Assessing the influence of maternal nutrition-sensitive epigenetic signatures in the *POMC* and *PAX-8* genes on health-related outcomes in offspring

Dr Toby Candler, B.Sc., MB BS, MRCPCH

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

2022

Department of Population Health

Faculty of Epidemiology and Public Health

London School of Hygiene and Tropical Medicine, University of London

Supervisor: Dr Matt Silver

Associate Supervisor: Prof Andrew Prentice

Funded by: Medical Research Council UK

Research Group/Affiliation: Nutrition Theme, MRC The Gambia at London School of Hygiene and Tropical Medicine

Statement of own work

I, Toby P Candler, declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

11th July 2022

Acknowledgements

I am extremely grateful to my supervisor, Dr Matt Silver for his kind, calm, and generous approach to supervision. Despite us being on different continents for much of the PhD, he still provided invaluable advice and guidance, continuous support, and patience. Matt's immense knowledge and plentiful experience is inspirational and I have learnt a *great* deal from our discussions. He remains pivotal in encouraging me to follow my own interests and academic ideas and I will be forever grateful.

I would like to thank my associate supervisor Professor Andrew Prentice for giving me the opportunity to come to The Gambia. Andrew fondly welcomed us to Keneba and made us feel at home in West Kiang quickly. It was great working closely on nutrition theme affairs which provided a wonderful opportunity to learn about strategic matters pertinent to the research group.

I would like to thank the 'frontline' *POMC* and *PAX8* study research team which included field supervisors, field workers, nurses, drivers, cooks, data managers, and laboratory staff. All of the team rose to the challenge of visiting nearly a 1000 participants every month in their villages. The calmness and professionalism you all showed in running the studies is commendable. I would like to specifically mention Lamin Sangyang and Kebba Bajo (field supervisors) who led their teams impeccably and are a credit to MRC The Gambia. Mustapha Saidy, Edrisa Sinjanka, Sainey Beyai, my nursing colleagues, managed to exceed all expectations in doing so many blood draws in a short time period. Thank you to Ebrima Danso and Ebrima Bah, Scientific officers in Keneba lab, who taught this lab- naïve clinician how to pipette and made me feel like I could justify wearing a lab coat. Thank you to Abdul Sessay, Genomics lab manager, for the warm welcome he gave me into his team. Thank you to Abdoulie Kanteh and Ebrima Bah in the genomic team for their calmness and sense of humour during many a late night in the lab.

Thank you to all the collaborators for their support, guidance, and expertise. Noah Kessler was incredibly helpful with genomic bioinformatic support and in numerous discussions about genetics and epigenetics. Dr Peter Kühnen has been a major collaborator from the beginning. It has been wonderful to have another endocrinologist to discuss ideas with and I am very grateful for Peter and his colleague Lara Lechner in supporting me in running methylation analysis in his lab in Berlin. I would like to thank Professor Marian Ludgate for encouraging me in my efforts to get the *PAX8* grant and for supporting the research proposal from its inception. Professor Rob Waterland has been very supportive of my work and with support from his lab we were really able to push forward the *PAX8* story into new areas. Professors Kate Ward and Ann Prentice were very supportive of using the DXA

facilities and were crucial in getting me trained on the DXA machine and provided advice regarding data interpretation. Dr Landing Jarjou, who is a fountain of knowledge on all things Keneba, provided me with astute counsel and advice to do with my studies and life in West Kiang.

Most importantly I would like to thank my family who though out this journey have been an unwavering source of encouragement, support, and inspiration. In particular a heartfelt thanks goes to my wife Liz, who has been a constant support and sanctuary through the ups and downs of grant applications, through my PhD and importantly maintaining a smile and a laugh whilst we juggled the inter-continental relocation of three small children! To Henry, Sam and Arthur thank you for being the incredible little boys that you are. I hope your early life experience in The Gambia will stay with you. I can also assure you that Daddy's "Gambia book" is done for now. Thank you to Liz's family, especially Mary and David, for supporting the family during the whole process. A huge thank you to my Mum for encouraging me to follow my interests, teaching me perseverance and being supportive when we moved to West Africa!

Finally, I would like to thank the study participants and their families for their time and trust. Your continued commitment to research is inspiring and an example to us all.

Abstract

A key challenge in nutritional epigenetics is to link early pregnancy exposures with offspring epigenetic alterations and later phenotype or health outcomes. Epigenetic processes such as DNA methylation can influence gene expression and the periconceptional period is a particularly sensitive developmental window for establishing DNA methylation patterns. In this respect metastable epialleles (MEs) are of particular interest since their distinctive methylation patterns suggest establishment in very early embryonic development that may be sensitive to environmental factors. Furthermore, ME methylation patterns are tissue-independent making them promising candidates for the study of epigenetic developmental programming in humans using easily accessible tissues such as blood.

Using a candidate gene approach, the research presented in this thesis characterised how early periconceptional exposures influence DNA methylation at two MEs at the *POMC* (Proopiomelanocortin) and *PAX8* (Paired Box 8) genes and how methylation levels at these genes are associated with phenotype.

Seasonality in The Gambia is associated with profound changes in multiple environmental factors and *POMC* and *PAX-8* methylation is associated with Gambian season of conception (SoC) and mothers' early pregnancy nutrition. POMC is a key regulator of satiety and energy balance and *POMC* hypermethylation is associated with obesity. PAX-8 is a thyroid transcription factor implicated in thyroid gland development and differentiation, important processes for children's growth and neurocognitive development.

In a prospective, year-long study of seasonally-driven weight and adiposity changes in Gambian mothers and children, a number of key associations with *POMC* methylation were identified. Firstly, *POMC* methylation was higher in those conceived in the rainy season and associated with maternal periconceptional amino acid concentrations. Secondly, higher methylation at *POMC* was associated with lower amplitude of mothers' fat mass index change across the year. Thirdly, for both mothers and children, there was no association between *POMC* methylation and measures of appetite or satiety.

Using a recall by epigenotype study design in Gambian children, *PAX8* hypermethylation was associated with lower free thyroxine (FT4, a thyroid hormone) and smaller thyroid volume. Furthermore, increased FT4 (still within the population reference range) was associated with lower

fat mass and bone mineral density. *PAX8* methylation was also associated with maternal periconceptional levels of key one carbon metabolites homocysteine, cysteine, B6 and B12.

In summary, this research highlights important phenotypic associations with DNA methylation at two human MEs with potential implications for epigenetic programming, developmental biology, and public health. Our demonstration that mothers' diet in early pregnancy influences offspring DNA methylation which in turn is associated with a particular phenotype, suggests that targeted dietary interventions that positively influence the offspring's epigenome could play a role in improving health outcomes in future generations.

Table of contents

List of Tables

List of Figures

Chapter 1 Background to thesis

Summary of the chapter

In this chapter I present the background of the thesis including introducing the DOHaD (developmental origins of health and disease) model. I provide a primer on epigenetic control of gene expression and introduce the concept of how prenatal maternal nutritional and environmental factors can influence the offspring's epigenetic landscape and postnatal phenotype. I summarise the rationale for focussing on the study of *POMC* and *PAX8* genes and provide an overview of their biological role. I detail the study hypothesis, research questions, aims and objectives. I describe my role with respect to each element of the thesis and provide details on the research output, PhD timeline and funding.

1.1 Introduction

1.1.1 Developmental Origins of Health and Disease (DOHaD)

Over three decades ago, David Barker proposed a link between low birth weight, prematurity and intrauterine growth restriction and adverse health outcomes such as type 2 diabetes mellitus, hypertension, and cardiovascular disease in adult life[1,2]. The 'Barker hypothesis' paved the way for an expanding scientific field, that has produced burgeoning evidence of a link between prenatal adversity and later health and disease outcomes.

Much of the evidence for such a link has stemmed from epidemiological longitudinal cohort studies such as the Dutch Hunger Winter Cohort (DHWC)[3]. The DHWC utilises a period during the Second World War where daily food intake fell below 1000 calories a day to examine how maternal exposure to famine could influence the health of offspring. By studying children born to mothers exposed to this famine, researchers identified that exposure in early gestation (compared to late or mid gestation) represented the greatest risk for adverse health outcomes in offspring. Early gestational famine exposure has been associated with increased risk of schizophrenia and depression, adverse lipid profile, and coronary artery disease[4,5]. These epidemiological studies fall short of identifying a specific biological mechanism linking exposure to outcome, but have been the basis of the DOHaD paradigm[6]. In summary, this model postulates that environmental and uterine factors can influence the foetus at key developmental windows, so called 'developmental plasticity'. Furthermore, at some point in development this plasticity diminishes, and the influence of the early stimuli 'programmes' the individual towards a phenotype in later life. Developmental plasticity and responsiveness to the *in utero* environment may present a survival advantage by the organism being better prepared for the anticipated *ex utero* environment[7]. However, a mismatch between *in utero* and *ex utero* environments could lead to deleterious health consequences[8]. In recent years, epigenetic processes are thought to be act as a 'biological conduit' through which early life exposures can influence later offspring phenotype.

1.1.2 Epigenetic regulation of gene expression

Epigenetics has been described as the study of mitotically heritable changes in gene expression that occur without changes in DNA sequence[9,10]. Epigenetic processes, including DNA methylation,

histone modification, chromatin remodelling and RNA-based mechanisms can affect gene expression[11]. Epigenetic processes are essential for cell differentiation and maintenance of cellular identity, control of imprinted genes and X-chromosome inactivation[12,13].

DNA methylation is widely studied in animals and humans and occurs primarily at cytosine-guanine (CG) dinucleotides (CpG methylation). CpG islands (genomic regions with high CpG density) are often located near promoter regions of genes, are invariably sites of transcription initiation and the majority are unmethylated[13]. DNA methylation is usually associated with condensed heterochromatin and subsequent gene silencing or reduced expression[14].

Two periods of DNA demethylation occur during embryonic development (see Figure 1.1). Firstly, there is global demethylation of the genome of proliferating primordial germ cells before new DNA methylation landscapes are established in the germ cell precursors. A second erasure of DNA methylation occurs after fertilisation. After blastocyst implantation, the DNA methylation landscape is thereafter re-established. The periconceptional and prenatal environment undergoes significant epigenetic reprogramming and represents a key window for epigenetic alterations.

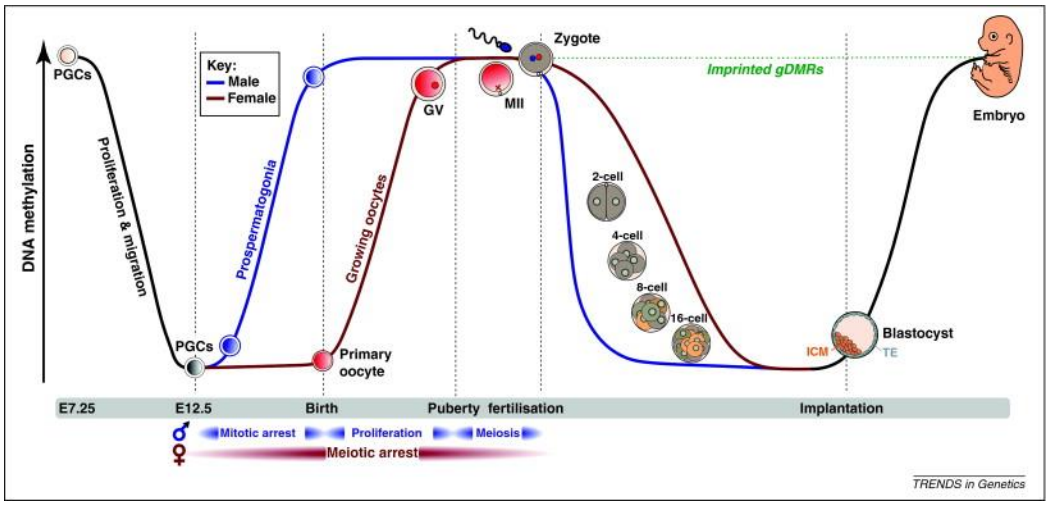


Figure 1.1 DNA methylation changes during developmental epigenetic reprogramming. Source: Reproduced with permission. Smallwood SA, Kelsey G. De novo DNA methylation: A germ cell perspective. Trends Genet. 2012 [15]

1.1.3 Metastable epialleles and the influence of periconceptional environment

Metastable epialleles (MEs) are variably expressed in isogenic individuals dependent on epigenetic modifications which are particularly sensitive to periconceptional environmental influences. As such, they represent epigenetic loci that demonstrate tissue concordant (systemic) methylation patterns within individuals[16], but show significant interindividual variation. The methylation patterns are established early in embryonic development[17] (before gastrulation), largely independent of genotype[16,17], may be influenced by maternal diet around conception[18–22] and are often close to transposable elements[21]. Variably Methylated Regions (VMR) are genomic regions that show substantial interindividual variation in methylation. Methylation variation at these genomic regions and loci may contribute to phenotypic (e.g. body weight) variation between individuals by altering gene expression and be an important contribution to an individual's health and disease risk[23]. MEs are of particular interest as i) it is possible to pinpoint the timing of the environmental exposure as the methylation pattern is established in the very early embryo (i.e. periconceptional period) and ii) due to the systemic DNA methylation pattern across different tissue types they enable analysis of methylation in accessible tissue as a proxy for the tissue of interest (e.g. leucocyte methylation (easily accessible) provides a proxy measure for methylation at the arcuate nucleus of the hypothalamus (difficult to measure to live human subjects)).

The most robust evidence for MEs comes from the Agouti variable yellow (A^{vy}) [21] and Axin-fused ($Axin^{Fu}$) mouse models[24]. With A^{vy} as an example, isogenic A^{vy} mice show variable agouti expression with the gene expression and subsequent phenotype dependent on the epigenetic state. The degree of DNA methylation at a cryptic promoter within an IAP (intracisternal A particle (IAP)) upstream of the *agouti* gene, influences *agouti* gene expression. A maternal diet preconceptionally (and through pregnancy) rich in methyl groups such as folate, B12, choline and betaine gives rise to increased methylation at the IAP in A^{vy} offspring[21]. The degree of IAP methylation influences the phenotype, such that hypomethylation leads to increased ectopic agouti expression producing an obese mouse with yellow fur, whereas hypermethylation is associated with reduced expression and a lean mouse with brown or mottled fur (see Figure 1.2).

MEs are well established in murine models with increasing evidence for human MEs [16,18,23,25–30].

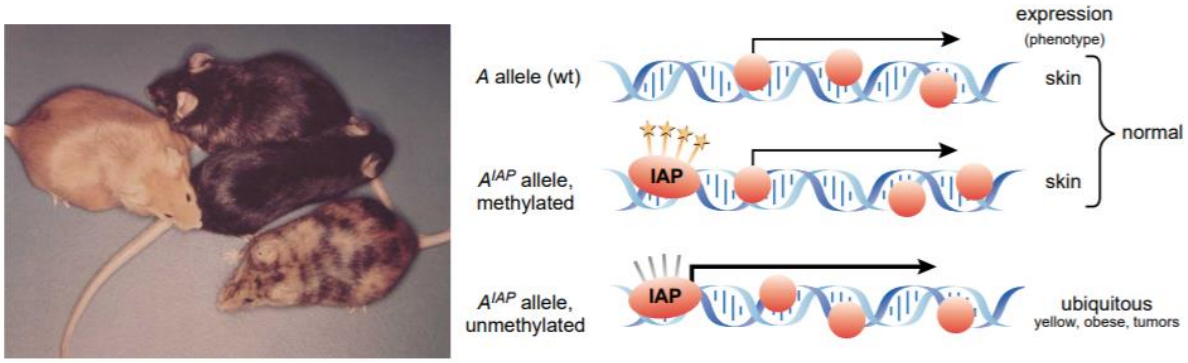


Figure 1.2 Agouti Mouse Model demonstrating the influence of differential methylation on phenotype. Source: Jaenisch R, Bird A. Epigenetic regulation of gene expression : how the genome integrates intrinsic and environmental signals. Nature Genetics, 2003. [12]

1.1.4 One-Carbon Metabolism

As demonstrated in the example of the A^{vy} mouse, periconceptional nutrition and importantly circulating levels of one-carbon metabolites are thought to be an important factors determining the methylation pattern at MEs. A set of interlocking pathways, collectively known as one-carbon metabolism (see Figure 1.3), provide methyl groups for methylation reactions including the methylation of cytosine bases and histone tails that in turn influence gene expression. These methylation reactions are controlled by methyltransferases that act on methyl groups produced by the conversion of S-adenosyl methionine (SAM) to S-adenosyl homocysteine (SAH). Periconceptional levels of one-carbon metabolites, namely folate, B12, choline and betaine, were shown to influence methylation at the IAP close to the Agouti gene. Hoyo et al [31] showed that maternal concentrations of folate in early pregnancy were associated with offspring's DNA methylation and birth weight, and a recent review illustrated the importance of prenatal exposure to one-carbon related nutrients and vitamins on DNA methylation in multiple genes in humans[26].

