Childhood location correlates with epigenetic age and methylation stability in British-Bangladeshi migrants

Reinhard Stöger[1*], Minseung Choi[2], Gregory Leeman[1#], Richard D. Emes[3,4], Khurshida Begum[5], Philippa Melamed[6], Gillian R. Bentley[5,7]

1 School of Biosciences, University of Nottingham, LE12 5RD, United Kingdom
2 School of Medicine, Stanford University, USA
3 School of Veterinary Medicine and Science, University of Nottingham, LE12 5RD, United Kingdom.
4 Advanced Data Analysis Centre, University of Nottingham, LE12 5RD, United Kingdom.
5 Department of Anthropology, Durham University, Durham, United Kingdom
6 Faculty of Biology, Technion-Israel Institute of Technology, Haifa, Israel
7 Wolfson Research Institute for Health and Wellbeing, Durham University, Durham, United Kingdom

E-Mail addresses: RS: reinhard.stoger@nottingham.ac.uk; MS: minseung@stanford.edu; GL gregoryleeman@outlook.com; RDE: richard.emes@nottingham.ac.uk; KB: khurshida.begum@qmul.ac.uk; PM: philippa@tx.technion.ac.il; GRB: g.r.bentley@durham.ac.uk

***Corresponding author:** Reinhard Stöger, E-mail: reinhard.stoger@nottingham.ac.uk

Abstract

Background

Migration from one environment to another often causes marked changes in developmental conditions. Here we compare epigenetic ageing and stability of the epigenetic maintenance system among British-Bangladeshi women who grew up in Bangladesh (adult migrants), where there are higher pathogen loads and poorer health care, to second-generation Bangladeshis who grew up in the UK. In our previous studies of these migrants, those who spent their childhoods in Bangladesh also had lower levels of reproductive hormones and a shorter reproductive lifespan compared to those who grew up in the UK, suggesting life history trade-offs during development. In the present study, we hypothesised that women who grew up in Bangladesh would have *i*) an older epigenetic/biological age compared to the women with a childhood in the UK and *ii*) that differences in the pace of epigenetic ageing might also be reflected by altered stability of DNA methylation marks.

Results

Illumina EPIC array methylation data from buccal tissue was used to establish epigenetic age estimates from 15 adult migrants and 11 second-generation migrants, aged 18-35 years. Using residuals from linear regression of DNA methylation-based biological age (DNAm age) on the chronological age, the results showed significant differences (p=0.016) in epigenetic age estimates: women whose childhood was in Bangladesh are on average 6.02 (± 2.34) years older, than those who grew up in London. We further investigated the efficiency of the epigenetic maintenance system which purportedly is reflected by epigenetic clocks. Methylation states of CpGs at the *LHCGR/LHR* locus, which contributes to Horvath's multi tissue epigenetic clock were evaluated. Based on the Ratio of Concordance Preference (RCP) approach that uses double-stranded methylation data, we find that maintenance of epigenetic information is more stable in women who grew up in Bangladesh.

Conclusions

The work supports earlier findings that adverse childhood environments lead to phenotypic life history trade-offs. The data indicate that childhood environments can induce subtle changes to the epigenetic maintenance system that are detectable long after exposure occurred. The implication of such a finding warrants further investigation as it implies that a less flexible epigenetic memory system established early in life could reduce the capacity to respond to different environmental conditions in adult life.

Keywords

Childhood, migrants, epigenetic age, RCP values, epigenetic stability, DNA methylation, accelerated ageing, Bangladesh, UK.

Background

Reproductive lifespans vary among individuals. Genetic variants associated with these complex traits, which include timing of puberty, age at first birth and age at menopause are closely related to fitness and undergo purifying selection [1,2]. The genetic architecture of reproductive ageing has been investigated largely in women of European ancestry. However, a limited number of studies in other populations suggests shared genetic underpinnings of these reproductive phenotypes, albeit with noticeable variations in effect allele frequencies and effect estimates in women of different ethnic groups [3–5]. Environmental exposures likely contribute to variations in heritability estimates and the phenotypic heterogeneity detected within and across different ethnic populations [6,7].

Our earlier work identified strong correlations between childhood environmental conditions and adult reproductive function [8–10]. In particular, Bangladeshi women who migrated as young adults to London, have lower levels of reproductive steroids when compared to British-Bangladeshi women who moved to the UK prior to the age of eight and women who were born in London to first-generation Bangladeshi immigrants [7,9–11]. An upbringing in Bangladesh is generally associated with a shortened reproductive lifespan, while its duration is longer for women with Bangladeshi ancestry, whose childhoods were spent in London [9]. Timing of reproductive functions across the life course correlates with the rate of ageing in other body systems [12].

Geographically and culturally the British-Bangladeshis women in these studies have a comparable background. They are all ethnic Bengalis and originally stem from a relatively affluent middle-class population in the northeast of Bangladesh and now live in East London. A possible environmental factor that distinguishes between the two childhood locations is the exposure to higher and recurrent infectious disease loads in Bangladesh [13–15]. Indeed, by mimicking early-life immune challenges in a mouse model, we replicated some of the distinct reproductive phenotypes characteristic of women with a childhood in Bangladesh, in including delayed onset of puberty lower ovarian reserve [16].

At the cellular level, environmental factors influence the chromatin state of the genome [17]. Stored as epigenetic information, cells have the capacity to retain some memory of past developmental and environmental conditions [18]. Methylation of genomic DNA is part of the epigenetic information storage system in mammalian cells where it is primarily confined to cytosines of CpG dinucleotides [19]. Methylation levels of discrete CpG sites have been

used to develop remarkably accurate estimators of age. Such 'epigenetic clocks' link developmental and maintenance processes to biological ageing [reviewed in [20]]. Pace of ageing can vary and result in a mismatch between chronological and biological age of an individual [21].

Here, we explore the possible association between chronological age, biological ageing and an epigenetic maintenance system in Bangladeshi women of prime reproductive age (18-35 years old). The women of this study live within the same ethnic community in London but can be divided into two groups: those with a childhood in the UK, and those with a childhood in Sylhet, a city in the northeast region of Bangladesh. Using buccal cell DNA from these London-based Bangladeshi women, we recently identified genome-wide, altered DNA methylation levels between the two groups [16]. Since these DNA methylation measurements were generated on the MethylationEpic array platform, we re-examined the data using 'Horvath's clock', a multi-tissue age-estimator with a robust relationship between chronological age and DNA methylation-based (DNAm) age [22].

Results and Discussion

Accelerated DNAm Age measured with Horvath's epigenetic clock

We find that the correlation between chronological age and DNAm age does not differ significantly between women who grew up the UK ('UK' group; n=11) and women who grew up in Bangladesh ('Bangladesh' group; n=15). That is, chronological age affects DNAm age in a similar way in both groups (Additional file 1). However, regression analysis showed that the y-intercepts of the UK and Bangladeshi groups differ significantly (p=0.0083) / Additional file 1). This suggested that a childhood in Bangladesh correlates with DNAm Age predictions that differ noticeable when compared with epigenetic age estimates for women of the UK group.

The tick rate of epigenetic clocks is increased by many different environmental factors, including psychological traumas, smoking, asthma, alcohol, infections and hormonal changes following menopause [23–26]. Such acceleration of epigenetic age is best measured by residuals obtained from regressing DNAm age on chronological age [22]. Indeed, the pace of epigenetic ageing is accelerated in women with a childhood in Bangladesh and overall differs significantly from the UK group (p=0.016) (Figure 1). This altered pace of biological ageing is consistent with our previous observations that women who grow up in Bangladesh have a shorter reproductive lifespan and chronically lower levels of reproductive hormones [9,13,15]; reviewed in [7]. Although our finding of accelerated

epigenetic ageing rests on a small number of sampled individuals, it highlights the limited utility of epigenetic clocks as a tool to determine the age – and consequently eligibility considerations - of asylum-seekers [27].

