MCM2 promotes symmetric inheritance of modified histones during DNA replication

Petryk, Nataliya; Dalby, Maria; Wenger, Alice; Stromme, Caroline B.; Strandsby, Anne; Andersson, Robin; Groth, Anja

MCM2 promotes symmetric inheritance of modified histones during DNA replication

Nataliya Petryk[1,2‡], Maria Dalby[3‡], Alice Wenger[1,2], Caroline B. Stromme[1,4], Anne Strandsby[1,5], Robin Andersson[3]*, Anja Groth[1,2]*.

‡These authors contributed equally to this work.

[1]Biotech Research and Innovation Centre (BRIC) and [2]Novo Nordisk Foundation Center for Protein Research (CPR), Faculty of Health and Medical Sciences, University of Copenhagen, 2200 Copenhagen, Denmark.

[3]The Bioinformatics Centre, Department of Biology, Faculty of Science, University of Copenhagen, 2200 Copenhagen, Denmark.

[4]Present address: Novo Nordisk A/S, 2860 Soborg, Denmark.

[5]Present address: Statens Serum Institut, Virus & Microbiological Special Diagnostics, 2300 Copenhagen S, Denmark.

*Correspondence to: R.A. (robin@binf.ku.dk) and A.G. (anja.groth@bric.ku.dk).

One sentence summary: MCM2 is required for recycling of parental histones to the lagging strand.

Abstract: During genome replication, parental histones are recycled to newly replicated DNA with their post-translational modifications (PTMs). It remains unknown whether sister chromatids inherit modified histones evenly. Here, we measured histone PTM partition to sister chromatids in embryonic stem cells. We found that parental histones H3-H4 segregate to both daughter DNA strands with a weak leading strand bias, skewing partition at topologically associating domain (TAD) borders and enhancers proximal to replication initiation zones. Segregation of parental histones to the leading strand increased markedly in cells with histone-binding mutations in MCM2, part of the replicative helicase, exacerbating histone PTM sister chromatid asymmetry. This work reveals how histones are inherited to sister chromatids and identifies a mechanism for how symmetric cell division is ensured by the replication machinery.

Histone PTMs contribute to the establishment and maintenance of epigenetic chromatin states that regulate transcriptional programs during development (*1, 2*), but the mechanisms that ensure transmission of histone PTM patterns to daughter cells remain unclear. Chromatin is disrupted upon replication fork passage and nucleosomes are rapidly reassembled on newly synthesized DNA through recycling of evicted parental histones and *de novo* deposition of new histones (*3, 4*). The recycling of modified parental histones is a critical step in histone PTM transmission (*5*) and early studies suggested that parental histones segregate randomly to both daughter DNA strands (*6, 7*). However, it remains open whether histone PTM inheritance is truly symmetric and how parental histones are segregated to the leading and lagging strand of the replication fork. Multiple replication origins are used to replicate large metazoan chromosomes and replication fork directionality (RFD) and leading/lagging strand replication therefore alternates along chromosomes (*8, 9*). Potential biases in segregation of modified parental histones during

2

replication will thus result in a specific pattern of sister chromatid asymmetry. We therefore investigated the distribution of parental and new histones on sister chromatids and linked this to RFD in order to understand histone segregation.

We developed SCAR-seq (Sister Chromatids After Replication by DNA sequencing) to track histone recycling and *de novo* deposition genome-wide (Fig. 1A, Methods). We differentiated old and new histones H4 by di-methylation at lysine 20 (H4K20me2, Fig. S1A), exclusively present on >80% of old H4 in nascent chromatin (*5, 10*), and acetylation at lysine 5 (H4K5ac) present on >95% of new H4 (*3, 11*). Mouse embryonic stem cells (mESCs) were EdU-labeled and nascent mono-nucleosomes carrying H4K20me2 or H4K5ac were purified sequentially by chromatin immunoprecipitation (ChIP) and streptavidin-capture of biotinylated EdU-labeled DNA. The new and parental DNA strands were separated (Fig. S1B) and sequenced strandedly to score genome-wide sister chromatid histone partition (Fig. 1A, S1C).

