0.7 M sorbitol, and incubated with gentle shaking at 30°C for 20 min. Spheroplasts were collected by centrifugation (1,000 g, 5 min, RT) and washed twice at 4°C with cold permeabilization buffer (20 mM Pipes-KOH, pH 6.8, 150 mM K-Acetate, 2 mM Mg- Acetate, 0.4 M sorbitol. The final pellet was resuspended in 1ml cold permeabilization buffer containing 10%(v/v)DMSO. 100µl aliquots were placed in 1.5 ml microfuge tubes and frozen slowly above liquid N_2 and stored at –80°C.

MNase-seq

0.4 x 109 semi-intact cells were digested with micrococcal Nuclease (MNase), 1.5 unit at 37ºC for 30min with 3mM CaCl2. The reactions were stopped by addition of EDTA to a final concentration of 0.02 M and subsequently incubated with RNase A (0.1 mg) for 4h at 37ºC and further treated with Proteinase K at 37ºC o/n. DNA was purified using phenol–chloroform extraction and concentrated by ethanol precipitation.

The percentage of mononucleosomal DNA fragments was examined by 2% agarose gels. Furthermore, the integrity and size distribution of digested fragments were determined using the microfluidics-based platform Bioanalyzer (Agilent) prior to library preparations following Illumina standard protocol. The short-insert paired-end libraries for MNase sequencing were prepared with PCR free protocol using KAPA

Library Preparation kit (Roche). In short, 2.0 micrograms of Micrococcal nuclease (MNase) digested genomic DNA from *S. cerevisiae* was end-repaired, adenylated and ligated to Illumina platform compatible adaptors with dual indexes (Integrated DNA Technologies). The adaptor-modified end library was size selected and purified with AMPure XP beads (Agencourt, Beckman Coulter). The final libraries were quantified by Kapa Library Quantification Kit for Illumina platforms (Roche).

The libraries were sequenced using TruSeq SBS Kit v4-HS (Illumina), in paired-end mode with a read length of 2x76bp following the manufacturer's protocol. Images analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (1.18.66.3).

Nucleosome calling

MNase-seq paired-end reads were mapped to yeast genome (sacCer3, Apr. 2011) using Bowtie 59 aligner, allowing a maximum of 2 mismatches and maximum insert size of 500 bp. Output BAM files were imported in R 60 58 and quality control was performed with htSeqTools package to remove PCR artifacts 61. Filtered reads were processed with nucleR package 62 as follows: mapped fragments were trimmed to 50bp maintaining the original center and transformed to reads per million. Then, noise was filtered through Fast Fourier Transform, keeping 2% of the principal components, and peak calling was performed using the parameters: peak width 147 bp, peak detection threshold 35%, maximum overlap of 80 bp, dyad length 50 bp. Nucleosome calls were considered well- positioned when nucleR peak width score and height score were higher than 0.6 and 0.4, respectively, and fuzzy otherwise.

Nucleosome Dynamics

NucDyn R package 44 was used to find changes in nucleosome organization between control and methylation induced samples. P-values quantifying the nucleosome change were obtained running NucDyn with the following parameters: maximum difference of 70, maximum length of 140, minimum number of reads to report a shift of 3, shifts threshold of 0.1, indels minimum number of reads to report evictions and inclusions (indels) of 3, indels threshold of 0.05.

GENOMIC ANNOTATION

Data was annotated from the UCSC gene track that contains 6692 genes. We discarded genes that are described as "Putative" or "Dubious" and genes located in the mitochondrial chromosome. We used gene lengths to normalize methylation proportions, nucleosome coverages and CpG density partitioning each gene in 137 bins (each bin has on average 10 bp since the mean length of yeast genes is 1369 bp).

GENE EXPRESSION

mRNA library preparation and sequencing

The RNASeq libraries were prepared from total RNA (extracted by the standard hot phenol protocol) as follows. Total RNA quality and quantity were assessed using Qubit® RNA HS Assay (Life Technologies) and RNA 6000 Nano Assay on a Bioanalyzer 2100 (Agilent). The RNASeq libraries were prepared following KAPA Stranded mRNA-Seq Illumina® Platforms Kit (Roche) following the manufacturer´s recommendations. Total RNA (500ng) was enriched for the polyA mRNA fraction and fragmented by divalent metal cations at high temperature. In order to achieve the directionality, the second strand cDNA synthesis was performed in the presence of dUTP. The blunt-ended double stranded cDNA was 3´adenylated and Illumina platform compatible adaptors with unique dual indexes and unique molecular identifiers (Integrated DNA Technologies) were ligated. The ligation product was amplified with 15 PCR cycles and the final library was validated on an Agilent 2100 Bioanalyzer with the DNA 7500 assay (Agilent).

The libraries were sequenced on HiSeq2500 (Illumina) using TruSeq SBS Kit v4-HS (Illumina), in paired-end mode with a read length of 2x76bp following the manufacturer's protocol. Images analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (1.18.66.3). We generated over 20 million paired-end reads for each sample in a fraction of a sequencing lane.

RNA-seq data processing and analysis

RNA-seq reads were mapped against the yeast reference genome (Sacharomyces_cerevisiae.R64–1–1+plasmid) using STAR version 2.5.2a 63 with ENCODE parameters. Quantification of annotated genes (ensembl release 87) was done using RSEM version 1.2.28 64 with default options. Heatmaps with the top differentially expressed genes was perform with the pheatmap R package. Differential expression between conditions was performed with DESeq2 version 1.18 65 with default parameters.

3D GENOME STRUCTURE

Hi-C libraries

The protocol was performed as previously described 66 with a few modifications. 100 ml of yeast culture were crosslinked with 3% formaldehyde during 20 min and quenched with Glycine 125mM during 5 min at RT. Cells were crushed during 30 min in liquid nitrogen and the chromatin was digested with *Hind*III. The DNA overhangs were filled- in with dNTP including Biotin–14-dATP, and the resulting blunt end were ligated. After ligation, samples were purified with phenol:chloroform and DNA was precipitated with ethanol.

The paired-end Hi-C-sequencing libraries were prepared with KAPA Library Preparation kit (Roche) with some modifications. The biotin marked and de-crosslinked DNA was sheared to a size of 300–500bp on Covaris™ LE220 (Covaris) focused-ultrasonicator. The fragmented DNA was end-repaired, adenylated and the biotin-tagged DNA was pulled down using the Dynabeads™ MyOne™ Streptavidin C1 beads (Thermo Fisher Scientific). The biotinylated fragments were ligated to Illumina platform compatible adaptors with unique dual indexes and unique molecular identifiers (Integrated DNA Technologies) and enriched by 12 PCR cycles by KAPA HiFi PCR Kit (Roche).

The libraries were sequenced on HiSeq2500 (Illumina) using TruSeq SBS Kit v4-HS (Illumina), in paired-end mode with a read length of 2x76bp following the manufacturer's protocol. Images analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (1.18.66.3).

Hi-C data processing and normalization

We processed Hi-C data using TADbit 67 (*https://github.com/3DGenomes/tadbit)* for quality control, mapping and filtering. First, quality control was performed with the FastQC protocol implementation in TADbit. Then, reads were mapped to the reference yeast genome (sacCer3, Apr. 2011) with a fragment-based strategy. Afterwards, non- informative contacts (self-circle, dangling-end, error, duplicated and random breaks) identified by TADbit were filtered-out, obtaining 32–37 million valid interactions per experiment. Off-target contacts (neither end of the read mapped to one of the capture regions) were also discarded (full details of the number of excluded reads are given in Supplementary Table S5). Finally, contact matrices were created from valid reads at 5 kb resolution with the corresponding TADbit module, and low frequency bins were removed.

Contact matrices were transformed to *.hic* format for visualization in Juicebox 68 using the *pre* command, and normalized with the Balanced method 69.

Differential Hi-C analysis was performed using the R/Bioconductor package diffHic 70. The mapped Hi-C data were filtered and the differential interaction analysis between the control and methylated samples (using the two replicates for each treatment) was performed using the procedure recommended in the diffHic manual.