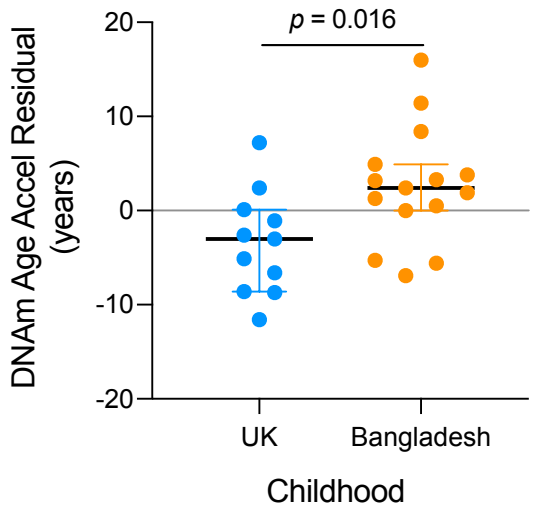


Fig. 1 Differences in pace of epigenetic ageing

Plot of DNAm Age Accel Residuals, with each data point representing an individual. The colour indicates the corresponding dataset: blue = childhood in UK, orange = childhood in Bangladesh. The median is indicated by a horizontal line with upper and lower hinges representing the 25th and 75th percentiles. A positive or negative value indicates that the estimated epigenetic/biological age of the sample is higher or lower, respectively, than expected based on chronological age.

Epigenetic stability of a clock locus

The tick rate of Horvath's epigenetic clock is thought to reflect the rate at which work is done to maintain epigenetic stability [20,22]. It is possible to infer epigenetic stability by analysing double-stranded DNA methylation data with a new metric, Ratio of Concordance Preference (RCP) [28]. We used the RCP metric to estimate epigenetic stability at the *Luteinizing Hormone/Choriogonadotropin Receptor* (*LHCGR/LHR*) gene, which plays an important role in reproductive function. The *LHCGR* locus contains a CpG site, which contributes to Horvath's DNAm Age clock [22].

We find that RCP estimates are generally higher for the 'Bangladeshi' group of women (Figure 2). Higher RCP estimates indicate higher levels of epigenetic stability [28]. That is, the methylation states of CpGs at *LHCGR* locus are more often identical on the two strands

of individual DNA molecules of 'Bangladeshi' individuals when compared to 'UK' individuals. We note that the RCP estimates are based on a relatively small number of data points (Additional file 1), yet they are sufficient to indicate subtle differences in the workings of the epigenetic maintenance system between two groups of women who appear to age at different rates.

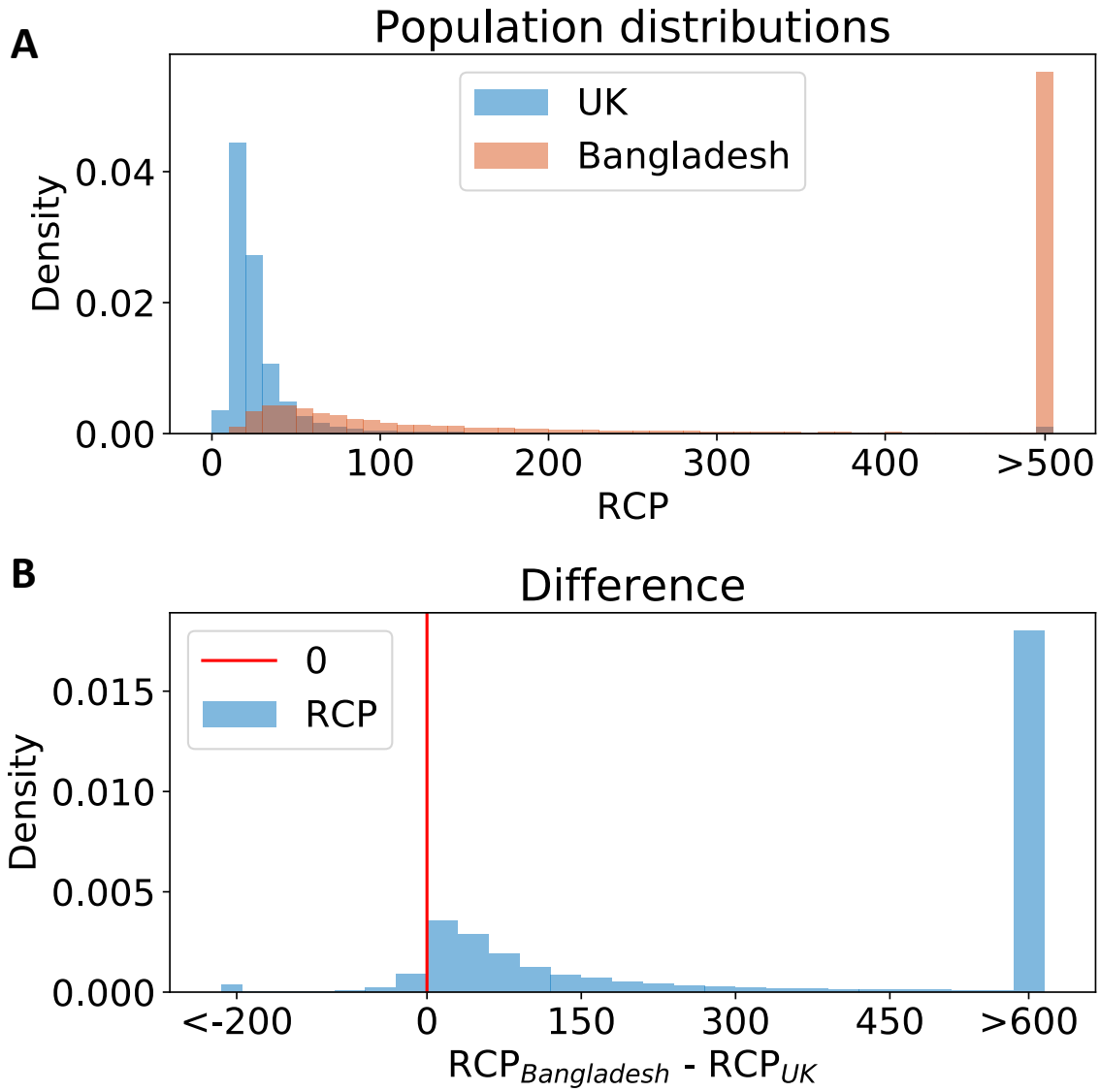


Fig. 2 Inferences of DNA methylation stability differ between UK and Bangladeshi samples at the epigenetic-clock associated *LHCGR/LHR* locus.

