Trends in Genetics

CellPress OPEN ACCESS

Review

Metastable epialleles in humans

Maria Derakhshan,[1] Noah J. Kessler,[2] Garrett Hellenthal,[3,*] and Matt J. Silver[1,4,*]

First identified in isogenic mice, metastable epialleles (MEs) are loci where the extent of DNA methylation (DNAm) is variable between individuals but correlates across tissues derived from different germ layers within a given individual. This property, termed systemic interindividual variation (SIV), is attributed to stochastic methylation establishment before germ layer differentiation. Evidence suggests that some putative human MEs are sensitive to environmental exposures in early development. In this review we introduce key concepts pertaining to human MEs, describe methods used to identify MEs in humans, and review their genomic features. We also highlight studies linking DNAm at putative human MEs to early environmental exposures and postnatal (including disease) phenotypes.

Highlights

Putative human metastable epialleles (MEs) can be identified by screening for systemic interindividual variation suggesting that stochastic DNA methylation is established before germ layer differentiation.

Putative human MEs appear to be influenced but not determined by genetic variation, and are associated with particular classes of retrotransposon, reminiscent of their murine counterparts.

Several putative human MEs have been linked to both environmental exposures in early development and to phenotype. ME methylation has been associated with outcomes relating to cancer, glucose metabolism, and thyroid function in later life.

Putative human MEs are therefore useful for exploring how stochastic and environmental effects in early development can influence disease risk in later life via epigenetic mechanisms.

The ME property of correlation across diverse tissue types allows associations with phenotypes and exposures to be studied in easily accessible tissues.

Murine MEs

The term **ME** (see Glossary) was initially coined to refer to an allele that can exist in multiple epigenetic states independently of the underlying genomic sequence ('epialleles') and that can probabilistically switch between these states ('metastable') [1]. Although this original definition refers to a generic epigenetic state, in practice the term ME is applied specifically to DNAm states. Methylation at MEs can be consistent across most tissues, indicating that DNAm is established in the early embryo.

The identification of MEs in inbred mice typically involves a search for loci exhibiting **SIV** of DNAm. SIV is characterised by DNAm states that correlate across tissues derived from different germ layers (endoderm, mesoderm, and ectoderm) within a given individual ('systemic'), and by DNAm states that vary between individuals. This property is attributed to stochastic DNAm establishment in the early embryo before germ layer differentiation.

In mice, MEs have been identified at intracisternal A particle (IAP) elements, a class of long terminal repeat (LTR) retrotransposon [2]. In the paradigm *Agouti viable yellow* (A^{vy}) and *Axinfused* ($Axin^{Fu}$) models, IAP methylation correlates with the expression of nearby genes, resulting in inbred littermates that show a spectrum of fur colours and degrees of tail kinking, respectively. The distribution of offspring phenotypes can be shifted by supplementing the maternal diet with methyl donors [3–5] and by exposing mothers to endocrine disruptors such as genistein and bisphenol A (BPA) [6–8] or to toxins such as ethanol and phthalates [9,10] during early gestation. Several reviews of murine MEs have been published [1,11,12].

Although MEs appear to be rare in mice [2,11], their ability to drive phenotypic variation in the absence of genetic variation, including susceptibility to metabolic diseases in later life [6,12], has motivated efforts to identify similar methylation states in humans. We provide an overview of the key studies that have investigated **putative human MEs**. Specifically, we describe the methods used to identify MEs in humans, and review evidence on their genomic context, sensitivity to the early developmental environment, and associations with disease. The qualified term 'putative human ME' is used to highlight the difficulty of discounting the influence of genetic variation in studies of genetically heterogeneous human subjects.

[1]London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[2]Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK
[3]UCL Genetics Institute, University College London, London WC1E 6BT, UK
[4]Medical Research Council (MRC) Unit The Gambia at the London School of Hygiene and Tropical Medicine, Fajara, Banjul, The Gambia

*Correspondence:
g.hellenthal@ucl.ac.uk (G. Hellenthal) and
matt.silver@lshtm.ac.uk (M.J. Silver).

Identifying putative human MEs

Systemic interindividual variation

The identification of putative human MEs requires DNAm data from tissues, derived from different germ layers from multiple individuals, in order to identify CpGs that show (i) high interindividual variation of DNAm, and (ii) low variation across tissues within each individual (Figure 1). Such loci exhibit SIV, in contrast to the majority of CpGs which show low methylation variability between individuals and/or high inter-tissue variation as a result of the celltype- and tissue-specificity of DNAm [13,14]. The term 'systemic' refers to the fact that DNAm states correlate across a diverse range of cell types and tissues soma-wide. Supporting this, inter-tissue correlations at SIV loci have been confirmed in additional tissues beyond those used in SIV screens [15,16]. However, some SIV loci do show celltype- and lineage-specific effects [17–20], a phenomenon that has also been observed at the *Axin*^{Fu} murine ME where DNAm in tail differs from that in other tissues [5]. It is important to note that,although the observation of SIV in inbred mice is sufficient to identify a locus as being an ME, this is not the case in humans. Because of the genetic variation that is present in any group of humans [excluding monozygotic (MZ) twin pairs], the variation in DNAm observed at SIV loci may be due to differences in genotype rather than being genuine epialleles.

To date, six independent SIV screens have been carried out, each varying in CpG coverage, sample size, number of tissues, and methylation profiling platform (summarised in Table 1). SIV loci have been identified at both the single-CpG [20,21] and regional levels [22–24], the latter being important because regions of contiguous correlated DNAm might be of greater functional relevance [25] (Table 1). Although most SIV loci have been identified in adult samples, several loci have been validated in foetal tissues [15], reinforcing the notion that their methylation states are established in early life (Figure 1).

There is strikingly little overlap between SIV loci identified in independent studies to date (Figure 2). This may be attributable in part to the different criteria used to define SIV loci (Table 1). Another factor is the small sample size of each screen, as the average power to detect SIV-CpGs in a multi-tissue dataset with four individuals has been estimated to be <60% [26]. For example, the three-tissue SIV screen with the highest sample size to date of ten individuals [24] did not detect sufficient interindividual variation at the well-established SIV locus *VTRNA2-1* [22]. This locus has

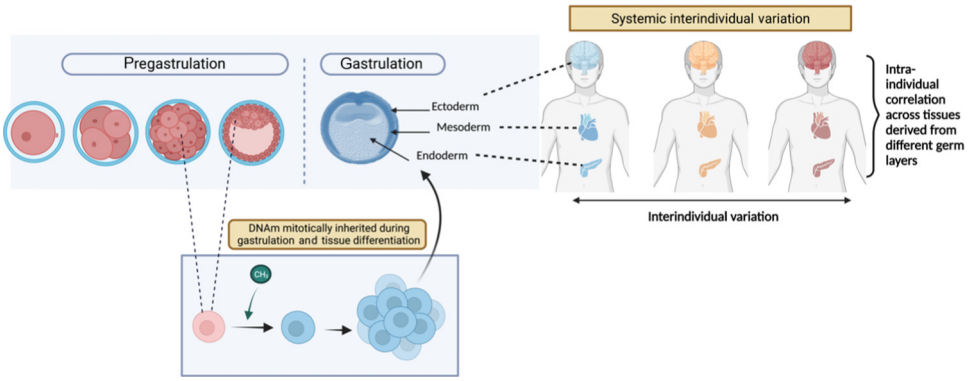

Figure 1. Model linking systemic interindividual variation (SIV) to the establishment of DNA methylation (DNAm) in the pregastrulation embryo. SIV loci show methylation differences between individuals but correlate across tissues derived from different germ layers of the same individual. Tissues are coloured by the methylation status of the SIV-CpG. This property suggests that variable DNAm states at SIV loci are established before germlayer differentiation (gastrulation), and are thereafter mitotically inherited through gastrulation and tissue differentiation. Figure generated with BioRender.com.

Glossary

Correlated regions of systemic interindividual variation (CoRSIVs): genomic regions comprising multiple CpGs that show correlated patterns of DNA methylation (DNAm) indicative of SIV.

Developmental origins of health and disease (DOHaD): the hypothesis first proposed by David Barker [106] and colleagues that environmental exposures experienced in early development can influence disease risk in later life.

Metastable epiallele (ME): a SIV locus identified in inbred mice that is indicative of stochastic methylation establishment in the early embryo independently of genetic variation.

Parent-of-origin-specific methylation (PofOm): the methylation state of an allele is dependent on the parent it was inherited from.

Putative human metastable epialleles: SIV loci identified in humans for which there is evidence that methylation variability is not determined by genotype alone.

Systemic interindividual variation (SIV): loci with this property show DNAm variation between individuals but correlation across tissues derived from different germ layers of the same individual.