Hi-C-based chromatin 3D structure

High resolution Hi-C data at 5 kb was used to obtain the 3D structure, conformation and dynamics of entire yeast chromosomes and their context inside the nucleus. The Hi-C technique provides interaction contacts between DNA fragments. The interaction counts or frequencies between two *loci i* and *j* (f_{ij}) can be converted to spatial 3D distances between those *loci* (d_{ij}) by an inverse relationship (equation 1),

$d_{ij} = \gamma \ / \ f_{ij}^{\alpha}$ (1)

where γ represents the scale of the structure and is usually taken to match experimental distances between selected genomic regions, and the precise value of α depends on the organism under study, the genomic distance, and the resolution of the Hi-C map and needs to be fitted 71, 72, 73. In the present work, γ was taken for the model to match the size of the cell nucleus measured by confocal microscopy and α was fitted to maximize the correlation between experimental and modeled contact maps.

Since Hi-C interaction counts are known to present several biases, such as mappability of fragments, GC content, and fragment length, they were normalized using iterative correction and eigenvector decomposition 74. Finally, the output of the conversion procedure was a matrix containing equilibrium distances ($r0$) for the different interacting *loci.* To remove the background noise, a cutoff of two times the median of all trans contacts (i.e., between *different* chromosomes) was applied to the HiC contact map to define *interacting regions.*

The chromosome model was built as a chain of beads, each bead representing a genomic region that corresponds to a bin from the Hi-C map. Spatial equilibrium distances were obtained from equation 1 as explained above. The distances between interacting beads

(r) were restrained near their equilibrium length during the simulations by penalizing with a harmonic potential (equation 2) when approaching at shorter distances or moving away at longer distances than the equilibrium. A tolerance of one bead radius was applied, thus resulting in a flat-welled parabola potential (equation 2),

$E = k(r\ r')^2$ (2)

where $r' = r_0\ r_{bead}$ for $r < r_0\ r_{bead,}$ $r' = r$ when r lies within r_0 +/- r_{bead} and $r' = r_0 + r_{bead}$

when $r > r_0 + r_{bead.}$

To ensure proper connectivity of the fiber, consecutive beads were also bound by a harmonic potential but with a force constant five orders of magnitude stronger than that applied to interacting non-consecutive beads. An excluded volume was defined for each bead by a standard Lennard-Jones potential with equilibrium distance equal to one bead radius and a soft energy well. Additional repulsive restraints were added for *non interacting* beads, forced to remain at a distance longer than the maximum equilibrium distance obtained from equation 1. The initial structure of the chromosome fiber was varied between an extended conformation and a random localization of initially unbound beads in different replicas. The system was allowed to sample the conformational space using pmemd simulation engine for GPU from Amber 18 package. Different conformations of the fibers were determined by attraction and repulsion forces arising from the distance restraints between beads.

In the end, an ensemble of structures was obtained by selecting the minimum number of snapshots minimizing the number of experimental restraint violations (equilibrium distances input). This method yields a population of structures with different conformations, which in average, but not individually, reproduce experimental Hi-C maps derived from population of cells with variable chromatin structure 75.

The ensemble was built in the following way. First, sampled structures with more restraint violations than the mean restraint violations were discarded. Then, the structure with less restraint violations was selected. Considering only the restraints violated by the selected structure, the structure fulfilling more of these restraints was kept. The procedure was repeated iteratively, always considering the restraints that were not fulfilled by any of the previously selected structures. Iterations were stopped when there was no structure left fulfilling new restraints.

Chromatin coarse-grained model at the nucleosome level

The starting point for the 3D chromatin model at the nucleosome level is the coverage of the MNase-seq signal obtained using NucleR software 62. Different families bearing nucleosomes in locations compatible with the MNase-seq experiment are derived by deconvolution of the coverage signal by using a composite Gaussian approximation. For each of the resulting families (compatible with the Mnase-seq signal and DNA/histone stoichiometry) an ideal 3D chromatin structure is prepared and further simulated by a coarse-grained Monte Carlo sampling approach with flexible linkers and rigid nucleosomes. Linker DNA is represented at the base pair level by a pseudo-harmonic potential expressed in helical parameters (rise, slide, shift, twist, roll, tilt) 76. Debye Huckel electrostatics and excluded volume potentials were added to avoid overlaps (exact details of the simulation procedure will be described elsewhere). The results of the different simulations are clustered to select the minimum number of nucleosome

structural families that makes physical sense and that together reproduce MNase-seq experiments.

Declarations

Data Availability

WGBS, RNA-seq and Mnase-seq and Hi-C raw data have been submitted to the European Nucleotide Archive (ENA) under accession number E-MTAB–9258, E-MTAB–9195, E- MTAB–9259 and E-MTAB–9257 respectively. Nanopore data have been submitted to NCBI under accession number (TO BE PROVIDED). Programs for reconstruction and visualization of chromatin structure (from nucleosome to entire chromatin) are available at the H2020 MuG virtual research environment (*https://www.multiscalegenomics.eu/)*

Acknowledgments

We thank all the former and actual members of the EBL for technical support and helpful discussions. We also want to thank Ron Schuyler, Mike Goodstadt, François Serra and David Castillo (CRG-CNAG), F.Posas, E.Nadal and F.Azorin (IRB) for fruitful discussions. We are grateful to Dr. Jessie Colin for providing various expression vectors and to Dr Jan Fronck, Dr Shen Li and Dr Jia-Lei Hu for providing DNMT1, DNMT3a, DNMT3b and DNMT3L cDNA. This work has been supported by the Spanish Ministry of Science (BIO2012–32868), the Catalan SGR, the Instituto Nacional de Bioinformática, the European Research Council (ERC_SimDNA) and the BioExcel and MuG VRE H2000 projects. MO is an ICREA Academia Fellow. The work of SH was supported by the Spanish Ministry of Science (PGC2018–099640-B-I00). AEC is funded by ISCIII /MINECO (PT17/0009/0019) and co-funded by FEDER

Contributions

MO and IBH designed the study; ML, RL, DBe, IBH and NV performed *in vivo* and *in vitro* experiments; DBe performed HPLC/MS analysis; DBu, OF, AE-C and SCH performed data analysis; DBu, JPA and PDD performed *in silico* molecular simulations; JB, MG and IG realized the sequencing; IBH, SCH and MO wrote the manuscript.

References

1. Bird AP, Wolffe AP. Methylation-induced repression–belts, braces, and chromatin. *Cell* **99**, 451-454 (1999).

2. Heyn H, Esteller M. DNA methylation profiling in the clinic: applications and *Nat Rev Genet* **13**, 679-692 (2012).

3. Holz-Schietinger C, Matje DM, Reich NO. Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *J Biol Chem* **287**, 30941-30951 (2012).

4. Winkelmann J*, et al.* Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum Mol Genet* **21**, 2205-2210 (2012).

5. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**, 204-220 (2013).

6. Orlanski S*, et al.* Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc Natl Acad Sci U S A* **113**, 5018-5023 (2016).

7. Miranda TB, Jones PA. DNA methylation: the nuts and bolts of repression. *Journal of cellular physiology* **213**, 384-390 (2007).

8. Carmona FJ*, et al.* A comprehensive DNA methylation profile of epithelial-to- mesenchymal transition. *Cancer research* **74**, 5608-5619 (2014).

9. Kulis M*, et al.* Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat Genet* **44**, 1236-1242 (2012).

10. Mayol G*, et al.* DNA hypomethylation affects cancer-related biological functions and genes relevant in neuroblastoma pathogenesis. *PLoS One* **7**, e48401 (2012).

11. Subramaniam D, Thombre R, Dhar A, Anant S. DNA Methyltransferases: A Novel Target for Prevention and Therapy. *Frontiers in oncology* **4**, 80 (2014).

12. Heyn H*, et al.* Whole-genome bisulfite DNA sequencing of a DNMT3B mutant patient. *Epigenetics : official journal of the DNA Methylation Society* **7**, 542-550 (2012).

13. Klutstein M, Nejman D, Greenfield R, Cedar H. DNA Methylation in Cancer and *Cancer research* **76**, 3446-3450 (2016).

14. Siegfried Z, Eden S, Mendelsohn M, Feng X, Tsuberi BZ, Cedar H. DNA methylation represses transcription in vivo. *Nat Genet* **22**, 203-206 (1999).