A) Ratio of Concordance Preference (RCP) is a metric that infers stability/flexibility of methylation states at matching CpG sites (CpG dyads) on the parent and daughter strand of individual DNA molecules, without assuming any specific enzymatic mechanisms of DNA methylation. Flexibility, indicated by RCP values near 1, indicates that the methylation system has no preference for either concordance or discordance of the methylation state at

a CpG dyad and follows the random model. High RCP values - with the extreme approaching infinity - indicate high stability, where epigenetic maintenance systems have complete preference for concordant methylation states of CpG dyads (they are either methylated or unmethylated); none, or very few CpG dyads are hemi-methylated. Shown are the RCP distributions taken from bootstrap samples, weighing each individual evenly within each group (UK = blue, Bangladesh = orange). The sampled population of double-stranded DNA molecules – and the corresponding methylation states of CpG dyads - revealed a clear preference for a more stable epigenetic maintenance system in operation at the *LHR* locus in women with a childhood in Bangladesh, when compared to inferred RCP values for the samples from women with a childhood in London. **B)** Testing if the bootstrap samples of RCP differences (Bangladesh vs UK childhood) are significantly differnt. The red line is set at 0. The p-value is derived from this as the proportion of samples to the left of 0 (a one-tailed test to examine whether Bangladesh RCPs are significantly greater than UK RCP values). Two-tailed p-value is the double of that amount. p= 0.026 (one tailed); p= 0.052 (two tailed)

Conclusions

The results of our study support a large body of work demonstrating phenotypic plasticity in response to environments encountered during early life. A childhood in Bangladesh measurably accelerates epigenetic/biological ageing in women, when compared to women of same chronological age (18 -35 yrs) and ethnicity, who were born and brought up in London, UK. The multi tissue epigenetic clock is thought to register the workings of developmental and epigenetic maintenance systems linking these processes with the life course [20,22]. Our study is one of the first to test if differences in function of the epigenetic maintenance system can be linked with epigenetic age estimators. The findings indicate that subtle differences in the stability of epigenetic states are indeed associated with biological ageing and opens a new line of investigation.

Methods

DNA methylation data and establishment of DNAm Age

Genome-wide cytosine methylation levels were established using the Illumina HumanMethylationEPIC BeadChip Array following isolation of genomic DNA from buccal cells DNA (DNeasy Blood & Tissue Kit (Qiagen). Multidimensional scaling (MDS) plots

179 indicated that no significant batch effects were skewing the MethylationEPIC BeadChip data
180 sets. The data were processed with the Bioconductor/minfi package. CpG probes associated
181 with known SNPs were removed, as were those with a detection probability of <0.01. Probes
182 on both X and Y chromosomes were retained. Methylation beta values (0-1) were
183 normalized by SWAN. The methylation data set (GSE133355 study) is accessible on the
184 Gene Expression Omnibus (GEO) data platform at:
185 https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133355

186 **Determination of DNAm Age and age acceleration**
187 A file with the beta values obtained from the Illumina HumanMethylationEPIC BeadChip
188 Array work (see above) was used to establish the epigenetic age (DNAm age and
189 AgeAccelerationResiduals) with Horvath's method [22]. The underlying algorithms are
190 available through the online DNA methylation calculator (http://dnamage.genetics.ucla.edu/).
191

192 **Generation of bisulfite hairpin data / methylation states of CpG dyads at the**
193 ***LHCGR/LHR* locus**
194 We have previously described in detail the concept and procedure of generating
195 authenticated, non-redundant double-stranded DNA methylation data [19,29,30]. In brief,
196 genomic sequence information surrounding the *LHR* clock-CpG site [one of the 353 CpG
197 sites contributing to Horvath's clock [22]; Illumina cluster ID cg12351433 / chr2:48982957-
198 48982957 / UCSC Genome Browser (GRCh37/hg19)] was used to identify suitable
199 restriction recognition sites to generate 3', or 5'-overhangs, respectively, for the ligation of
200 UMI-barcoded hairpin linkers. Specifically, restriction enzymes StyI or BstXI (New England
201 Biolabs) were used. Combinations of the following primers were used to amplify hairpin-
202 linked, bisulfite converted DNA:
203 bsLHR-R1 5'-RCAAATCAAAACAAAACAAACTC-3';
204 bsLHR-R2 5'-CACTAAACACTATCRCAAATCAAAAC-3';
205 bsLHR-F1 5'-TAGTAGGAAGGAGGTTATTGG-3';
206 bsLHR-F2 5'-GTAGGTTAAGGTAGAGTAGATTTAG-3';
207 bsLHR-F3 5'-GAATTGGGTTTTTGCGGTTTGTTAG-3'.
208 Further information of the hairpin-concept and of the barcoded and batch-stamped hairpin
209 linkers (Eurofins Genomics) are provided in Additional file 1.
210

211 **Processing of the sequencing data:** Fold is a web application for the analysis of the output
212 of hairpin-bisulphite sequencing data. Specifically, the programme reconstructs, visualises,

and generates statistics on the double-stranded CpG methylation patterns of the original cohort of DNA molecules. This is achieved by first 'realigning' the top and bottom strand of the molecule about the hairpin, in which the programme attempts to manage 'PCR slippage', and other sequencing errors. Then algorithm then identifies and categorises CpG dyads, which is possible due to the previous bisulphite conversion of unmethylated cytosine to uracil (and so recognised as tyrosine when sequenced). For example, fully methylated dyads are those regions in where the reconstructed top strand is C-G and the bottom is G-C. Similarly, fully unmethylated dyads are those where the top is T-G and the bottom is G-T. In addition, the programme calculates a metric: 'Ratio of Concordance Principle' which quantifies the concordance of methylation between the top and bottom strands of the DNA molecule (0=complete discordance, 1=random concordance, inf=complete concordance). This metric represents the preference of the summation of epigenetic mechanisms of the cell to either maintain or obscure methylation patterns of the DNA in the cells at the time the sample was taken. The functions of Fold was written in R and the web application is written in PHP. The live web application can be found at http://www.gregoryleeman.com/fold, and the repository can be found at https://github.com/gregoryleeman/fold.

Analysis and comparison of RCPs at the *LHCGR/LHR* locus

RCP values are based on double-stranded DNA methylation data derived from sequences of individual hairpin bisulfite PCRs products. RCP analyses were done following the procedures described in [28] with the small addition of bootstrapping individuals within each population. The additional step in the procedure helps to address the possibility of uneven sampling from a larger population. The analysis procedures in brief are described below. Each population RCP distribution was drawn through hierarchical bootstrap sampling. For each of 20,000 bootstrap samples, individuals in each population were sampled with replacement, and double stranded DNA sequences of each of the sampled individuals were in turn sampled with replacement. Dyad counts were then normalised such that each individual had the same number of dyads. The normalised dyad counts were then summed, corrected for failed bisulfite conversions (rate of 0.0039, measured empirically) and inappropriate conversions (rate of 0.017, estimated as described in [31] [Genereux et al., 2008]), and used to compute the RCP value. A bootstrap sample of the RCP difference was computed by taking the difference of the RCP values sampled for the two populations.

For one-tailed comparison tests, with which we examine directional differences, we determined the p-value as the proportion of bootstrap-difference samples to the left of 0. For

two-tailed tests, with which we can detect differences in any direction, we determined the p-value as twice the smaller proportion of the bootstrap difference samples on either side of 0.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the Gene Expression Omnibus (GEO) data platform https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133355

All data generated or analysed during this study are included in this published article and its supplementary information file.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the Biotechnology and Biological Science Research Council (BBSRC) and the Economic and Social Research Council (ESRC) grant ES/N000471/1(to GB, PM and RS).

Authors' contributions

RS: Conceptualisation, Funding acquisition, Experimental work, Analysis, Resources, Supervision, Data curation, Project administration, Writing – original draft, Writing – review & editing.

MC: Analysis, review & editing

GL: Analysis, Coding

RDE: Analysis, Data curation

KB: Experimental work, Resources

PM: Funding acquisition, Project administration, Writing - review & editing.

GRB: Conceptualisation, Funding acquisition, Project administration, Writing - review & editing.