Table 1. Summary of screens to detect systemic interindividual variation (SIV) in humans[a]

SIV screen	Tissues	Number of samples	Ethnicity	Platform	Regional/CpG-level	Method	Number of SIV loci identified
Waterland *et al.* [29]	PBLs (M), hair follicles (Ec)	8	Caucasian	MSAM	Regional	Parallel two-tissue cohybridisation screen in which genomic DNA is digested with a methylation-sensitive restriction endonuclease and amplified by PCR. SIV regions were defined as genomic intervals with concordant interindividual differences in both tissues	13
Harris *et al.* [20]	PBLs (M), colonic mucosa (En)	10	Not reported	Illumina 450K	CpG	SIV definition:(i) One person has the highest methylation value in each tissue and another individual has the lowest value in each tissue (ii) The average interindividual methylation range (across tissues) is ≥0.1 and ≤0.6 (iii) The absolute Pearson correlation between tissues is ≥0.63	1776
Silver *et al.* [22][b]	PBLs (M), hair follicles (Ec)	2	Caucasian	WGBS	Regional	200 bp genomic bins containing at least six CpGs with a SIV 'index' (SIVI) of ≥20; the SIVI was designed to maximise interindividual variation that is consistent across both tissues ([22] for further details)	109
van Baak *et al.* [21]	Gall bladder (En), abdominal aorta (M), sciatic nerve (Ec)	4	Not reported	Illumina 450K	CpG	SIV definition:(i) Average interindividual variation at least threefold greater than average inter-tissue variation(ii) Interindividual methylation range ≥0.2	1042
Kessler *et al.* [23][b]	Small bowel (En), PBLs (M), hair follicles (Ec)	5	Four Caucasian, one Caucasian/ African American	WGBS	Regional	CpGs were first tested for SIV and then clustered into regions defined as having at least four SIV-CpGs and at least twice as many SIV-CpGs as non-SIV-CpGs SIV definition:(i) Interindividual variation that is at least threefold greater than inter-tissue variation(ii) Interindividual methylation range ≥0.15	687
Gunasekara *et al.* [24]	Thyroid (En), heart (M), brain (Ec)	10	Caucasian	WGBS	Regional	100 bp genomic bins first grouped into correlated regions and then tested for SIV to generate a set of **CoRSIVs** SIV definition:(i) Inter-tissue correlation ≥0.71(ii) Interindividual methylation range ≥0.2	9926

[a]Abbreviations: CoRSIVs, correlated regions of systemic interindividual variation; Ec, ectoderm; En, endoderm; M, mesoderm; MSAM, methylation-specific amplification microarray; WGBS, whole-genome bisulphite sequencing.
[b]The samples used in the screens by Silver *et al.* [22] and Kessler *et al.* [23] partially overlap.

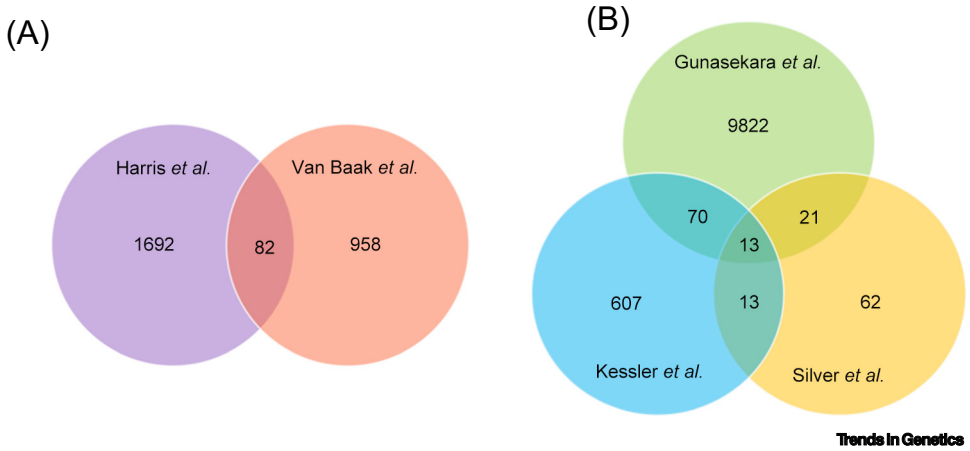

Figure 2. Overlap between systemic interindividual variation (SIV) loci identified in independent studies. (A) Venn diagram showing the overlap between SIV-CpG identified in two separate Illumina450K screens (van Baak *et al.* [21] and Harris *et al.* [20]). (B) Venn diagram showing the overlap between SIV regions identified in three whole-genome bisulphite sequencing (WGBS) screens (Silver *et al.* [22], Kessler *et al.* [23], and Gunasekara *et al.* [24]). Note that the total number of regions in each screen may not exactly match those given in Table 1 owing to lifting over of genome assemblies.

a bimodal methylation distribution, with ~25% of individuals (hypomethylated), in contrast to the more common 50% methylation pattern. We note that the large interindividual differences seen at this and similar loci make them strong candidates for SIV, but their population distributions may make them difficult to detect in small samples. Interestingly, *VTRNA2-1* and >75% of SIV-CpGs reported in multiple screens from Table 1 exhibit high interindividual variation in several larger (N >20) datasets that span different tissue types and ethnicities (but with only one tissue sampled per dataset) [26]. This suggests that investigation of datasets with larger sample sizes may result in more consistent SIV sets. Furthermore, it suggests that some SIV loci exhibit high methylation variability in diverse ethnicities beyond the largely Caucasian datasets that have been explored.

Influence of genotype

Genetic variation is a key driver of interindividual DNAm differences [27,28]. This raises the possibility that SIV measured in genetically heterogeneous human populations is driven by cross-tissue genetic effects instead of by stochastic (alternatively described as random or probabilistic) and/or environmental effects in the preimplantation embryo. Multiple approaches have been used to assess the influence of genetics at SIV loci, including analysis of MZ twins, examination of methylation quantitative trait locus/loci (mQTL) effects, and filtering of SNP-discordant ME regions.

MZ twin discordance and epigenetic 'supersimilarity'

MZ twin datasets can be leveraged to determine whether SIV loci show discordance between MZ twins, indicative of methylation variation that is not linked to genetic variation, thereby positioning these SIV loci as candidate MEs [20,29,30]. Of the SIV-CpGs reported by Harris *et al.*, 14% showed significant methylation differences (>10%) in peripheral blood leukocyte (PBL) samples between MZ twins [20]. These SIV loci were found to be tenfold enriched in CpGs whose differential methylation between MZ twins was as high as that between unrelated individuals in another study [31]. This effect was observed in both whole blood and adipose tissue, and these CpGs also showed high variability in 2-year-olds and no evidence of epigenetic drift, in other words no increase in methylation variance with age. This suggests that stochastic methylation at these loci arises in early life and is maintained through childhood [31].

Although discordance in inbred mice may reflect DNAm differences that were established at any time during development, an advantage of exploring MZ discordance is that it narrows down the time frame of likely DNAm establishment to after MZ twinning events in the early embryo. By contrast, van Baak *et al.* focussed on loci that show excessive concordance in methylation between MZ twins, which may be attributable to establishment before MZ twin cleavage in the very early embryo [21]. Focussing on the top 10% most variable CpGs in adipose tissue samples, van Baak *et al.* defined 'epigenetic supersimilarity' (ESS) CpGs [21] as loci that show methylation concordance between MZ twins that is more than twice that seen between dizygotic (DZ) cotwins. Given that MZ twins share approximately double the amount of genetic sequence compared to DZ twins, methylation similarity between MZ twins at ESS CpGs is higher than would be expected from their shared genetic identity [21].

Overall, ~40% of the SIV loci identified by van Baak *et al.* overlapped with an ESS CpG [21], suggesting not only that genetics plays a limited role at these SIV loci but also that these methylation states may be established in the cleavage-stage embryo. The timing of methylation establishment at these loci could be further refined using DNAm data from MZ twins that split at earlier and later stages in development, for example, dichorionic MZ twins and monochorionic diamniotic MZ twins [32].

Methylation quantitative trait loci

An alternative approach to investigating genetic influence at SIV-CpGs is to screen for mQTL. SIV-CpGs are reported to be enriched for mQTL effects [21–24], in agreement with findings that CpGs that covary between blood and brain (indicative of SIV) are also enriched for mQTL [33,34]. However, there is some evidence that *cis* mQTL explain only a moderate proportion (<25% in [23]) of methylation variance at SIV-CpGs [21,23], as well as at tissue- and ethnicity-independent 'hypervariable' CpGs that show evidence of establishment in the early embryo [26]. An exception is the SIV regions identified by Gunasekara *et al.*(Table 1) which were associated with large *cis* mQTL effects that explained a median of 76% methylation variance [24]. It is unclear whether their approach, which first selects correlated genomic regions before testing for SIV (Table 1), enriches for *cis* genetic effects compared to other methods for detecting SIV loci that do not focus on regions with multiple highly correlated CpGs.