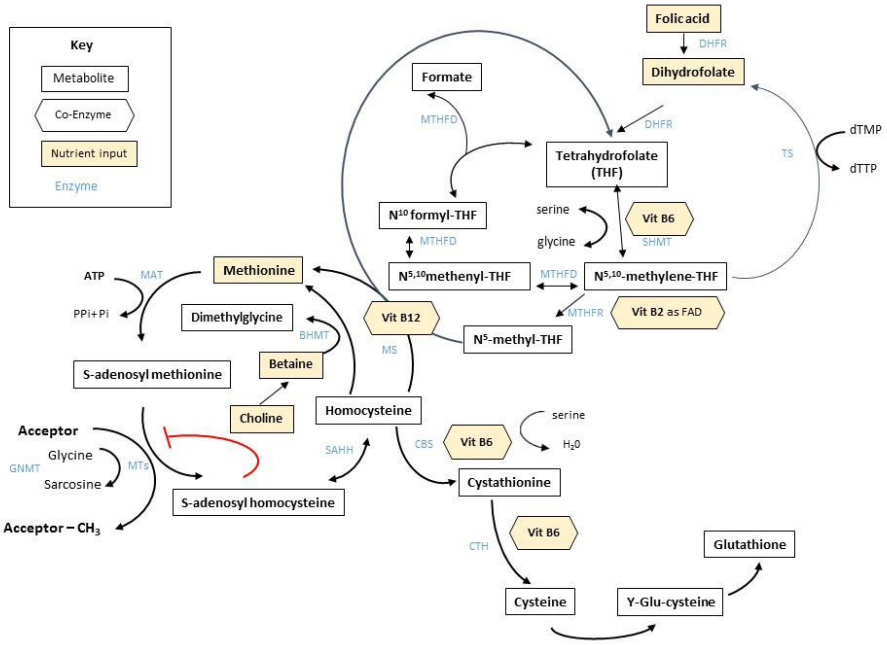


Figure 1.3 An overview of one-carbon metabolism. BHMT, Betaine Homocysteine MethylTransferase; CBS, Cystathionine-Beta-Synthase; CTH, Cystathionine Gamma-Lyase; DHFR, Dihydrofolate Reductase; dTMP, Deoxythymidine Monophosphate; dTTP, Deoxythymidine Triphosphate; FAD, Flavin Adenine Dinucleotide; GNMT, Glycine N-MethylTransferase; MAT, Methionine AdenosylTransferase; MS, Methionine Synthase; MT, Methyl Transferases; MTHFD, MethyleneTetraHydroFolate Dehydrogenase; MTHF, MethyleneTetraHydroFolate Reductase; SAHH, S-Adenosyl Homocysteine Hydrolase; SHMT, Serine HydroxyMethylTransferase; TS, Thymidylate Synthase. Source: Reproduced with permission under Creative Commons, CC BY-NC-ND. James PT, Silver MJ, Prentice AM. Epigenetics, nutrition, and infant health. In: The Biology of the First 1,000 Days [32]

1.1.5 Season of conception (SoC) in The Gambia, the periconceptional environment and offspring DNA methylation

Seasonal changes in climatic conditions, such as those observed in The Gambia, present an opportunity to observe an experiment of nature whereby conceptions, pregnancies and births are randomised to occur against different environmental (especially nutritional) backgrounds.

The SoC in The Gambia has been shown to be associated with later health consequences. Moore et al, demonstrated that those conceived in the rainy season were up to 10 times more likely to die prematurely in young adulthood compared to those conceived in the dry season[33]. This striking

periconceptional seasonal effect on adult mortality has driven further research to characterise the underlying mechanisms and biological pathways.

Waterland et al was first to demonstrate that the SoC in The Gambia was associated with differential DNA methylation in mid-childhood at putative human MEs[30]. Thirty children (mean 8.9 years [SE=0.5]) conceived in each season (Wet=August-September, Dry=March to May) between 1991-1998 had DNA methylation measured at 5 putative human MEs (including *PAX8*). DNA methylation was higher in rainy season conceptions across all 5 MEs (p=0.0001).

Exploring why the nutritionally challenged rainy or 'hungry' season was associated with higher DNA methylation led to research explore if seasonal differences in one-carbon metabolites in early pregnancy were driving the observed methylation differences. Dominguez-Salas et al demonstrated that dietary intake and maternal circulating levels of one-carbon metabolites (riboflavin, folate, choline, and betaine) varied throughout the year[34]. There were seasonally driven differences in SAM/SAH ratio with a 17.6% higher SAM:SAH ratio in rainy season blood draws.

Further work by Dominguez-Salas et al, established a link between seasonally-driven differences in maternal circulating one-carbon metabolites and offspring DNA methylation at several MEs[22]. Circulating levels of one-carbon metabolites from early pregnancy were back extrapolated to the time of conception using seasonal trends determined by biomarker levels found in a non-pregnant indicator group (a group of women who had monthly blood draws and assessment of nutritional intake). This study recapitulated the previous finding that DNA methylation in infants (mean [SD] age of blood sampling =3.6 months, [0.9 months]) conceived in the rainy season was significantly higher than those conceived in the dry season across 6 putative human MEs. Additionally, concentrations of homocysteine, cysteine and B6 were associated with lower offspring mean methylation, whereas B2 was associated with a higher offspring mean methylation across the 6 MEs.

James et al, found significantly higher mean methylation in Gambian rainy season conceptions across 50 human ME loci[35] in children aged 2 years of age. Furthermore, this study highlighted that biomarker-offspring methylation relationships may be different between the seasons. A significant SoC interaction was found for maternal levels of folate, homocysteine, and choline.

Silver et al, highlighted a further 100 candidate human MEs and again demonstrated higher methylation in offspring conceived in the rainy season[18]. The top SoC region mapped to *VTRNA2-1* gene with riboflavin, methionine and dimethylglycine associated with higher offspring *VTRNA2-1* methylation.

Maternal circulating homocysteine levels have been consistently identified as a negative predictor of ME methylation in offspring. James et al, demonstrated that nutritional supplements reduced plasma homocysteine[36] in a randomly controlled trial in non-pregnant Gambian women. This study confirmed that dietary manipulation could influence metabolic pathways associated with offspring DNA methylation and has led the way for potential targeted interventions in the future. The EMPHASIS study demonstrated that preconception nutritional supplementation extending into pregnancy can influence DNA methylation at loci in offspring aged between 7-9 years[37]. The study examined the effect of separate supplementation trials in two ethnically disparate cohorts (from Gambia and India) on offspring DNA methylation. Despite no discernible DNA methylation differences seen in the Indian cohort, there were 6 differentially methylated positions and an enrichment for MEs and imprinted regions in the nutritional intervention group (consisting of a daily UNIMMAP (United Nations International Multiple Micronutrient Antenatal Preparation)) in The Gambia.

Regions of the genome displaying significant inter-tissue methylation correlation and significant interindividual variations (CoRSIVs) have been shown to be especially sensitive to the SoC compared to control regions or regions displaying tissue specific methylation in Gambian children[29]. A significant proportion of CoRSIVs exhibit metastability and are influenced by nutrition in early pregnancy with significant overlap between known MEs and CoRSIVs[29] and furthermore, Gambia SoC sensitive genomic regions show an enrichment for MEs[16] and CoRSIVs[29]. A study of 233 Gambian 2 year olds demonstrated the predicted methylation maxima for CoRSIVs corresponded with the Gambian rainy season (July-September) and predicted methylation minima corresponded to the dry season (January – April); consistent with previous studies of higher methylation at MEs in rainy season conceptions.

More recently, Silver et al identified 259 SoC sensitive loci that were replicated in two independent Gambia cohorts of children[38]. Some, but not all loci associated with SoC persisted into mid-childhood. These loci were enriched for MEs. Interindividual variation in methylation and sensitivity to SoC appeared to be influenced but not determined by genotype in cis suggesting a possible genotype-environment interaction.

The challenge going forward is to link these nutritionally driven epigenetic alterations with a subsequent postnatal phenotype.

Environmentally driven epigenetic modifications are emerging as a leading candidate mechanism that may contribute to human health and disease. There are multiple lines of evidence, that *POMC* (*Proopiomelanocortin*) and *PAX-8* (*Paired Box 8*) are putative human metastable epialleles [18,20,22,30].

The phenotypic consequences of variation in DNA methylation at nutritionally sensitive regions of *POMC* and *PAX8* forms the basis of this thesis.

1.1.6 *POMC*

POMC is a key mediator of energy balance

POMC is a key component of the melanocortin system[39]; a complex network of systemic signals and neural pathways that regulate food intake and energy balance (Figure 1.4). *POMC* neurons in the ARC of the hypothalamus integrate peripheral signals such as leptin[40], glucose[41] and insulin[42], and regulate energy balance by inducing satiety and increasing energy expenditure[43]. Satiety is mediated via the actions of α- and β-MSH on melanocortin 4 receptors (MC4R) in the paraventricular nucleus (PVN) of the hypothalamus[44]. Perturbations of the melanocortin system can lead to disorders of energy balance such as obesity. For example, individuals with bi-allelic loss of function mutations in *POMC* demonstrate early hyperphagia, severe obesity (due to α-/β-MSH deficiency) and central adrenal insufficiency [45]. It has been demonstrated that heterozygote variant carriers have an increased risk of developing obesity without adrenal insufficiency [46,47] suggesting a gene dosage effect on energy balance.

Chapter 3 provides a comprehensive published literature review on epigenetic regulation of *POMC* and its implications for metabolic health and obesity. The most robust evidence of a link between *POMC* methylation and obesity was reported by Kühnen et al[48]. This study was the first to examine the relationship between *POMC* DNA methylation and obesity in humans. In a case control study comparing 71 obese and 36 normal weight children, they reported a significant difference in average peripheral blood cell (PBC) *POMC* DNA methylation at a VMR overlapping the boundary of intron2/exon3 (average methylation 25% in normal weight individuals vs. 40% obese, $p<0.001$). This

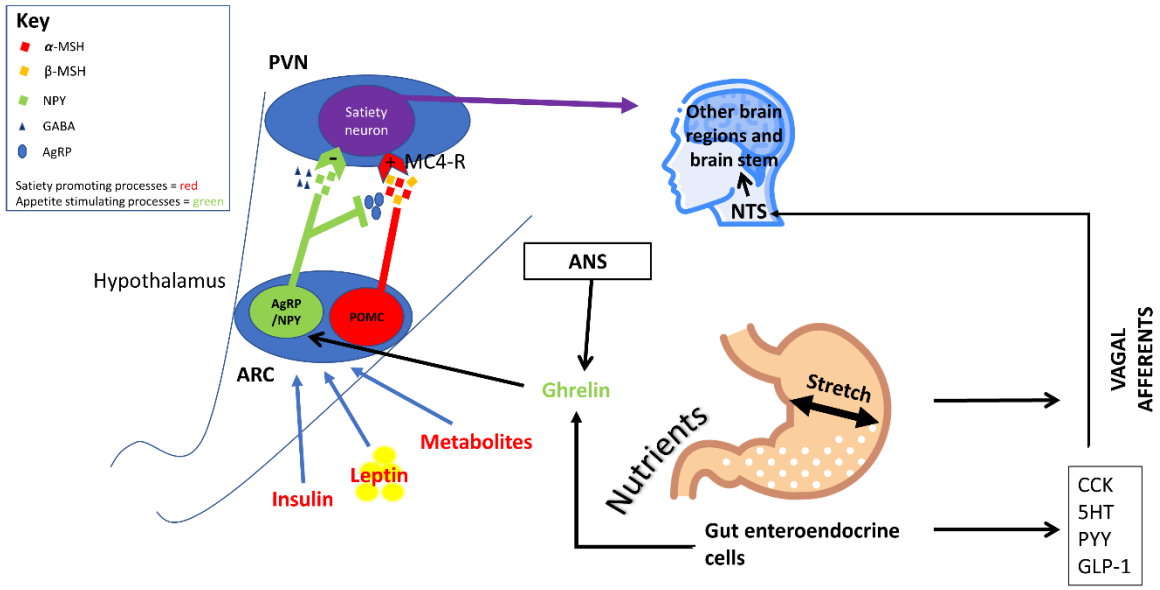


Figure 1.4 The melanocortin system and the control of appetite and satiety. At the level of the hypothalamus, appetite and satiety regulating neurons (POMC, AgRP/NPY) of the arcuate nucleus (ARC) send projections to the paraventricular nucleus (PVN). **The** anorectic POMC expressing neurons are responsive to systemic signals including leptin, insulin, and metabolites (such as glucose). ***α*** and β-MSH (derived from POMC) mediates the satiety signal via the action on MC4R. AgRP/NPY expressing neurons respond to ghrelin (which is predominantly under the control of the autonomic nervous system (ANS)), a hormone released by enteroendocrine cells that acts to increase appetite. AgRP antagonises the action of ***α***-MSH at MC4R, whilst the neurotransmitters NPY and GABA convey an orexigenic signal via PVN neurons. Meal termination (satiation) is brought about via activation of vagal afferents from stomach stretch receptors and nutrient-induced release of enteroendocrine factors (CCK, 5HT, PYY, GLP-1). The vagal afferents send projections to the NTS (nucleus tractus solitarii) to bring about meal termination. Key: POMC; Proopiomelanocortin, AgRP; Agouti-related peptide, NPY; neuropeptide-Y, ***α***-MSH; alpha-melanocyte stimulating hormone, β-MSH; beta-melanocyte stimulating hormone, MC4R; melanocortin 4 receptors, GABA; gamma-aminobutyric acid, NTS; nucleus tractus solitarii, CCK; cholecystokinin, 5HT; 5-hydroxytryptamine, PYY; Peptide YY, GLP-1; glucagon-like peptide 1. Reproduced with permission under Creative Commons, CC BY-NC-ND from Epigenetic regulation of POMC; implications for nutritional programming, obesity and metabolic disease. Toby Candler, P. Kühnen, A.M. Prentice, M. Silver. Frontiers in Neuroendocrinology. July 2019. DOI: 10.1016/j.yfrne.2019.100773

finding was replicated in a second case control study in adults with comparable results[19]. An association between *POMC* hypermethylation at the VMR and individual's BMI was also

demonstrated in MSH neurons, where a 10% increase in methylation was associated with a 2.8kg/m^2 increase in BMI[19]. The hypermethylation at *POMC* VMR decreased histone acetyltransferase P300 binding, leading to reduced expression of PBC *POMC* [19].

