View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Apollo

Early life nutrition modulates the epigenetic state of specific rDNA genetic variants in mice

Michelle L. Holland1*†, Robert Lowe1*, Paul W. Caton2, Carolina Gemma1, Guillermo Carbajosa1, Amy F. Danson1, Asha A. M. Carpenter3, Elena Loche3, Susan E. Ozanne3, Vardhman K. Rakyan1†

1 The Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, United Kingdom

2Division of Diabetes and Nutritional Sciences, King's College London, 150 Stamford Street, London SE1 9NH, United Kingdom

3University of Cambridge Metabolic Research Laboratories and MRC Metabolic Diseases Unit, Institute of Metabolic Science, Addenbrookes Hospital, Cambridge, CB2 0QQ, United Kingdom

*These authors contributed equally to this work.

† Corresponding authors.

A sub-optimal early life environment, due to poor nutrition or stress during pregnancy can influence life-long phenotypes in the progeny. Epigenetic factors are thought to be key mediators of these effects. We show that protein restriction in mice from conception until weaning induces a linear correlation between growth restriction and DNA methylation at ribosomal DNA (rDNA). This epigenetic response remains into adulthood and is restricted to rDNA copies associated with a specific genetic variant within the promoter. Related effects are also found in models of maternal high-fat or obesogenic diets. Our work identifies environmentally induced epigenetic dynamics that are dependent on underlying genetic variation, and establishes rDNA as a genomic target of nutritional insults.

Exposure to an adverse *in utero* environment can have a long-lasting influence on adult phenotypes in mammals, a process termed ‘developmental programming’ (*1, 2*). Consequently, there is great interest in identifying the molecular mechanisms that underlie developmental programming; and in this regard, modulation of the epigenome has emerged as a potentially key contributing factor (*3, 4*).

To explore epigenetic mechanisms involved in developmental programming, we employed a maternal protein restriction model (*5*). Inbred C57BL/6J mice were mated, and G0 females assigned to either a protein restricted (PR; 8% protein) or control diet (C; 20% protein; Table S1) until their G1 offspring were weaned. Only male G1s were studied in detail (n = 146). From weaning onwards, both ‘G1-PR’ and ‘G1-C’ were kept on control diet until sacrifice at 16-20 weeks. Consistent with previous work, G1-PR males were ~25% lighter than G1-C at weaning (*5*) (Fig 1A; $P = 2 \times 10^{-6}$). PRs also displayed reduced spontaneous locomotor activity (Fig S1) and reduced glucose-stimulated insulin secretion (Fig S2).

Several studies have shown that developmental programming can perturb DNA methylation profiles (*1*). We used reduced representation bisulfite sequencing (RRBS) to generate genome-scale, single base resolution DNA methylomes for 8 G1-PR and 8 G1-C mice, initially focussing on sperm as it can be isolated to a high degree of purity. After genome-wide correction, we identified a single 1916 bp differentially methylated region (DMR) hypermethylated in G1-PR males, that mapped to *Rn45s* on chromosome 17 (mm10, Table S2). Further analysis revealed that *Rn45s* displays 98% homology to the 973-2883 bp region of the ribosomal DNA (rDNA) consensus (Fig 1B). rDNA is excluded from genome assemblies because of its multi-copy nature.

We therefore re-mapped the RRBS data to the consensus sequence for mouse rDNA (BK000964) and confirmed extensive hypermethylation in PR sperm across the entire promoter and coding regions (~13.5 kb) (Fig 1B). Directly correlating weaning weight with rDNA levels revealed that G1-PR displayed a significantly greater negative correlation between weaning weight and DNA methylation compared to G1-C (Wilcoxon rank sum test; $P < 2.2 \times 10^{-16}$).(Fig 1C). This correlation was not confounded by weight or age at death (Fig S3).