Acknowledgements

We thank Kamila Derecka for technical support, Ian C. W. Hardy for statistical advice and Steve Horvath for information on DNAm age analysis.

References

1. Day FR, Ruth KS, Thompson DJ, Lunetta KL, Pervjakova N, Chasman DI, et al. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. Nat Genet. Nature Publishing Group; 2015;47:1294–303.

2. Zeng J, De Vlaming R, Wu Y, Robinson MR, Lloyd-Jones LR, Yengo L, et al. Signatures of negative selection in the genetic architecture of human complex traits. Nat Genet. Nature Publishing Group; 2018;50:746–53.

3. Fernández-Rhodes L, Malinowski JR, Wang Y, Tao R, Pankratz N, Jeff JM, et al. The genetic underpinnings of variation in ages at menarche and natural menopause among women from the multi-ethnic Population Architecture using Genomics and Epidemiology (PAGE) Study: A trans-ethnic meta-analysis. Thameem F, editor. PLoS One. Public Library of Science; 2018;13:e0200486.

4. Horikoshi M, Day FR, Akiyama M, Hirata M, Kamatani Y, Matsuda K, et al. Elucidating the genetic architecture of reproductive ageing in the Japanese population. Nat Commun. Nature Publishing Group; 2018;9:1–9.

5. Carty CL, Spencer KL, Setiawan VW, Fernandez-Rhodes L, Malinowski J, Buyske S, et al. Replication of genetic loci for ages at menarche and menopause in the multi-ethnic Population Architecture using Genomics and Epidemiology (PAGE) study. Hum Reprod. Oxford University Press; 2013;28:1695–706.

6. Parent AS, Teilmann G, Juul A, Skakkebaek NE, Toppari J, Bourguignon JP. The Timing of Normal Puberty and the Age Limits of Sexual Precocity: Variations around the World, Secular Trends, and Changes after Migration. Endocr. Rev. Endocr Rev; 2003. p. 668–93.

7. Bar-Sadeh B, Rudnizky S, Pnueli L, Bentley GR, Stöger R, Kaplan A, et al. Unravelling the role of epigenetics in reproductive adaptations to early-life environment. Nat. Rev. Endocrinol. Nature Research; 2020. p. 519–33.

8. Magid KS, Uddin Ahamed F, Lawson DW, Chatterton RT, Bentley GR. Effects of adult migration on male salivary testosterone. Am J Hum Biol. 2006;18:262.

9. Begum K, Muttukrishna S, Sievert L, Sharmeen T, Murphy L, Chowdhury O, et al. Ethnicity or environment: Effects of migration on ovarian reserve among Bangladeshi women in the UK. Fertil Steril. 2015;under revi.

10. Nunez-De La Mora A, Bentley GR, Choudhury OA, Napolitano DA, Chatterton RT. The impact of developmental conditions on adult salivary estradiol levels: why this differs from progesterone? Am J Hum Biol. 2007/10/25. 2008;20:2–14.

11. Houghton LC, Cooper GD, Booth M, Chowdhury OA, Troisi R, Ziegler RG, et al.

Childhood environment influences adrenarcheal timing among first-generation Bangladeshi migrant girls to the UK. Sear R, editor. PLoS One. 2014/10/14. 2014;9:e109200.

12. Perry JRB, Murray A, Day FR, Ong KK. Molecular insights into the aetiology of female reproductive ageing. Nat. Rev. Endocrinol. Nature Publishing Group; 2015. p. 725–34.

13. Núñez-de la Mora A, Chatterton RT, Choudhury OA, Napolitano DA, Bentley GR, Nunez-de la Mora A, et al. Childhood conditions influence adult progesterone levels. Fisk NM, editor. PLoS Med. 2007/05/17. 2007;4:e167.

14. Siddique AK, Baqui AH, Eusof A, Zaman K. 1988 floods in Bangladesh: pattern of illness and causes of death. J Diarrhoeal Dis Res. 1991;9:310–4.

15. Murphy L, Sievert L, Begum K, Sharmeen T, Puleo E, Chowdhury O, et al. Life course effects on age at menopause among Bangladeshi sedentees and migrants to the UK. Am J Hum Biol. 2012/11/24. 2013;25:83–93.

16. Bar-Sadeh B, Pnueli L, Begum K, Leeman G, Emes RD, Stöger R, et al. Early-life environment programs reproductive strategies through epigenetic regulation of SRD5A1 1. bioRxiv. Cold Spring Harbor Laboratory; 2020;2020.09.16.299560.

17. Zhu J, Adli M, Zou JY, Verstappen G, Coyne M, Zhang X, et al. Genome-wide Chromatin State Transitions Associated with Developmental and Environmental Cues. Cell. 2013;152:642–54.

18. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003;33 Suppl:245–54.

19. Laird CD, Pleasant ND, Clark AD, Sneeden JL, Hassan KMA, Manley NC, et al. Hairpin-bisulfite PCR: Assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. Proc Natl Acad Sci U S A. National Academy of Sciences; 2004;101:204–9.

20. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet. Nature Publishing Group; 2018;19:371–84.

21. Khan SS, Singer BD, Vaughan DE. Molecular and physiological manifestations and measurement of aging in humans. Aging Cell. Blackwell Publishing Ltd; 2017;16:624–33.

22. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013/10/22. BioMed Central Ltd; 2013;14:R115.

23. Dhingra R, Nwanaji-Enwerem JC, Samet M, Ward-Caviness CK. DNA Methylation Age—Environmental Influences, Health Impacts, and Its Role in Environmental Epidemiology. Curr. Environ. Heal. reports. Springer; 2018. p. 317–27.

24. Quach A, Levine ME, Tanaka T, Lu AT, Chen BH, Ferrucci L, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging (Albany NY). 2017;9:419–

353 46.
354 25. Austin MK, Chen E, Ross KM, McEwen LM, Maclsaac JL, Kobor MS, et al. Early-life
355 socioeconomic disadvantage, not current, predicts accelerated epigenetic aging of
356 monocytes. Psychoneuroendocrinology. 2018;97:131–4.
357 26. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic
358 biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10:573–91.
359 27. Editorial. Molecular test of age highlights difficult questions. 2018;561.
360 28. Choi M, Genereux DP, Goodson J, Al-Azzawi H, Allain SQ, Simon N, et al. Epigenetic
361 memory via concordant DNA methylation is inversely correlated to developmental potential
362 of mammalian cells. Barsh GS, editor. PLOS Genet. 2017;13:e1007060.
363 29. Miner BE, Stöger RJ, Burden AF, Laird CD, Hansen RS. Molecular barcodes detect
364 redundancy and contamination in hairpin-bisulfite PCR. Nucleic Acids Res. 2004;32.
365 30. Stöger R. Hairpin-Bisulfite PCR. Methods Mol Biol. NLM (Medline); 2021;2198:287–99.
366 31. Genereux DP, Johnson WC, Burden AF, Stöger R, Laird CD. Errors in the bisulfite
367 conversion of DNA: modulating inappropriate- and failed-conversion frequencies. Nucleic
368 Acids Res. 2008/11/06. Oxford University Press; 2008;36:e150.
369
370

Additional file 1: Chronological age vs DNAm Age

Estimates of DNA methylation age (DNAm Age)