POMC is a putative human ME

There is evidence that methylation at the *POMC VMR* appears i) sensitive to maternal diet in early pregnancy, ii) set in early embryonic period with systemic methylation pattern across tissue layers, iii) largely independent of genotype and iv) associated with transposable elements. These features are consistent with MEs and the POMC VMR is thus a putative human ME.

Mother-child paired blood samples from a Gambian cohort demonstrated an association between early pregnancy one-carbon metabolite concentrations in maternal plasma and offspring PBC *POMC* methylation[19]. Specifically, a significant negative correlation for SAH and positive correlations with betaine and the ratio of SAM to SAH at a VMR (intron2/exon3 boundary) of the *POMC* gene. Offspring DNA methylation was also associated with SoC, with lower DNA methylation at the *POMC VMR* in children conceived in the dry season compared to those conceived in the rainy season[19].

POMC methylation appears to be set very early in embryonic development (before separation of the germ layers at gastrulation) as post-mortem samples demonstrated *POMC* DNA methylation was highly correlated across tissues originating from different germ cell layers e.g. brain (ectoderm) and kidney (mesoderm)[19].

Methylation at the *POMC* VMR is thought to be largely independent of genotype, at least in *cis*, with similar methylation patterns across genetically diverse cohorts[19].

POMC methylation is associated with the presence of neighbouring transposable elements. Kühnen et al, observed three Alu elements in intron 2[49] of the *POMC* gene (see Figure 1.5) and proposed that Alu elements drives hypermethylation in this region, drawing a parallel with the IAP retrotransposon in the A^{vy} mouse[50].

Importantly, *POMC* methylation appears stable with age, suggesting that associations with postnatal phenotypes are potentially not driven by reverse causation effects i.e. phenotype affecting methylation [19,51].

There is provisional evidence from animal and human studies to suggest *POMC* epigenetic marks may be transmitted across generations and mediated via the paternal line. Firstly, evidence from animal models on the effect of fetal alcohol exposure on *POMC* epigenetic marks suggests the

potential for transgenerational epigenetic inheritance via the male germline[52,53]. Secondly, evidence from human family trios demonstrates a significant correlation between offspring PBC *POMC* methylation and paternal, but not maternal *POMC* methylation[19]. Therefore, there could be additional contribution to offspring's embryonic methylation pattern from the father in addition to an effect from the intrauterine or maternal environment.

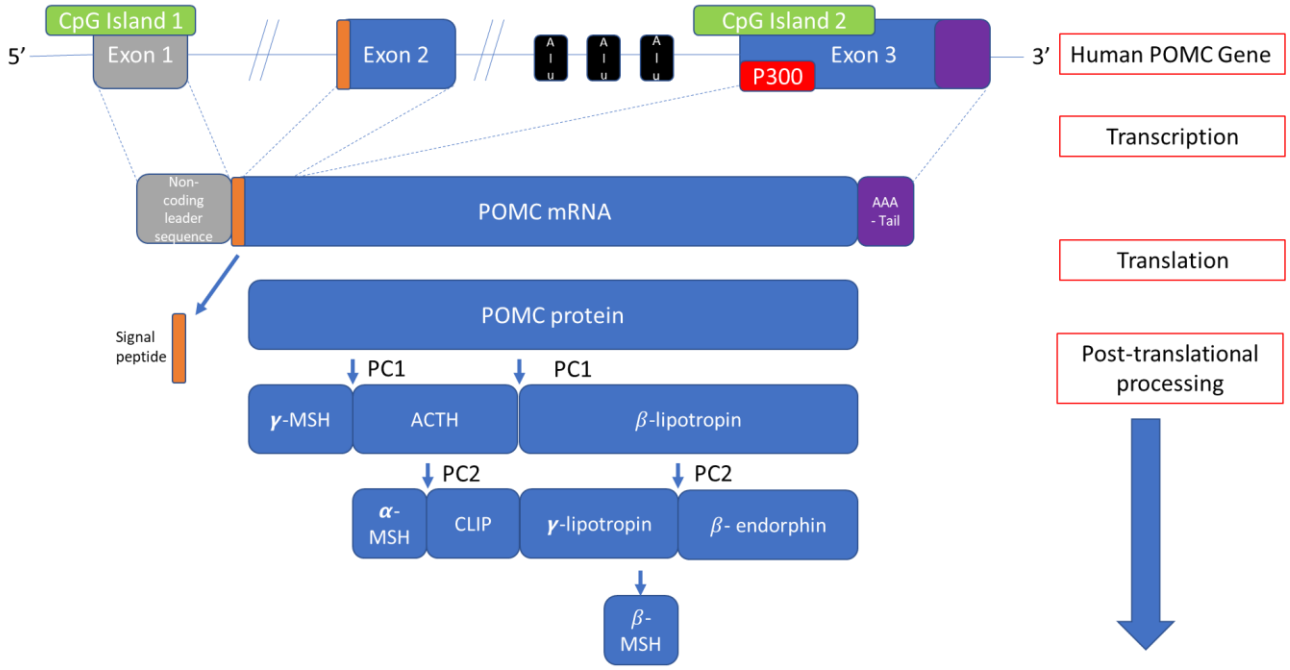


Figure 1.5 Human POMC Gene; transcription, translation, and post-translational processes. The human POMC gene consists of 3 exons and 2 large introns. There are two CpG islands related to the POMC gene; the first in the promoter region and second over the boundary of intron2/exon3. Exon 1 (87 bp) contains no coding sequence but produces a short leader sequence that binds the ribosome at the start of translation. Exon 2 (152bp) gives rise to a small signal peptide and forms the N terminal end of the POMC peptide. Exon 3 (835bp) produces the majority of the POMC peptide as well the signal for the addition of the poly-A tail. Key: CpG; cytosine-guanine dinucleotide, Alu; Alu element, P300; P300 complex binding domain, POMC; Proopiomelanocortin, PC1; Prohormone convertase 1, PC2; Prohormone convertase 2, -MSH; -melanocyte stimulating hormone, ACTH; Adrenocorticotropic hormone, CLIP; corticotropin-like intermediate peptide, AAA-tail; poly-A tail. Reproduced with permission under Creative Commons, CC BY-NC-ND from Epigenetic regulation of POMC; implications for nutritional programming, obesity and metabolic disease. Toby Candler, P. Kühnen, A.M. Prentice, M. Silver. Frontiers in Neuroendocrinology. July 2019. DOI: 10.1016/j.yfrne.2019.100773

1.1.7 PAX-8

***PAX-8* is important for thyroid gland differentiation and function**

Thyroid hormones contribute to a wide range of physiological processes and have important health outcomes related to cognition, bone health and growth, cardiovascular function, and metabolism[54,55].

The hypothalamic-pituitary-thyroid axis controls the production of thyroid hormone[56,57] and is summarised in Figure 1.6. Thyrotropin releasing hormone (TRH), released from the hypothalamus promotes release of thyrotropin (TSH) from the anterior pituitary gland. TSH acts via the TSH receptor on the surface of thyroid follicular cells to promote iodide uptake, thyroid hormone secretion and thyroid gland growth and differentiation. Thyroid hormone production starts with active uptake of iodide (I-) from the circulation by sodium-Iodide symporter on the surface of thyroid follicular cells. Thyroglobulin is a protein synthesised by thyroid follicular cells and has numerus tyrosine residues. Both thyroglobulin and iodide are transported into the thyroid lumen where thyroid peroxidase enzyme acts to oxidise I- to iodine (I2) and binds I2 to tyrosine residues by a process known as organification. Tyrosine residues can be either singularly or doubly iodinated, forming Monoiodotyrosine (MIT) and Diiodotyrosine (DIT) respectively. Coupling of MIT and DIT forms triiodothyronine (T3), and coupling of two DIT residues form thyroxine (T4) by the action of thyroid peroxidase. The iodinated thyroglobulin is brought back into the follicular cell by endocytosis whereby T3 and T4 are released by lysosomal breakdown of thyroglobulin. The thyroid follicular cell transports the hormones into the circulation. The thyroid gland produces more T4 compared to T3 by an order of 4:1 with the majority bound to transport proteins. Extra-thyroidal conversion of T4 to T3 occurs by the action of deiodinases (D1 and D2). The action of thyroid hormone is mediated by the thyroid hormone nuclear receptor which has a higher affinity for T3 than T4. Thyroid hormone acts as a transcriptional regulator by binding to response elements in gene promotors with a wide range of physiological actions. T3 and T4 exert negative feedback on TRH and TSH to main equilibrium in the axis.

Clinical sequelae of severe perturbations of thyroid hormone production are well documented[58], but variability of thyroid function within what is considered the normal range could itself influence the individual's phenotype and disease susceptibility e.g. cardiovascular disease or osteoporosis.

A recent review[55] concluded that higher TSH (thyrotropin)/lower free T4 (free thyroxine) was associated with poorer cardiovascular, metabolic and pregnancy outcomes. Lower TSH/higher free

T4 was associated with increased risk of osteoporosis and fracture[55]. The relationship between thyroid function and weight has been described, with higher levels of TSH associated with increased BMI, and free T4 negatively associated with BMI demonstrated in adults[59,60] and children[61]. Elevated serum TSH in otherwise well individuals is correlated with lipoprotein concentrations[62] and intra-individual variation in free T3 (free tri-iodothyronine)) accounts for approximately 20% of variation in serum HDL cholesterol and apolipoprotein[63], both known risk factors for cardiovascular disease.

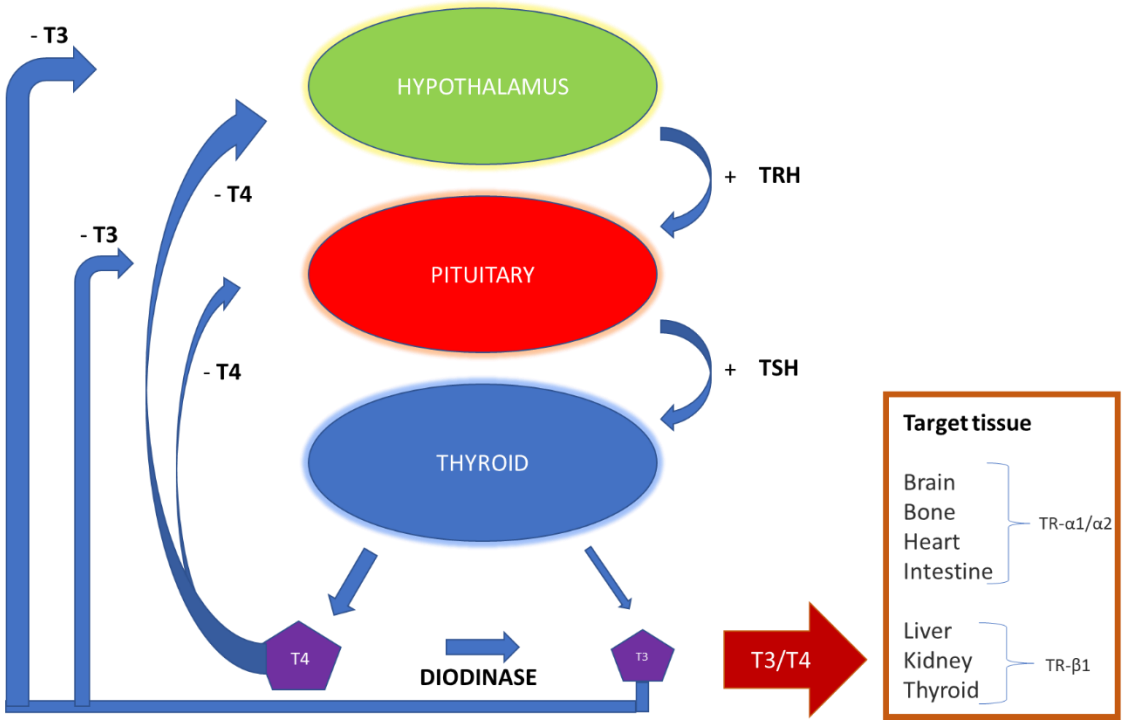


Figure 1.6 The hypothalamic-pituitary-thyroid axis. Key: T3= triiodothyronine, T4=thyroxine, TRH=thyrotropin releasing hormone, TSH= thyrotropin, TR-α= Thyroid receptor alpha, TR-β=Thyroid receptor beta. - = inhibitory signal (negative feedback) + = stimulatory signal.

There is a narrow intra-individual variation in thyroid function in adults[64,65] and children[66] and a genetic 'set point' has been suggested. Each of the genes associated with thyroid function only contribute a small amount to the variability of hormone concentrations, suggesting that other genes and/or epigenetic factors may explain much of the reported heritability[67,68]. The first epigenome-wide association study concerning thyroid function was recently published, and explored DNA methylation in over 1,400 14-17 year old European children[69]. Of loci reaching epigenome-wide significance: two differentially methylated positions (DMPs) were associated with TSH and 6 associated with free T3. There were no DMP associated with free T4. The potential for an epigenetic contribution to thyroid hormone variation is still little explored.

PAX-8 protein is one of four known thyroid transcription factors (TTF) involved in thyroid development and function (others include NKX2-1, FOXE1, HHEX). *PAX-8* and other TFFs contribute to differentiation of the thyroid gland but also act to regulate the expression of proteins needed for thyroid hormone production and storage (e.g. thyroglobulin, thyroid peroxidase). Therefore, TFF are also important in regulating the function of the differentiated thyroid gland[70]. *PAX-8* has been described as the "master regulator" by regulating the activity of other TFFs (HHEX and FOXE1) as well as transcriptional activation of thyroglobulin and thyroid peroxidase (essential for thyroid hormone production).

PAX-8 knockout mice demonstrate thyroid hypoplasia, low birth weight and growth retardation[71]. In humans, CHT (congenital hypothyroidism) can be a consequence of mutations in *PAX-8* (including heterozygote mutations)[727374]. CHT is generally not inherited, and 98% of cases are non-familial[75] with a high discordance rate (92%) in monozygotic twins[76]. This further suggests that epigenetic or unknown genetic mechanisms play a role in thyroid dysgenesis.

A negative correlation between *PAX-8* DNA methylation and gene expression has been reported in colorectal carcinoma[77] and *PAX-8* has been reported to be hypomethylated in human cartilage in those with osteoarthritis[78]. Importantly, significant differences in *PAX-8* methylation have also been described between monozygotic twins discordant for CHT (personal communication, Dr Peter Kuhnen). *PAX-8* gene is a prime candidate to explain variation in thyroid function due to its importance both in thyroid gland development and in regulation of the differentiated thyroid gland.

PAX-8 is a putative human ME

There is evidence that methylation at the *PAX-8* gene is i) sensitive to maternal diet in early pregnancy, ii) set in early embryonic period with systemic methylation pattern across tissue layers consistent with MEs and therefore *PAX-8* gene is a putative human ME.

PAX-8 has been shown to be sensitive to the periconceptional nutritional environment[18,20,30]. In Gambian children, the DNA methylation at *PAX-8* was significantly higher in those conceived in the rainy season[18,30], and mean methylation across 6 MEs (including *PAX8*) was influenced by maternal concentrations of one-carbon metabolites[22]. In Bangladesh, significant hypermethylation at *PAX-8* was reported following gestational famine exposure[20]. Maternal folate supplementation has been associated with differential *PAX-8* methylation in offspring with methylation differences seen into adulthood[79] and maternal preconception micronutrient supplementation was nominally associated with differential methylation at *PAX8* in Gambian children[37].

In studies employing a two-tissue (peripheral blood cell and hair follicle) parallel methylation screen, *PAX-8* demonstrated systemic methylation with significant interindividual differences[18,30]. In cadaveric samples from an adult Vietnamese population, methylation patterns were concordant across liver (endoderm), kidney (mesoderm) and brain (ectoderm) tissues (i.e. all three germ layers) yet demonstrated significant interindividual methylation differences[30]. The findings demonstrate a systemic methylation pattern that is likely established before germ cell separation i.e. early in embryonic development before gastrulation.