15. Kulis M, Queiros AC, Beekman R, Martin-Subero JI. Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer. *Biochimica et biophysica acta* **1829**, 1161-1174 (2013).

16. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from *Nat Rev Genet* **9**, 465-476 (2008).

17. Jiang C, Pugh BF. Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* **10**, 161-172 (2009).

18. Perez A*, et al.* Impact of methylation on the physical properties of *Biophysical journal* **102**, 2140-2148 (2012).

19. Portella G, Battistini F, Orozco M. Understanding the connection between epigenetic DNA methylation and nucleosome positioning from computer simulations. *PLoS Comput Biol* **9**, e1003354 (2013).

20. Collings CK, Anderson JN. Links between DNA methylation and nucleosome occupancy in the human genome. *Epigenetics & chromatin* **10**, 18 (2017).

21. Choy JS, Wei S, Lee JY, Tan S, Chu S, Lee TH. DNA methylation increases nucleosome compaction and rigidity. *J Am Chem Soc* **132**, 1782-1783 (2010).

22. Chodavarapu RK*, et al.* Relationship between nucleosome positioning and DNA methylation. *Nature* **466**, 388-392 (2010).

23. Cokus SJ*, et al.* Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**, 215-219 (2008).

24. Gowher H, Stockdale CJ, Goyal R, Ferreira H, Owen-Hughes T, Jeltsch A. De novo methylation of nucleosomal DNA by the mammalian Dnmt1 and Dnmt3A DNA methyltransferases. *Biochemistry* **44**, 9899-9904 (2005).

25. Felle M, Hoffmeister H, Rothammer J, Fuchs A, Exler JH, Langst G. Nucleosomes protect DNA from DNA methylation in vivo and in vitro. *Nucleic Acids Res* **39**, 6956-6969 (2011).

26. Huff JT, Zilberman D. Dnmt1-Independent CG Methylation Contributes to Nucleosome Positioning in Diverse Eukaryotes. *Cell* **156**, 1286-1297 (2014).

27. Morselli M*, et al.* In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. *Elife* **4**, e06205 (2015).

28. Pedersen JS*, et al.* Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. *Genome Res* **24**, 454-466 (2014).

29. Kelly TK, Liu Y, Lay FD, Liang G, Berman BP, Jones PA. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res* **22**, 2497-2506 (2012).

30. Capuano F, Mulleder M, Kok R, Blom HJ, Ralser M. Cytosine DNA methylation is found in Drosophila melanogaster but absent in Saccharomyces cerevisiae, Schizosaccharomyces pombe, and other yeast species. *Analytical chemistry* **86**, 3697-3702 (2014).

31. Duan Z*, et al.* A three-dimensional model of the yeast genome. *Nature* **465**, 363- 367 (2010).

32. Kim S*, et al.* The dynamic three-dimensional organization of the diploid yeast *Elife* **6**, (2017).

33. Lazar-Stefanita L*, et al.* Cohesins and condensins orchestrate the 4D dynamics of yeast chromosomes during the cell cycle. *EMBO J* **36**, 2684-2697 (2017).

34. Rutledge MT, Russo M, Belton JM, Dekker J, Broach JR. The yeast genome undergoes significant topological reorganization in quiescence. *Nucleic Acids Res* **43**, 8299-8313 (2015).

35. Swygert SG, *et al.* Condensin-Dependent Chromatin Compaction Represses Transcription Globally during Quiescence. *Molecular cell* **73**, 533-546 e534 (2019).

36. Bulkowska U, Ishikawa T, Kurlandzka A, Trzcinska-Danielewicz J, Derlacz R, Fronk J. Expression of murine DNA methyltransferases Dnmt1 and Dnmt3a in the yeast Saccharomyces cerevisiae. *Yeast* **24**, 871-882 (2007).

37. Hu JL, Zhou BO, Zhang RR, Zhang KL, Zhou JQ, Xu GL. The N-terminus of histone H3 is required for de novo DNA methylation in chromatin. *Proc Natl Acad Sci U S A* **106**, 22187-22192 (2009).

38. Finnegan AI, *et al.* Epigenetic engineering of yeast reveals dynamic molecular adaptation to methylation stress and genetic modulators of specific DNMT3 family members. *Nucleic Acids Res*, (2020).

39. Gari E, Piedrafita L, Aldea M, Herrero E. A set of vectors with a tetracycline- regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. *Yeast* **13**, 837-848 (1997).

40. Gong T, *et al.* Both combinatorial K4me0-K36me3 marks on sister histone H3s of a nucleosome are required for Dnmt3a-Dnmt3L mediated de novo DNA methylation. *J Genet Genomics*, (2020).

41. Ooi SK, *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**, 714-717 (2007).

42. Otani J, Nankumo T, Arita K, Inamoto S, Ariyoshi M, Shirakawa M. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO reports* **10**, 1235-1241 (2009).

43. Weinberg DN, *et al.* The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. *Nature* **573**, 281-286 (2019).

44. Buitrago D, *et al.* Nucleosome Dynamics: a new tool for the dynamic analysis of nucleosome positioning. *Nucleic Acids Res* **47**, 9511-9523 (2019).

45. Nadal-Ribelles M, *et al.* Hog1 bypasses stress-mediated down-regulation of transcription by RNA polymerase II redistribution and chromatin *Genome Biol* **13**, R106 (2012).

46. Nocetti N, Whitehouse I. Nucleosome repositioning underlies dynamic gene expression. *Genes Dev* **30**, 660-672 (2016).

47. Guidi M, *et al.* Spatial reorganization of telomeres in long-lived quiescent *Genome Biol* **16**, 206 (2015).

48. Wang R, Kamgoue A, Normand C, Leger-Silvestre I, Mangeat T, Gadal O. High resolution microscopy reveals the nuclear shape of budding yeast during cell cycle and in various biological states. *J Cell Sci* **129**, 4480-4495 (2016).

49. Jin Q, Trelles-Sticken E, Scherthan H, Loidl J. Yeast nuclei display prominent centromere clustering that is reduced in nondividing cells and in meiotic prophase. *J Cell Biol* **141**, 21-29 (1998).

50. Noma K, Cam HP, Maraia RJ, Grewal SI. A role for TFIIIC transcription factor complex in genome organization. *Cell* **125**, 859-872 (2006).

51. Belton JM*, et al.* The Conformation of Yeast Chromosome III Is Mating Type Dependent and Controlled by the Recombination Enhancer. *Cell reports* **13**, 1855-1867 (2015).

52. Miele A, Bystricky K, Dekker J. Yeast silent mating type loci form heterochromatic clusters through silencer protein-dependent long-range interactions. *PLoS Genet* **5**, e1000478 (2009).

53. Bagci H, Fisher AG. DNA demethylation in pluripotency and reprogramming: the role of tet proteins and cell division. *Cell Stem Cell* **13**, 265-269 (2013).

54. Kulis M*, et al.* Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet* **47**, 746-756 (2015).

55. Shen L*, et al.* A single amino acid substitution confers enhanced methylation activity of mammalian Dnmt3b on chromatin DNA. *Nucleic Acids Res* **38**, 6054- 6064 (2010).

56. Friso S, Choi SW, Dolnikowski GG, Selhub J. A method to assess genomic DNA methylation using high-performance liquid chromatography/electrospray ionization mass spectrometry. *Analytical chemistry* **74**, 4526-4531 (2002).

57. Merkel A*, et al.* gemBS: high throughput processing for DNA methylation data from bisulfite sequencing. *Bioinformatics* **35**, 737-742 (2019).

58. Schlenstedt G, Hurt E, Doye V, Silver PA. Reconstitution of nuclear protein transport with semi-intact yeast cells. *J Cell Biol* **123**, 785-798 (1993).

59. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).

60. Team RDC. R: A Language and Environment for Statistical Computing.). R Foundation for Statistical Computing (2011).

61. Planet E, Attolini CS, Reina O, Flores O, Rossell D. htSeqTools: high-throughput sequencing quality control, processing and visualization in R. *Bioinformatics* **28**, 589-590 (2012).