Definition of a 'putative human ME'

Together, these findings suggest that,although some reported SIV loci may be under strong genetic control, others appear not to be. This has led to a proposal that the definition of an ME in genetically heterogeneous human populations should be extended to include SIV loci at which methylation is 'influenced but not determined' by genotype [23]. Apart from studies using MZ twin data, comprehensive examination of long-range and multi-locus genetic effects on methylation variability at SIV loci is challenging because of factors such as the limited SNP coverage of genotype arrays and a lack of power to reliably estimate SNP effect sizes, including *trans* associations. The relative influences of genetic, environmental, and stochastic effects at SIV loci therefore remain unclear, and quantification of the genetic contribution will require large-scale mQTL studies. We use the term 'putative human ME' to distinguish these loci from murine MEs where the influence of genotype can be ruled out. In this review we consider that SIV loci reported in any of the studies listed in Table 1 are candidates for putative human MEs.

Genomic context

Characterising genomic features that distinguish putative human MEs from other loci may give insights into the molecular mechanisms that underpin the establishment of variable DNAm

states in early development. Notable features that have been associated with putative human MEs include proximal transposable elements (TEs) and regions of **parent-of-origin-specific methylation (PofOm)**.

Transposable elements

TEs are mobile genetic elements that comprise ~50% of mammalian genomes [35]. They can be categorised into DNA transposons which transpose through a cut-and-paste mechanism, and retrotransposons which copy via an RNA intermediate [36]. When activated, TEs can harm the host though several mechanisms, including genome damage via transposition [37–39] and activation of aberrant expression of nearby genes [39–41]. TEs are epigenetically silenced in early development to prevent their activation during the global wave of demethylation that occurs shortly after fertilisation [40,42–46]. Krüppel associated box (KRAB) zinc-finger proteins (KZFPs) silence TEs in embryonic stem cells by recruiting the cofactor KAP1 (KRAB-associated protein) which acts as a scaffold for proteins that deposit DNAm and histone H3 lysine 9 trimethylation (H3K9me3) marks [47,48].

Murine MEs have been associated with IAPs [2], an evolutionarily young class of rodent-specific endogenous retroviruses (ERVs) – retrotransposons that contain LTRs. The evidence suggests that stochastic methylation establishment at murine MEs may result from competition between KZFP/KAP1 binding that promotes DNAm and CTCF (CCCTC-binding factor) and ZF (zinc finger) CxxC domain-containing proteins that prevent DNAm [49,50].

Although IAPs are not present in the human genome, putative human MEs are enriched for proximity to ERVs [23,49]. Putative human MEs are also enriched for proximity to non-LTR-containing long interspersed nuclear elements (LINEs) [22] and have been associated with primate-specific *Alu*-type short interspersed nuclear elements (SINEs) [51,52]. Several studies have also reported an enrichment of putative MEs in subtelomeric regions [23,26] which are known to contain many repeat sequences and TEs.

Putative human MEs are enriched for proximal KAP1 and CTCF binding sites [23], suggesting that mechanisms driving methylation variability at these loci may in some cases be similar to those at murine MEs. The association with CTCF is also noteworthy because it has been linked to stochastic switching between epigenetic states in humans [53].

Parent-of-origin specific methylation

Putative human MEs are also enriched in regions with evidence of PofOm, at which methylation differs between the maternal and the paternally inherited alleles [54,55]. PofOm can arise from differential DNAm between gametes that is maintained during epigenetic reprogramming events in the early embryo. In accord, putative human MEs are enriched for regions that are hypermethylated in the oocyte relative to sperm (termed oocyte germline differentially methylated regions, DMRs) [54]. Several putative MEs also overlap with imprinted genes [21,54], at which PofOm regulates monoallelic expression of the maternal or paternal allele in the offspring. Methylation variability at imprinted genes is noteworthy because of the established role of imprinted genes in foetal growth and development [56,57].

These associations raise the possibility that DNAm variation at some putative human MEs is caused by incomplete maintenance of PofOm in the early embryo, whereby stochastic and/or environmental effects could cause a gain of methylation at the normally unmethylated allele and/or loss of methylation at the methylated allele. Methylation variability could also arise due to the variable establishment of gamete-specific DNAm states that evade reprogramming at

fertilisation [58]. The environmental lability of regions of PofOm is further supported by growing evidence that the early developmental nutritional environment can influence DNAm at imprinted genes [59–63].

The link between putative human MEs and PofOm is strengthened by their enrichment for proximity to binding sites for ZFP57 [23], a KZFP that maintains PofOm in early embryogenesis [64]. Individuals homozygous for a *ZFP57* mutation that is associated with dysregulation of imprinting maintenance showed an altered methylation status at putative MEs compared to wild-type individuals [23], indicating that ZFP57 may directly influence putative ME methylation. Interestingly, *ZFP57* promoter methylation is associated with season of conception (SoC) in Gambian 2-year-olds [22] and with periconceptional and gestational maternal plasma folate during the late stages of pregnancy [65–67], suggesting that *ZFP57* DNAm differences associated with the early nutritional environment may be linked to DNAm changes at ZFP57 targets [54]. In addition, methylation variance at *ZFP57* has been linked to a proximal genetic variant [21], and allele-specific methylation at this locus has been attributed to an upstream polymorphic CTCF binding site [21,68], suggesting that ZPF57 may mediate both genetic and environmental effects on DNAm at regions of PofOm and at some putative human MEs [21].

Influence of the environment

Examination of the potential influence of the early developmental environment on putative human MEs has involved observational studies of cohorts subjected to naturally occurring exposures [21,23,29,30,54,69], retrospective studies of cohorts subject to gestational exposures such as famine [70,71] or smoking [15], randomised supplementation trials [72,73], and *in vitro* approaches [70,74,75]. Although environmental lability is not a prerequisite for metastability [76], evidence of sensitivity to the early environment supports the notion that DNAm variability at putative human MEs is established in early development and is not explained by genetic variation alone.

Periconceptional exposures

DNAm states established in the periconceptional period may show heightened sensitivity to environmental perturbations through the extensive reprogramming of the epigenome and rapid rates of DNA synthesis that occur after fertilisation [77]. A natural experiment in rural Gambia provides an interesting setting to investigate the effects of periconceptional environment on methylation at putative human MEs. In this region of Gambia, a community of subsistence farmers experiences differences in environmental exposures, including diet and infection levels, between the rainy ('hungry') and dry ('harvest') seasons [78]. Methylation levels at putative MEs are higher in individuals conceived during the rainy season versus the dry season, a finding that has been robustly replicated across several studies and cohorts [22,29,30,69]. This effect is observed in both blood and hair follicle tissues [22,30], suggesting that perturbed DNAm states in the early embryo can be maintained in different cell types.

Sensitivity to the periconceptional environment is further supported by a study that investigated the influence of reproductive technologies (ART) on ME methylation [74]. Estill *et al.* [74] examined DNAm states at 22 putative MEs covered by the Illumina 450K array in bloodspots of newborns conceived naturally and through different ART methods, including intrauterine insemination (IUI) and intracytoplasmic sperm injection (ICSI) with fresh or frozen embryos. The majority of putative MEs (19/22) were differentially methylated between individuals conceived in at least one conception-type comparison [74]. However, the association with ART was not replicated in recent epigenome-wide association studies (EWAS) with large sample sizes [79,80].

Later gestation and postnatal exposures

Methylation states at putative human MEs have also been studied in individuals exposed to maternal gestational famine lasting at least 7 months in rural Bangladesh [71]. Although this targeted study was relatively small, threeof 16 putative MEs (*PAX8*, *PRDM-9*, and *ZFP57*), that were previously associated with SoC in Gambians, were also associated with exposure to gestational famine. Other studies have found associations between several putative ME loci and maternal gestational diabetes (*CYP2E1*) [81], maternal tobacco smoke exposure (*PAX8*) [15], and prenatal exposure to alcohol (*POMC*) [82], antidepressants (*CYP2E1*) [83], or perfluoroalkyl substances (*ZFP57*) [84]. Although the exact timing of the exposure effect on DNAm is not pinpointed in these studies, there is evidence that MEs are sensitive to maternal folic acid supplementation at a gestational age of 16 weeks (mid-gestation) (*VTRNA2-1*, *PAX8*) [85], suggesting that environmental sensitivity at MEs may extend beyond the periconceptional period.

Several putative MEs have also been associated with postnatal environmental exposures including smoking (*POMC*, *VTRNA2-1*) [86,87], psychosocial deprivation in childhood (*CYP2E1*) [88], and exposure to pesticides (*VTRNA2-1*) [89], air pollution (*VTRNA2-1*) [90], ortoxic aromatics (*CYPE21*) [91,92]. Because these studies were carried out in a single tissue type (largely blood), the extent to which these effects are tissue-specific is unknown.

Parental and germline exposures

Exposures before conception may also influence methylation patterns in offspring if environmentally sensitive DNAm states established in gametes are able to evade the global epigenetic reprogramming that occurs after fertilisation [93,94]. As an example, the putative ME *VTRNA2-1* is an imprinted gene that is usually hypomethylated on the paternal allele and variably methylated on the maternal allele [19,58]. These methylation states show evidence of establishment in the oocyte, and *VTRNA2-1* methylation on the maternal allele in offspring has been associated with maternal age and alcohol consumption during the oocyte maturation period [58], suggesting that environmentally labile methylation states in the oocyte may be maintained in the zygote.