1.2 Seasonality

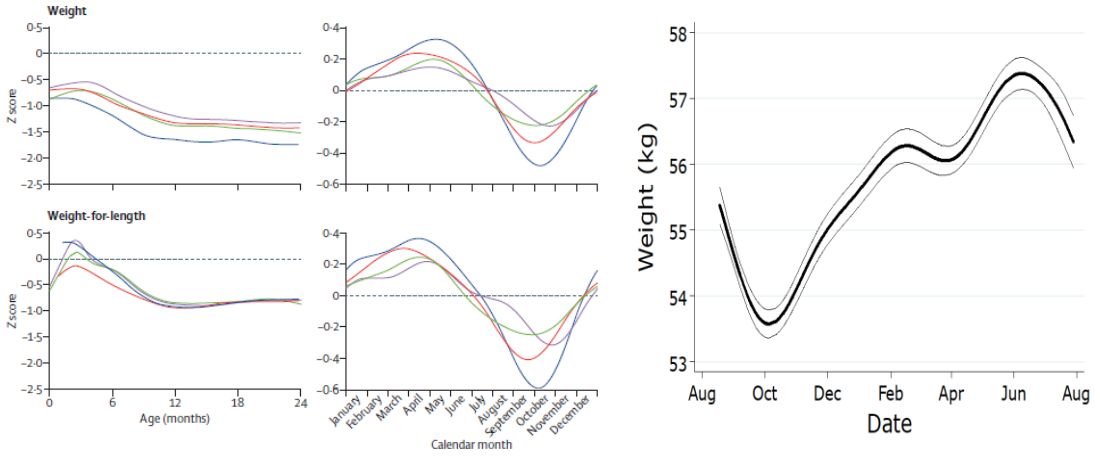
1.2.1 Seasonality in West Kiang, The Gambia

The climate in West Kiang produces a well-established Gambian seasonal 'experiment of nature', whereby a sharply delineated bimodal pattern of rainfall creates profound contrasts in the availability and composition of food, work patterns among adults, and infectious disease and growth patterns in children[34,80,81]. The stark environmental result of the changing seasons is demonstrated in Figure 1.7. A rainy ('hungry') season runs from July-October, coinciding with increased agricultural workload, depleted food supplies and higher prevalence of infectious disease[82,83]. The dry ('harvest') season runs from October-June, with the main harvest occurring between February-April. Thus conceptions, pregnancies and births are randomised to occur against different environmental (especially nutritional) backgrounds. This provides a scientific model able to interrogate the impact of the peri-conceptional environment on offspring methylation at a range of MEs[22] including the *POMC* and *PAX-8* genes.

1.2.2 Seasonality and weight change

The seasons bring significant fluctuations in weight in both women and children (Figure 1.8). Across the year, October sees a nadir of weight in both mothers and children, and the peak of weight is seen in June for mothers and March-April for children. In Gambian women on average weight fluctuates by nearly 4 kg between June and October[34].

Figure 1.7 Aerial photographs of Keneba Field Station, WK, The Gambia demonstrating the change in environment between dry and rainy seasons. Source: Reproduced with permission from Andrew Prentice



a) **b)**

Figure 1.8 a) WHO 2006[84] Weight and weight-for-length z score fluctuations in Gambian children under 2 years of age across the year. Blue line=1976-79, Red line=1980-89, Green line=1990-99, Purple line=2000-12. Source: Reproduced with permission under CC BY licence. Nabwera HM, Fulford AJ, Moore SE, Prentice AM. Growth faltering in rural Gambian children after four decades of interventions: a retrospective cohort study. Lancet Glob Heal. 2017 [81]. **b) Weight fluctuations in non-pregnant adult Gambian women across the year.** Thick line=arithmetic mean. Thin line=95% CI. Source: Reproduced with permission under CC BY licence. Dominguez-Salas P, Moore SE, Cole D, et al. DNA methylation potential: Dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. Am J Clin Nutr. 2013 [34].

1.3 PhD rationale

The focus of this PhD is to explore how the periconceptional environment influences DNA methylation and how these alterations in DNA methylation are related to phenotype. I focussed on two putative human MEs, *POMC* and *PAX8,* that are implicated in hormonal pathways related to energy balance and thyroid development and function respectively. There is evidence that at both genes the methylation state is influenced by the maternal periconceptional environment including season of conception.

I ran two prospective studies referred to as the *POMC* and *PAX8* studies. For both studies, I use the season of conception as a periconceptional exposure to help understand how early embryonic environment may be associated with offspring DNA methylation at metastable regions of *POMC* and *PAX8.* By understanding factors such as intergenerational effects and maternal nutrition associated with *POMC* and *PAX8* methylation I hope to gain insight into how we might focus interventions to improve health outcomes for future generations.

***POMC* study**

Low and middle-income countries (LMICs) are predicted to be the hardest hit by the projected increases in obesity and T2D. Due to poorly resourced health care in LMICs, the case-fatality rates for obesity and T2D are much higher than in affluent nations and constitute a major burden for health budgets both now, and increasingly so in the future.

Body weight and body mass index (BMI) are highly heritable, and yet identified genetic variants so far explain just over 20% of individual body weight variability[85]. It is postulated that environmentally-driven heritable epigenetic modifications may contribute to the development of obesity.

POMC is a critical mediator of energy balance. Studies have shown that increased DNA methylation at a variably methylated region (VMR) at the *POMC* intron2/exon3 boundary and is associated with BMI in children and adults[4827]. This VMR has been independently identified in European and Gambian cohorts. Due to the intra-individual stability of DNA methylation at the *POMC* VMR (i.e. not tissue specific methylation) between peripheral blood cells (PBC) and hypothalamic tissue (where POMC neurons exert their effect), PBC are an appropriate proxy to measure hypothalamic *POMC* VMR methylation[27].

Importantly in a Gambian cohort, methylation at the *POMC* VMR is thought to be modified by a mother's nutritional status around conception[22]. Finally, increased *POMC* methylation is present

directly after birth and is stable longitudinally. Together, these observations position the *POMC* VMR as a metastable epiallele (ME)[17], whose methylation state is established in the very early embryo, is influenced by maternal nutrition, and has the potential to influence later phenotype. Previous study has observed that offspring methylation at the *POMC* VMR is correlated with paternal but not maternal methylation in blood[27].

Obesity generally develops gradually in adults and children due to a relatively small imbalance in energy intake and expenditure[86], and therefore to observe the influence of *POMC* methylation on energy balance in a prospective study in individuals developing obesity would take many years.

In the Gambia (like many low and middle income countries), there are seasonally driven variations in individual's weight and body composition due to changes in food availability and infectious disease burden through the year and the influence of methylation at the *POMC* VMR on these weight changes is unknown. The seasonally driven shift in energy balance provides an exciting opportunistic experimental model to interrogate the effect of *POMC* methylation on weight gain and loss over a year.

To test the central hypothesis that *POMC* hypermethylation (by dampening satiety) will promote more rapid weight gain in the harvest season and protect against weight loss in the hungry season, a year-long prospective study measuring weight and body composition monthly in Gambian women and children was conducted. I further hypothesised that *POMC* methylation would influence energy balance by altering eating behaviour. To test this I developed a controlled appetite test where participants had an assessment of satiety and appetite.

***PAX8* study**

Thyroid hormones contribute to a wide range of physiological processes associated with important health outcomes important for childhood development including cognition, bone health and growth. It has been postulated that variability of thyroid function within 'normal range' could itself influence the individual's phenotype and disease susceptibility e.g. cardiovascular disease, osteoporosis [55].

An individual's thyroid function is thought to be governed by a strong genetic component, with heritability reported between 32-65% (TSH), 32-65% (free T4), and 23-67% (free T3)[67]. The narrow intra-individual variation in thyroid function in children [66] could be explained by a genetic 'set point'. Known genetic variants only contribute a small amount to the variability of hormone concentrations, suggesting that other genes and/or epigenetic factors explain much of reported heritability [6768]. Epigenetic contribution to thyroid function is little explored. *PAX8* is a prime candidate gene to explain variation in thyroid function.

PAX8 and other thyroid transcription factors contribute to differentiation of the thyroid gland and act to regulate the expression of proteins needed for thyroid hormone production and storage [70]. *PAX8* -/- mice demonstrate thyroid hypoplasia, low birth weight and growth retardation [71]. CHT (congenital hypothyroidism) can be caused by mutations in *PAX8* (including in the heterozygous state) [727374]. Ninety-eight percent of CHT cases are non-familial [75] with a high discordance rate (92%) in monozygotic (MZ) twins [76] suggesting that epigenetic or unknown genetic mechanisms play a role in CHT.

A negative correlation between *PAX8* methylation and gene expression has been reported in colorectal carcinoma[77] and *PAX8* is hypomethylated in human cartilage with osteoarthritis[78]. Importantly, significant differences in *PAX8* methylation are reported between MZ twins discordant for CHT (personal communication, Dr Peter Kuhnen).

In The Gambia, significantly higher *PAX8* methylation has been reported in those conceived in the rainy season[1830]. Higher *PAX8* methylation was associated with early gestational famine exposure in Bangladesh[20]. Maternal folate supplementation has been associated with differential *PAX8* methylation in offspring with methylation differences seen into adulthood[79]. Maternal preconception micronutrient supplementation was nominally associated with differential methylation at *PAX8* in Gambian children[37]. Maternal tobacco smoking has also been associated with higher *PAX8* methylation in offspring [87].

To explore the association between *PAX8* methylation and thyroid phenotype a recall by epigenotype study recruiting the top and bottom centile for *PAX8* methylation was performed. At the end of the *POMC* study these selected participants were assessed for thyroid function and size. Furthermore, to understand how changes in free T4 may be associated with phenotype these subjects had an assessment of body composition and bone mineral density.

1.4 PhD aims and objectives

Overall aim: To characterise phenotypic sequelae related to maternal nutrition-sensitive epigenetic signatures in the *POMC* and *PAX-8* genes.

The PhD covers two separate studies that focus on the two genes: *POMC* and *PAX8.* The study titles, research questions, and hypotheses are outlined below.

1.4.1 *POMC* study

In summary, this study combined i) a year-long prospective study whereby weight and body composition was measured monthly in Gambian women and children, ii) a test of satiety around a controlled eating episode and iii) an assessment of nutritional, seasonal, intergenerational, and genetic influences on *POMC* methylation.

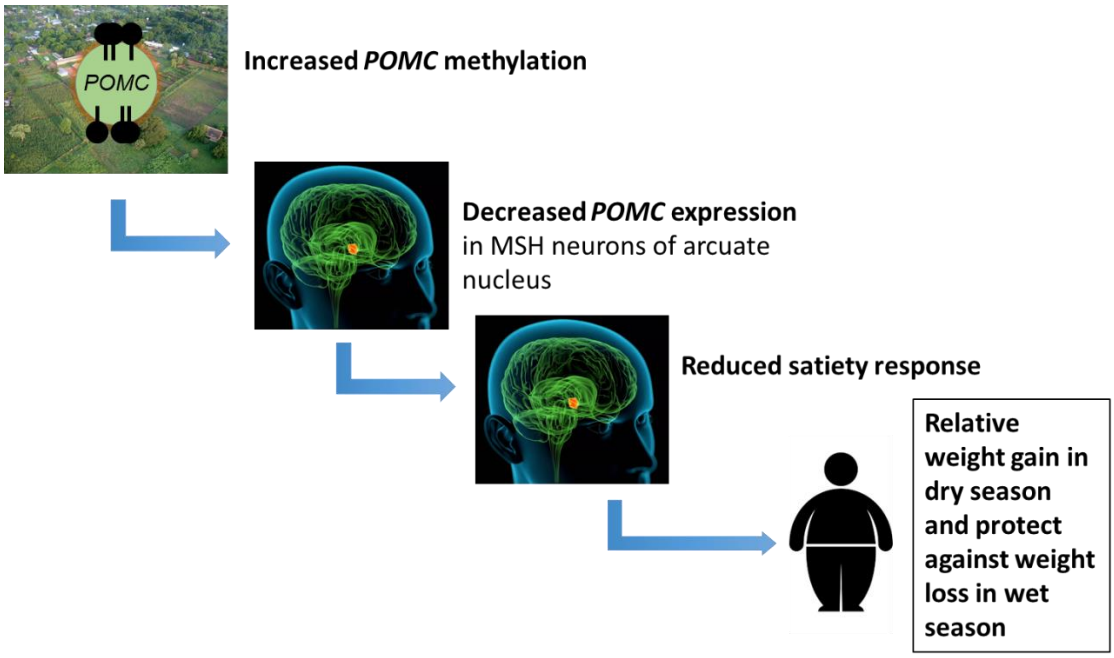


Figure 1.9 Summary figure of POMC study 'Hypothesis 1'

Hypothesis 1

POMC methylation is inversely associated with POMC expression. Lower POMC expression (associated with higher *POMC* methylation) will be associated with a diminished satiety signal and thus will promote more rapid weight gain in the harvest (dry) season and protect against weight loss in the hungry (rainy) season (see Figure 1.9).

Research Question 1 & 2

How does methylation at the *POMC* VMR affect seasonally-driven changes in weight regulation and adiposity?

How does methylation at the *POMC* VMR affect satiety?

Hypothesis 2

Offspring *POMC* VMR methylation is influenced by maternal periconceptional nutritional status and is correlated with paternal, but not maternal, PBC methylation.

Research Question 3

What are the basic nutritional and intergenerational factors associated with methylation at the *POMC* VMR?

1.4.2 *PAX-8* study

In summary (see Figure 1.10), this study selected children from the top and bottom centiles for *PAX8* methylation (recall by epigenotype design) and assessed their thyroid phenotype (thyroid size and hormone levels). A DXA scan was performed to assess adiposity and bone mineral density.

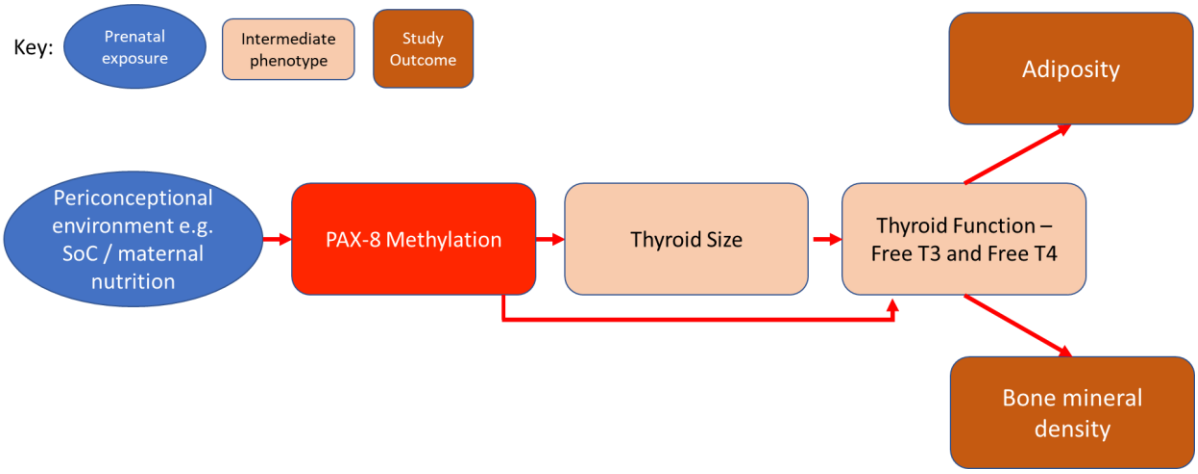


Figure 1.10 Summary figure of PAX8 study hypotheses

Hypothesis 1

Higher *PAX8* methylation (by reducing expression of PAX8) will negatively affect follicular cell development in the thyroid gland resulting in lower thyroid hormone production and a smaller thyroid gland.