62. Flores O, Orozco M. nucleR: a package for non-parametric nucleosome positioning. *Bioinformatics* **27**, 2149-2150 (2011).

63. Dobin A*, et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15- 21 (2013).

64. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

65. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

66. Belaghzal H, Dekker J, Gibcus JH. Hi-C 2.0: An optimized Hi-C procedure for high- resolution genome-wide mapping of chromosome conformation. *Methods* **123**, 56-65 (2017).

67. Serra F, Bau D, Goodstadt M, Castillo D, Filion GJ, Marti-Renom MA. Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. *PLoS Comput Biol* **13**, e1005665 (2017).

68. Durand NC*, et al.* Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* **3**, 99-101 (2016).

69. Rao SS*, et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).

70. Lun AT, Smyth GK. diffHic: a Bioconductor package to detect differential genomic interactions in Hi-C data. *BMC Bioinformatics* **16**, 258 (2015).

71. Adhikari B, Trieu T, Cheng J. Chromosome3D: reconstructing three-dimensional chromosomal structures from Hi-C interaction frequency data using distance geometry simulated annealing. *BMC Genomics* **17**, 886 (2016).

72. Varoquaux N, Ay F, Noble WS, Vert JP. A statistical approach for inferring the 3D structure of the genome. *Bioinformatics* **30**, i26-33 (2014).

73. Zhang Z, Li G, Toh KC, Sung WK. 3D chromosome modeling with semi-definite programming and Hi-C data. *J Comput Biol* **20**, 831-846 (2013).

74. Imakaev M*, et al.* Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat Methods* **9**, 999-1003 (2012).

75. Tjong H*, et al.* Population-based 3D genome structure analysis reveals driving forces in spatial genome organization. *Proc Natl Acad Sci U S A* **113**, E1663-1672 (2016).

76. Walther J, Dans PD, Balaceanu A, Hospital A, Bayarri G, Orozco M. A multi-modal coarse grained model of DNA flexibility mappable to the atomistic level. *Nucleic Acids Res* **48**, e29 (2020).

Tables

				All Contexts		CpG Contexts			Non-CpG Contexts		
DNMT expressed	Hours of induction	State of the culture	Avg. meth	No. Cyt	Frac. > 0	Avg. meth Cyt	No. Frac. > 0		Avg. meth Cyt	No. Frac. > 0	
DNMT1, 3a, 3L	30hrs	Exponential - Not synchronized	0.75%	3066478	*2.74%*	3.14%	525937	15.60%	0.16%	2538540	0.08%
DNMT1, 3b, 3L	30hrs	Exponential - Not synchronized	0.70%	2584192	2.65%	*2.77%*	463127	14.36%	0.13%	2118519	0.10%
DNMT3a, 3b, 3L	30hrs	Exponential - Not synchronized	0.49%	2889913	2.03%	1.97%	502806	11.46%	0.11%	2385079	0.04%
DNMT1, 3a, 3b	30hrs	Exponential - Not synchronized	0.56%	3066743	1.86%	2.08%	524612	10.73%	0.18%	2539830	0.03%
None	30hrs	Exponential - Not synchronized	0.15%	2732598	0.00%	0.13%	464513	0.00%	0.16%	2265439	0.00%
All 4 DNMT1	27.5hrs	Exponential - Not synchronized	2.03%	2810320	6.66%	8.55%	506621	34.88%	0.24%	2301426	0.45%
All 4 DNMT*2*	27.5hrs	Exponential - Not synchronized	2.13%	2937375	6.90%	9.14%	522749	36.43%	0.26%	2412449	0.51%
All 4 DNMT1	24hrs	Exponential- Synchronized in G1	1.75%	4427610	6.73%	9.72%	676754	41.0%	0.26%	3750856	0.54%
All 4 DNMT*2*	24hrs	Exponential- Synchronized in G1	1.51%	4425531	6.20%	8.36%	676458	38.1%	0.24%	3749073	0.44%
All 4 DNMT1	>72hrs	Saturation	5.12%	4425775	14.7%	27.1%	676923	67.9%	1.11%	3748852	5.09%
All 4 DNMT*2*	>72hrs	Saturation	4.75%	4427957	13.3%	25.2%	677239	63.4%	1.00%	3750718	4.22%

1Samples corresponding to replica1 [2] Samples corresponding to replica2

Table 1 : Average methylation in CpG and non-CpG context. (Avg. meth : Average of methylation in all cytosine (col 4), in cytosine only in CpG

context (col 5) or in non-CpG context (col 6); No. Cyt : Total number of cytosines for which the methylation status can be determined; Frac>0 :

% of cytosine with a methylation level >0)

Figures

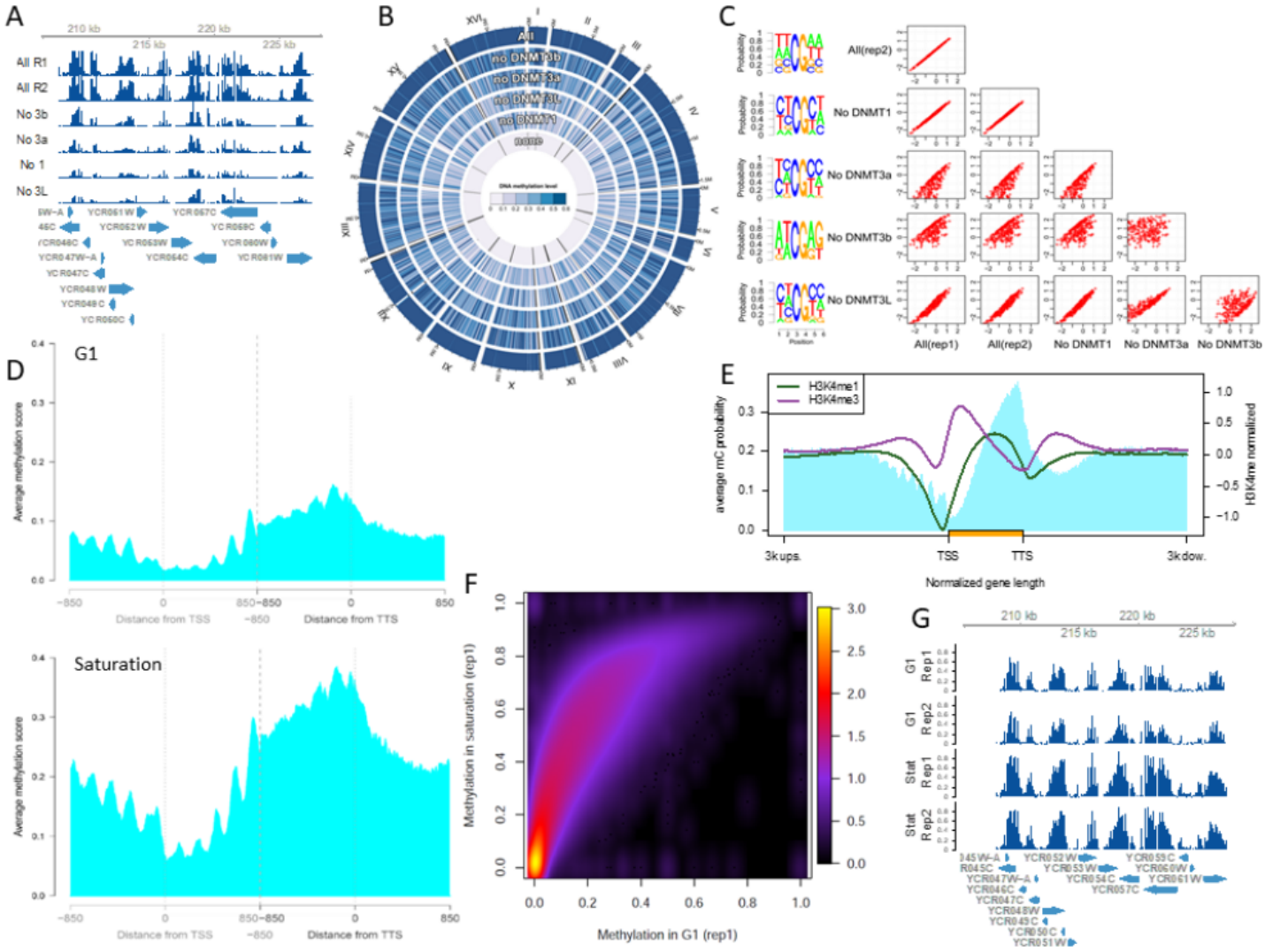

Figure 1

Methylation pattern across several samples and along the gene body. (A) The pattern of methylation is conserved in all samples as illustrated for this 20kb region of chromosome III (208135..227458) where the level of methylation at each position is represented for two samples with the four DNMTs expressed and four samples with each combination of 3 DNMTs. (B) Circular plot comparing DNA methylation for the samples in (A) and a control without methylation (None). Methylation levels decrease when one DNMT is missing, the strongest effect being without DNMT1 and the mildest effect when DNMT3b is not present. (C) Correlation between the sequence effect on methylation status among samples with different combinations of three DNMTs. Motif effects are estimated from logistic regression (for details see Materials and Methods) and correlation plots are produced for each pair of samples. Motifs have nearly the same effect in the two replicates with all DNMTs induced (correlation coefficient is cor=0.999) and when DNMT1 (cor=0.996) or DNMT3L (cor=0.954) are