A further example is provided by a putative ME at the *POMC* (proopiomelanocortin) gene where DNAm states measured in offspring correlated with those in paternal somatic cells [51]. However, putative human MEs including *POMC* show marked hypomethylation in sperm [23,95], suggesting that paternal transmission of this epigenetic state does not occur via DNAm. This raises the possibility that other factors such as sperm RNAs [96] and histone modifications or chromatin architecture [97] may mediate any paternal inheritance of DNAm status at *POMC.*

Nutritional exposures and 1-carbon metabolism

ME methylation has been associated with nutritional biomarkers involved in 1-carbon (1C) metabolism. 1C metabolic pathways use diet-derived methyl donors such as folate, choline, and betaine, and cofactors such as vitamins B2, B6, and B12, to synthesise *S*-adenosyl methionine (SAM). DNA methyltransferases (DNMTs) transfer the methyl group from SAM to cytosine to establish and maintain DNAm [98,99].

Blood plasma concentrations of key 1C metabolism biomarkers measured in pregnant Gambian women have been back-extrapolated to the time of conception to give an estimate of periconceptional 1C metabolism biomarker levels [30]. The concentrations of several 1C biomarkers were found to be higher in the rainy season, and were predictive of DNAm at several putative MEs [30,51,69,100]. Methylation status at the putative ME *PAX8* is altered in offspring born to mothers who were periconceptionally supplemented with micronutrients that provide methyl groups and cofactors involved in 1C pathways [72], and *PAX8* methylation has also been

associated with maternal periconceptional levels of amino acids linked to 1C metabolism [18,101,102]. These findings suggest that 1C metabolites may contribute to Gambian SoC associations at putative human MEs. 1C metabolism may also underpin the association between tobacco smoke exposure and ME methylation [15] because tobacco exposure has been associated with reduced 1C metabolism micronutrients (folate, B12, and B6) [15,103].

The link to 1C metabolism aligns with similar findings in A^{vy} mouse models in which offspring born to dams supplemented with B12, folate, and choline during pregnancy show altered methylation profiles at the *Agouti* IAP [104,105]. However, the 1C metabolic pathway is complex, making elucidation of the influence of specific 1C metabolites on DNAm at putative MEs in human populations challenging. These effects are also likely to be context-specific, as evidenced by the finding from a Gambian study of an interaction between SoC and maternal 1C biomarkers that influence offspring ME methylation [69].

We note that several studies reporting associations with environmental exposures employed small sample sizes, particularly where genome-wide DNAm data were analysed, reinforcing the need for replication in independent studies.

MEs and the developmental origins of health and disease

The **developmental origins of health and disease (DOHaD)** hypothesis posits that environmental exposures in early life can impact on health outcomes in later life [106]. Barker and colleagues formulated this hypothesis based on observations that low birth weight is associated with adverse health outcomes in adults, such as high blood pressure and cardiovascular disease [106–109]. The evidence suggests that epigenetic mechanisms may underpin some DOHaD observations [105,110]. For example, maternal and foetal exposures during gestation, such as to dietary folate [60,66,85,111], famine [59,79,112–115], alcohol [116–118], aflatoxins [119], and smoking [120–123], have been associated with offspring DNAm and, in several cases, DNAm changes induced by such exposures were associated with offspring phenotype [112,114,117].

Although environmental lability it not a defining feature of MEs, putative human MEs are particularly useful for exploring DOHaD effects because the timing of the establishment of DNAm states is presumed to be limited to the pregastrulation embryo. Furthermore, DNAm states can be easily studied in accessible tissues (e.g., blood, saliva). This is because DNAm states are consistent across tissues within an individual, allowing population-scale sampling of DNAm at human MEs.

ME methylation and disease

Prospective studies in longitudinal cohorts have indicated that methylation at putative human MEs in accessible tissues such as blood predicts health outcomes including cancer (*VTRNA2-1*, *SPATC1L*, *DUSP22*, *ZFP57*, *ORL13*, *HCG4B*, and *PF4*) [21], body mass index (BMI; *VTRNA2-1*) [124], bodyweight (*MEST*) [75], and thyroid function (*PAX8*) [18] in later life. Several well-characterised putative human MEs show evidence of stability with age in longitudinal cohorts, supporting their potential utility in disease prognosis. For example, DNAm at *VTRNA2-1* is stable over a 10 year period from childhood to early adolescence [22] and over 25 years in adulthood [125], *POMC* DNAm is stable from birth to at least 12 years of age [51,126], and *PAX8* DNAm is stable from ages 7 to 17 years [18]. These three putative human MEs are associated with multiple health-related phenotypes, despite being under-represented on the commonly used Illumina methylation arrays (Table 2).

DNAm at the putative MEs *CYP2E1* and *DUSP22* (dual specificity phosphatase 22) measured in four immune cell blood types is associated with active and erosive rheumatoid arthritis,

Table 2. Examples of putative human MEs at which DNA methylation variation has been associated with both phenotypes and early developmental exposures.[a]

Putative human ME	Phenotype/exposure		Tissue	*N*[b]	Refs
VTRNA2-1 *(alias nc886)*	Phenotype	Childhood BMI	Whole blood	442	[134]
		Glucose and insulin levels in adolescence, adiposity in childhood	Whole blood	1654	[125]
		Lung cancer, mature B cell neoplasm	Whole blood	351[c], 435[c]	[21]
		Preterm birth	Whole blood	92	[148]
		HCC and tumour aggressiveness	HCC tumour and normal tissues	92[c]	[149]
		Breast cancer	Whole blood	1738	[150]
	Exposure	Gambian season-of-conception	Whole blood, hair follicle	120	[22]
		Gestational folate supplementation	Saliva	111	[85]
		Maternal age and maternal socioeconomic status	Whole blood	1646[c], 1594[c]	[125]
		Maternal age and preconceptual alcohol use	Whole blood	1132[c]	[58]
POMC	Phenotype	BMI and obesity risk	MSH-positive neurons, whole blood	82[c], 228[c]	[51]
		Adolescent depressive disorder	Whole blood	30[c]	[151]
	Exposure	Gambian season of conception and maternal 1C metabolites	MSH-positive neurons, whole blood	144[c]	[51]
		Prenatal alcohol exposure	Saliva	107 males[c], 87 females[c]	[82]
PAX8	Phenotype	Child thyroid volume and thyroid hormone levels	Whole blood	118[c]	[18]
		Sleep duration	Saliva	28[c]	[152]
		Sperm concentration, morphology, motility	Sperm	65[c]	[153]
	Exposure	Gambian season of conception	Whole blood	50[c], 120	[22,29]
		Maternal periconceptional amino acid levels		303[c]	[18]
		Gestational famine	Whole blood	61[c]	[71]
		Gestational folate supplementation	Saliva	111	[85]
		Maternal smoking	Cord blood	236[c]	[15]
CYP2E1	Phenotype	Autism	Placenta	41	[142]
		Active rheumatoid arthritis	$CD14^+$ monocytes, $CD4^+$ naïve T cells	58	[127]
		Parkinson's disease	Brain	24	[154]

Table 2. (continued)

Putative human ME	Phenotype/exposure		Tissue	*N*[b]	Refs
	Exposure	Maternal gestational diabetes	Cord blood	3677	[81]
		Psychosocial deprivation in early childhood	Buccal cells	65	[88]
		Prenatal antidepressant exposure	Cord blood white blood cells	43	[83]
DUSP22	Phenotype	Schizophrenia	Whole blood, PFC	214	[70]
		Erosive rheumatoid arthritis	$CD14^+$ monocytes, $CD19^+$ B cells, $CD4^+$ naïve T cells, $CD4^+$ memory T cells	58, 57, 56, 58	[127]
		Alzheimer's disease	Hippocampus	20	[155]
		Mature B cell neoplasm, urothelial cell carcinoma	Whole blood	862[c]	[21]
	Exposure	Gestational famine	Whole blood, PFC	79	[70]
		ART	NA	137[c]	[74]
SPATC1L	Phenotype	Colorectal cancer, prostate cancer	Whole blood	834[c], 863[c]	[21]
		BMI *z*-score across childhood	Placenta	426[c]	[140]
		Post-traumatic stress disorder	Blood	554	[156]
	Exposure	ART	NA	137[c]	[74]
MEST	Phenotype	BMI at age 1 year, and longitudinal weight gain	Cord blood	408	[75]
		Adult obesity	Blood	74[c]	[157]
		BMI *z*-score across childhood	Placenta	426[c]	[140]
		Birthweight	Placenta	211[c]	[141]
	Exposure	Bisphenol-A	Cord blood	408	[141]
		Maternal gestational diabetes	Cord blood	211[c]	[157]
		Periconceptional micronutrient supplementation	Cord blood	58[c]	[73]
PLAGL1	Phenotype	Estimated foetal weight and weight at age 1 year	Cord blood	254[c]	[158]
		Sperm concentration morphology motility	Sperm	65[c]	[153]
	Exposure	Maternal alcohol, vitamin B2, vitamin B12	Cord blood	254[c]	[158]
		Maternal erythrocyte folate levels	Cord blood	438[c]	[61]
ZFP57	Phenotype	Colorectal cancer	Whole blood	834[c]	[21]
	Exposure	Gambian season of conception	Whole blood	120	[22]
		Gestational famine	Whole blood	61[c]	[71]