In the C57BL/6J genome, rDNA is comprised of hundreds of copies in large arrays on chromosomes 12, 15, 18, and 19, but only a subset are actively transcribed (*6*). Silenced copies are methylated at a CpG site located 133 bp upstream of the *45S-rRNA* transcriptional start site (Fig 1C), and this prevents binding of the transcription factor UBF and assembly of RNA polymerase I (*7*). We therefore focussed on CpG-133 in the rest of the study using high-throughput sequencing (>1000X coverage) of bisulfite PCR amplicons (henceforth termed bisPCR-seq). BisPCR-seq analysis of the same samples profiled by RRBS revealed strong concordance between the two methods (Fig S4, $\tau = 0.77$, $P = 1 \times 10^{-5}$).

As rDNA copies within a single genome are genetically polymorphic (*8*), we designed the bisPCR-seq amplicon targeting CpG-133 to simultaneously assay previously documented genetic variation at position -104 (C or A, Fig 2A. Note: this variant does not overlap a CpG site) (*9*). CpG-133 methylation levels were substantially lower for the C-variant relative to the A-variant (Fig 2A) and there was no interaction between C-variant associated CpG-133 methylation and weaning weight in G1-PR or G1-C sperm (Fig S5). On the other hand, CpG-133 methylation levels of A-variant

rDNA (which we denote as CpG-133^A) were negatively correlated with weaning weight (Fig 2B; τ = -0.43, P = 0.017). Figure 2B incorporates additional males (9 G1-PR and 7 G1-C from litters not represented in the RRBS data), reinforcing the negative correlation between weaning weight and total CpG-133 methylation observed in the RRBS dataset. BisPCR-seq analysis of *in vitro* methylated samples confirmed that there was no amplification bias associated with either variant (Fig S6). We also confirmed sperm purity by analysis of several parentally imprinted regions (Fig S7). Analysis of liver using BisPCR-seq revealed a strong correlation with sperm within individual G1-C (Fig S8; τ = 0.72, P = 0.00028) or G1-PR animals (Fig S8; τ = 0.54, P = 0.0041). Liver CpG-133^A methylation was negatively correlated with weaning weight in G1-PR (τ = -0.46, n = 24, P = 0.0016) but not in G1-C (n = 26) (Fig 2C). Collectively, these data demonstrate that PR exposure induces not just rDNA hypermethylation, but also a linear relationship between a phenotypic outcome (weaning weight) and CpG-133^A methylation in sperm and liver, which is maintained into adulthood.

Further exploration of the bisPCR-seq data revealed inter-individual variation in the relative copy number of rDNA harbouring the A-variant at position -104, even in an inbred genetic background. This underlying copy number variation (which we denote as '%A' i.e. the percentage of A-variant reads relative to total coverage for this amplicon) was positively correlated between sperm and liver of both G1-C (Fig S9; τ = 0.77, P = 7 x 10^{-5}) and G1-PR animals (Fig S9; τ = 0.73, P = 3.7 x 10^{-5}). The accuracy of the bisPCR-seq derived estimates of %A were confirmed by whole genome re-sequencing of 6 individuals (Fig S10; τ = 1, P = 0.0028). Furthermore, CpG-133^A methylation correlated positively with %A in G1-PR sperm (Fig 2D; τ = 0.71, P = 1.9 x 10^{-5}) and liver (Fig S11; τ = 0.31, P = 0.034), but not in G1-C sperm (Fig 2D) or liver (Fig S11). Therefore, early life

PR induces an interdependence between underlying variation in the relative abundance of a specific genetic variant of rDNA and methylation state of this variant at a functionally relevant CpG site.