removed. In contrast, the estimated effect of some motifs on methylation probability is different in samples lacking DNMT3a (cor=0.793) or DNMT3b (cor=0.631). The left panels show the sequence logo of the motifs preferentially methylated in each sample. (D,F,G) Comparison of methylation pattern in samples in exponential phase and at saturation (D) Average methylation level around TSS and TTS (850 bp upstream and downstream from each point). (E) Superposition of the pattern of DNA methylation and the pattern of H3K4 methylation along the average gene body (from 3kb upstream TSS to 3kb downstream TTS). DNA methylation preferentially occurs where H3K4 is not methylated. (F) Heatmap showing the correlation between methylation probabilities in samples in G1 and at saturation. (G) The pattern of methylation is conserved in samples in G1 and samples at saturation as illustrated for this 20Kb region of chromosome III (208135..227458) where the level of methylation at each position is represented for 2 replicas of each condition.

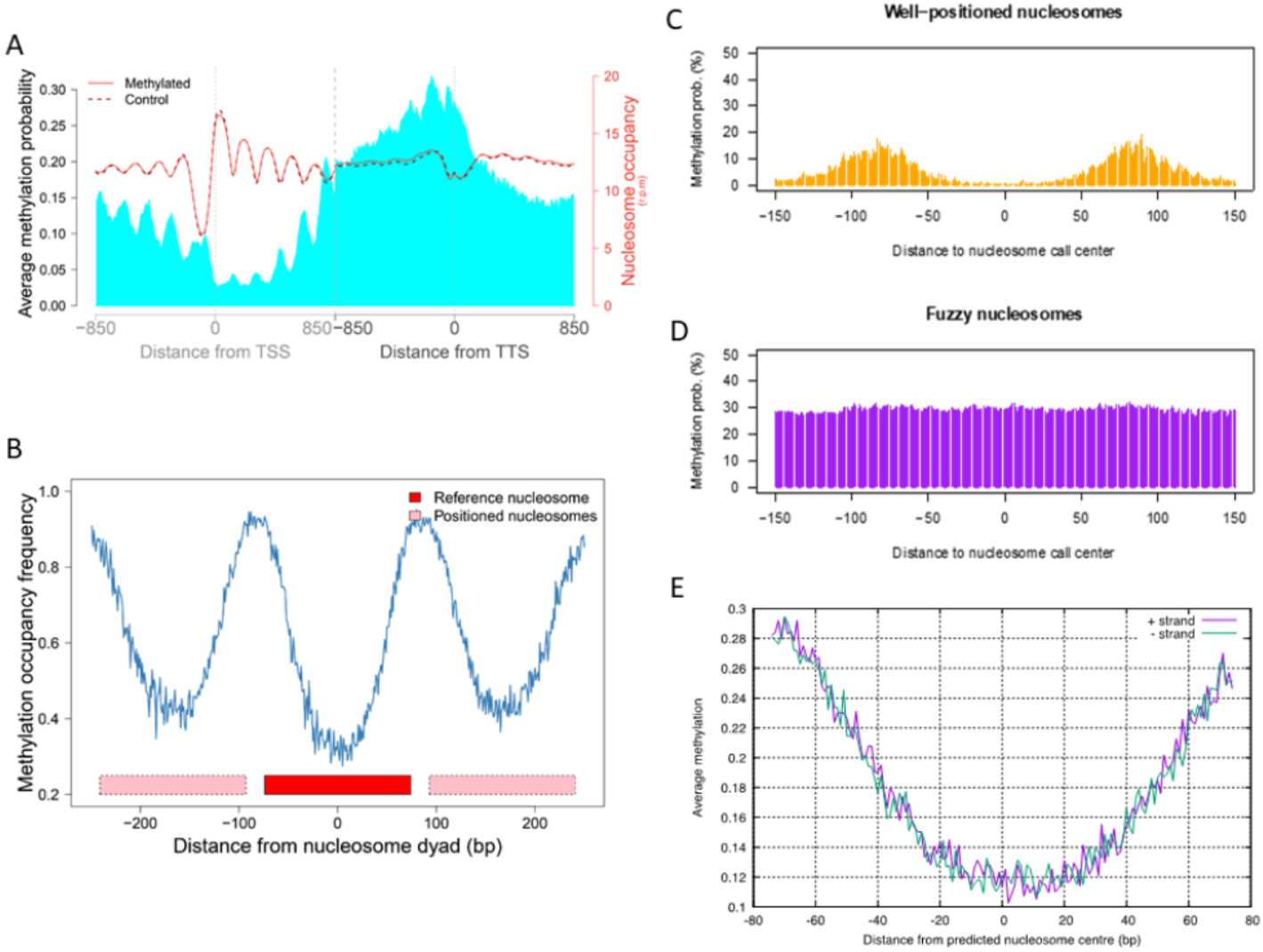

Figure 2

Correlation between DNA methylation and nucleosome coverage genome wide (A) Nucleosome positioning (in red) in a sample before (dashed lane) and after (plain lane) induction of methylation and average methylation probabilities (in blue). Plots are built around TSS and TTS (850 bp upstream and downstream from each point). Average nucleosome positioning does not change drastically upon methylation. (B) Percentage of CpG with methylation probability above 0.01 around well positioned (W) nucleosomes. Nucleosome calls were considered well-positioned when nucleR peak width score and height score were higher than 0.6 and 0.4, respectively. DNA methylation is anticorrelated with nucleosome occupancy in W nucleosomes. (C,D) Average methylation probability around nucleosome call center (150 bp upstream and downstream) for (C) Well positioned and (D) Fuzzy nucleosomes. (E) Average methylation probability per strand around nucleosome call center (75 bp upstream and downstream) of wellpositioned nucleosomes