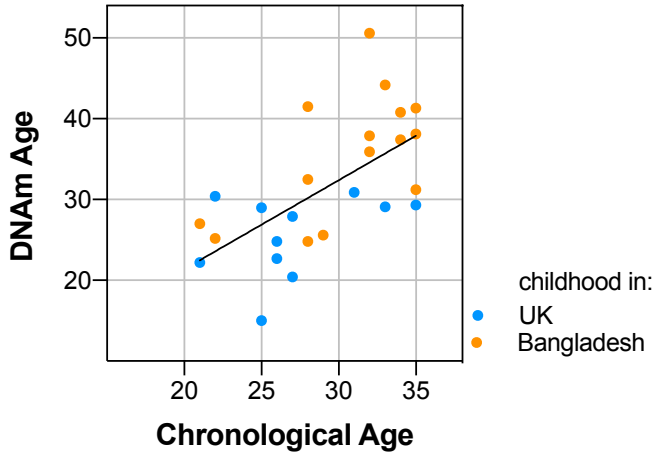
The online calculator (https://dnamage.genetics.ucla.edu/) was used to generate estimates of DNAm Age, AgeAccelerationDiff (= DNAmAge-Age), and AgeAccelerationResidual (= the recommended age acceleration measure based on Horvath's linear regression model [Horvath S (2013) DNA methylation age of human tissues and cell types. *Genome Biol 14*(10):R115 PMID: 24138928].

id	Sample ID	Childhood location	Age at time of sample collection	DNAmAge	AgeAcceleration Diff	AgeAcceleration Residual
1	X201247480032_R01C01	uk	22	30.4	8.4	7.2
5	X201247480032_R04C01	uk	25	29.0	4.0	2.4
8	X201247480032_R07C01	uk	27	27.9	0.9	-1.1
14	X201247480031_R03C01	uk	31	30.9	-0.1	-2.6
15	X201247480031_R04C01	uk	33	29.1	-3.9	-6.6
22	X201247480036_R03C01	uk	35	29.3	-5.7	-8.7
23	X201247480036_R04C01	uk	26	24.8	-1.2	-3.0
24	X201247480036_R05C01	uk	27	20.4	-6.6	-8.6
25	X201247480036_R06C01	uk	25	15.0	-10.0	-11.6
30	X201247480034_R03C01	uk	21	22.2	1.2	0.1
31	X201247480034_R04C01	uk	26	22.7	-3.3	-5.1
2	X201247480032_R02C01	bangladesh	34	37.4	3.4	0.5
4	X201247480032_R03C01	bangladesh	32	50.6	18.6	16.0
12	X201247480031_R01C01	bangladesh	35	41.3	6.3	3.2
13	X201247480031_R02C01	bangladesh	21	27.0	6.0	4.9
16	X201247480031_R05C01	bangladesh	33	44.2	11.2	8.4
17	X201247480031_R06C01	bangladesh	28	41.5	13.5	11.4
18	X201247480031_R07C01	bangladesh	35	31.2	-3.8	-6.9
19	X201247480031_R08C01	bangladesh	29	25.6	-3.4	-5.6
20	X201247480036_R01C01	bangladesh	32	35.9	3.9	1.3
21	X201247480036_R02C01	bangladesh	28	24.8	-3.2	-5.3
26	X201247480036_R07C01	bangladesh	32	37.9	5.9	3.3
27	X201247480036_R08C01	bangladesh	35	38.1	3.1	0.0
28	X201247480034_R01C01	bangladesh	34	40.8	6.8	3.8
32	X201247480034_R05C01	bangladesh	28	32.5	4.5	2.4
36	X201247480034_R07C01	bangladesh	22	25.2	3.2	1.9

Prism GraphPad 8 analysis / Chronological age vs DNAm Age:

The slopes of the regression lines between the 'UK' group (n=11) and the 'Bangladesh' group (n=15) are not significantly different (F = 1.136; DFn = 1, DFd = 22 p=0.29). Therefore, a single slope for data of the entire cohort can be established: the pooled slope equals 0.8128 (Additional Fig 1a). However, the intercepts are significantly different (F = 8.341. DFn = 1, DFd = 23, p=0.0083)

Additional Fig 1a



Additional Fig 1a: Plot of predicted methylation age (Horvath clock) against chronological age**.** The scatter plot shows DNA methylation age vs. chronological age vs. and the line in which DNA methylation age was regressed on chronological age using both, the 'UK' and the 'Bangladesh' data sets. 'UK' = Bangladeshi women who grew up in London, UK (blue); 'Bangladesh' = Bangladeshi women who grew up in Sylhet, Bangladesh. Each data point represents an individual, with the colour indicating the corresponding dataset.

Genstat analysis / Chronological age vs DNAm Age:

Genstat analysis of covariance yielded similar results to those obtained by Prism GraphPad 8 analysis, in that Bangladeshi women with a childhood in the UK and Bangladeshi women with a childhood in Bangladesh are affected by chronological age the same way - the two slopes are the same based on a minimal adequate mode (Additional Fig 1b). Parameters needed to reconstruct the lines are indicated in yellow:

Estimates of parameters

Parameter	estimate	s.e.	t(23)	t pr.
Constant	10.78	8.35	1.29	0.209
Country UK	-7.19	2.49	-2.89	0.008
chronological_age	0.813	0.269	3.02	0.00

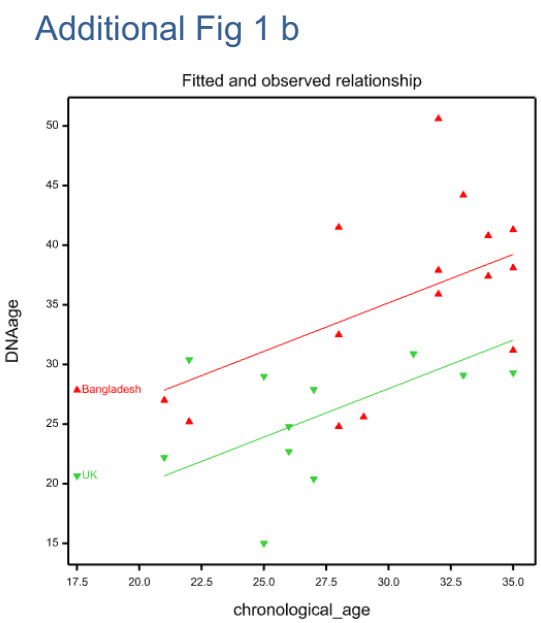
Parameters for factors are differences compared with the reference level:

Factor	Reference level
Country	Bangladesh

Accumulated analysis of variance

Change	d.f.	s.s.	m.s.	v.r.	F pr.	
+ chronological_age	1	660.09	660.09	19.57	<.001	
+ Country	1	282.95	282.95	8.39	0.008	
+ chronological_age.Country	1	38.32	38.32	1.14	0.298	
Residual	22	741.94	33.72			
- chronological_age.Country	-1	-38.32	38.32	1.14	0.298	interaction not significant
- Country	-1	-282.95	282.95	8.39	0.008	country significant
+ Country	1	282.95	282.95	8.39	0.008	country put back into model
- chronological_age	-1	-309.57	309.57	9.18	0.006	age significant
+ chronological_age	1	309.57	309.57	9.18	0.006	age put back into model, so that the final model can be plotted.
Total		25	1723.29		68.93	

Additional Fig 1b: The minimal adequate model (Genstat analysis):



DNA sequence surrounding the *LHR*-clock CpG site analysed by hairpin bisulfite PCR

The CpG site [clock CpG] at the *Luteinizing Hormone/Choriogonadotropin Receptor (LHCGR/LHR)* locus, which contributes to 'Horvath's clock' (Horvath, 2013) was PCR-amplified following hairpin linker-ligation and sodium bisulfite conversion. The hairpin PCR products also captured methylation information of flanking CpGs (highlighted in red): four CpGs with the StyI-hairpin linker approach and eight CpGs with the BstXI-hairpin linker approach, respectively:

BstXI

CAGCTTCAGCAGCTGCAGCGCCGAGAACCGCTGCTTCATGGCCGGCGAACTGGGCTTCTGCGGCTTGCCAGTGTC

StyI

TTGGACGGCCTCTGAGTGCGGCCCGCACCCCTCTCCCCGCCCCTTGGCTGCCCTTCCACACACCCTTCCCTGCCG

clock CpG

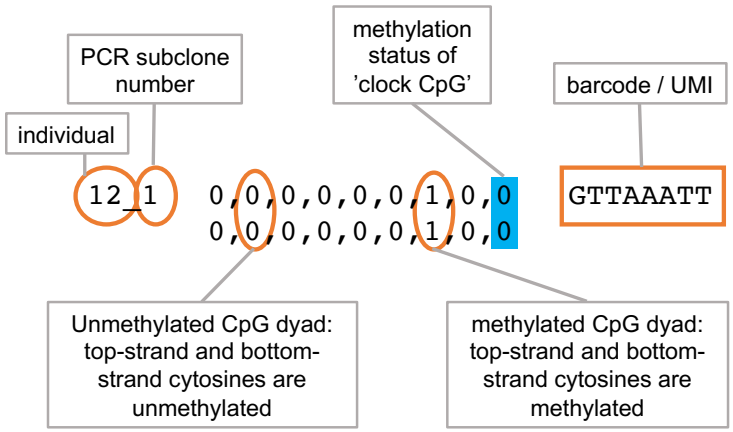
GCCCGCCCCTGCCCTCCCCCTCTTACCGCGCACCCCGCTGAGTCTGCTCTGCCTTGACCTGCGACAGTGCCCAGT

BstXI- and StyI-hairpin linkers used in this study

Name:	Sequence:	individual	converted
hLHR-hp-III	ACAGTGCADDDDDDDDTGCACTGTtgtc	#12	ATAGTGTADDDDDDDDTGTATTGTtgtT
hLHR-hp-1	ACATGGCADDDDDDDDTGCCATGTtgtc	#16	ATATGGTADDDDDDDDTGTTATGTtgtT
hLHR-hp-2	ATCGTGCADDDDDDDDTGCACGATtgtc	#17	ATTGTGTADDDDDDDDTGTATGATtgtT
hLHR-hp-3	AAGGTGCADDDDDDDDTGCACCTTtgtc	#19	AAGGTGTADDDDDDDDTGTATTTTtgtT
hLHR-hp-4	AACCTGCADDDDDDDDTGCAGGTTtgtc	#22	AATTTGTADDDDDDDDTGTAGGTTtgtT
hLHR-hp-5	TAGCACGTDDDDDDDDACGTGCTAtgtc	#13	TAGTATGTDDDDDDDDATGTGTTAtgtT
hLHR-hp-6	TTACACGTDDDDDDDDACGTGTAAtgtc	#15	TTATATGTDDDDDDDDATGTGTAAtgtT
hLHR-hp-7	TTCCACGTDDDDDDDDACGTGGAAtgtc	#23	TTTTATGTDDDDDDDDATGTGGAAtgtT
hLHR-hp-8	TTTCACGTDDDDDDDDACGTGAAAtgtc	#29	TTTTATGTDDDDDDDDATGTGAAAtgtT
hLHR-hp-9	TTGAACGTDDDDDDDDACGTTCAAtgtc	#30	TTGAATGTDDDDDDDDATGTTTAAtgtT
hLHR-hp-10	TTGTACGTDDDDDDDDACGTACAAtgtc	#31	TTGTATGTDDDDDDDDATGTATAAtgtT
StyI			
hLHR-hp-11	caagACATGGCADDDDDDDDTGCCATGT	#5/6	taagATATGGTADDDDDDDDTGTTATGT
hLHR-hp-12	caagATCGTGCADDDDDDDDTGCACGAT	#7	taagATTGTGTADDDDDDDDTGTATGAT
hLHR-hp-13	caagAAGGTGCADDDDDDDDTGCACCTT	#8	taagAAGGTGTADDDDDDDDTGTATTTT
hLHR-hp-14	caagAACCTGCADDDDDDDDTGCAGGTT	#10	taagAATTTGTADDDDDDDDTGTAGGTT
hLHR-hp-15	caagTAGCACGTDDDDDDDDACGTGCTA	#25	taagTAGTATGTDDDDDDDDATGTGTTA
hLHR-hp-16	caagTTACACGTDDDDDDDDACGTGTAA	#16	taagTTATATGTDDDDDDDDATGTGTAA
hLHR-hp-17	caagTTCCACGTDDDDDDDDACGTGGAA	#17	taagTTTTATGTDDDDDDDDATGTGGAA
hLHR-hp-18	caagTTTCACGTDDDDDDDDACGTGAAA	#18 / #27	taagTTTTATGTDDDDDDDDATGTGAAA
hLHR-hp-19	caagTTGAACGTDDDDDDDDACGTTCAA	#20 / #28	taagTTGAATGTDDDDDDDDATGTTTAA
hLHR-hp-20	caagTTGTACGTDDDDDDDDACGTACAA	#36 / #32	taagTTGTATGTDDDDDDDDATGTATAA

Double-stranded methylation data of individual DNA molecules derived from buccal cells

The data shown on page 5 are processed, in that the methylation status of matching CpG sites of the top- and bottom strands (CpG dyads) are indicated as methylated (=1), or unmethylated (=0). Information of individual, double-stranded DNA molecules is displayed as follows:



Hairpin methylation data / Bangladesh-childhood

12_1	0,0,0,0,0,0,1,0,0 0,0,0,0,0,0,1,0,0	GTTAAATT
12_2	0,0,0,0,0,0,0,0,1 0,0,0,0,0,0,0,0,1	GTATTATT
12_3	0,0,0,0,0,0,0,0,1 0,0,0,0,0,0,0,0,1	TGGTTTTA
12_4	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	GGTTGAGG
12_5	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	ATTTGTTG
12_6	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TAGGAGTG
12_7	0,0,0,1,0,0,0,0,0 0,0,0,0,0,0,0,0,0	GGTTTGAA
12_8	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	GGTGTGTT
12_9	0,0,0,0,0,0,1,0,0 0,0,0,0,0,0,1,0,0	AGATGTAA
12_10	0,0,0,1,0,0,0,0,0 0,0,0,1,0,0,0,0,0	TTGGAGGG
12_11	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	ATTTAGTT
12_12	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TGTAAAGA
12_13	0,0,1,1,0,0,0,0,0 0,1,1,1,1,0,0,0,0	TGTTTTTT
12_14	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,1	ATTTTAGG
12_15	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TTGGATTT
12_16	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TTGAAAGG
12_17	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	GGGTTAAG
13_1	0,0,0,0,1,1,0,0,0 0,0,0,0,0,0,0,0,0	AAGAGAGT
13_2	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TTGTTATA
13_3	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TAAATGTA
13_4	0,0,0,1,0,0,0,0,1 0,0,0,1,0,0,0,0,1	TTAGATGG
13_5	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TGATATAA
13_6	0,0,0,0,0,0,0,1,0 0,0,0,0,0,0,0,0,0	GGGTTGGG
13_7	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	GATTGATA
13_8	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	GGGTTAAG
13_9	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TAGTATAT
13_10	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	AAAAATAA
13_11	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	ATAGGAAG
13_12	0,0,0,0,0,0,0,0,1 0,0,0,0,0,0,0,0,1	TGGGATGA
13_13	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TATGTTTG
13_14	0,0,0,0,1,1,0,0,0 0,0,0,0,1,1,0,0,0	GAGAAATG
13_15	0,1,0,0,0,0,0,0,0 0,1,0,0,0,0,0,0,0	TTTATTGT
13_16	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TTTGTAGA
13_17	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	ATGTTAGT
13_18	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	ATGAGGGG
13_19	0,0,0,0,0,0,1,0,0 0,0,0,0,0,0,0,0,0	GTTTGAAG
19_1	0,0,0,0,0,0,0,0,0 0,1,0,0,0,0,0,0,0	GTGAATAT
19_2	0,0,0,0,0,0,0,0,0 0,0,0,0,0,0,0,0,0	TAGGGAAT