rDNA copies that lack methylation at CpG-133 have the potential to be transcriptionally active (*7*). As most methylation is localized to A-variant rDNA, both the level of methylation at CpG-133^{A} and the relative abundance of this variant (i.e. %A) will contribute towards transcriptional competency. This interaction can be represented as the percentage of total rDNA copies that are both A-variant and unmethylated at CpG-133 (which we denote as '%A^{UN}'. Note: %A^{UN} is different to simply considering the percentage of CpG-133^{A} that is unmethylated). As expected, %A^{UN} correlates between the sperm and liver of G1-C and G1-PR mice (Fig S12). To confirm the functional significance of %A^{UN}, we analyzed a regulatory non-coding RNA (promoter associated RNA; *pRNA*) that spans the rDNA promoter (Fig 3A). *pRNA* is transcribed from early replicating and unmethylated rDNA copies (*10*). It functions in *trans* to recruit nucleolar chromatin remodelling complex and DNA methyltransferase to silenced rDNA copies (*11*). Using RT-qPCR, we generated a *pRNA*-derived amplicon spanning the genetic polymorphism at position -104, and determined the percentage of A-variant reads after high throughput sequencing (pRNA(%A)). The pRNA(%A) reads in liver were consistently and positively correlated with %A^{UN} (Fig 3B), but not %A, (Fig S13). Therefore, %A^{UN} is indicative of transcriptional competency at rDNA.

The *45S-rRNA* is co-transcriptionally cleaved at position +650 within the 5' external transcribed spacer, and the first 650 nt is then rapidly degraded (*12*). We assessed the abundance of the nascent, uncleaved *45S-rRNA* precursor via RT-qPCR targeting the first 650 nt. In the liver of G1-C, *45S-*

rRNA abundance did not correlate with CpG-133^{A} methylation, %A, or %A^{UN} (Fig 3C, Fig S14). In PR males, *45S-rRNA* levels did not correlate with CpG-133^{A} methylation or %A, but correlated positively with %A^{UN} (Fig 3C; $\tau = 0.52$, $P = 0.021$, Fig S14). Therefore, PR exposure induces a correlation between transcriptional competency and *45S-rRNA* levels.

As rDNA expression is sensitive to nutrient availability (*13*), the types of effects we describe could be a conserved feature of other nutritional developmental programming models. We identified a recent study in which the authors fed C57BL/6J G0 females a low-fat (LF) or high-fat (HF) diet from 3 weeks prior to pregnancy up till when the male G1 offspring were weaned on to a LF diet till sacrifice at 9 weeks (*14*). Their RRBS analysis of G1 liver did not identify any maternal diet-induced DNA methylation differences. We mapped their raw sequencing reads to rDNA and found that early life exposure to HF induces CpG-133^{A} hypermethylation in the G1s (Fig 4A, $P = 0.0098$), again CpG-133^{C} showed lower methylation levels that were not affected by diet. Unfortunately, there were insufficient mice to examine correlations between %A and methylation or weaning weight. Next, we generated bisPCR-seq data for G1 male C57BL/6J mice from a recent study of maternal obesogenic diet (*15*) (elevated fat and sugar content). G0 females were fed either control or obesogenic diet 6 weeks prior to mating until the G1 offspring were weaned onto control diet and sacrificed at 6 months. G1 males exposed *in utero* to obesogenic diet showed hypermethylation at CpG-133^{A} (Fig 4B; $P = 0.017$).

Recently, Shea et al., (2015) reported a study in which they exposed male C57BL/6J mice to one of three different diets (PR, HF, or caloric restriction) post-weaning (*16*). They identified substantial inter-individual genetic and methylomic variability at rDNA, but no consistent diet

induced effects. Although part of the reason for the discrepant conclusions could be that they didn't discriminate between the 'A' or 'C' genetic variants, the more likely explanation is differences in developmental timing of the dietary insults as we analysed exposures spanning only the period between conception and weaning. Previous human epidemiological and animal studies suggest that early life is a critical time when exposures can have a long-term phenotypic effects on the offspring (*17*).

In summary, we have described an example of a mammalian 'epiallele' whose epigenetic state is influenced by an interaction between the underlying genotype and early life environment, and this correlates with transcriptional and phenotypic outcomes. A schematic model of the effects we describe is presented in Fig S15. Our work, in combination with previous demonstrations in flies and yeast (*18, 19*), identifies rDNA as a genomic target of various nutritional insults that is conserved amongst non-mammalian and mammalian models. Exploration of such interactions at rDNA in humans could provide novel insights into the molecular basis of some complex phenotypes and diseases.