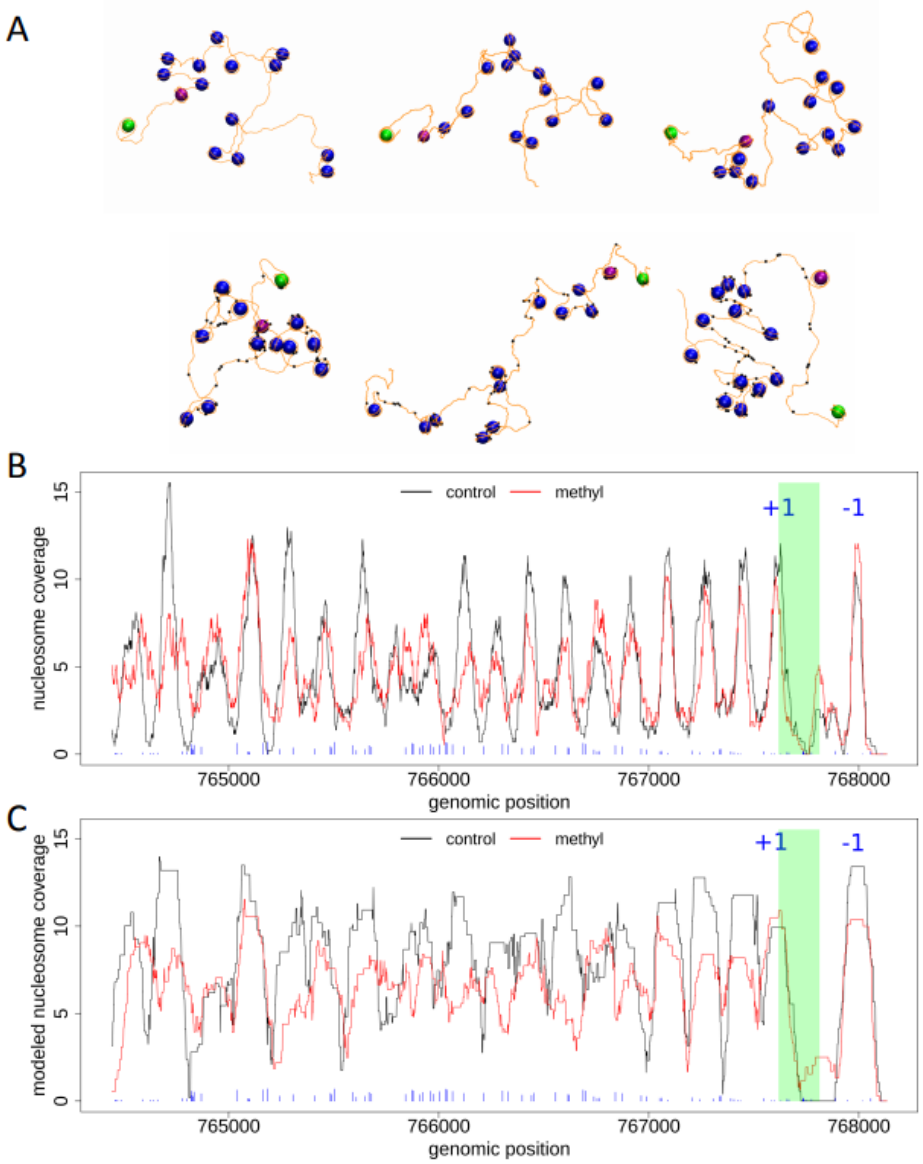

Figure 3

Modeled nucleosome fibers from MNase-seq data (A) Modeled nucleosome fiber for (top panel) control and (lower panel) methylated cells. DNA base pairs are depicted as orange lines and nucleosomes as blue spheres (+1 and -1 nucleosomes are depicted in purple and green, respectively). Methylated regions (signal above 10% coverage) are highlighted as black balls. Experimental (B) and modeled (C) nucleosome position coverage for DUG2 gene (chromosome II). The TSS is highlighted in green and nucleosomes -1 and +1 are indicated. The blue sticks in the bottom correspond to methylated regions with their height proportional to methylation extent.

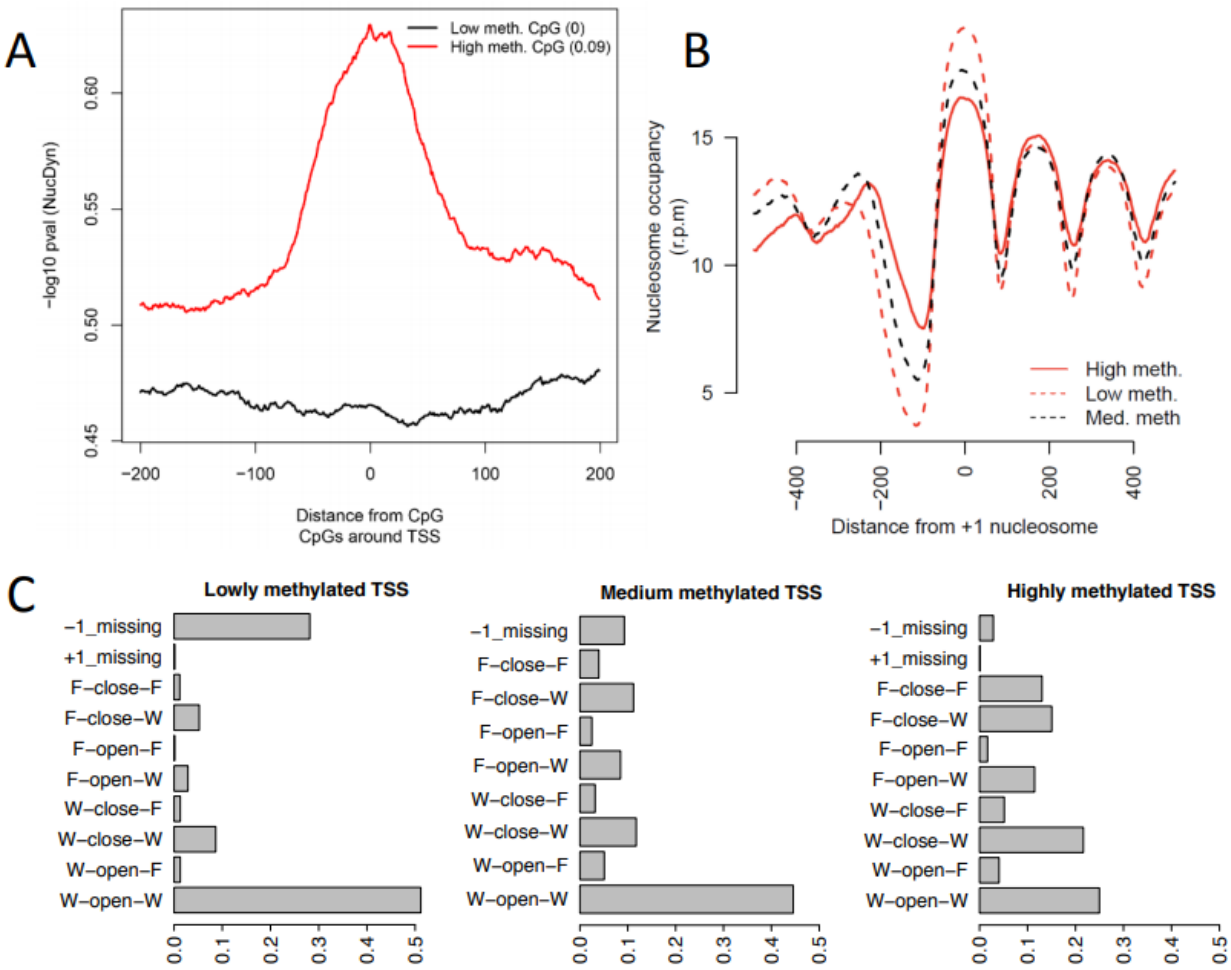

Figure 4

Nucleosome dynamics upon methylation at promoters. (A) NucDyn score around highly methylated and lowly methylated CpGs at promoters. (B) Nucleosome coverage around +1 nucleosome for genes with highest (top 10%), lowest (bottom 10%) or medium (45-55%) level of CpG methylation around promoters. (C) Nucleosome architecture around promoters according to their methylation level.