(continued on next page)

Table 2. (continued)

Putative human ME	Phenotype/exposure		Tissue	*N*[b]	Refs
		Gestational folate supplementation	$CD4^+$ and antigen-presenting cells from cord blood	23	[66]
		Prenatal perfluoroalkyl substance exposure	Cord blood	380[c]	[84]

[a]Abbreviations:1C, 1-carbon; ART, artificial reproductive technology; BMI, body mass index; *CYP2E1*, cytochrome P450 2E1; *DUSP22*, dual specificity protein phosphatase 22; HCC, hepatocellular carcinoma; *MEST,* mesoderm-specific transcript; MSH, melanocyte-stimulating hormone, NA, not available; *PAX8*, paired box8; PFC, prefrontal cortex; *PLAGL1*, pleomorphic adenoma gene-like 1; *POMC*, proopiomelanocortin; *SPATC1L,* spermatogenesis and centriole associated 1-like; *VTRNA2-1*, *Vault RNA2-1*; ZFP57, zinc finger protein 57.
[b]*N* refers to the total number of individuals (some values are combined across sample categories; e.g., for case–control studies).
[c]Indicates studies that used a candidate gene approach; the other studies employed epigenome-wide association studies (EWAS).

respectively [127]. Furthermore, *DUSP22* methylation measured in whole blood and brain tissue is associated with schizophrenia risk [70], and *VTRNA2-1* methylation was the most significant reported DMR associated with orofacial clefts [128]. However, despite the potential early establishment of ME methylation states, it should be noted that cross-sectional EWAS are prone to confounding and there is the potential for reverse causation effects, making it difficult to infer causal relationships [129]. We note that some EWAS assume that DNAm has a continuous distribution, leading to the possibility of false findings in cases where methylation values are multimodal [130].

MEs as mediators of environmental influences on phenotype

Methylation at several MEs has been associated with gene expression in various cell/tissue types [21,52,75,123,125], and methylation at many putative MEs has been associated with both environmental exposures and phenotypes, using either a candidate gene or EWAS approach (see Table 2 for examples). Although many of these associations are from studies with relatively small sample sizes that require replication, some patterns are beginning to emerge. In particular, joint analysis of exposure, methylation, expression, and phenotype data is beginning to elucidate potential causal pathways through which putative human MEs could mediate the effects of environmental exposures on phenotypes.

For example, hypomethylation at a putative ME in the *MEST* gene measured in cord blood is associated with gestational exposure to BPA, a chemical agent found in polycarbonate plastics, and with BMI at ages 1 and 6 years among the same individuals [75]. Notably, *MEST* expression is upregulated upon BPA exposure of human mesenchymal stem cells *in vitro,* resulting in increased adipogenesis [75]. These findings indicate that hypomethylation of *MEST* in response to prenatal BPA exposure may increase obesity risk in later life by increasing *MEST* expression.

A second example is POMC*,* a key component of the satiety-regulating melanocortin signalling pathway [131]. Hypermethylation of a variably methylated region at the *POMC* intron2/exon3 boundary in blood is associated with conception in the rainy ('hungry') season in Gambian children and with increased obesity risk and BMI in European children and adults [51,132]. *POMC* methylation is associated with BMI in melanocyte-stimulating hormone (MSH)-positive neurons from the hypothalamus that are involved in regulating the satiety response [51]. Taken together with evidence that hypermethylation of the intron2/exon3 boundary is associated with decreased *POMC* expression [132], these findings suggest that hypermethylation of *POMC* in early development may increase obesity risk by decreasing *POMC* gene expression in MSH-positive neurons [51].

PAX8, a transcription factor that regulates thyroid cell differentiation and function, provides another notable example [133]. *PAX8* hypermethylation is observed in Gambian infants and 2-year-old children conceived during the rainy season [22,29], and predicts decreased thyroid volume and decreased levels of the thyroid hormone free thyroxine (T4) in mid-childhood [18]. Lower free T4 levels were in turn associated with increased adiposity and bone mineral density in the same cohort. This raises the possibility that *PAX8* hypermethylation in early development may increase the propensity to develop metabolic disorders in later life via downregulation of thyroid hormones, although a direct causal relationship could not be established in this study [18].

Finally, at *VTRNA2-1*, methylation is lower in Gambian individuals conceived during the dry relative to the rainy season [22]. Lower methylation at this locus has also been associated with increased BMI and adiposity in childhood [134], as well as with increased expression of *VTRNA2-1* RNA in European cohorts [125]. These RNAs can in turn predict glucose levels in childhood and adulthood, thus providing a plausible causal pathway through which *VTRNA2-1* methylation might influence metabolism in later life [125]. *VTRNA2-1* is also a putative tumour suppressor that inhibits protein kinase RNA-activated (PKR) [123], and increased methylation of this small non-coding RNA gene is associated with decreased survival in patients with leukaemia [135], lung [136], or oesophageal [137] cancers.

Taken together, these studies position putative human MEs as interesting candidates for exploring epigenetic mechanisms linking the early developmental environment to disease risk in later life [113]. It is possible that some MEs may exert their effects on gene expression in early development (rather than in postnatal tissues) and influence later disease risk by altering developmental trajectories. The study of ME methylation in the placenta may offer particular insights in this regard because of its role in regulating foetal growth and development [138,139]. A study analysing the association between placental DNAm and BMI identified several MEs annotated to growth and metabolism genes [140]. Placental methylation status of MEs has also been linked to offspring weight and autism spectrum disorder [141,142].

Developmental plasticity as an adaptive mechanism

The association between several environmentally sensitive MEs and low birth weight (Table 2) aligns with the thrifty phenotype hypothesis first proposed by Hales and Barker [143]. This postulates that reduced foetal growth and altered metabolism can be a survival strategy for offspring exposed to high levels of nutritional deprivation in early life by focussing on the growth of key organs instead of on tissues such as pancreas and muscle [143,144]. However, when offspring transition from an energy-poor to an energy-rich environment, these adaptations (e.g., reduced insulin secretion) will be maladaptive, rendering individuals at increased risk of developing metabolic disorders [143].

By contrast, the predictive adaptive response (PAR) hypothesis suggests that developmental plasticity *in utero* primes the developing embryo to its future environment, without necessarily having an immediate effect on phenotype (e.g., reduced foetal growth) in early life [145,146]. In cases where the postnatal environment is different from the anticipated environment, this 'prediction' will be incorrect and result in increased disease risk.

The association between the putative MEs, *POMC* and *PAX8*, and metabolic outcomes in later life may reflect PAR, whereby offspring born to mothers with compromised nutritional status would be epigenetically programmed to conserve energy and store fat in later life despite showing no clear phenotypes at birth.

It has been suggested that these adaptive mechanisms may be evolutionarily maintained; in other words, that epigenetic variability at environmentally responsive loci is under genetic control. This

would provide a heritable mechanism to allow phenotypic plasticity in response to rapidly changing local environments [147]. The putative MEs *ZFP57* and *PAX8* offer tentative evidence of this phenomenon, in that they appear to be environmentally responsive, with DNAm variance associated with a specific genetic variant [18,21].

Concluding remarks and future perspectives

Putative human MEs exhibit interindividual variation that is not completely explained by genotype. They also exhibit consistent methylation across tissues, which (i) suggests establishment in the early embryo before tissue differentiation, and (ii) indicates that DNAm levels measured in accessible tissues are likely to reflect those in inaccessible, but phenotypically potentially more relevant, tissues. These characteristics, together with reported associations between ME methylation, early environmental exposures, and postnatal disease traits, make them interesting candidates for studying the epigenetic mechanisms that underpin the DOHaD.

The study of human MEs is in its infancy, and many open questions remain. Larger cohorts with DNAm data spanning tissues derived from multiple germ layers and from diverse ancestries will provide a more definitive list of putative human MEs and will give insight into the extent to which these features are conserved in the human genome. The molecular mechanisms driving ME establishment in early development (in humans and in other organisms) are unclear, and causal pathways linking ME methylation to diverse exposures and phenotypic traits remain to be elucidated. This will require integration of multiomic data from longitudinal epidemiological cohorts coupled with *in vitro* and *in vivo* functional studies using cell and animal models, although we note that many putative human MEs do not align with those reported in rodents.

Finally, more detailed analysis of the influence of genetic variation will be necessary to fully elucidate the factors driving DNAm variation at MEs. Further insights may be gained from analysis of long-read sequencing data to map repetitive regions of the genome that are implicated in epigenetic silencing of TEs in early development (see Outstanding questions). Progress on all these fronts will ultimately advance our understanding of the relevance of this intriguing epigenetic phenomenon to human health and disease.