19_3	0,0,0,0,0,0,1,1,1	GTGGGGTG
	0,0,0,0,0,0,1,1,1	
19_4	0,0,0,0,0,0,0,0,0	GAAATAGT
	0,0,0,0,0,0,0,0,0	
19_5	0,0,0,0,0,0,1,0,1	GTGTGTAT
	0,0,0,0,0,0,1,0,1	
19_6	0,0,0,0,0,0,0,0,0	TGATGTGA
	0,0,0,0,0,0,0,0,0	
19_7	0,0,0,0,0,0,0,0,0	ATAGTAAT
	0,0,0,0,0,0,0,0,0	
19_8	0,0,0,0,0,0,0,0,0	ATATGTTC
	0,0,0,0,0,0,0,0,0	
19_9	1,1,0,0,0,0,0,0,0	AGTTGAGG
	1,1,0,0,0,0,0,0,0	
19_10	0,0,0,0,0,0,0,0,0	TTTAAAAA
	0,0,0,0,0,0,0,0,0	
19_11	0,0,0,0,0,0,0,0,0	GATAATTT
	0,0,0,0,0,0,0,0,0	
19_12	0,0,0,0,0,0,0,0,0	GTTGTTGT
	0,0,0,0,0,0,0,0,0	
19_13	0,0,0,0,0,0,0,0,0	TTGTGTAT
	0,0,0,0,0,0,0,0,0	
19_14	0,1,1,0,0,0,0,0,0	TTGGGTAT
	0,1,1,0,0,0,0,0,0	
19_15	0,0,0,0,0,0,1,0,0	ATTTATTT
	0,0,0,0,0,0,1,0,0	
19_16	0,0,0,1,0,0,0,0,0	GTGGAGGT
	0,0,0,1,0,0,0,0,0	
19_17	0,0,0,0,0,0,0,0,1	TAAGTAGG
	0,0,0,0,0,0,0,0,1	
19_18	0,0,0,0,0,1,0,0,0	AATGTTGT
	0,0,0,0,0,1,0,0,0	
19_19	0,0,0,0,0,0,0,0,0	TTAAGGGT
	0,0,0,0,0,0,0,0,0	
19_20	0,0,0,0,0,0,0,0,0	TTAAAGGT
	0,0,0,0,0,0,0,0,0	
19_21	0,0,0,0,0,0,0,0,0	TAGATTTG
	0,0,0,0,0,0,0,0,0	
19_22	0,0,0,0,0,0,0,0,0	GGGATTTT
	0,0,0,0,0,0,0,0,0	
19_23	0,0,0,0,0,0,0,0,0	AGATTGTG
	0,0,0,0,0,0,0,0,0	
19_24	0,0,0,0,0,0,0,0,0	TTAATAAT
	0,0,0,0,0,0,0,0,0	
20_1	0,0,1,1,1	AGAGATGG
	0,0,1,1,1	
20_2	0,0,1,1,1	GAAGGATA
	0,0,1,1,1	
20_3	0,0,1,1,1	TAGGTAGG
	0,0,1,1,1	
20_4	0,0,0,0,0	GGGGATGG
	0,0,0,0,0	
20_5	0,0,0,0,0	AAATGTGG
	0,0,0,0,0	
20_6	0,0,0,0,0	GTAGTTTA
	0,0,0,0,0	
20_7	0,0,1,0,1	ATAAAGAG
	0,0,1,0,1	
20_8	1,1,1,0,1	GTGAGGAG
	1,1,1,0,1	
20_9	0,0,0,0,0	GATGGGGT
	0,0,1,0,1	
20_10	0,0,0,0,0	TATGAGTG
	0,0,0,0,0	
20_11	0,0,0,0,0	AAAAATTG
	0,0,0,0,0	
21_1	1,1,1,1,1	ATGAGGAT
	1,1,1,1,1	
21_2	1,1,1,1,1	GGTGATTG
	1,1,1,1,1	
21_3	1,1,1,1,1	TGGGATAA
	1,1,1,1,1	
21_4	1,1,1,1,1	GGATGTGA
	1,1,1,1,1	
21_5	1,1,1,1,1	TAAGGGGG
	1,1,1,1,1	
36_1	0,1,0,1,0	GGAGAAGT
	0,1,0,1,0	

36_2	0,0,0,0,0	AGGGAGTT
	0,0,0,0,0	
36_3	0,0,0,0,0	GGAGAGGA
	0,0,0,0,0	
36_4	0,0,0,0,0	GGAGAGGG
	0,0,0,0,0	
36_5	0,0,0,0,0	AGAGTAGT
	0,0,0,0,0	
36_6	0,0,0,0,0	AGGGAGGA
	0,0,0,0,0	
36_7	0,0,0,0,0	AAATTTAG
	0,0,0,0,0	
36_8	0,0,0,0,0	GGGTAATA
	0,0,0,0,0	
36_9	0,0,0,0,0	ATATGGAG
	0,0,0,0,0	
36_10	0,0,0,0,0	AAGAAAAA
	0,0,0,0,0	
36_11	0,0,0,0,0	ATAAGAAA
	0,0,0,0,0	
36_12	0,0,0,0,1	AAGGGGTA
	0,0,0,0,1	

Hairpin methylation data / Bangladesh-childhood

5_1	0,0,0,0,0	AGTTGTTT
	0,0,0,0,0	
5_2	0,0,1,0,0	GGATGATA
	0,0,1,0,0	
5_3	1,1,1,1,1	GATAGTTG
	1,1,1,1,1	
5_4	1,1,1,1,1	GTTTAGGT
	1,1,1,0,1	
5_5	0,0,0,0,1	TGGGGTGA
	0,0,0,0,0	
15_1	0,0,0,0,0,0,0,0,0	ATATGAGA
	0,0,0,0,0,0,0,0,0	
15_2	0,0,0,0,0,0,0,0,0	ATGGGAGA
	0,0,0,0,0,0,0,0,0	
15_3	0,0,0,0,0,0,0,0,0	GATTGTAA
	0,0,0,0,0,0,0,0,0	
15_4	0,0,0,0,0,0,1,0,1	ATAGGGAA
	0,0,0,0,0,1,1,0,1	
15_5	0,0,0,0,0,0,1,0,0	TTGAAGAT
	0,0,0,0,0,0,1,0,1	
15_6	0,0,0,0,0,0,0,0,0	TGGTTAAA
	0,0,0,0,0,0,0,0,0	
15_7	0,0,0,0,0,0,0,0,0	TAAAGAAA
	0,0,0,0,0,0,0,0,0	
15_8	0,0,0,0,0,0,0,0,0	TATTTTTT
	0,0,0,0,0,0,0,0,0	
15_9	0,0,0,0,0,0,0,0,0	TGTATGTA
	0,0,0,0,0,0,0,0,0	
15_10	0,0,1,0,0,0,0,0,0	AAGTAAAG
	0,0,0,0,0,0,0,0,0	
15_11	1,1,0,0,0,0,1,0,0	AAAATTGA
	1,1,0,0,0,0,1,0,1	
15_12	0,0,0,0,0,0,0,0,0	GTGTGGGG
	0,0,0,0,0,0,0,0,0	
15_13	0,1,0,0,0,0,0,0,0	TAGTTGGA
	0,1,0,0,0,0,0,0,0	
15_14	1,1,1,1,1,0,0,0,1	AATGTAAA
	1,1,1,1,0,0,0,0,1	
15_15	0,0,0,0,0,0,0,0,0	TTGGTAGG
	0,0,0,0,0,0,0,0,0	
15_16	0,0,0,0,1,0,0,0,1	GTATTATG
	0,0,0,0,0,0,0,1,0,1	
30_1	0,0,0,0,0,0,0,0,0	ATAAGGGA
	0,0,0,0,0,0,0,0,0	
30_2	0,0,0,0,0,0,0,0,1	TTTTTAGA
	0,0,0,0,0,0,0,0,0	
30_3	0,0,0,0,0,0,0,0,1	TTGTGATA
	0,0,0,0,0,0,0,0,1	
30_4	0,0,0,0,0,0,0,0,0	TATAATTA
	0,0,0,0,0,0,0,0,0	