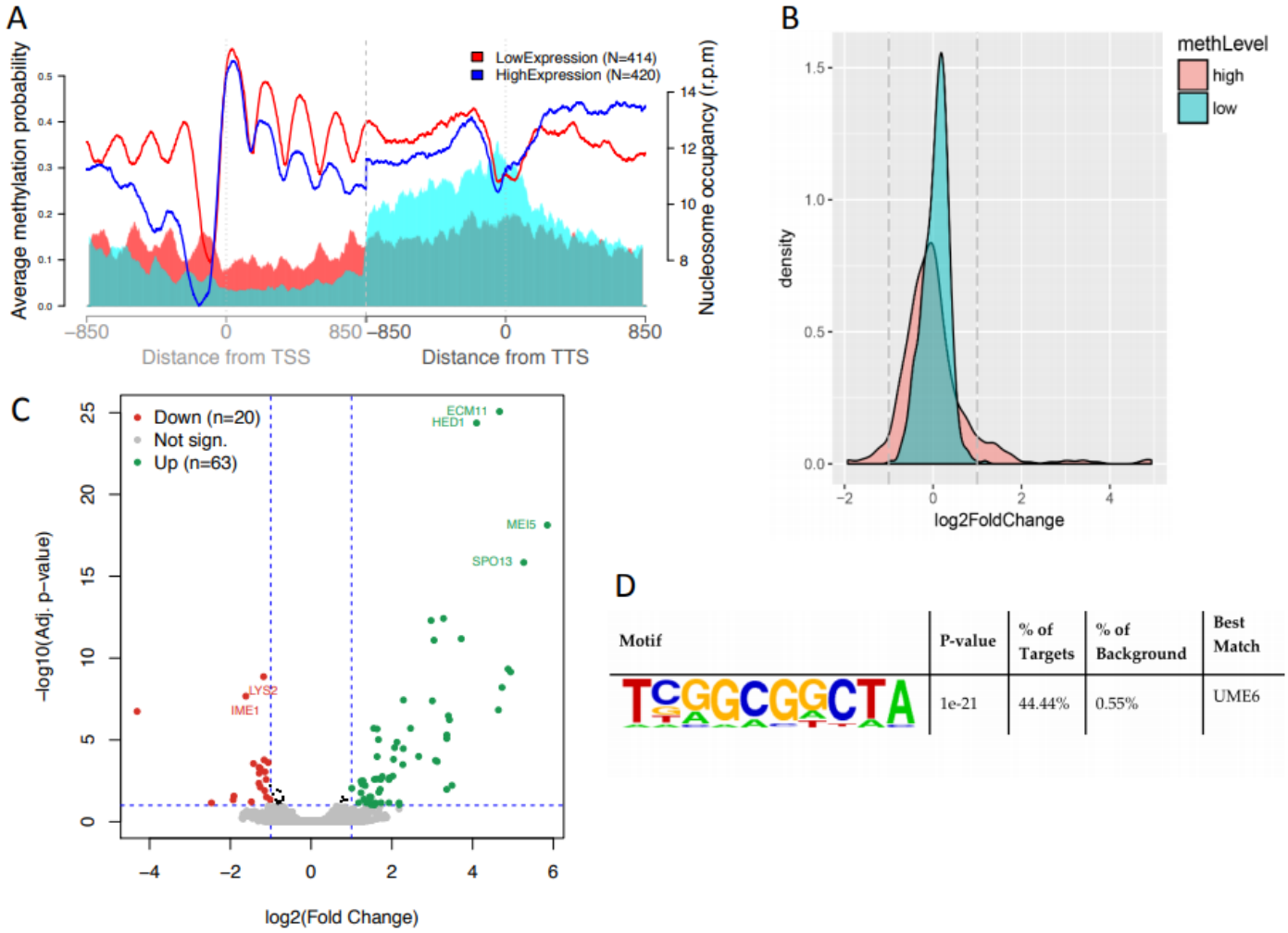

Figure 5

Correlation of DNA methylation, nucleosome positioning and gene expression. (A) Methylation probability and nucleosome coverage (solid lines) around genes with expression in the top 10% (blue) or bottom 10% (red) expression level. (B) Log2 of the fold change in expression for genes with highly and lowly methylated promoters. (C) Gene expression difference between methylated and not methylated samples. RNA level difference is plotted on the x-axis and the Adj. p-value on the y-axis. Downregulated (20 genes) and upregulated (63 genes) genes are shown in red and green, respectively. The genes with the highest changes are highlighted. (D) Promoter motif enrichment for genes with highly methylated promoters and increase in expression level.

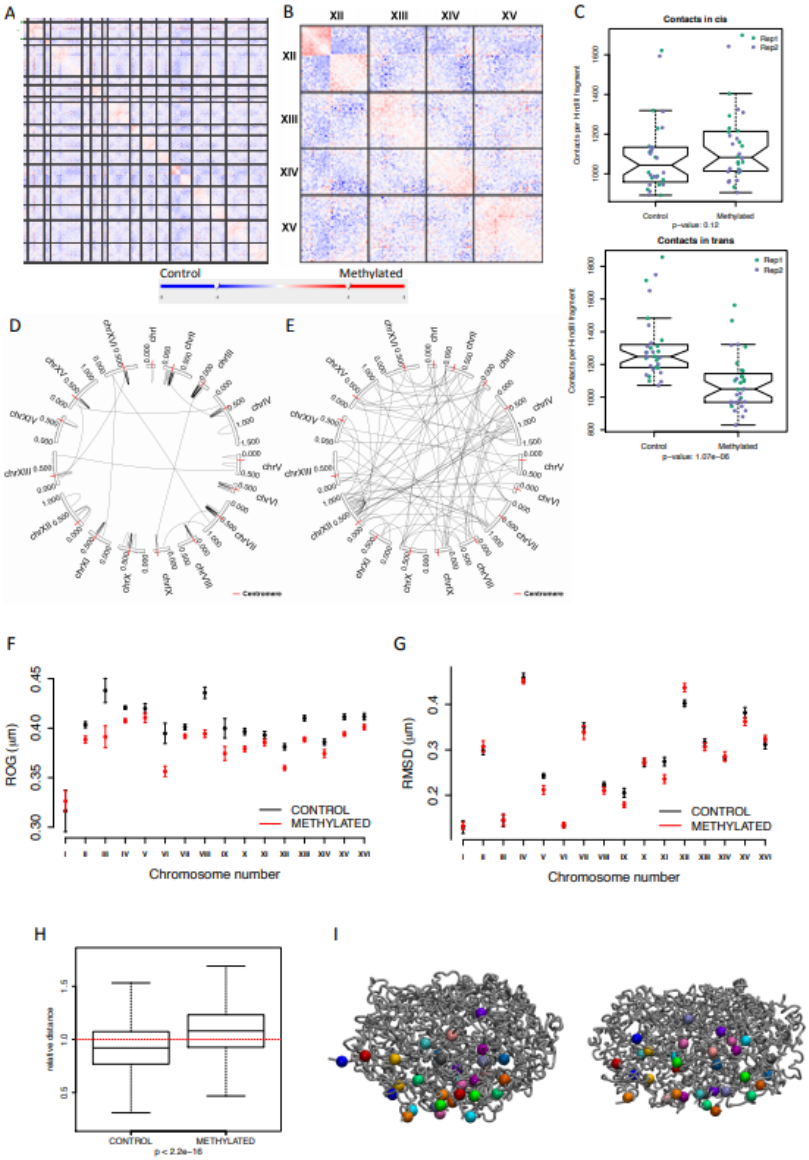

Figure 6

Effect of DNA methylation on 3D genome structure (A,B) Differential contact frequencies in control and methylation induced samples in replica 1 for (A) whole genome and (B) focus on four chromosomes. Blue indicates interaction with a higher frequency in the non-methylated control sample and red indicates interactions with a higher frequency in the methylated samples. (C) Comparison of contact frequencies between control and methylated Hi-C samples in cis (+/- 50Kb from the centromere, top panel) and in trans (lower panel). (D,E) Circos plots displaying the significant (FDR<0.5) differential interactions identified with diffHiC: (D) gained interactions (log2FC>1) are clustered around the centromeres (red tick marks) and (E) lost interactions (log2FC< -1) preferentially occur between chromosomes. Structural changes measured on the ensemble of structures obtained with our restraint-based 3D model for each chromosome: (F) Mean radius of gyration computed around the centromeres (+/-100Kb) and (G) flexibility of each chromosome measured by the RMSD for the control (black) and methylated (red) samples. (H) Relative distance (distance divided by mean distance in both conditions) between all telomeres in the ensemble of 3D structures. (I) Whole genome 3D model for a representative structure from the ensemble for the control (Left) and the methylated sample (Right). All chromosomes are represented as grey tubes, and telomeres are highlighted as spheres. The pair of telomeres of each chromosome is represented in a specific color.