Outstanding questions

The true prevalence of putative human MEs in the human genome is unknown because approaches to identify them have relied on rare and difficult-to-acquire multi-tissue datasets with small sample sizes.

The relative influences of genetic, stochastic, and environmental effects on DNAm at putative human MEs, and the molecular mechanisms underpinning their establishment, remain to be fully elucidated.

Further investigation of the causal nature of links between ME methylation and later health outcomes is required.

The extent to which putative human MEs are selectively maintained in the human genome is unknown.

Acknowledgments

We would like to thank two anonymous reviewers for their helpful comments, as well as our funders, the Wellcome Trust (224575/Z/21/Z to G.H.), the Biotechnology and Biological Sciences Research council (BB/M009513/1 to M.D.), and the Medical Research Council (MC-A760-5QX00, MR/M01424X/1, and MR/T032863/1 to M.S.).

Declaration of interests

No interests are declared.

References

1. Rakyan, V.K. *et al.* (2002) Metastable epialleles in mammals. *Trends Genet.* 18, 348–351
2. Kazachenka, A. *et al.* (2018) Identification, characterization, and heritability of murine metastable epialleles: implications for non-genetic inheritance. *Cell* 175, 1259–1271
3. Wolff, G.L. *et al.* (1998) Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. *FASEB J.* 12, 949–957
4. Cooney, C.A. *et al.* (2002) Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. *J. Nutr.* 132, 2393S–2400S
5. Waterland, R.A. *et al.* (2006) Maternal methyl supplements increase offspring DNA methylation at Axin fused. *Genesis* 44, 401–406
6. Dolinoy, D.C. *et al.* (2006) Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. *Environ. Health Perspect.* 114, 567–572
7. Dolinoy, D.C. *et al.* (2007) Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. *Proc. Natl. Acad. Sci. U. S. A.* 104, 13056–13061
8. Anderson, O.S. *et al.* (2012) Epigenetic responses following maternal dietary exposure to physiologically relevant levels of bisphenol A. *Environ. Mol. Mutagen.* 53, 334–342
9. Kaminen-Ahola, N. *et al.* (2010) Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. *PLoS Genet.* 6, e1000811
10. Neier, K. *et al.* (2019) Perinatal exposures to phthalates and phthalate mixtures result in sex-specific effects on body weight,

organ weights and intracisternal A-particle (IAP) DNA methylation in weanling mice. *J. Dev. Orig. Health Dis.* 10, 176–187
11. Oey, H. *et al.* (2015) Genetic and epigenetic variation among inbred mouse littermates: identification of inter-individual differentially methylated regions. *Epigenetics Chromatin* 8, 54
12. Wolff, G.L. *et al.* (1999) Physiological consequences of ectopic agouti gene expression: the yellow obese mouse syndrome. *Physiol. Genomics* 1999, 151–163
13. Byun, H.M. *et al.* (2009) Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. *Hum. Mol. Genet.* 18, 4808–4817
14. Slieker, R.C. *et al.* (2013) Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics Chromatin* 6, 26
15. Joglekar, R. *et al.* (2022) Maternal tobacco smoke exposure is associated with increased DNA methylation at human metastable epialleles in infant cord blood. *Environ. Epigenet.* 8, dvac005
16. Gunasekara, C.J. *et al.* (2023) Systemic interindividual epigenetic variation in humans is associated with transposable elements and under strong genetic control. *Genome Biol.* 24, 2
17. Marzi, S.J. *et al.* (2016) Tissue-specific patterns of allelically-skewed DNA methylation. *Epigenetics* 11, 24–35
18. Candler, T. *et al.* (2021) DNA methylation at a nutritionally sensitive region of the PAX8 gene is associated with thyroid volume and function in Gambian children. *Sci. Adv.* 7, eabj1561
19. Marttila, S. *et al.* (2022) Methylation status of VTRNA2-1/nc886 is stable across populations, monozygotic twin pairs and in majority of tissues. *Epigenomics* 14, 1105–1124
20. Harris, R.A. *et al.* (2013) Human metastable epiallele candidates link to common disorders. *Epigenetics* 8, 157–163
21. Van Baak, T.E. *et al.* (2018) Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol.* 19, 2
22. Silver, M.J. *et al.* (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* 16, 118
23. Kessler, N.J. *et al.* (2018) Establishment of environmentally sensitive DNA methylation states in the very early human embryo. *Sci. Adv.* 4, eaat2624
24 Gunasekara, C.J. *et al.* (2019) A genomic atlas of systemic interindividual epigenetic variation in humans. *Genome Biol.* 20, 105
25. Michels, K. *et al.* (2013) Recommendations for the design and analysis of epigenome-wide association studies. *Nat. Methods* 10, 949–955
26. Derakhshan, M. *et al.* (2022) Tissue-and ethnicity-independent hypervariable DNA methylation states show evidence of establishment in the early human embryo. *Nucleic Acids Res.* 50, 6735–6752
27. Bell, J.T. *et al.* (2011) DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* 12, R10
28. Villicaña, S. and Bell, J.T. (2021) Genetic impacts on DNA methylation: research findings and future perspectives. *Genome Biol.* 22, 127
29. Waterland, R.A. *et al.* (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.* 6, e1001252
30. Dominguez-Salas, P. *et al.* (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* 5, 3746
31. Planterose Jiménez, B. *et al.* (2021) Equivalent DNA methylation variation between monozygotic co-twins and unrelated individuals reveals universal epigenetic inter-individual dissimilarity. *Genome Biol.* 22, 18
32. van Dongen, J. *et al.* (2021) Identical twins carry a persistent epigenetic signature of early genome programming. *Nat. Commun.* 12, 5618
33. Hannon, E. *et al.* (2021) Assessing the co-variability of DNA methylation across peripheral cells and tissues: implications for the interpretation of findings in epigenetic epidemiology. *PLoS Genet.* 17, e1009443
34. Islam, S.A. *et al.* (2019) Integration of DNA methylation patterns and genetic variation in human pediatric tissues help inform EWAS design and interpretation. *Epigenetics Chromatin* 12, 1
35. Lander, E.S. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
36. Bourque, G. *et al.* (2018) Ten things you should know about transposable elements. *Genome Biol.* 19, 199
37. Hancks, D.C. and Kazazian, H.H. (2016) Roles for retrotransposon insertions in human disease. *Mob. DNA* 7, 9
38. Huang, C.R.L. *et al.* (2012) Active transposition in genomes. *Annu. Rev. Genet.* 46, 651–675
39. Kazazian, H.H. (2011) Mobile DNA transposition in somatic cells. *BMC Biol.* 9, 62
40. Quenneville, S. *et al.* (2012) The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development. *Cell Rep.* 2, 766–773
41. Babaian, A. and Mager, D.L. (2016) Endogenous retroviral promoter exaptation in human cancer. *Mob. DNA* 7, 24
42. Rowe, H.M. *et al.* (2013) De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. *Development* 140, 519–529
43. Wolf, G. *et al.* (2015) The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses. *Genes Dev.* 29, 538–554
44. Rowe, H.M. *et al.* (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463, 237–240
45. Wiznerowicz, M. *et al.* (2007) The Krüppel-associated box repressor domain can trigger de novo promoter methylation during mouse early embryogenesis. *J. Biol. Chem.* 282, 34535–34541
46. Wolf, D. and Goff, S.P. (2009) Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* 458, 1201–1204
47. Iyengar, S. and Farnham, P.J. (2011) KAP1 protein: an enigmatic master regulator of the genome. *J. Biol. Chem.* 286, 26267–26276
48. Ecco, G. *et al.* (2017) KRAB zinc finger proteins. *Dev. (Cambridge)* 144, 2719–2729
49. Costello, K.R. *et al.* (2021) Sequence features of retrotransposons allow for epigenetic variability. *eLife* 10, e71104
50. Elmer, J.L. *et al.* (2021) Genomic properties of variably methylated retrotransposons in mouse. *Mob. DNA* 12, 6
51. Kühnen, P. *et al.* (2016) Interindividual variation in DNA methylation at a putative POMC metastable epiallele is associated with obesity. *Cell Metab.* 24, 502–509
52. Kuehnen, P. and Krude, H. (2012) Alu elements and human common diseases like obesity. *Mob. Genet. Elem.* 2, 197–201
53. Onuchic, V. *et al.* (2018) Allele-specific epigenome maps reveal sequence-dependent stochastic switching at regulatory loci. *Science* 361, eaar3146
54. Silver, M.J. *et al.* (2022) Environmentally sensitive hotspots in the methylome of the early human embryo. *eLife* 11, e72031
55. Zink, F. *et al.* (2018) Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. *Nat. Genet.* 50, 1542–1552
56. Piedrahita, J.A. (2011) The role of imprinted genes in fetal growth abnormalities. *Birth Defects Res. A Clin. Mol. Teratol.* 91, 682–692
57. Moore, G.E. *et al.* (2015) The role and interaction of imprinted genes in human fetal growth. *Philos. Trans. R. Soc. B Biol. Sci.* 370, 20140074
58. Carpenter, B.L. *et al.* (2021) Oocyte age and preconceptual alcohol use are highly correlated with epigenetic imprinting of a noncoding RNA (nc886). *Proc. Natl. Acad. Sci. U. S. A.* 118
59. Heijmans, B.T. *et al.* (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. U. S. A.* 105, 17046–17049
60. Steegers-Theunissen, R.P. *et al.* (2009) Periconceptional maternal folic acid use of 400 μg per day is related to increased methylation of the IGF2 gene in the very young child. *PLoS One* 4, e7845
61. Hoyo, C. *et al.* (2011) Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. *Epigenetics* 6, 928–936
62. Lee, H.S. *et al.* (2014) Dietary supplementation with polyunsaturated fatty acid during pregnancy modulates DNA methylation at IGF2/H19 imprinted genes and growth of infants. *Physiol. Genomics* 46, 851–857
63. Monk, D. *et al.* (2019) Genomic imprinting disorders: lessons on how genome, epigenome and environment interact. *Nat. Rev. Genet.* 20, 235–248
64. Takahashi, N. *et al.* (2015) ZFP57 and the targeted maintenance of postfertilization genomic imprints. *Cold Spring Harb. Symp. Quant. Biol.* 80, 177–187