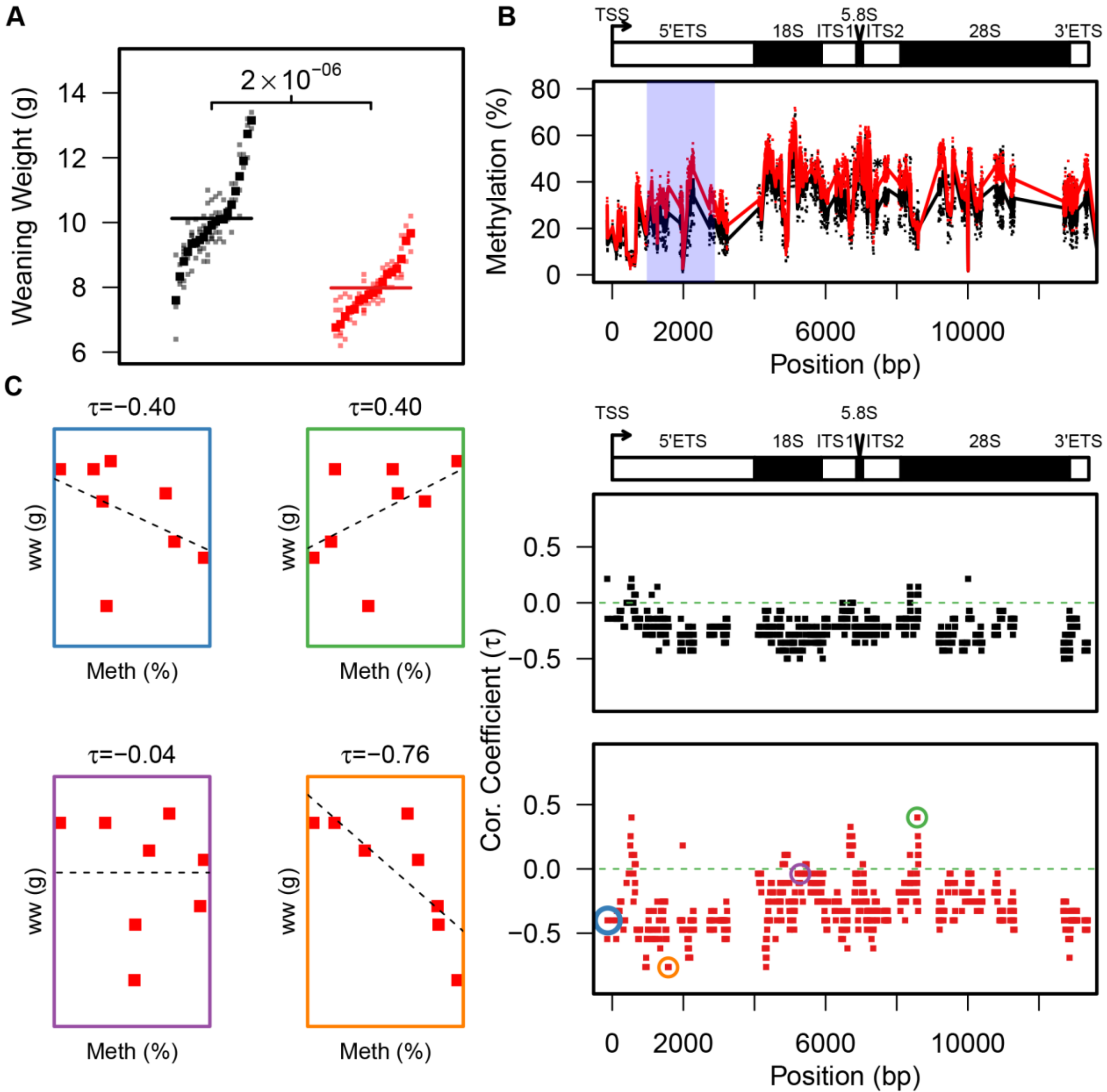

Figure 1: Maternal protein restriction (PR) induces a correlation between rDNA methylation and weaning weight. (A) Weaning weight of G1-PR males (red, 62 individual animals from 17 different litters) was reduced compared to G1-C (black, 84 individual males from 20 different litters) (t-test, $P = 2 \times 10^{-6}$ using litter means, and $P < 2.2 \times 10^{-6}$ using individual mice). Small points represent individual animals, larger squares represent mean of a given G1 litter. (**B**) RRBS analysis of rDNA in G1 sperm shows that PRs (n = 8) are hypermethylated compared to Controls (n = 8). The line represents mean methylation, and points represent individual mice. The rDNA schematic