To determine locally which sister chromatid was replicated preferentially by leading strand, we measured RFD by Okazaki fragment sequencing (OK-seq, Methods) (*8*). Replication initiation zones (n = 2,844) (Fig. S2A) were comparable to those in human (*8*) and *C. elegans* (*12*), ranging in size and efficiency (Fig. S2B, C) and were mostly intergenic (Fig. S2D), enriched in enhancer-associated features (H3K27ac, H3K4me1, p300 occupancy, DNase I hypersensitive sites) and flanked by active genes (H3K36me3, H3K4me3) (Fig. S2E). Around initiation zones, the partition of old and new H4 showed a weak reciprocal shift with H4K20me2 and H4K5ac skewed towards leading and lagging strand replication, respectively (Fig. 1B-D and Fig. S3A-C). The partition amplitude was considerably lower than RFD, arguing that old histones segregate to both strands but not entirely symmetrically. Analysis of the parental DNA strands showed the complementary partition shift (Fig. S3D, E), excluding an effect of EdU on partition

measurements. The partition skew was most pronounced around highly efficient initiation zones (Fig. S3F), indicating that DNA replication drives the observed sister chromatid asymmetries. Histone partition skew also tracked with RFD at higher genomic scales (Fig. S4A, B), for example across replication units with early replicating borders and late replicating centers termed U-domains (*8, 9, 13*). Together, these results demonstrate that parental histones segregate to both arms of the replication fork with a slight preference for the leading strand, while *de novo* deposition has a comparable bias towards the lagging strand.

Replication timing is related to chromosome organization in TADs (*9, 14, 15*). TAD borders (*16*) are enriched in initiation zones (Fig. S5A) (*8, 13*) and showed a reciprocal histone partition skew (Fig. 2A, B and Fig. S5B). B compartment TADs (transcriptionally inactive) displayed stronger RFD and partition shifts than transcriptionally active A compartment TADs (*17*) (Fig. 2A), possibly due to increased internal initiation within active TADs ($P < 2.2 \times 10^{-16}$, odds ratio 2.2, Fisher's exact test) or an effect of transcription. To investigate partition asymmetries over genes, we tracked histone H3 tri-methylated at lysine K36 (H3K36me3) (Fig. S5C) present on parental histones in gene bodies (*5, 18, 19*). H3K36me3 partition skewed moderately towards leading strand replication consistent with H4K20me2 (Fig. 2A and Fig. S5C-E), and was stronger over active genes and co-directional with transcription (Fig. 2C) (*8*). Further, the correlation of chromatin interaction directionality with histone partitioning was weaker than with RFD (Fig. 2B), suggesting that while histone partitioning is driven by RFD it can be affected by transcription. Active enhancers often coincided with initiation zone centers and promoters tended to be flanking (*20*) (Fig. S5F), suggesting that enhancer activity might affect partitioning in neighboring regions. Indeed, both RFD and histone partition asymmetry were higher around

active enhancers (*21*) and super-enhancers that control cell-type specific genes (*22*) (Fig. 2D, Mann-Whitney U test, $P < 1.1 \times 10^{-16}$) .

MCM2, part of the replicative helicase, is proposed to recycle parental histone H3-H4 via its N-terminal histone-binding domain (HBD) (*11, 23-25*). Using genome editing, we mutated two critical residues in the HBD (MCM2-2A; Y81A Y90A) (*24, 25*) (Fig. S6A) that disrupt histone-binding (Fig. S6B (*25*)) without affecting cell cycle progression (Fig. S6C). Strikingly in MCM2-2A mutants, old and new histone partition were strongly skewed towards leading and lagging strands, respectively, generating partition ratios similar to RFD in amplitude and pattern (Fig. 3A-C, Fig. S6D and S7). Moreover, partition of new and old histones showed strong anti-correlation in MCM2-2A (Fig. 3D and S7D), indicating segregation to opposite strands. H3K36me3 occupancy was not altered in MCM2-2A cells (Fig. S8A, B), indicating that histone partitioning rather than recycling was perturbed. The association between H3K36me3 partition and transcriptional directionality was reduced in MCM2-2A cells (Fig. S8C-E), further indicating increased replication-driven sister chromatid asymmetry in parental and new histones. Sister chromatid asymmetry was also strongly increased at TAD borders, around enhancers (Fig. S8F, G) and across important developmental loci (*e.g.* the *Hox* clusters, Fig. S9) in MCM2-2A cells, which thus provides a model to address histone PTM inheritance in development. The high correlation between H4K20me2 partition in MCM2-2A cells and RFD prompted us to map H4K20me2 partition breakpoints. They showed strong co-localization with initiation zones mapped by OK-seq (Fig. 3E, Methods), suggesting H4K20me2 SCAR-seq in MCM2-2A as a method to map replication dynamics.