30_5	0,0,0,0,1,0,0,0	GAGTAATT
	0,0,0,0,1,0,0,0	
30_6	0,0,0,0,0,0,0,0	TGAGTGAG
	0,0,0,0,0,0,0,0	
30_7	0,0,0,0,1,0,0,0	GGTGAATG
	0,0,0,0,0,0,0,0	
30_8	0,0,0,0,0,0,0,0	TGGAAATA
	0,0,0,0,0,0,0,0	
30_9	0,0,0,0,0,0,0,0	TTTTTGGA
	0,0,0,0,0,0,0,0	
30_10	0,0,0,0,0,0,1,0,1	GGTAAGTA
	0,0,0,0,1,0,1,0,1	
30_11	0,0,0,0,0,0,1,0,1	GATAAAAG
	0,0,0,0,0,0,0,0,1	
30_12	0,0,0,1,1,0,0,0,0	TTAATAGG
	0,1,0,1,0,0,0,1,0	
30_13	0,0,0,0,0,0,0,0,0	AAGGTTTG
	0,0,0,0,0,0,0,0,0	
30_14	0,0,0,0,0,0,0,0,0	AGTTGGAG
	0,0,0,0,0,0,0,0,0	
30_15	0,0,0,0,0,0,0,0,0	TAAAGTAG
	0,0,0,0,0,0,0,0,0	
30_16	0,0,0,0,0,0,1,0,1	TTTTTAGA
	0,0,0,0,0,0,0,0,0	
30_17	0,0,0,0,0,0,0,0,0	TTAATTGA
	0,0,0,0,0,0,0,0,0	
30_18	0,0,0,0,0,0,0,0,0	TGGTTAAT
	0,0,0,0,0,0,0,0,0	
30_19	0,0,0,1,0,0,0,0,0	ATGTTGTG
	0,0,0,0,0,0,0,0,0	
30_20	0,0,0,0,0,0,0,0,0	TGAAAGAT
	0,0,0,0,0,0,0,0,0	
30_21	0,0,0,0,0,0,0,0,0	AGTAAGTG
	0,0,0,0,0,0,0,0,0	
30_22	0,0,0,0,0,0,0,0,0	TTTTTAGA
	0,0,0,0,0,0,0,0,0	
30_23	0,0,0,0,0,0,0,0,1	AAATGGGT
	0,0,0,0,0,0,0,0,0	
31_1	0,1,0,0,0,1,0,0,0	GGATTAGA
	0,1,0,0,0,1,0,0,0	
31_2	0,0,0,0,0,1,1,1,1	TGAGTAAA
	0,0,0,0,0,1,1,1,1	
31_3	0,0,0,1,0,0,1,1,1	TAGAGAGA
	0,0,0,1,0,0,1,1,1	
31_4	0,0,0,0,0,0,0,0,0	TTTATAAT
	0,0,0,0,0,0,0,0,0	
31_5	1,1,0,0,0,0,1,0,0	AGATATAG
	1,1,0,0,0,0,1,0,0	
31_6	0,0,0,0,0,0,0,0,0	AAATAAAA
	0,0,0,0,0,0,0,0,0	
6_1	0,0,0,0,0	ATTGTTGA
	0,0,0,0,0	
6_2	0,0,1,0,0	TTGGGAAT
	0,0,0,0,0	
6_3	0,0,0,0,0	TTAAGATG
	0,0,0,0,0	
6_4	0,1,0,1,1	GTAGGGGG
	0,1,0,1,1	
6_5	0,0,0,0,0	GGTAGGAT
	0,0,0,0,0	
8_1	0,0,1,1,0	TATTGGAG
	0,0,1,1,0	
8_2	0,0,0,0,0	GTTTGTAT
	0,0,0,0,0	
8_3	0,0,0,0,0	GTGTGTGT
	0,0,0,0,0	
8_4	0,0,0,0,0	TAGGGGAT
	0,0,0,0,0	
8_5	0,0,1,0,1	TATTAAAG
	0,0,1,0,1	
8_6	0,0,0,0,0	GTAGGTGG
	0,0,0,1,0	
8_7	1,0,1,1,1	GTAGAGGG
	1,0,1,1,1	
8_8	0,0,0,0,0	AAAATATA
	0,0,0,0,0	
8_9	0,0,0,0,0	GTTTTTGT
	0,0,0,0,0	

8_10	0,0,0,0,0	GAGGTTAT
	0,0,0,0,1	
8_11	0,0,0,0,0	TGTAGGAA
	0,0,0,0,0	
8_12	0,0,1,1,1	AATTTGTT
	0,0,1,1,1	
8_13	0,0,0,0,0	AAGTTTAT
	0,0,0,0,0	
8_14	0,0,0,0,0	TGGATTTG
	0,0,0,0,0	
8_15	0,0,0,0,0	TATTAGGT
	0,0,1,1,1	
8_16	0,0,0,0,0	TATTGGAT
	0,0,0,0,0	
8_17	0,0,0,0,0	AATTGGGT
	0,0,0,0,0	
8_18	0,0,0,0,0	TGTTGTAA
	0,0,0,0,0	
8_19	0,0,1,0,1	AGGATATA
	0,0,1,0,1	
8_20	0,0,1,0,0	AGTTAGGT
	0,0,1,0,0	
8_21	0,0,0,0,0	TTAGTAAG
	0,0,0,0,0	
8_22	0,0,1,1,1	GAATTTGT
	0,0,1,1,1	
8_23	0,0,0,0,0	GAAATGGA
	0,0,0,0,0	
8_24	0,0,0,0,0	GTTGTTAA
	0,0,0,0,0	
8_25	1,0,1,1,1	TTTTGGGT
	0,0,0,0,0	
8_26	0,0,0,0,0	GATAAGGG
	0,0,0,0,0	
8_27	0,0,0,0,0	GGAAGTTG
	0,0,0,0,0	
8_28	0,1,1,1,1	AGTAATGT
	0,1,1,1,1	
8_29	0,0,0,0,0	AAATGGGT
	0,0,0,0,0	
8_30	0,0,0,0,0	AGTTTGGT
	0,0,0,0,0	
8_31	0,0,0,0,0	GAGATTTT
	0,0,0,0,0	