Research Question 1

Does differential methylation of *PAX8* gene affect thyroid gland volume and/or thyroid hormone production?

Hypothesis 2

Higher *PAX8* (by lowering thyroid hormone production) will influence adiposity and bone mineral density.

Research Question 2

Does differential methylation of *PAX8* gene (by affecting thyroid function) have an association with adiposity and/or bone mineral density?

Hypothesis 3

PAX8 DNA methylation is set early in embryological development and is influenced by the mother's nutritional status around the time of conception.

Research Question 3

What are the basic nutritional and intergenerational factors associated with methylation at the *PAX8*?

1.5 PhD outline

The thesis is presented as a combination of published papers and distinct chapters in a narrative format. The published papers are included as final accepted submitted manuscripts with the format dictated by the accepting journals. The thesis covers 8 chapters outlined below.

Chapter 1: An introduction to DOHaD, epigenetics and metastable epialleles. An overview of the biological role of POMC and PAX8. A summary of the melanocortin system and thyroid axis. An introduction to the aims, objectives, hypothesis, and research questions. A summary of the thesis outline, my contribution, funding, and PhD timeline.

Chapter 2: An introduction to the research setting, cohorts and datasets used. An outline of the *POMC* and *PAX8* study activity.

Chapter 3: A published manuscript of a comprehensive literature review entitled 'Epigenetic regulation of POMC; implications for nutritional programming, obesity and metabolic disease' published in Frontiers of Neuroendocrinology (PMID: 31344387).

Chapter 4: The methods of the yearlong prospective study examining maternal and child weight are reported. The results of the modelling testing the association between seasonal weight and adiposity changes and *POMC* methylation are presented.

Chapter 5: The methods of the appetite test used in a subset of children and mothers are reported. The results exploring an association between *POMC* methylation and appetite and satiety outcomes are presented.

Chapter 6: Using periconceptional maternal nutritional biomarker data, maternal nutritional influences on *POMC* methylation are explored. SoC effects are also modelled. Using data from family trios from the *POMC* study intergenerational influences on POMC methylation are explored.

Chapter 7: A published manuscript comprising all of the work pertinent to the *PAX8* study entitled 'DNA Methylation at a nutritionally-sensitive region of the *PAX8* gene is associated with thyroid volume and function in Gambian children' published in Science Advances (PMID: 34739318).

Chapter 8: A concluding chapter summarising key findings, exploring limitations and challenges faced. Potential for future directions are also outlined.

1.6 Candidate's involvement

I joined the nutrition theme at MRC The Gambia in November 2017 as the nutrition theme coordinator based at Keneba field station, West Kiang, The Gambia. Professor Andrew Prentice and Dr Matt Silver had written a proposal for some aspects of the *POMC* study to examine how seasonally driven changes in children and mother's weight may be affected by methylation at the *POMC VMR*. The initial grant application for this project was unsuccessful, though a proportion of the proposed work was funded by MRC The Gambia core nutrition theme funding. I was successful in securing an MRC Clinical Research Training Fellowship in November 2018 which supported my

salary and further elements of the *POMC* study. I joined the team with a specific objective to develop the project and set up and run the study from April 2018. I was given freedom to shape additional elements of the study included in this thesis. I had the original idea to perform appetite testing (chapter 5) and to include adiposity measures (chapter 4) into the study design. I was given the freedom to develop the appetite testing and took advice from Prof John Blundell (University of Leeds) when designing the appetite test protocol (chapter 5).

I had the original idea to explore how *PAX8* methylation may influence thyroid function and volume in children for the *PAX8* study. I refined the study design and analysis plan with my supervisors (Dr Silver and Prof Prentice) and Prof Ludgate (University of Cardiff, advisory committee member). I wrote the grant application and secured additional funding for this study through a British Society of Paediatric Endocrinology and Diabetes Research and Innovation award 2018.

For chapter 3 (literature review), I had the idea to write a literature review on epigenetic regulation of *POMC* and its implications for obesity and metabolic health. I set the search criteria, performed the literature review on my own, and wrote the initial manuscript. I had support from Dr Silver regarding structure and the direction of the article's narrative. All the authors reviewed and approved the manuscript.

For chapter 4 I discussed the analysis plan and the statistical modelling with Dr Matt Silver and Professor Tim Cole (University College London). They both provided feedback and suggestions for development on review of initial data. I performed all the analysis myself in R including the Cosinor modelling. I led the prospective study on the ground in Keneba field station. I conducted the Leptin ELISA tests after training from Ebrima Danso and Ebrime Bah (Scientific Officers, Keneba field station). I developed the lab protocol for the POMC PCR and helped perform this in the genomics laboratory in Fajara (under the supervision of Dr Abdul Sessay). We struggled to perform the pyrosequencing on the Pyromark Q48 in Fajara despite over 6 months of trying; largely due to technical issues with the new machine (Pyromark Q48), an issue which has been reported across many centres (e.g. by colleagues from University of Southampton and University of Birmingham). Due to the COVID-19 pandemic and my relocation to the UK, the pyrosequencing was conducted in Berlin Germany in Dr Peter Kühnen laboratory (University of Berlin, advisory committee member).

For chapter 5, I designed the pilot and the refined appetite/satiety test for the study subset. I had some advice from Professor Prentice and Professor Blundell when refining the final study design. I designed the analysis plan with some statistical advice from Prof Cole and Dr Silver. I performed all the analysis myself in R including the linear mixed effects models.

For chapter 6, I performed all the analysis in R including Fourier analysis. I had advice on statistical modelling from Dr Silver and I was kindly provided with the back extrapolated maternal nutritional biomarker data set (from MDEG2) by Dr Phillip James (LSHTM).

For chapter 7 (*PAX8* study paper), I led the study design with Dr Silver and implemented the activity on the ground in Keneba. I organised the thyroid ultrasound training in Keneba and performed all of the ultrasound scans. I organised the blood sample preparation and sample shipment from MRC The Gambia to the University of Cardiff where the thyroid function analysis took place. I performed all of the data analysis for the paper, though I had support with identifying the participants from Noah Kessler (University of Cambridge) and thus I remained blinded to the study group when performing the ultrasound. Noah Kessler also identified the genotype from methyl-seq data and performed the TCGA methylation-expression analysis. I collaborated with Prof Waterland (Baylor College of Medicine, Texas USA)once the initial results were analysed to explore the potential of using the GTEx biobank to examine PAX8 gene expression and methylation relationships. This led to the GTEx results being used in the paper. I worked with Dr Kate Ward to be trained on DXA scanning and scan quality control. She made suggestion for the analysis plan used for the DXA measures. Dr Phil James kindly shared the MDEG2 maternal biomarker data which I analysed further in relation to *PAX8* methylation. Dr Silver supported the project throughout and helped with particular guidance with analysis of the biomarkers and *PAX8* methylation and the genotype analysis. All the authors reviewed and approved the manuscript.

Many aspects of the PhD involved collaboration with other scientists within and outside the MRC The Gambia, nutrition theme. The members of my PhD advisory committee and key collaborators are summarised in Annex 1.1.

1.7 PhD publications and additional output

Published papers:

Epigenetic regulation of POMC; implications for nutritional programming, obesity and metabolic disease. Toby Candler, P. Kühnen, A.M. Prentice, M. Silver. Frontiers in Neuroendocrinology. July 2019. DOI: 10.1016/j.yfrne.2019.100773 (Chapter 3)

DNA Methylation at a nutritionally sensitive region of the *PAX-8* gene is associated with thyroid volume and function in Gambian children. T Candler, NJ Kessler, CJ Gunasekara, KA Ward, P James, E Laritsky, MS Baker, R Dyer, R Elango, D Jeffries, RA Waterland, S Moore, M Ludgate, AM Prentice, MJ Silver. November 2021. doi: 10.1126/sciadv.abj1561 (Chapter 7).

Textbook chapter:

Chapter 8: Twin and family studies on epigenetics and obesity. Twin and Family studies of Epigenetics. Toby Candler, Peter Kuhnen, Andrew Prentice, Matt Silver. Elsevier. ISBN: 9780128209516. Editors: John Hopper, Shuai Li (August 2021).

1.8 PhD timeframe

My PhD was registered with the London School of Hygiene and Tropical Medicine in April 2018. The first 1 year of my PhD I was registered part time as I had an additional scientific management and leadership role as the nutrition theme coordinator for MRC The Gambia. My role included chairing the Nutrition Scientific Administration Meeting convening representatives from data, lab, clinic, operations, logistics teams to plan and share information about research activity. I was an active member of the International Nutrition Group senior management team who focussed more on strategic plans e.g. developing a mission/vision statement. I was seconded to the National Health Service from March 2020 to September 2020 on the request of Severn Post Graduate Education, Health Education England due to the COVID-19 pandemic (see annex 1.2). I wrote much of my thesis whilst in full time work as a consultant paediatric endocrinologist at Bristol Royal Hospital for Children.

1.9 PhD funding

MRC The Gambia Nutrition Theme core funding (MC-A760-5QX00) paid for my salary and travel costs (April 2018 to April 2019), field worker salary costs and genomic investigations (DNA extraction, genotyping, pyrosequencing associated costs). I was awarded an MRC Clinical Research Training Fellowship (£242,088 (30-months duration), MR/S006516/1) in April 2019 which covered

my salary and travel, biochemical laboratory recharges, transport and Keneba clinic and data contributions. I received £10,000 to support the *PAX-8* study from British Society of Paediatric Diabetes and Endocrinology (BSPED) Research and Innovation award in November 2018.

1.10 References

1. Barker DJ. The fetal and infant origins of adult disease. *BMJ*. 1990;301(6761):1111. doi:10.1136/bmj.301.6761.1111

2. Barker DJ, Osmond C, Law CM. The intrauterine and early postnatal origins of cardiovascular disease and chronic bronchitis. *J Epidemiol Community Health*. 1989;43(3):237-240. doi:10.1136/jech.43.3.237

3. Lumey LH, Stein AD, Kahn HS, et al. Cohort Profile: The Dutch Hunger Winter Families Study. *Int J Epidemiol*. 2007;36(6):1196-1204. doi:10.1093/ije/dym126

4. Roseboom T, de Rooij S, Painter R. The Dutch famine and its long-term consequences for adult health. *Early Hum Dev*. 2006;82(8):485-491. doi:10.1016/j.earlhumdev.2006.07.001

5. Roseboom TJ, Painter RC, van Abeelen AFM, Veenendaal MVE, de Rooij SR. Hungry in the womb: what are the consequences? Lessons from the Dutch famine. *Maturitas*. 2011;70(2):141-145. doi:10.1016/j.maturitas.2011.06.017

6. Heindel JJ, Vandenberg LN. Developmental origins of health and disease: a paradigm for understanding disease cause and prevention. *Curr Opin Pediatr*. 2015;27(2):248-253. doi:10.1097/MOP.0000000000000191

7. Gluckman PD, Hanson MA. Living with the past: Evolution, development, and patterns of disease. *Science (80-)*. 2004;305(5691):1733-1736. doi:10.1126/science.1095292

8. Gluckman PD, Hanson MA, Spencer HG. Predictive adaptive responses and human evolution. *Trends Ecol Evol*. 2005;20(10):527-533. doi:10.1016/j.tree.2005.08.001

9. Riggs AD, Porter TN. Overview of Epigenetic Mechanisms. In: *Epigenetic Mechanisms of Gene Regulation*. ; 1996:29-45.

10. Dupont C, Armant DR, Brenner CA. Epigenetics: definition, mechanisms and clinical perspective. *Semin Reprod Med*. 2009;27(5):351-357. doi:10.1055/s-0029-1237423

11. Gibney ER, Nolan CM. Epigenetics and gene expression. *Heredity (Edinb)*. 2010;105(1):4-13.

12. Jaenisch R, Bird A. Epigenetic regulation of gene expression : how the genome integrates intrinsic and environmental signals. 2003;33(march):245-254. doi:10.1038/ng1089

13. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes Dev*. 2011. doi:10.1101/gad.2037511

14. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484-492.

15. Smallwood SA, Kelsey G. De novo DNA methylation: A germ cell perspective. *Trends Genet*. 2012. doi:10.1016/j.tig.2011.09.004

16. Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. *Sci Adv*. 2018;4(7):eaat2624. doi:10.1126/sciadv.aat2624

17. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. *Trends Genet*. 2002;18(7):348-351. doi:10.1016/S0168-9525(02)02709-9

18. Silver MJ, Kessler NJ, Hennig BJ, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol*. 2015;16:118. doi:10.1186/s13059-015-0660-y

19. Kühnen P, Handke D, Waterland RA, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016;24(3):502-509. doi:10.1016/j.cmet.2016.08.001

20. Finer S, Iqbal MS, Lowe R, et al. 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*. 2016;6(11):e011768.

21. Waterland R a, Jirtle RL. Transposable Elements : Targets for Early Nutritional Effects on Epigenetic Gene Regulation. *Mol Cell Biol*. 2003;23(15):5293-5300. doi:10.1128/MCB.23.15.5293

22. Dominguez-Salas P, Moore SE, Baker MS, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat Commun*. 2014;5:3746. doi:10.1038/ncomms4746

23. Alan Harris R, Nagy-Szakal D, Kellermayer R. Human metastable epiallele candidates link to common disorders. *Epigenetics*. 2013;8(2):157-163. doi:10.4161/epi.23438

24. Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG. Maternal methyl supplements increase offspring DNA methylation at Axin fused. *Genesis*. 2006;44(9):401-406. doi:10.1002/dvg.20230

25. Khulan B, Cooper WN, Skinner BM, et al. Periconceptional maternal micronutrient supplementation is associated with widespread gender related changes in the epigenome: a study of a unique resource in the Gambia. *Hum Mol Genet*. 2012;21(9):2086-2101. doi:10.1093/hmg/dds026

26. James P, Sajjadi S, Tomar AS, et al. Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism. *Int J Epidemiol*. 2018;47(6):1910-1937. doi:10.1093/ije/dyy153

27. Kühnen P, Handke D, Waterland R a, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016:1-8. doi:10.1016/j.cmet.2016.08.001

28. Dolinoy DC, Das R, Weidman JR, Jirtle RL. Metastable epialleles, imprinting, and the fetal origins of adult diseases. *Pediatr Res*. 2007;61(5 PART 2 SUPPL.):30-37. doi:10.1203/pdr.0b013e31804575f7

29. Gunasekara CJ, Scott CA, Laritsky E, et al. A genomic atlas of systemic interindividual epigenetic variation in humans. *Genome Biol*. 2019;20(1):105. doi:10.1186/s13059-019-1708-1

30. Waterland RA, Kellermayer R, Laritsky E, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet*. 2010;6(12):1-10. doi:10.1371/journal.pgen.1001252

31. Hoyo C, Daltveit AK, Iversen E, et al. Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics*. 2014;9(8):1120-1130. doi:10.4161/epi.29332

32. James PT, Silver MJ, Prentice AM. Epigenetics, nutrition, and infant health. In: *The Biology of the First 1,000 Days*. ; 2017. doi:10.1201/9781315152950

33. Moore SE, Cole TJ, Poskitt EME, et al. Season of birth predicts mortality in rural Gambia. *Nature*. 1997;388(6641):434. doi:10.1038/41245

34. Dominguez-Salas P, Moore SE, Cole D, et al. DNA methylation potential: Dietary intake and

blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am J Clin Nutr*. 2013;97(6):1217-1227. doi:10.3945/ajcn.112.048462

35. James PT, Dominguez-Salas P, Hennig BJ, Moore SE, Prentice AM, Silver MJ. Maternal One-Carbon Metabolism and Infant DNA Methylation between Contrasting Seasonal Environments: A Case Study from The Gambia. *Curr Dev Nutr*. 2019;3(1). doi:10.1093/cdn/nzy082