In summary, SCAR-seq revealed that parental histone segregation is almost symmetrical with a weak inherent preference for the leading strand (Fig. 3F, left), creating modest sister chromatid asymmetries that might be mitotically transmitted as new histones acquire PTMs with slow kinetics (*5*). Importantly, MCM2 histone chaperone activity promotes balanced segregation of old histones to leading and lagging strand, thereby ensuring inheritance of histone-based information to both sister chromatids. This is consistent with MCM2 chaperoning old histones (*11, 23*) and cryo-EM data placing the MCM2 HBD in front of the fork (*26, 27*). We envisage that MCM2 recycles parental histones to the lagging strand (Fig. 3F, right), while a separate pathway deposits parental histones on the leading strand. In this vein, it is conceivable that histone segregation can be regulated during development to drive asymmetric cell fates (*28*).

References and Notes

1. C. D. Allis, T. Jenuwein, The molecular hallmarks of epigenetic control. *Nat Rev Genet* **17**, 487-500 (2016).
2. R. P. Halley-Stott, J. B. Gurdon, Epigenetic memory in the context of nuclear reprogramming and cancer. *Brief Funct Genomics* **12**, 164-173 (2013).
3. C. Alabert, A. Groth, Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* **13**, 153-167 (2012).
4. E. I. Campos, J. M. Stafford, D. Reinberg, Epigenetic inheritance: histone bookmarks across generations. *Trends Cell Biol* **24**, 664-674 (2014).
5. C. Alabert *et al.*, Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev* **29**, 585-590 (2015).
6. V. Pospelov, G. Russev, L. Vassilev, R. Tsanev, Nucleosome segregation in chromatin replicated in the presence of cycloheximide. *J Mol Biol* **156**, 79-91 (1982).
7. A. T. Annunziato, Split decision: what happens to nucleosomes during DNA replication? *J Biol Chem* **280**, 12065-12068 (2005).
8. N. Petryk *et al.*, Replication landscape of the human genome. *Nat Commun* **7**, 10208 (2016).
9. O. Hyrien, in *The Initiation of DNA Replication in Eukaryotes,* D. Kaplan, Ed. (Springer, 2016), pp. 65-85.
10. G. Saredi *et al.*, H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex. *Nature* **534**, 714-718 (2016).
11. Z. Jasencakova *et al.*, Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol Cell* **37**, 736-743 (2010).

12. E. Pourkarimi, J. M. Bellush, I. Whitehouse, Spatiotemporal coupling and decoupling of gene transcription with DNA replication origins during embryogenesis in C. elegans. *Elife* **5**, (2016).

13. A. Baker *et al.*, Replication fork polarity gradients revealed by megabase-sized U-shaped replication timing domains in human cell lines. *PLoS Comput Biol* **8**, e1002443 (2012).

14. B. D. Pope *et al.*, Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**, 402-405 (2014).

15. J. R. Dixon *et al.*, Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380 (2012).

16. B. Bonev *et al.*, Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572 (2017).

17. E. Lieberman-Aiden *et al.*, Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science* **326**, 289-293 (2009).

18. A. Loyola, T. Bonaldi, D. Roche, A. Imhof, G. Almouzni, PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell* **24**, 309-316 (2006).

19. J. C. Black, C. Van Rechem, J. R. Whetstine, Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell* **48**, 491-507 (2012).

20. C. Cayrou *et al.*, The chromatin environment shapes DNA replication origin organization and defines origin classes. *Genome Res* **25**, 1873-1885 (2015).

21. R. Andersson *et al.*, An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455-461 (2014).

22. W. A. Whyte *et al.*, Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319 (2013).

23. A. Groth *et al.*, Regulation of replication fork progression through histone supply and demand. *Science* **318**, 1928-1931 (2007).

24. M. Foltman *et al.*, Eukaryotic replisome components cooperate to process histones during chromosome replication. *Cell Rep* **3**, 892-904 (2013).

25. H. Huang *et al.*, A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. *Nat Struct Mol Biol* **22**, 618-626 (2015).

26. R. Georgescu *et al.*, Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *Proc Natl Acad Sci U S A* **114**, E697-E706 (2017).

27. M. E. Douglas, F. A. Ali, A. Costa, J. F. X. Diffley, The mechanism of eukaryotic CMG helicase activation. *Nature* **555**, 265-268 (2018).

28. V. Tran, C. Lim, J. Xie, X. Chen, Asymmetric division of Drosophila male germline stem cell shows asymmetric histone distribution. *Science* **338**, 679-682 (2012).