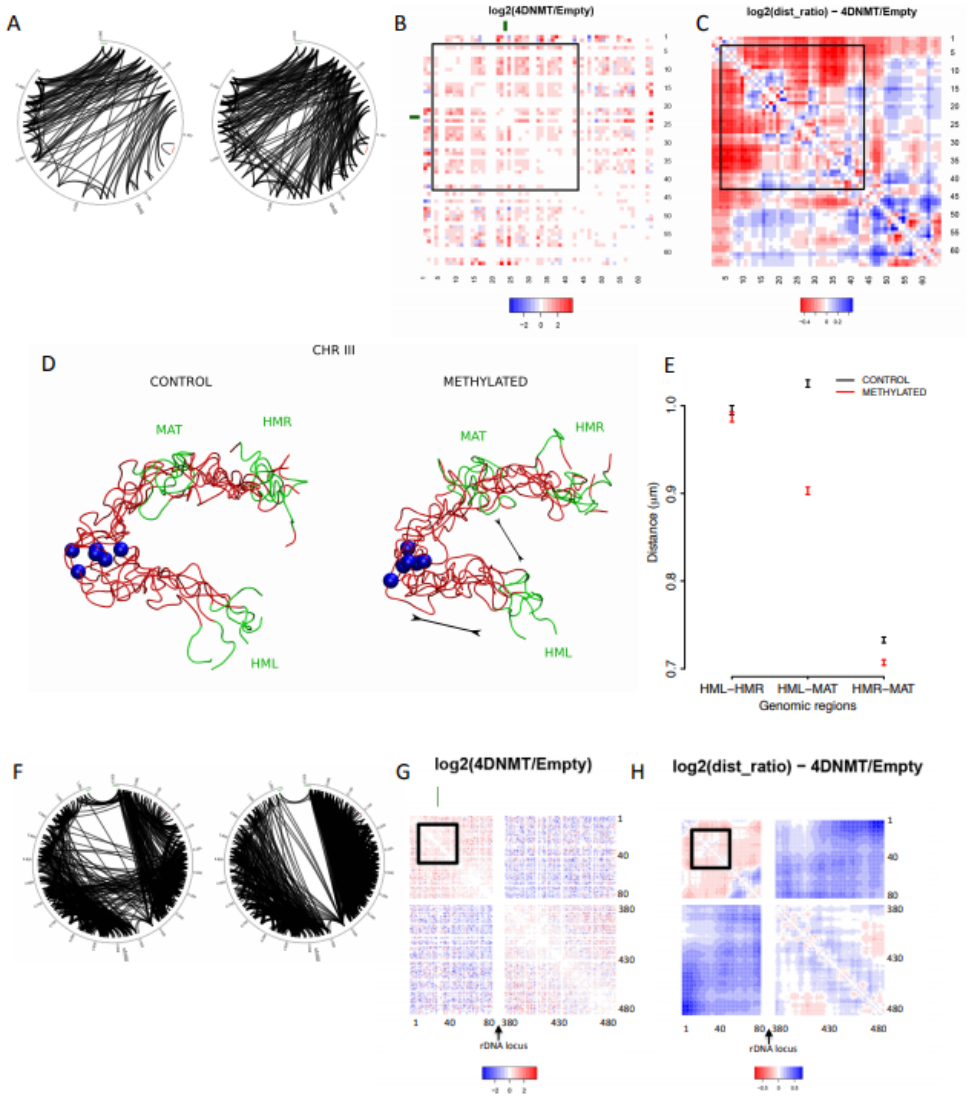

Figure 7

Chromosome conformation changes under DNA methylation. (A) Circos diagrams of significant interactions in chromosome III for the control (left) and methylated (right) samples. (B,C) Heatmaps displaying the log2 ratio (Methylated/Control) of (B) the contact frequencies and (C) the distances in the ensemble for chromosome III. Blue indicates interaction with a higher frequency or shorter distance in the non-methylated control sample and red indicates interactions with a higher frequency or shorter distance in the methylated samples. (D) Average distances between matting type loci in the ensemble of structures for chromosome III (E) 3D representative structures for chromosome III highlighting the distance between MAT and HML loci shown in green. The centromere is represented as a blue sphere. (F) Circos diagrams of significant interactions in chromosome XII for the control (left) and methylated (right) samples. (G.H) Log2 ratio (Methylated/Control) of (G) the contact frequencies and (H) the distances in the ensemble for chromosome XII.

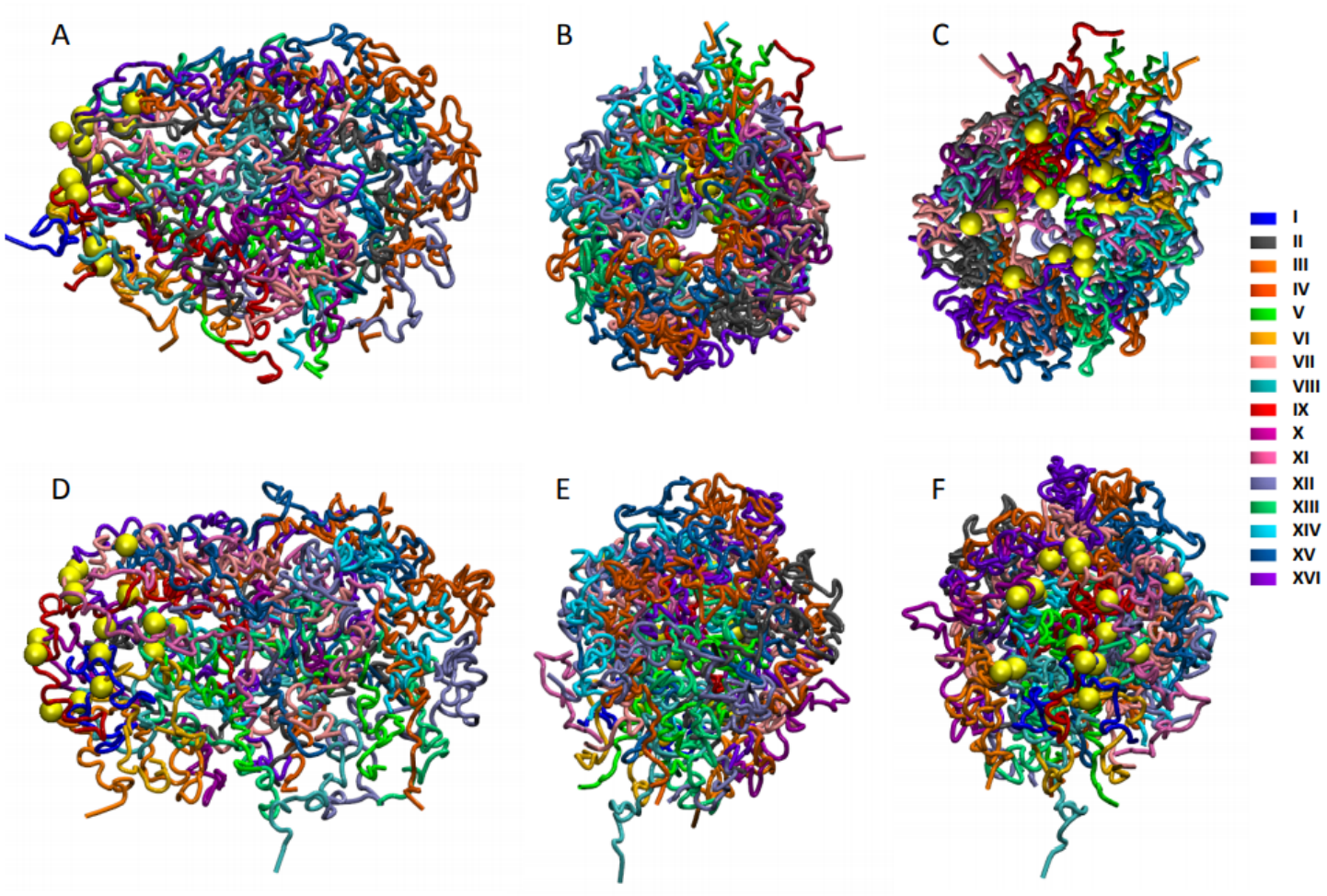

Figure 8

Whole genome 3D model. Three views are provided representing three different angles for (A-C) the control samples and (D-F) the methylated sample. Each chromosome is represented in a specific color. Centromeres are represented in yellow. The rDNA in chromosome XII is not represented due to the lack of information provided by the Hi-C in repeated regions

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupTableS1.xlsx
- supplementaltableS2S5.pdf
- SuppFigures20200502.pdf
- Supplmatandmeth.pdf