36. James PT, Jawla O, Mohammed NI, et al. A novel nutritional supplement to reduce plasma homocysteine in nonpregnant women: A randomised controlled trial in The Gambia. *PLoS Med*. 2019;16(8):e1002870. doi:10.1371/journal.pmed.1002870

37. Saffari A, Shrestha S, Issarapu P, et al. Effect of maternal preconceptional and pregnancy micronutrient interventions on children's DNA methylation: Findings from the EMPHASIS study. *Am J Clin Nutr*. 2020;112(4):1099-1113. doi:10.1093/ajcn/nqaa193

38. Silver MJ, Saffari A, Kessler NJ, et al. Environmentally sensitive hotspots in the methylome of the early human embryo. *Elife*. 2022;11:e72031. doi:10.7554/eLife.72031

39. Mountjoy KG. Pro-Opiomelanocortin (POMC) Neurones, POMC-Derived Peptides, Melanocortin Receptors and Obesity: How Understanding of this System has Changed Over the Last Decade. *J Neuroendocrinol*. 2015;27(6):406-418. doi:10.1111/jne.12285

40. Balthasar N, Coppari R, McMinn J, et al. Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. *Neuron*. 2004;42(6):983-991. doi:10.1016/j.neuron.2004.06.004

41. Parton LE, Ye CP, Coppari R, et al. Glucose sensing by POMC neurons regulates glucose homeostasis and is impaired in obesity. *Nature*. 2007;449(7159):228-232. doi:10.1038/nature06098

42. Qiu J, Zhang C, Borgquist A, et al. Insulin excites anorexigenic proopiomelanocortin neurons via activation of canonical transient receptor potential channels. *Cell Metab*. 2014;19(4):682-693. doi:10.1016/j.cmet.2014.03.004

43. Toda C, Santoro A, Kim JD, Diano S. POMC Neurons: From Birth to Death. *Annu Rev Physiol*. 2017;79(1):209-236. doi:10.1146/annurev-physiol-022516-034110

44. Landeghem F Van, Deimling A Von, Escher F, et al. A role for b -melanocyte-stimulating hormone in human body-weight regulation. 2006;(February):141-146. doi:10.1016/j.cmet.2006.01.007

45. Farooqi IS, Keogh JM, Yeo GSH, Lank EJ, Cheetham T, O'Rahilly S. Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene. *N Engl J Med*. 2003;348(12):1085-1095. doi:10.1056/NEJMoa022050

46. Farooqi IS, Drop S, Clements A, et al. Heterozygosity for a POMC-null mutation and increased obesity risk in humans. *Diabetes*. 2006;55(9):2549-2553. doi:10.2337/db06-0214

47. Krude H, Biebermann H, Schnabel D, et al. Obesity Due to Proopiomelanocortin Deficiency: Three New Cases and Treatment Trials with Thyroid Hormone and ACTH4-10. *J Clin Endocrinol Metab*. 2003;88(10):4633-4640. doi:10.1210/jc.2003-030502

48. Kuehnen P, Mischke M, Wiegand S, et al. An alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet*. 2012;8(3):1-12. doi:10.1371/journal.pgen.1002543

49. Kuehnen P, Krude H. Alu elements and human common diseases like obesity. *Mob Genet Elements*. 2012;2:4:197-201.

50. Morgan HD, Sutherland HGE, Martin DIK, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet*. 1999;23(3):314-318. doi:10.1038/15490

51. Yoo JY, Lee S, Lee HA, et al. Can proopiomelanocortin methylation be used as an early predictor of metabolic syndrome? *Diabetes Care*. 2014;37(3):734-739. doi:10.2337/dc13-1012

52. Govorko D, Bekdash RA, Zhang C, Sarkar DK. Male germline transmits fetal alcohol adverse effect on hypothalamic proopiomelanocortin gene across generations. *Biol Psychiatry*. 2012;72(5):378-388. doi:10.1016/j.biopsych.2012.04.006

53. Mead EA, Sarkar DK. Fetal alcohol spectrum disorders and their transmission through genetic and epigenetic mechanisms. *Front Genet*. 2014;5(JUN):1-10. doi:10.3389/fgene.2014.00154

54. Boelaert K, Franklyn JA. Thyroid hormone in health and disease. *J Endocrinol*. 2005;187(1):1-15. doi:10.1677/joe.1.06131

55. Taylor PN, Razvi S, Pearce SH, Dayan CM. Clinical review: A review of the clinical consequences of variation in thyroid function within the reference range. *J Clin Endocrinol Metab*. 2013;98(9):3562-3571. doi:10.1210/jc.2013-1315

56. Pleasure JR, Pleasure D, Pleasure SJ. 133 - Trophic Factor, Nutritional, and Hormonal Regulation of Brain Development. In: Polin RA, Abman SH, Rowitch DH, Benitz WE, Fox

WWBT-F and NP (Fifth E, eds. Elsevier; 2017:1326-1333.e3. doi:https://doi.org/10.1016/B978-0-323-35214-7.00133-5

57. Ortiga-Carvalho TM, Chiamolera MI, Pazos-Moura CC, Wondisford FE. Hypothalamus-Pituitary-Thyroid Axis. *Compr Physiol*. 2016;6(3):1387-1428. doi:10.1002/cphy.c150027

58. Hanley P, Lord K, Bauer AJ. Thyroid Disorders in Children and Adolescents: A Review. *JAMA Pediatr*. 2016;170(10):1008-1019. doi:10.1001/jamapediatrics.2016.0486

59. Knudsen N, Laurberg P, Rasmussen LB, et al. Small differences in thyroid function may be important for body mass index and the occurrence of obesity in the population. *J Clin Endocrinol Metab*. 2005;90(7):4019-4024. doi:10.1210/jc.2004-2225

60. Mosso L, Mart??nez A, Rojas MP, et al. Early pregnancy thyroid hormone reference ranges in Chilean women: the influence of body mass index. *Clin Endocrinol (Oxf)*. 2016;85(6):942-948. doi:10.1111/cen.13127

61. Taylor PN, Richmond R, Davies N, et al. Paradoxical Relationship Between Body Mass Index and Thyroid Hormone Levels: A Study Using Mendelian Randomization. *J Clin Endocrinol Metab*. 2015;101(2):730-738. doi:10.1210/jc.2015-3505

62. Duntas LH. Thyroid Disease and Lipids. *Thyroid*. 2002;12(4):287-293. doi:10.1089/10507250252949405

63. Comuzzie AG, Blangero J, Mahaney MC, et al. Triiodothyronine exerts a major pleiotropic effect on reverse cholesterol transport phenotypes. *Arterioscler Thromb Vasc Biol*. 1996;16(2):289-293. doi:10.1161/01.ATV.16.2.289

64. Andersen S. Narrow Individual Variations in Serum T4 and T3 in Normal Subjects: A Clue to the Understanding of Subclinical Thyroid Disease. *J Clin Endocrinol Metab*. 2002;87(3):1068-1072. doi:10.1210/jc.87.3.1068

65. Ankrah-Tetteh T, Wijeratne S, Swaminathan R. Intraindividual variation in serum thyroid hormones, parathyroid hormone and insulin-like growth factor-1. *Ann Clin Biochem*. 2008;45(2):167-169. doi:10.1258/acb.2007.007103

66. Oladipo O, Nenninger DA, Parvin CA, Dietzen DJ. Intraindividual variability of thyroid function tests in a pediatric population. *Clin Chim Acta*. 2010;411(15-16):1143-1145. doi:10.1016/J.CCA.2010.03.030

67. Panicker V. Genetics of thyroid function and disease. *Clin Biochem Rev*. 2011;32(4):165-175.

68. Porcu E, Medici M, Pistis G, et al. A Meta-Analysis of Thyroid-Related Traits Reveals Novel Loci and Gender-Specific Differences in the Regulation of Thyroid Function. *PLoS Genet*. 2013;9(2). doi:10.1371/journal.pgen.1003266

69. Lafontaine N, Campbell PJ, Castillo-Fernandez JE, et al. Epigenome-Wide Association Study of Thyroid Function Traits Identifies Novel Associations of fT3 With KLF9 and DOT1L. *J Clin Endocrinol Metab*. January 2021. doi:10.1210/clinem/dgaa975

70. Ruiz-Llorente S, de Pau EC, Sastre-Perona A, et al. Genome-wide analysis of Pax8 binding provides new insights into thyroid functions. *BMC Genomics*. 2012;13(1):147. doi:10.1186/1471-2164-13-147

71. Mansouri A, Chowdhury K, Gruss P. Follicular cells of the thyroid gland require Pax8 gene function. *Nat Genet*. 1998;19(1):87-90. doi:10.1038/ng0598-87

72. Macchia PE, Lapi P, Krude H, et al. PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis. *Nat Genet*. 1998;19(1):83-86. doi:10.1038/ng0598-83

73. Komatsu M, Takahashi T, Takahashi I, Nakamura M, Takahashi I, Takada G. Thyroid dysgenesis caused by PAX8 mutation: The hypermutability with CpG dinucleotides at codon 31. *J Pediatr*. 2001;139(4):597-599. doi:10.1067/mpd.2001.117071

74. Montanelli L, Tonacchera M. Genetics and phenomics of hypothyroidism and thyroid dys- and agenesis due to PAX8 and TTF1 mutations. *Mol Cell Endocrinol*. 2010;322(1-2):64-71. doi:10.1016/j.mce.2010.03.009

75. Castanet M, Lyonnet S, Bonaiti-Pellie C, Polak M, Czernichow P, Leger J. Familial forms of thyroid dysgenesis among infants with congenital hypothyroidism. *N Engl J Med*. 2000;343(6):441-442. doi:10.1056/NEJM200008103430614

76. Perry R, Heinrichs C, Bourdoux P, et al. Discordance of monozygotic twins for thyroid dysgenesis: implications for screening and for molecular pathophysiology. *J Clin Endocrinol Metab*. 2002;87(9):4072-4077. doi:10.1210/jc.2001-011995

77. Center BITGDA. Correlation between mRNA expression and DNA methylation. 2013. doi:doi:10.7908/C1PR7T9G

78. Alvarez-Garcia O, Fisch KM, Wineinger NE, et al. Increased DNA Methylation and Reduced Expression of Transcription Factors in Human Osteoarthritis Cartilage. *Arthritis Rheumatol*. 2016;68(8):1876-1886. doi:10.1002/art.39643

79. Richmond RC, Sharp GC, Herbert G, et al. 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*. 2018;47(3):928-937. doi:10.1093/ije/dyy032

80. Moore SE. Early-Life Nutritional Programming of Health and Disease in the Gambia. *Ann Nutr Metab*. 2017;70(3):179-183. doi:10.1159/000456555

81. Nabwera HM, Fulford AJ, Moore SE, Prentice AM. Growth faltering in rural Gambian children after four decades of interventions: a retrospective cohort study. *Lancet Glob Heal*. 2017;5(2):e208-e216. doi:10.1016/S2214-109X(16)30355-2

82. Prentice A, Whitehead R, Roberts S, Paul A. Long-term energy balance in child-bearing Gambian women. *Am J Clin Nutr*. 1981;34(12):2790-2799.

83. Moore SE, Cole TJ, Collinson AC, Poskitt EM, McGregor IA, Prentice AM. Prenatal or early postnatal events predict infectious deaths in young adulthood in rural Africa. *Int J Epidemiol*. 1999;28(6):1088-1095.

84. WHO GMGRS. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. *World Heal Organ*. 2006. doi:10.4067/S0370-41062009000400012

85. Locke A, Kahali B, Berndt S, Justice A, Pers T. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):197-206. doi:10.1038/nature14177.Genetic

86. Hall KD, Sacks G, Chandramohan D, et al. Quantification of the effect of energy imbalance on bodyweight. *Lancet*. 2011. doi:10.1016/S0140-6736(11)60812-X

87. Joglekar R, Grenier C, Hoyo C, Hoffman K, Murphy SK. Maternal tobacco smoke exposure is associated with increased DNA methylation at human metastable epialleles in infant cord blood. *Environ Epigenetics*. 2022;8(1):dvac005. doi:10.1093/eep/dvac005

Chapter 2 Research setting, cohorts used and POMC and PAX8 study overview

Summary of the chapter

In this chapter, I describe the research setting of The Medical Research Council The Gambia at The London School of Tropical Medicine (MRCG at LSHTM) Keneba Field Station. I describe the characteristics of the cohorts used and give an overview of the *POMC* and *PAX8* studies.

2.1 Research setting

The MRC The Gambia at LSHTM field station in Keneba, West Kiang (WK) has been operational since 1950 (see Figure 1). Research activities cover 36 villages with a collective population of ~14,000 people who are mostly subsistence farmers and identify as being from the Mandinka tribe[1,2]. Longitudinal growth studies have been studying four 'core' villages in WK for over four decades. The field station has long supported nutrition research and has accrued a cadre of staff expertise and equipment necessary for complex nutrition and epigenetic focussed studies. For example, Keneba Field station houses two Dual Energy X-Ray Absorptiometry (DXA) scanners, field workers specifically experienced in dietary assessment and a laboratory able to support DNA extraction and sample processing.

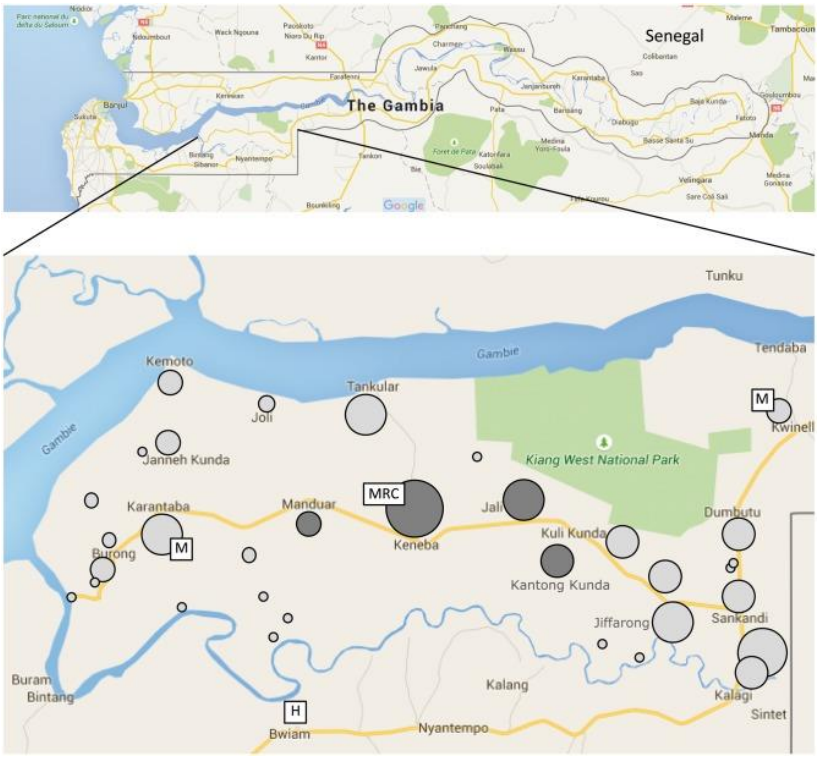


Figure 2.1 Location of MRC The Gambia Keneba Field Station with surrounding villages in WK, The Gambia. Source: Reproduced with permission. Hennig BJ, Unger SA, Dondeh BL, Hassan J, Hawkesworth S, Jarjou L, Jones KS, Moore SE, Nabwera HM, Ngum M, Prentice A, Sonko B, Prentice AM, Fulford AJ. Cohort Profile: The Kiang West Longitudinal Population Study (KWLPS)-a platform for integrated research and health care provision in rural Gambia. Int J Epidemiol. 2017[3]

2.2 A background to the cohorts used in this thesis

Children from the Early Nutrition and Immune Development (ENID) trial and their mothers formed the participants for the *POMC* and *PAX-8* studies. Maternal data obtained in early pregnancy from mothers in the ENID trial and periconceptional biomarker data from a subset of these mothers enrolled in the Methyl Donors and Epigenetics 2 (MDEG2) study were used to explore nutritional and seasonal predictors of *POMC* and *PAX8* methylation described in later chapters.