shows the rRNA subunits, transcriptional start site (TSS), external transcribed spacer (ETS), internal transcribed spacer (ITS). The *Rn45S* regions identified in the initial RRBS analysis is 98% homologous to the region shaded blue. **(C)** The correlation coefficient (τ) between weaning weight (ww) and DNA methylation across the rDNA. Highlighted are examples of a positive correlation (green), close to zero (purple) and negative (orange). CpG-133 is circled in blue.

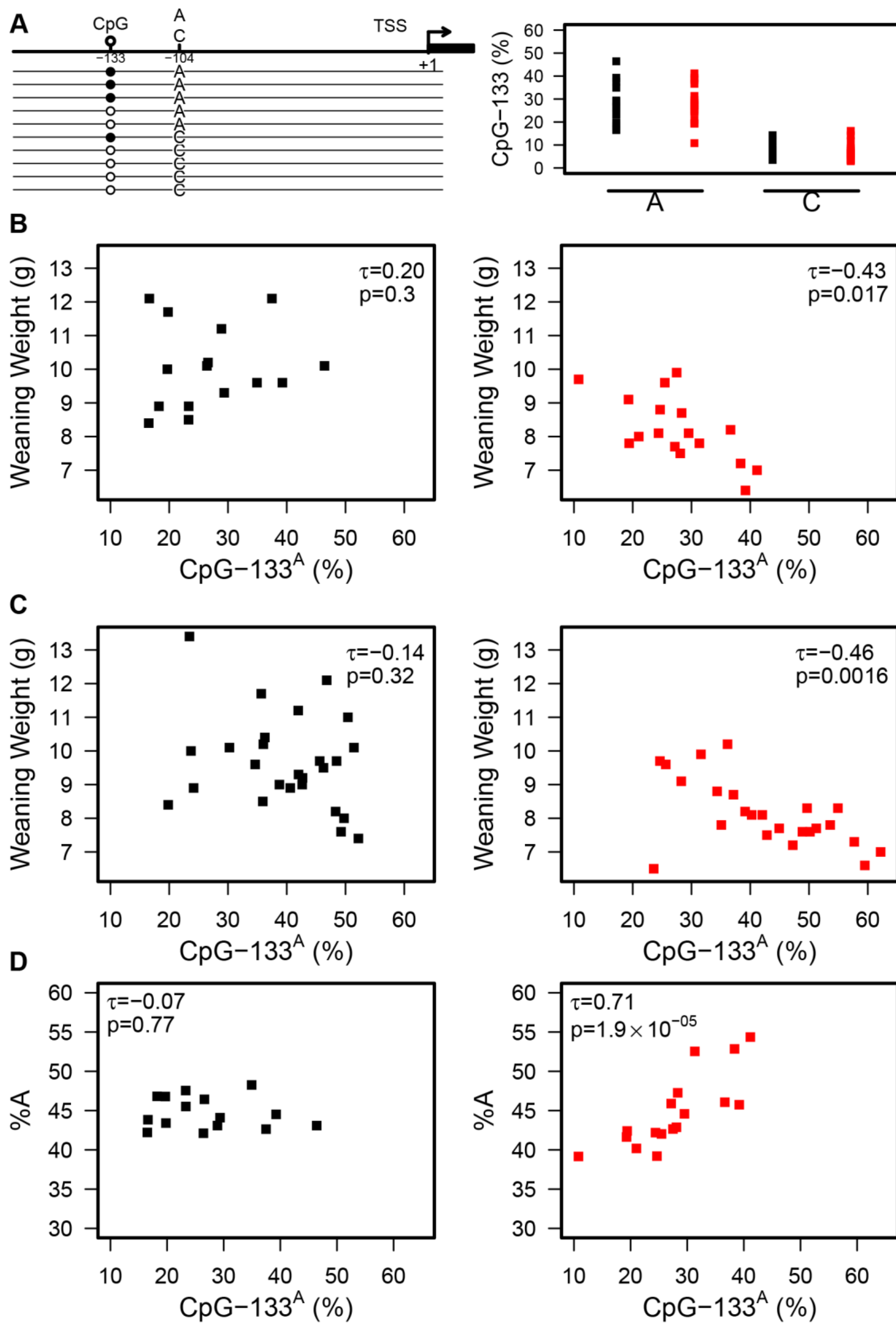
A
CpG
A
C
TSS
−133
−104
+1
CpG−133 (%)
A
C
B
Weaning Weight (g)
τ=0.20
p=0.3
τ=−0.43
p=0.017
CpG−133^A (%)
C
τ=−0.14
p=0.32
τ=−0.46
p=0.0016
D
%A
τ=−0.07
p=0.77
τ=0.71
p=1.9 × 10^−05