65. Gonseth, S. *et al.* (2015) Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. *Epigenetics* 10, 1166–1176
66. Amarasekera, M. *et al.* (2014) Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. *FASEB J.* 28, 4068–4076
67. Irwin, R.E. *et al.* (2019) A randomized controlled trial of folic acid intervention in pregnancy highlights a putative methylation-regulated control element at ZFP57. *Clin. Epigenetics* 11, 31
68. Do, C. *et al.* (2020) Allele-specific DNA methylation is increased in cancers and its dense mapping in normal plus neoplastic cells increases the yield of disease-associated regulatory SNPs. *Genome Biol.* 21, 153
69. James, P.T. *et al.* (2019) Maternal one-carbon metabolism and infant DNA methylation between contrasting seasonal environments: a case study from The Gambia. *Curr. Dev. Nutr.* 3, nzy082
70. Boks, M.P. *et al.* (2018) Genetic vulnerability to DUSP22 promoter hypermethylation is involved in the relation between in utero famine exposure and schizophrenia. *NPJ Schizophr.* 4, 16
71. Finer, S. *et al.* (2016) Is famine exposure during developmental life in rural Bangladesh associated with a metabolic and epigenetic signature in young adulthood? A historical cohort study. *BMJ Open* 6, e011768
72. Saffari, A. *et al.* (2020) Effect of maternal preconceptional and pregnancy micronutrient interventions on children's DNA methylation: findings from the EMPHASIS study. *Am. J. Clin. Nutr.* 112, 1099–1113
73. Cooper, W.N. *et al.* (2012) DNA methylation profiling at imprinted loci after periconceptional micronutrient supplementation in humans: results of a pilot randomized controlled trial. *FASEB J.* 26, 1782–1790
74. Estill, M.S. *et al.* (2016) Assisted reproductive technology alters deoxyribonucleic acid methylation profiles in bloodspots of newborn infants. *Fertil. Steril.* 106, 629–639
75. Junge, K.M. *et al.* (2018) MEST mediates the impact of prenatal bisphenol A exposure on long-term body weight development. *Clin. Epigenetics* 10, 58
76. Bertozzi, T.M. *et al.* (2021) Variably methylated retrotransposons are refractory to a range of environmental perturbations. *Nat. Genet.* 53, 1233–1242
77. Fleming, T.P. *et al.* (2018) Origins of lifetime health around the time of conception: causes and consequences. *Lancet* 391, 1842–1852
78. Moore, S.E. *et al.* (1999) Prenatal or early postnatal events predict infectious deaths in young adulthood in rural Africa. *Int. J. Epidemiol.* 28, 1088–1095
79. Tobi, E.W. *et al.* (2021) DNA methylation differences at birth after conception through ART. *Hum. Reprod.* 36, 248–259
80. Håberg, S.E. *et al.* (2022) DNA methylation in newborns conceived by assisted reproductive technology. *Nat. Commun.* 13, 1896
81. Howe, C.G. *et al.* (2020) Maternal gestational diabetes mellitus and newborn DNA methylation: findings from the pregnancy and childhood epigenetics consortium. *Diabetes Care* 43, 98–105
82. Sarkar, D.K. *et al.* (2019) Persistent changes in stress-regulatory genes in pregnant women or children exposed prenatally to alcohol. *Alcohol. Clin. Exp. Res.* 43, 1887–1897
83. Gurnot, C. *et al.* (2015) Prenatal antidepressant exposure associated with CYP2E1 DNA methylation change in neonates. *Epigenetics* 10, 361–372
84. Miura, R. *et al.* (2018) An epigenome-wide study of cord blood DNA methylations in relation to prenatal perfluoroalkyl substance exposure: the Hokkaido study. *Environ. Int.* 115, 21–28
85. Richmond, R.C. *et al.* (2018) The long-term impact of folic acid in pregnancy on offspring DNA methylation: follow-up of the Aberdeen Folic Acid Supplementation Trial (AFAST). *Int. J. Epidemiol.* 47, 928–937
86. Ambatipudi, S. *et al.* (2016) Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. *Epigenomics* 8, 599–618
87. Ehrlich, S. *et al.* (2012) Smoking, but not malnutrition, influences promoter-specific DNA methylation of the proopiomelanocortin gene in patients with and without anorexia nervosa. *Can. J. Psychiatry* 57, 168–176
88. Kumsta, R. *et al.* (2016) Severe psychosocial deprivation in early childhood is associated with increased DNA methylation across a region spanning the transcription start site of CYP2E1. *Transl. Psychiatry* 6, e830
89. van der Plaat, D.A. *et al.* (2018) Occupational exposure to pesticides is associated with differential DNA methylation. *Occup. Environ. Med.* 75, 427–435
90. van der Plaat, D.A. *et al.* (2019) Occupational exposure to gases/fumes and mineral dust affect DNA methylation levels of genes regulating expression. *Hum. Mol. Genet.* 28, 2477–2485
91. Jiménez-Garza, O. *et al.* (2017) Promoter methylation status in genes related with inflammation, nitrosative stress and xenobiotic metabolism in low-level benzene exposure: searching for biomarkers of oncogenesis. *Food Chem. Toxicol.* 109, 669–676
92. Jiménez-Garza, O. *et al.* (2015) CYP2E1 epigenetic regulation in chronic, low-level toluene exposure: relationship with oxidative stress and smoking habit. *Toxicol. Appl. Pharmacol.* 286, 207–215
93. Perez, M.F. and Lehner, B. (2019) Intergenerational and transgenerational epigenetic inheritance in animals. *Nat. Cell Biol.* 21, 143–151
94. Messerschmidt, D.M. *et al.* (2014) DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* 28, 812–828
95. Lambrot, R. *et al.* (2021) Whole-genome sequencing of H3K4me3 and DNA methylation in human sperm reveals regions of overlap linked to fertility and development. *Cell Rep.* 36, 109418
96. Chen, Q. *et al.* (2016) Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nat. Rev. Genet.* 17, 733–743
97. Radford, E.J. *et al.* (2014) In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. *Science* 345, 1255903
98. Fox, J.T. and Stover, P.J. (2008) Folate-mediated one-carbon metabolism. *Vitam. Horm.* 79, 1–44
99. Steegers-Theunissen, R.P. *et al.* (2013) The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. *Hum. Reprod. Update* 19, 640–655
100. Dominguez-Salas, P. *et al.* (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* 97, 1217–1227
101. Brosnan, M.E. *et al.* (2015) Division of labour: how does folate metabolism partition between one-carbon metabolism and amino acid oxidation? *Biochem. J.* 472, 135–146
102. Rios-Avila, L. *et al.* (2015) Metabolite profile analysis reveals association of vitamin B-6 with metabolites related to one-carbon metabolism and tryptophan catabolism but not with biomarkers of inflammation in oral contraceptive users and reveals the effects of oral contraceptives on these processes. *J. Nutr.* 145, 87–95
103. Tuenter, A. *et al.* (2019) Folate, vitamin B12, and homocysteine in smoking-exposed pregnant women: a systematic review. *Matern. Child Nutr.* 15, e12675
104. Waterland, R.A. and Jirtle, R.L. (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* 23, 5293–5300
105. Waterland, R.A. and Michels, K.B. (2007) Epigenetic epidemiology of the developmental origins hypothesis. *Annu. Rev. Nutr.* 27, 363–388
106. Barker, D.J.P. (2004) The developmental origins of adult disease. *J. Am. Coll. Nutr.* 23, 588S–595S
107. Barker, D.J. *et al.* (1989) Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. *Br. Med. J.* 298, 564–567
108. Barker, D.J.P. *et al.* (1993) Fetal nutrition and cardiovascular disease in adult life. *Lancet* 341, 938–941
109. Ravelli, A.C.J. *et al.* (1999) Obesity at the age of 50 y in men and women exposed to famine prenatally. *Am. J. Clin. Nutr.* 70, 811–816
110. Gluckman, P.D. *et al.* (2011) The role of developmental plasticity and epigenetics in human health. *Birth Defects Res. C Embryo Today* 93, 12–18
111. Joubert, B.R. *et al.* (2016) Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat. Commun.* 7, 10577
112. Tobi, E.W. *et al.* (2014) DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nat. Commun.* 5, 5592