2.2.1 ENID Trial

The ENID trial was a WK-wide pregnancy and infant supplementation trial[4]. The study started in January 2010 with the final child born in February 2014. In summary, women were recruited in early pregnancy (10-20 weeks) and randomised to receive either i) Iron-Folate (FeFol, standard care), ii) Multiple Micronutrient (MMN) supplement, iii) Energy, protein, and lipid supplement with FeFol, or iv) Energy, protein, and lipid supplement with MMN, for the remainder of their pregnancy. From 6 to 12 months of age, infants are further randomized to a lipid-based nutritional supplement, with or without additional MMN. A total of 875 women were randomised in pregnancy to one of the four study arms and 724 participants completed the first year of follow-up[5]. At the time of the *POMC* and *PAX8* studies, these children were in mid-childhood (5-8 years).

2.2.2 MDEG2 study

The MDEG2 study was established to examine an extended panel of maternal nutritional biomarkers and their influence on offspring DNA methylation. The MDEG2 participants were a subset of 350 ENID trial mother-children dyads and were previously selected for nutritional biomarker analysis on the stored samples close to the time of conception[6]. These women were selected from the ENID recruits to give an even distribution by month of booking and by the earliest gestational age i.e. nearest to time of conception. These data were previously analysed by Dr Phil James for his PhD thesis. The maternal biomarker dataset was provided by Dr James and all biomarkers were preadjusted for gestational age, maternal BMI, maternal age, inflammation (AGP) and then back

extrapolated to date of conception using previously described methods[7]. These data are used to assess maternal nutritional influences on both *POMC* and *PAX8* methylation.

2.3 POMC study overview

A year-long prospective study ran from April 2018 to April/May 2019 with the aim of characterising weight, and adiposity changes across the year in ENID-recruited mothers and children in WK. Participants were scheduled to undergo monthly anthropometry for weight, height, mid upper arm circumference (MUAC), skin fold thickness and bioimpedance for 13 months: across one full cycle of harvest-hungry-harvest seasons. A summary of study activity is presented in Figure 2.2 below.

Potential participants for the study needed to satisfy the inclusion and exclusion criteria.

Inclusion criteria:

- Enrolled in ENID trial (children) or parent of an enrolled child (adults).
- Resides predominantly in WK.

Exclusion criteria:

- Unwell on day of study participation.
- Pregnancy (at any time).

A total of 493 children were recruited from a potential 572 from the ENID trial. A total of 513 mothers were recruited from a potential of 691 mothers from the ENID trial. A subgroup of these mothers (n=118) and children (n=118) were recruited into a study subset (further details below). The study subset was recruited from the villages of Keneba, Jali, Kantong Kunda, Manduar and Tankular representing the 5 villages the closest distance to the field station.

2.3.1 Study timepoints and summary of activity

Study baseline (all participants): Participants attended Keneba field station from 16th April until 15th May 2018 (start of Ramadan). Five ml of fasted blood was taken. All children were measured in triplicate for weight, height, MUAC, skin fold thickness (MRCG@LSTHM SOP:NUTP.SOP.2009) and

had a measurement of bioimpedance using TANITA BC-418 MA body composition analyser (MRCG@LSHTM SOP:NUTP.SOP.2018).

Monthly Field visits (all participants): All participants had monthly (12 consecutive months from May 16th onwards) scheduled field visits for anthropometry including weight, height, MUAC, skin fold thickness and a measurement of bioimpedance. Two field teams visited participants in a fixed order to standardise the period between each measurement. 'Mopping' days were included to attempt to visit participants again if they were unavailable during a planned visit.

During each encounter, the mother was asked if she was pregnant and offered antenatal care accordingly. If pregnant, she would exit the study.

Subset activity: Three time-points were termed **baseline**, **midline** and **endline**. The timing of peaks of season was an estimate of peaks of weight loss and gain informed by previous data in both women and children (Figure 1.8). Ideally, the preferred peaks of season would be May and October. However, the study started in Mid-April to avoid Ramadan (May 16th 2018) which would have prevented blood draws and appetite testing.

At these time points, participants were scheduled a whole body DXA (Dual energy X-ray Absorptiometry) scan (Lunar Prodigy).

A measurement of appetite and satiety was taken at baseline/midline/endline. A further 5 ml fasted blood draw was taken at midline and endline for leptin (biochemical assessment of adiposity).

2.3.2 Ethics approval

Ethical approvals for the ENID trial and for this study were given by The Gambia Government/MRC Joint Ethics Committee (SCC1126v2, SCC1640, L2015.51v2, SCC1588) and LSHTM ethics committee (16439). Consent was gained by signature or thumb print from mothers for their own participation and that of their child. All data were anonymized before analysis. The ethics approvals for this study are found in annex 2.1.

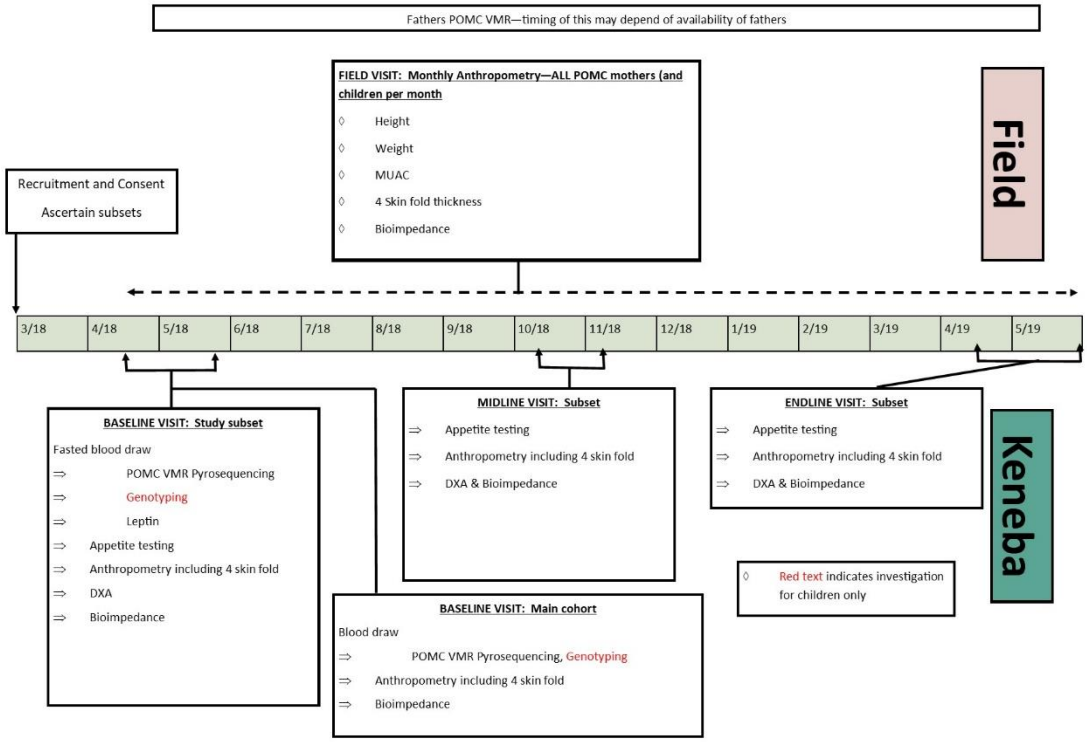


Figure 2.2 Summary of POMC prospective study of children and mothers.

2.4 PAX8 study overview

The study utilised a "recall by epigenotype" design whereby participants were selected from n=493 ENID study children enrolled in *POMC*, by taking the top ('hypermethylated') and bottom ('hypomethylated') quantiles for DNA methylation in the *PAX-8* ME region of interest (from stored DNA at age 2 years). Noah Kessler (collaborating bioinformatician) provided 125 potential participants for recruitment (n=64 low methylation group, n=61 high methylation group) with 118 (n=60 low methylation group, n=58 high methylation group) consenting to study activity. Potential participants for the study needed to satisfy the inclusion and exclusion criteria.

Inclusion criteria:

- Enrolled in *POMC* study.
- Identified in bottom or top quantiles for *PAX-8* methylation from ENID 2-year samples.

Exclusion criteria:

- Lost to follow up or withdrew from *POMC* study.
- Permanently moved from Kiang West.

2.4.1 Study timepoints and summary of activity

The PAX8 study activity coincided with *POMC* endline (see Figure 2.3). Participants attended Keneba field station where they had a 5ml morning fasted blood sample for thyroid function (thyroid stimulating hormone or thyrotropin (TSH), free thyroxine (FT4), triiodothyronine (FT3), thyroglobulin (Tg)), a urine sample for a measurement of urinary iodine, a thyroid ultrasound scan to assess thyroid volume, anthropometry measured (as for POMC study) and a whole body DXA scan.

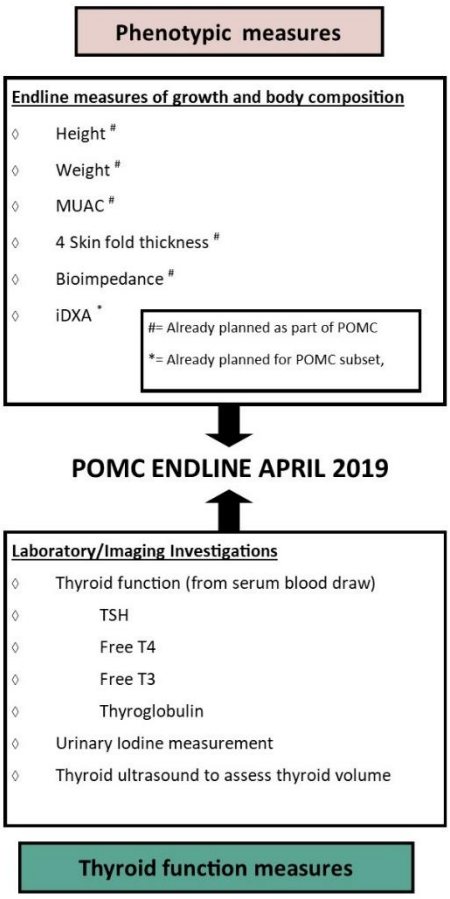


Figure 2.3. Summary of PAX8 study activity. Key: thyroid stimulating hormone or thyrotropin (TSH), free thyroxine (FT4), triiodothyronine (FT3), thyroglobulin (Tg), iDXA (Dual Energy X-Ray Absorptiometry), MUAC (mid upper arm circumference)

2.4.2 Ethics approval

Ethical approvals for the ENID trial and for this study were given by The Gambia Government/MRC Joint Ethics Committee (SCC1126v2, SCC1640, L2015.51v2, SCC1640) and LSHTM ethics committee (16280). Consent was gained by signature or thumb print from mothers for their own participation and that of their child. All data were anonymized before analysis. The ethics approvals for this study are found in annex 2.2.

2.5 References

1. Hennig BJ, Unger SA, Dondeh BL, et al. Cohort Profile: The Kiang West Longitudinal Population Study (KWLPS)-a platform for integrated research and health care provision in rural Gambia. *Int J Epidemiol*. 2017;46(2):e13. doi:10.1093/ije/dyv206

2. Rayco-Solon P, Moore SE, Fulford AJ, Prentice AM. Fifty-year mortality trends in three rural African villages. *Trop Med Int Health*. 2004;9(11):1151-1160. doi:10.1111/j.1365-3156.2004.01325.x

3. Hennig BJ, Unger SA, Dondeh BL, et al. Cohort Profile: The Kiang West Longitudinal Population Study (KWLPS)-a platform for integrated research and health care provision in rural Gambia. *Int J Epidemiol*. 2017;46(2):e13-e13. doi:10.1093/ije/dyv206

4. Moore SE, Fulford AJC, Darboe MK, Jobarteh ML, Jarjou LM, Prentice AM. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: The ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth*. 2012. doi:10.1186/1471-2393-12-107

5. Moore SE, Fulford AJC, Sosseh F, Nshe P, Darboe MK, Prentice AM. Thymic size is increased by infancy, but not pregnancy, nutritional supplementation in rural Gambian children: A randomized clinical trial. *BMC Med*. 2019. doi:10.1186/s12916-019-1264-2

6. James PT, Jawla O, Mohammed NI, et al. A novel nutritional supplement to reduce plasma homocysteine in nonpregnant women: A randomised controlled trial in The Gambia. *PLoS Med*. 2019;16(8):e1002870. doi:10.1371/journal.pmed.1002870

7. Dominguez-Salas P, Moore SE, Baker MS, et al. Maternal nutrition at conception modulates

DNA methylation of human metastable epialleles. *Nat Commun*. 2014;5:3746. doi:10.1038/ncomms4746

Chapter 3 Literature review: Epigenetic regulation of POMC, implications for nutritional programming, obesity, and metabolic disease

Summary of the chapter

In this chapter, I present a systematic narrative review summarising the evidence, in both animals and humans, of epigenetic regulation of *POMC* relevant to nutritional programming, energy balance, obesity and metabolic outcomes. The review provides evidence that *POMC* is sensitive to nutritional programming and is associated with a wide range of weight-related and metabolic outcomes. This chapter is presented as final accepted manuscript for in Frontiers of Neuroendocrinology (PMID: 31344387). The open access typeset version can be found at DOI: 10.1016/j.yfrne.2019.100773.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1801141	**Title**	Dr
First Name(s)	Toby		
Surname/Family Name	Candler		
Thesis Title	Assessing the influence of maternal nutrition-sensitive epigenetic signatures in the POMC and PAX-8 genes on health-related outcomes in offspring		
Primary Supervisor	Dr Matt Silver		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Frontiers of Neuroendocrinology		
When was the work published?	2019		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	n/a		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I had the idea to write a literature review on epigenetic regulation of POMC and its implications for obesity and metabolic health. I set the search criteria, performed the literature review on my own, and wrote the initial manuscript. I had support from Dr Silver regarding structure and the direction of the article's narrative. All the authors reviewed and approved the manuscript.