Figure 2: Diet induced methylation dynamics are restricted to a specific genetic variant of rDNA. **(A)** BisPCR-seq amplicons were generated to simultaneously analyse methylation at CpG-133 (methylation indicated by black circle) and genetic variation at position -104 (A or C) (left panel). CpG-133 methylation levels in sperm for each genetic variant is shown for G1-Control (black, n=15), and G1-PR (red, n=17). **(B)** In sperm, methylation levels at 'A' variant associated CpG-133 sites (CpG-133^{A}) and weaning weight are not correlated in G1-Control (black, n=15; $\tau = 0.20$, $P = 0.30$), but negatively correlated in G1-PR (red, n=17; $\tau = -0.43$, $P = 0.017$). **(C)** In liver**,** CpG-133^{A} methylation levels and weaning weight are not correlated in G1-Control (black, n=26; $\tau = -0.14$, $P = 0.32$), but negatively correlated in G1-PR (red, n=24; $\tau = -0.46$, $P = 0.0016$). **(D)** In sperm, CpG-133^{A} methylation levels are uncorrelated with the percentage of total rDNA copies with an A-variant (%A) in G1-Control (black, n=15; $\tau = -0.07$, $P = 0.77$), but positively correlated in G1-PR (red, n=17; $\tau = 0.71$, $P = 1.9 \times 10^{-5}$).