113. Tobi, E.W. *et al.* (2018) DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci. Adv.* 4, eaao4364
114. Shen, L. *et al.* (2019) Early-life exposure to severe famine is associated with higher methylation level in the IGF2 gene and higher total cholesterol in late adulthood: the Genomic Research of the Chinese Famine (GRECF) study. *Clin. Epigenetics* 11, 88
115. Wang, Z. *et al.* (2019) Early-life exposure to the Chinese Famine is associated with higher methylation level in the INSR gene in later adulthood. *Sci. Rep.* 9, 3354
116. Loke, Y.J. *et al.* (2018) Time-and sex-dependent associations between prenatal alcohol exposure and placental global DNA methylation. *Epigenomics* 10, 981–991
117. Frey, S. *et al.* (2018) Prenatal alcohol exposure is associated with adverse cognitive effects and distinct whole-genome DNA methylation patterns in primary school children. *Front. Behav. Neurosci.* 12, 125
118. Mandal, C. *et al.* (2017) Gestational alcohol exposure altered DNA methylation status in the developing fetus. *Int. J. Mol. Sci.* 18, 1386
119. Hernandez-Vargas, H. *et al.* (2015) Exposure to aflatoxin B1 in utero is associated with DNA methylation in white blood cells of infants in The Gambia. *Int. J. Epidemiol.* 44, 1238–1248
120. Joubert, B.R. *et al.* (2012) 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect.* 120, 1425–1431
121. Markunas, C.A. *et al.* (2014) Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environ. Health Perspect.* 122, 1147–1153
122. Richmond, R.C. *et al.* (2015) Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum. Mol. Genet.* 24, 2201–2217
123. Lee, K. *et al.* (2011) Precursor miR-886, a novel noncoding RNA repressed in cancer, associates with PKR and modulates its activity. *RNA* 17, 1076–1089
124. van Dijk, S.J. *et al.* (2015) Recent developments on the role of epigenetics in obesity and metabolic disease. *Clin. Epigenetics* 7, 1–13
125. Marttila, S. *et al.* (2021) Methylation status of nc886 epiallele reflects periconceptional conditions and is associated with glucose metabolism through nc886 RNAs. *Clin. Epigenetics* 13, 143
126. Yoo, J.Y. *et al.* (2014) Can proopiomelanocortin methylation be used as an early predictor of metabolic syndrome? *Diabetes Care* 37, 734–739
127. Mok, A. *et al.* (2018) Hypomethylation of CYP2E1 and DUSP22 promoters associated with disease activity and erosive disease among rheumatoid arthritis patients. *Arthritis Rheum.* 70, 528–536
128. Gonseth, S. *et al.* (2019) Epigenomic profiling of newborns with isolated orofacial clefts reveals widespread DNA methylation changes and implicates metastable epiallele regions in disease risk. *Epigenetics* 14, 198–213
129. Teschendorff, A.E. and Zheng, S.C. (2017) Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics* 9, 757–768
130. Ross, J.P. *et al.* (2022) Batch-effect detection, correction and characterisation in Illumina HumanMethylation450 and MethylationEPIC BeadChip array data. *Clin. Epigenetics* 14, 58
131. Mountjoy, K.G. (2015) Pro-opiomelanocortin (POMC) neurones, POMC-derived peptides, melanocortin receptors and obesity: how understanding of this system has changed over the last decade. *J. Neuroendocrinol.* 27, 406–418
132. Kuehnen, P. *et al.* (2012) An Alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet.* 8, e1002543
133. Fernandez, L.P. *et al.* (2015) Thyroid transcription factors in development, differentiation and disease. *Nat. Rev. Endocrinol.* 11, 29–42
134. van Dijk, S.J. *et al.* (2018) DNA methylation in blood from neonatal screening cards and the association with BMI and insulin sensitivity in early childhood. *Int. J. Obes.* 42, 28–35
135. Treppendahl, M.B. *et al.* (2012) Allelic methylation levels of the noncoding VTRNA2-1 located on chromosome 5q31.1 predict outcome in AML. *Blood* 119, 206–216
136. Cao, J. *et al.* (2013) DNA methylation-mediated repression of miR-886-3p predicts poor outcome of human small cell lung cancer. *Cancer Res.* 73, 3326–3335
137. Lee, H.S. *et al.* (2014) Epigenetic silencing of the non-coding RNA nc886 provokes oncogenes during human esophageal tumorigenesis. *Oncotarget* 5, 3472
138. Woods, L. *et al.* (2018) Regulation of placental development and its impact on fetal growth – new insights from mouse models. *Front. Endocrinol. (Lausanne)* 9, 570
139. Koukoura, O. *et al.* (2012) DNA methylation in the human placenta and fetal growth. *Mol. Med. Rep.* 5, 883–889
140. Clark, J. *et al.* (2019) Associations between placental CpG methylation of metastable epialleles and childhood body mass index across ages one, two and ten in the Extremely Low Gestational Age Newborns (ELGAN) cohort. *Epigenetics* 14, 1102–1111
141. Kappil, M.A. *et al.* (2015) Placental expression profile of imprinted genes impacts birth weight. *Epigenetics* 10, 842–849
142. Zhu, Y. *et al.* (2019) Placental DNA methylation levels at CYP2E1 and IRS2 are associated with child outcome in a prospective autism study. *Hum. Mol. Genet.* 28, 2659–2674
143. Hales, C.N. and Barker, D.J.P. (2001) The thrifty phenotype hypothesis: type 2 diabetes. *Br. Med. Bull.* 60, 5–20
144. Prentice, A.M. *et al.* (2005) Insights from the developing world: thrifty genotypes and thrifty phenotypes. *Proc. Nutr. Soc.* 64, 153–161
145. Bateson, P. *et al.* (2014) The biology of developmental plasticity and the predictive adaptive response hypothesis. *J. Physiol.* 592, 2357–2368
146. Gluckman, P.D. *et al.* (2005) Predictive adaptive responses and human evolution. *Trends Ecol. Evol.* 20, 527–533
147. Feinberg, A.P. and Irizarry, R.A. (2010) Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci. U. S. A.* 107, 1757–1764
148. You, Y.A. *et al.* (2021) Elevated methylation of the vault RNA2-1 promoter in maternal blood is associated with preterm birth. *BMC Genomics* 22, 528
149. Yu, M.C. *et al.* (2020) Differential hypermethylation of the VTRNA2-1 promoter in hepatocellular carcinoma as a prognostic factor: tumor marker prognostic study. *Int. J. Surg.* 79, 282–289
150. Dugué, P. *et al.* (2018) Heritable methylation marks associated with breast and prostate cancer risk. *Prostate* 78, 962–969
151. Zheng, D. *et al.* (2020) Epigenetic alterations of the promoter region of the POMC gene in adolescent depressive disorder patients with nonsuicidal self-injury behaviors. *Psychol. Res. Behav. Manag.* 13, 997
152. Plante, D.T. *et al.* (2021) PAX8/PAX8-AS1 DNA methylation levels are associated with objective sleep duration in persons with unexplained hypersomnolence using a deep phenotyping approach. *Sleep* 44, zsab108
153. Houshdaran, S. *et al.* (2007) Widespread epigenetic abnormalities suggest a broad DNA methylation erasure defect in abnormal human sperm. *PLoS One* 2, e1289
154. Kaut, O. *et al.* (2012) Genome-scale methylation analysis of Parkinson's disease patients' brains reveals DNA hypomethylation and increased mRNA expression of cytochrome P450 2E1. *Neurogenetics* 13, 87–91
155. Sanchez-Mut, J.V. *et al.* (2014) Promoter hypermethylation of the phosphatase DUSP22 mediates PKA-dependent TAU phosphorylation and CREB activation in Alzheimer's disease. *Hippocampus* 24, 363–368
156. Katrinli, S. *et al.* (2021) PTSD is associated with increased DNA methylation across regions of HLA-DPB1 and SPATC1L. *Brain Behav. Immun.* 91, 429–436
157. El Hajj, N. *et al.* (2013) Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes* 62, 1320–1328
158. Azzi, S. *et al.* (2014) Degree of methylation of ZAC1 (PLAGL1) is associated with prenatal and post-natal growth in healthy infants of the EDEN mother child cohort. *Epigenetics* 9, 338–345