29. The Encode Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57 (2012).

30. Q. L. Ying *et al.*, The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523 (2008).

31. Z. Tseng, T. Wu, Y. Liu, M. Zhong, A. Xiao, in *Cancer Genomics and Proteomics*. (Springer, 2014), pp. 11-22.

32. S. I. Presolski, V. P. Hong, M. G. Finn, Copper-Catalyzed Azide-Alkyne Click Chemistry for Bioconjugation. *Curr Protoc Chem Biol* **3**, 153-162 (2011).

33. P. V. Kharchenko, M. Y. Tolstorukov, P. J. Park, Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat Biotechnol* **26**, 1351-1359 (2008).

34.   F. Ramirez *et al.*, deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).

35.   E. Meshorer *et al.*, Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**, 105-116 (2006).

36.   Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

37.   G. Yu, L. G. Wang, Q. Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015).

38.   K. D. Brabanter, J. A. K. Suykens, B. D. Moor, Nonparametric Regression viaStatLSSVM. *J Stat Softw* **55**, (2013).

39.   J. M. Engreitz *et al.*, RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent Pre-mRNAs and chromatin sites. *Cell* **159**, 188-199 (2014).

40.   S. Rennie, M. Dalby, L. van Duin, R. Andersson, Transcriptional decomposition reveals active chromatin architectures and cell specific regulatory interactions. *Nat Commun* **9**, 487 (2018).

41.   I. Hiratani *et al.*, Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biol* **6**, e245 (2008).

42.   N. C. Durand *et al.*, Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst* **3**, 95-98 (2016).

supervised the project. N.P. and M.D. wrote the manuscript with input from A.G., R.A. and A.W. **Competing interests:** Authors declare no competing interests. **Data and material availability:** Data are deposited to NCBI GEO, accession GSE117274.

Supplementary Materials:

Materials and Methods

Figures S1-S9

Tables S1-S3

External Database S1

References (*30-42*)

Figure legends

Fig. 1. Parental histones segregate to both sister chromatids with a weak bias towards the leading strand.

(**A**) SCAR-seq technique. Partition of old/new histones is calculated as the proportion of Forward (red) and Reverse (blue) counts in genomic windows. (**B**) RFD and partition of H4K20me2 and H4K5ac at a genomic region. Initiation zone centers (black lines), active gene orientation (arrowheads) and active enhancers (bars) are shown. (**C**) Average RFD (blue) and partition of old (H4K20me2, orange) and new (H4K5ac, green) histones around initiation zones. (**D**) Partition at downstream (leading, dark shaded) and upstream edges (lagging, light shaded) of initiation zones with significant partition difference in each replicate (paired Wilcoxon signed-rank test, $P < 5.3 \times 10^{-14}$).

Fig. 2. TAD borders and genes flanking active enhancers show skewed histone PTM inheritance to sister chromatids.

(**A**) Average RFD (blue) and partition of H4K5ac (green), H4K20me2 (orange) and H3K36me3 (red) across scaled TADs split by compartment class (*16, 17*). (**B**) Spearman correlation of Hi-C directionality index (*16*) and RFD, histone PTM partition (colors as in **A**), transcriptional directionality measured by precision nuclear run-on (PRO-seq) (lilac) (*29*). (**C**) Density distribution of H3K36me3 partition and RFD in active (full) and inactive (dashed) forward (red) and reverse (blue) strand genes and intergenic regions (grey). (**D**) Average RFD and histone

10

PTM partition centered at enhancers (active, blue; inactive, light blue) (*21*) and super-enhancers (brown) (*22*).

Fig 3. MCM2 histone binding is required for parental histone recycling to the lagging strand.

(**A**) Average RFD (blue) and partition of old (H4K20me2, orange; H3K36me3, red) and new (H4K5ac, green) histones in WT (full) and MCM2-2A (dashed) around initiation zones. (**B**) RFD and histone PTM partition at a genomic region in WT and MCM2-2A. Initiation zone centers (black lines), active gene orientation (arrowheads) and active enhancers (bars) are shown. (**C**) Scatterplots of RFD and histone PTM partition in WT and MCM2-2A. Spearman's rank correlation coefficient in top left corner. (**D**) Scatterplot of H4K20me2 versus H4K5ac partition in MCM2-2A. Spearman's rank correlation coefficient shown in top right corner. (**E**) Fraction of initiation zones with nearest distance to predicted H4K20me2 partition breakpoints (orange) or random H4K20me2 bins (grey) in WT and MCM2-2A. Horizontal dotted line represents random mean fractions. (**F**) Model for segregation of parental histones H3-H4 in WT and MCM2-2A cells.