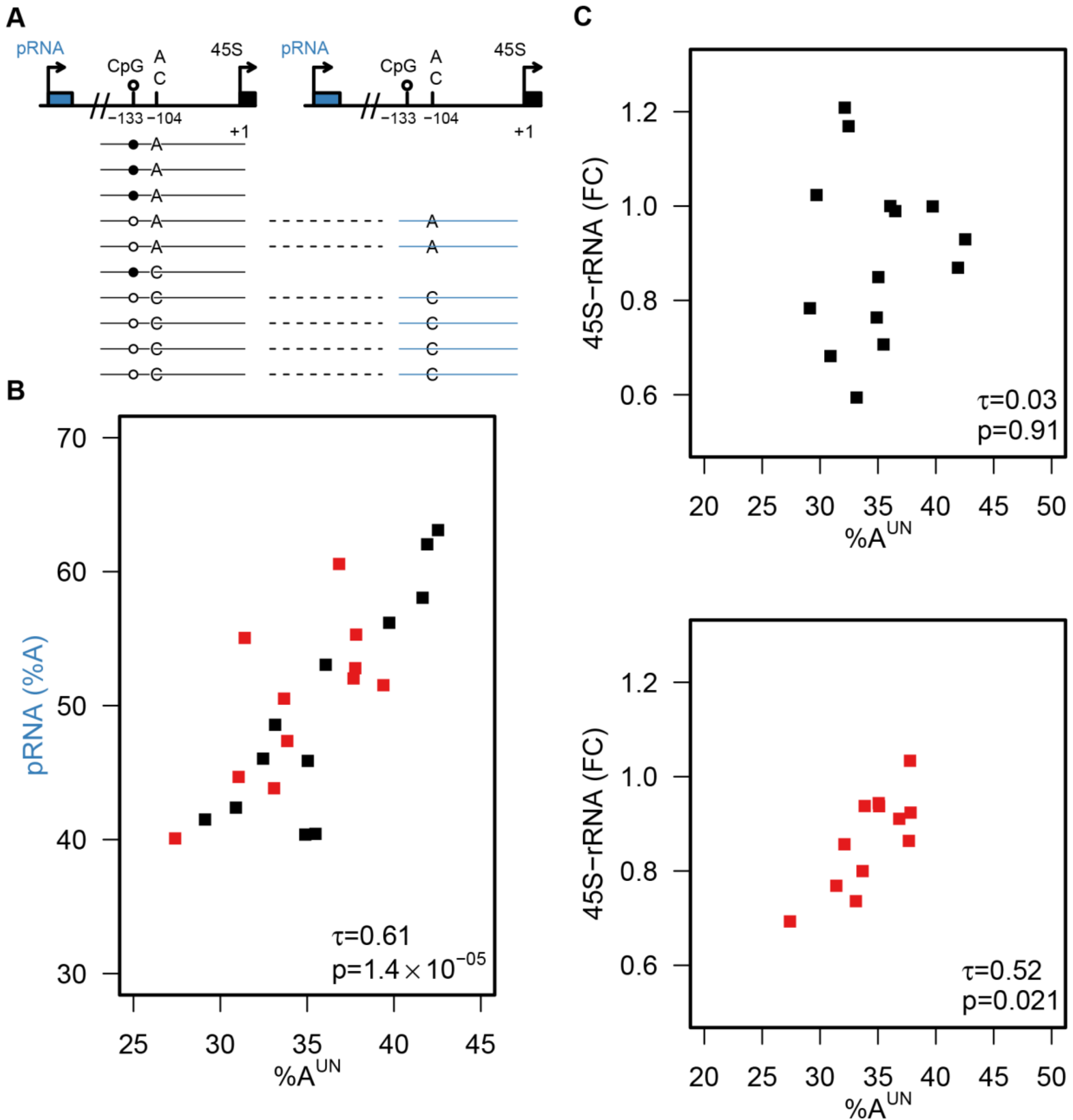

Figure 3. Functional consequences of altered rDNA dynamics. (A) *pRNA* is transcribed from early replicating rDNA copies (assumed to be unmethylated at CpG-133). Therefore, the percentage of *pRNA* reads that encode an A at position -104 (*pRNA*(%A); indicated in blue; right) should reflect the proportion of A-variant rDNA copies that are unmethylated at CpG-133 (%A^{UN}) **(B)** pRNA(%A) positively correlates with (%A^{UN}) in both G1-Control (black) and G1-PR (red) liver (total, n=23, $\tau = 0.61$, $P = 1.4 \times 10^{-5}$). **(C)** %A^{UN} is not correlated with the abundance of *45S-*

rRNA in liver of G1-Control (black; n=14, $\tau = 0.03$, $P = 0.91$), but is positively correlated in liver of G1-PR (red; n=12, $\tau = 0.52$, $P = 0.021$).

Figure 4. Maternal high fat or obesogenic diet induces hypermethylation at CpG-133^{A}. (A) RRBS raw sequencing reads (obtained from Cannon *et al*., 2014) were mapped to the rDNA consensus. G0 dams that were fed either a low (M^{L}) or high fat (M^{H}) diet prior to conception and up until the G1s were weaned. Data shown here is from the livers of 9 week old G1 males that were placed on a low fat diet from weaning up to sacrifice. n=10 for group. Note – we only re-analysed data from Cannon et al. for the dietary groups analogous to the design of our PR model

(B) G0 Dams were fed either a Control (C) or Obesogenic (O) diet 6 weeks prior to conception and up until the G1s were weaned. bsPCR-seq data shown here is from the livers of 6 month old G1 males that were placed on a control diet from weaning up to sacrifice (CC: n=7, OC: n=8).

References

1. D. S. Fernandez-Twinn, M. Constancia, S. E. Ozanne. *Seminars in cell & developmental biology*, (2015).
2. M. A. Hanson, P. D. Gluckman. *Physiological reviews* **94**, 1027-1076 (2014).
3. E. W. Tobi *et al.*, *Nature communications* **5**, 5592 (2014).
4. E. J. Radford *et al.*, *Science* **345**, 1255903 (2014).
5. Z. Vucetic *et al.*, *Neuroscience* **168**, 359-370 (2010).
6. V. G. Dev, R. Tantravahi, D. A. Miller, O. J. Miller. *Genetics* **86**, 389-398 (1977).
7. R. Santoro, I. Grummt. *Molecular cell* **8**, 719-725 (2001).
8. H. Tseng *et al.*, *PloS one* **3**, e1843 (2008).
9. Y. H. Shiao *et al.*, *PloS one* **6**, e22266 (2011).
10. C. Mayer, K. M. Schmitz, J. Li, I. Grummt, R. Santoro. *Molecular cell* **22**, 351-361 (2006).
11. R. Santoro, K. M. Schmitz, J. Sandoval, I. Grummt. *EMBO reports* **11**, 52-58 (2010).
12. S. Kass, N. Craig, B. Sollner-Webb. *Molecular and cellular biology* **7**, 2891-2898 (1987).
13. A. Murayama *et al.*, *Cell* **133**, 627-639 (2008).
14. M. V. Cannon *et al.*, *PloS one* **9**, e90335 (2014).
15. C. E. Aiken, J. L. Tarry-Adkins, N. C. Penfold, L. Dearden, S. E. Ozanne. *FASEB J* **30**, 1548-1556 (2016).
16. J. M. Shea *et al.*, *Developmental cell* **35**, 750-758 (2015).
17. N. El Hajj, E. Schneider, H. Lehnen, T. Haaf. *Reproduction* **148**, R111-120 (2014).
18. C. V. Jack *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **112**, 9674-9679 (2015).
19. J. C. Aldrich, K. A. Maggert, *PLoS genetics* **11**, e1005148 (2015).

Supplementary information

Supplementary file: Methods, Figure S1-S15, Table S1-S4.

Acknowledgements

This work was supported by the following grants and fellowships; Biotechnology and Biological Sciences Research Council, UK (BB/M012494/1) to VKR, and (BB/G00711/X/1) to VKR and CG. Research Council UK Academic Fellowship to MLH. RL is supported by EU-FP7 BLUEPRINT. British Heart Foundation (FS/12/64/30001) and MRC (MC_UU_12012/4) to SEO. This research utilised Queen Mary's MidPlus computational facilities, supported by QMUL Research-IT and funded by EPSRC grant EP/K000128/1. We thank King's College London FWB Genomics Centre and Barts and The London Genome Centre for performing high throughput sequencing.

Author contributions

MLH, CG, PWC, VKR, AAMC and EL performed all experiments. RL and GC conducted the bioinformatic analyses. VKR and SEO provided reagents and contributed to experimental design. MLH, RL, VKR conceived the study and prepared the manuscript. All authors discussed the results and interpretation and approved the final manuscript.

Data deposition statement

Whole Genome sequencing data and pRNA sequencing was submitted to BioProject under ID (PRJNA293403). RRBS, and BisPCR-seq were all submitted to the NCBI Gene Expression Omnibus under accession (GSE72610). All data will be linked under the BioProject ID.

Competing interest declaration

We declare no conflict of interest

Corresponding authors

- Prof. Vardhman Rakyan

  Email: v.rakyan@qmul.ac.uk

  Ph: +442078828140

- Dr. Michelle Holland

  Email: m.holland@qmul.ac.uk

  Ph: +442078823433