SECTION E

Student Signature	[redacted]
Date	10/07/2022

Supervisor Signature	[redacted]
Date	10/07/2022

Epigenetic regulation of *POMC*; implications for nutritional programming, obesity and metabolic disease

T Candler[1], P Kühnen[2], AM Prentice[1], M Silver[1]

1. MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, London, UK
2. Institute for Experimental Pediatric Endocrinology, Charité – Universitätsmedizin Berlin. Corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health; Berlin, Germany

Review article: For consideration by Frontiers in Neuroendocrinology

Corresponding author: Dr Toby Candler, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK, tcandler@mrc.gm

Declaration of interest: None

Abstract: 140 words

Main text Word count: 10,998 words

References: 131

Tables: 3

Figures: 3

Abstract

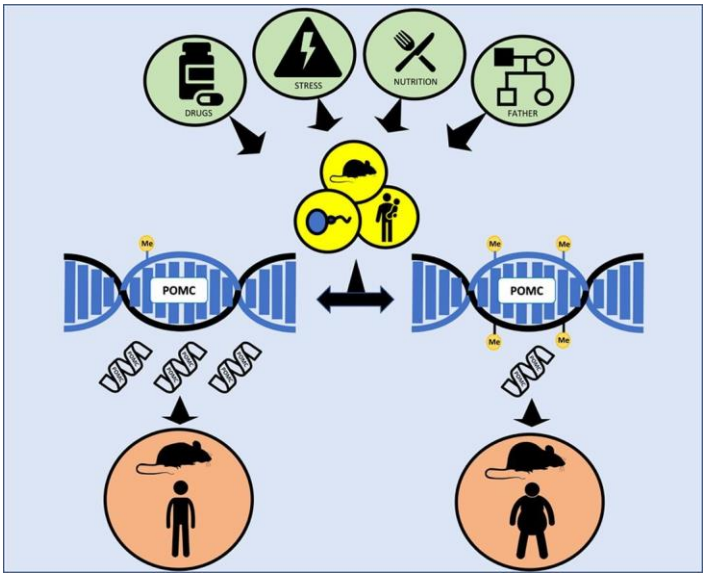
Proopiomelanocortin (*POMC*) is a key mediator of satiety. Epigenetic marks such as DNA methylation may modulate *POMC* expression and provide a biological link between early life exposures and later phenotype. Animal studies suggest epigenetic marks at *POMC* are influenced by maternal energy excess and restriction, prenatal stress and Triclosan exposure. Postnatal factors including energy excess, folate, vitamin A, conjugated linoleic acid and leptin may also affect *POMC* methylation. Recent human studies suggest *POMC* DNA methylation is influenced by maternal nutrition in early pregnancy and associated with childhood and adult obesity. Studies in children propose a link between *POMC* DNA methylation and elevated lipids and insulin, independent of body habitus. This review brings together evidence from animal and human studies and suggests that *POMC* is sensitive to nutritional programming and is associated with a wide range of weight-related and metabolic outcomes.

Keywords

Epigenetics, DNA Methylation, *POMC*, Obesity, Nutrition, DOHaD, Glucose, Insulin, Lipids, Transgenerational

Highlights

- *POMC* is a key mediator of satiety and gene mutations are associated with obesity
- Nutritional and environmental factors influence *POMC* epigenetic marks
- *POMC* DNA methylation is influenced by maternal nutrition in early pregnancy in humans
- Hypermethylation of a region of the *POMC* gene is associated with obesity in humans
- *POMC* hypermethylation is associated with higher fasting lipids and insulin in children



Abbreviations

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ACTH	adrenocorticotropic hormone
AgRP	agouti-related peptide
AN	anorexia nervosa
ARC	arcuate nucleus of the hypothalamus
A^{vy}	Agouti viable yellow
$Axin^{Fu}$	Axin-fused
BEC	buccal epithelial cells
BMI	body mass index
CAF	Cafeteria diet
C-C	Control diet – control diet (pre and postnatally)
CLA	Conjugated linoleic acid
CpG	Cytosine-phosphate-Guanine
CREB1	cAMP responsive element binding protein 1
DIO	diet induced obesity
DNMT1	DNA methyltransferase 1
DR	diet resistant
E2F	E2 Factor
ERRα	estrogen-related receptor alpha
FS-FS	high fat - high sucrose diet (pre and postnatally)
H3K9	Histone3 Lysine9
HA	high vitamin A
HC	high carbohydrate

HDAC	histone deacetylase
HDL	high density lipoprotein
HFD	high fat diet
HFol	high folate diet
HOMA-IR	homeostatic model assessment of insulin resistance
HV	high vitamin
IAP	intracisternal A particle
JAK	Janus Kinase
LA	linoleic acid
MBD1	methyl binding domain protein 1
MC4R	melanocortin 4 receptors
Mecp2	Methyl CpG Binding Protein
ME	metastable epiallele
MM	mother's milk
MSH	melanocyte stimulating hormone
nPE	neural *POMC* enhancer
NF-KB	nuclear factor K-B
NPY	neuropeptide-Y
PAR	Predictive adaptive response
PBC	peripheral blood cell
PNS	Prenatal stress
POMC	Proopiomelanocortin
PUFA	polyunsaturated fatty acids
PVN	paraventricular nucleus of the hypothalamus

RELA	v-rel reticuloendotheliosis viral oncogene homologue A (avian)
RV	recommended intake of vitamins
SAH	S-adenosyl homocysteine
SAM	S-adenosyl methionine
SETDB1	SET domain binding 1
SFA	saturated fatty acid
Sp1	Specificity Protein 1
SRY	sex determining region Y
STAT3	Signal transducer and activator of transcription 3
VMR	variably methylated region
Y^{NPAR}	non-paring region of the Y chromosome

Background

Proopiomelanocortin (*POMC*), a pro-hormone, gives rise to numerous active peptides and hormones with a wide range of physiological actions(1). In the brain, *POMC* is expressed in the arcuate nucleus (ARC) of the hypothalamus, the pituitary gland and the brain stem. Tissue-specific post-translational proteolysis of *POMC* gives rise to the active hormones ACTH (adrenocorticotropic hormone), α-, β– and γ- MSH (melanocyte stimulating hormone) and β – endorphin (see Figure 1). Different populations of *POMC* neurons produce different amounts of the active hormones dependent on the expression levels of prohormone convertases e.g. corticotrophs of the anterior pituitary produce predominantly ACTH, whereas melanotrophs of the hypothalamus produce predominantly α- and β-MSH(1)(2).

POMC is a key component of the melanocortin system(3); a complex network of systemic signals and neural pathways that regulate food intake and energy balance (see Figure 2). *POMC* neurons in the ARC of the hypothalamus integrate peripheral signals such as leptin(4), glucose(5) and insulin(6), and regulate energy balance by inducing satiety and increasing energy expenditure(1). Satiety is mediated via the actions of α- and β-MSH on melanocortin 4 receptors (MC4R) in the paraventricular nucleus (PVN) of the hypothalamus(7)(8). An opposing group of orexigenic neurons (Agouti-related peptide (*AgRP*) and neuropeptide-Y (*NPY*)) receive systemic inputs from ghrelin (released from enteroendocrine cells) and have the opposite action by increasing appetite and decreasing energy expenditure by antagonising MC4R and by direct inhibition of satiety neurons in the PVN(1)(9). Perturbations of the melanocortin system can lead to disorders of energy balance such as obesity. For example, individuals with bi-allelic loss of function mutations in *POMC* demonstrate early hyperphagia, severe obesity (due to α-/β-MSH deficiency) and central adrenal insufficiency (due to ACTH deficiency)(10). It has been demonstrated that heterozygote variant carriers have an increased risk of developing obesity but with no adrenal insufficiency(11)(12) suggesting a gene dosage effect on energy balance.

Epigenetic processes, including DNA methylation, histone modification, chromatin remodelling and RNA-based mechanisms can affect gene expression(13). DNA methylation is widely studied in animals and humans and occurs primarily at cytosine-guanine (CG) dinucleotides (CpG methylation). DNA methylation is usually associated with condensed heterochromatin and subsequent gene silencing or reduced gene expression(14). DNA methylation is mitotically heritable(15) and is influenced by genetic and environmental factors(16)(17)(18).

In humans, the *POMC* gene is found on chromosome 2p23 and consists of three exons. Exon 3 codes for the majority of translated mRNA (see Figure 1) and functionally-relevant peptides i.e. ACTH, α-/β-MSH, β – endorphin(19). The gene contains two CpG islands (areas of the genome with high CpG density), the first in the 5'-promoter region and the second at the boundary of intron2/exon3(20).

POMC expression is controlled by different and independent transcriptional enhancers in the pituitary and hypothalamus(21) (see figure 1). Of particular relevance to *POMC*'s role in energy balance is regulation of gene expression in the hypothalamus. In the hypothalamus, two proximal enhancers; neural *POMC* enhancer 1 (nPE1) and 2 (nPE2) regulate *POMC* expression. nPE1 contains a STAT3 (Signal transducer and activator of transcription – 3) binding site, important for leptin signalling. Leptin's transcriptional activation of *POMC* is dependent on the Janus kinase (JAK)-STAT pathway with STAT 3 and 5 especially relevant for mediating satiety(22). STAT activation of *POMC* transcription is also dependent on Specificity Protein 1 (Sp1) binding to a region in the *POMC* promoter(23). nPE2 is a shorter genomic region and contains a binding site for the estrogen-related receptor alpha (ERRα) (21,24). These enhancer and promoter binding regions are functionally conserved in rodents and humans(21,25,26). In both humans and mice, additional transcriptional enhancer activity has also been identified at the boundary of intron2/exon3. This region contains a histone acetyltransferase p300 complex binding site that is associated with gene activation(27,28). The majority of animal studies (see table 1 and 2) concentrate on the promoter region, but more recently, human studies have examined other regulatory regions.

Due to *POMC*'s involvement in multiple physiological processes, epigenetic changes at *POMC* are associated with a diverse range of phenotypes. In this review, we present evidence linking epigenetic regulation of *POMC* to nutritional programming, obesity, energy balance and metabolic outcomes.

Methods

A MEDLINE database search was conducted using the following search terms; (((epigen*) OR methylation)) AND (((*POMC*) OR Proopiomelanocortin) OR Pro-opiomelanocortin)) on 12th November 2018. This search produced 181 articles. The purpose of this systematic narrative review is to summarise the evidence, in both animals and humans, of epigenetic regulation of *POMC* relevant to nutritional programming, energy balance, obesity and metabolic outcomes. Therefore, only articles that referred to obesity, body weight or BMI (body mass index) or energy balance, adiposity, lipid or glucose metabolism were subsequently included. Articles concerning intergenerational or

transgenerational processes were also included as these are relevant to nutritional programming. Articles with no specific reference to *POMC* epigenetic processes (such as DNA methylation or histone modifications) were excluded. Reference sections of included studies were also reviewed, and additional relevant papers were included if not found during the initial MEDLINE search.

Animal studies

1. Prenatal exposures, epigenetic alterations in the *POMC* gene and offspring phenotype

Animal studies have allowed researchers to manipulate the fetal environment and examine subsequent postnatal phenotypes under controlled conditions. Fetal adaptation to the *in utero* environment may present a survival advantage by the organism being better prepared for the anticipated *ex utero* environment(29). For example, a consequence of fetal adaptation to prenatal energy restriction might be for offspring to be better adapted to postnatal nutritional scarcity by, for instance, maintaining a smaller body size and having greater energy efficiency and increased appetite. This could result in later obesity if there is a mismatch between *in utero* and *ex utero* environments and forms the basis of the predictive adaptive response (PAR) hypothesis (30). Alternatively, environmentally induced epigenetic changes may simply be a consequence of the exposure on normal epigenetic development(31). Either way, a better understanding of how epigenetic mechanisms can mediate links between prenatal exposures and postnatal phenotype will have important implications for human health. A number of animal studies have examined how altering the *in utero* nutritional and/or prenatal environment influences *POMC* epigenetic marks and the subsequent phenotype in the offspring (see Table 1).

1.1 Maternal energy excess

Maternal overnutrition, induced by feeding pregnant rats high energy diets, has been shown to lead to persistent changes in offspring phenotype with accompanying epigenetic changes (32)(33)(34).

Rat pups exposed to high fat diet (HFD) *in utero* have higher *POMC* DNA methylation and a shift towards an obese phenotype as adults. This was demonstrated in a study by Marco et al(33) who showed that, compared to controls, the offspring of mothers fed HFD in pregnancy had a higher birth weight, ate more postnatally and had higher body weight up to 110 days of age. Offspring from the HFD group demonstrated ARC *POMC* hypermethylation across the promoter region (with

significantly higher methylation in 5 CpGs) and this hypermethylation was maintained into adulthood suggesting that *POMC* methylation is a stable epigenetic mark. When these pups were subsequently exposed to an HFD challenge, they consumed more and gained more weight than offspring of controls. Despite differences in *POMC* methylation, there were no differences in *POMC* expression. However, the HFD group had higher levels of leptin though this did not lead to higher *POMC* expression as might be expected. Only female offspring were examined for the hormonal, POMC expression and methylation analyses. This study provides evidence of an association between *POMC* hypermethylation (influenced by prenatal HFD) and increased body weight in offspring.

A further study by Marco et al (35), replicated the earlier finding that offspring of HFD fed rats had higher body weight into adulthood and had higher ARC *POMC* promoter methylation (with significantly higher methylation in 3 CpGs) compared to controls, but went further by postulating a process of dual epigenetic silencing with higher levels of H3K9me2 (an epigenetic mark associated with transcriptional repression) at the *POMC* promoter in the ARC. These epigenetic changes were associated with lower *POMC* expression and higher body weight. Interestingly, lean control offspring had higher levels of 5-hydroxymethylcytosine (5hmC), an epigenetic mark predominantly found in the central nervous system that is associated with increased gene transcription(36)(37). Body weight was positively correlated with *POMC* promoter 5-methylcytosine (5mC) and negatively correlated with *POMC* promoter 5hmC levels. The study also demonstrated higher levels of methyl binding domain protein 1 (MBD1) binding in the *POMC* promoter and SET domain binding 1(SETDB1, a histone methyltransferase) in HFD offspring and showed that this complex of MBD1 and SETDB1 promoted histone methylation. There were lower levels of the MBD1-SETDB1 complex in lean controls, suggesting that higher levels of 5hmC inhibited binding at this region leading to greater expression of *POMC* and a stronger satiety signal. This study gives evidence of an association between levels of 5mC and 5hmC in the *POMC* promoter and offspring's body weight. It further supports a relationship between prenatal HFD exposure and increased levels of 5mC in the *POMC* promoter.

In those exposed to a high energy diet *in utero, a* high energy diet postnatally may have a modulatory effect on *POMC* methylation. Evidence for this came from Zheng et al(34) who produced two groups of rats; FS-FS (high fat, high sucrose diet fed to mothers and offspring from weaning) and C-C (control diet fed to mothers and offspring from weaning). The FS-FS group demonstrated higher body weight in adulthood, though with higher expression of *POMC* and lower average methylation of the *POMC* promoter compared to controls (mean methylation+/- SD = FS-FS 37.5±1.7 % vs C-C

46.3±3.5 %, p=0.03). Despite examining the same genomic region, this is different to the observation reported by Marco et al(33,35), where increased *POMC* methylation was reported in offspring exposed to HFD *in utero.* In the studies by Marco et al (33,35), pups were weaned onto a normal calorie diet (not high energy) and therefore the disparate postnatal diets may explain the difference in *POMC* methylation reported following high energy exposure *in utero*. Another potential explanation is that Zheng et al (34) studied male rats from weaning, in contrast to the use of female rats (for hormonal and POMC expression/methylation analysis) examined in Marco et al(33).

Further evidence of differential effects of pre- and postnatal dietary exposures on offspring *POMC* methylation was highlighted in a recent study by Ramamoorthy et al (32) who reported hypermethylation in the *POMC* promoter and enhancer regions in 3-week-old male pups from HFD fed mothers. This methylation pattern in the promoter (but not the enhancer) was conserved into adulthood and suggests a potential critical period of prenatal programming. In this study, post weaning HFD also led to hypermethylation in the *POMC* promoter of male pups from the control group, though there were no additive effects of post-weaning HFD on *POMC* methylation in pups exposed to HFD prenatally. This study gives further evidence that maternal HFD is associated with increased methylation at regulatory regions of *POMC* and that increased *POMC* methylation is associated with increased body weight and food intake into adulthood.

Other epigenetic pathways have been shown to be influenced by maternal HFD. Desai et al (38), showed lower expression of hypothalamic DNA methyltransferase 1 (DNMT1) and histone deacetylase (HDAC, a class of enzyme that removes acetyl groups from histones and is associated with gene suppression) in offspring exposed to HFD *in utero*. In this study, as expected, male offspring of HFD mothers, when milk fed by the same HFD mothers, developed markedly increased weight and adiposity compared to controls. On day 1 of life, those offspring from HFD fed mothers showed lower expression of hypothalamic DNMT1 and increased *AgRP* expression compared to controls. However, *POMC* expression was the same in both groups. By 6 months, hypothalamic levels of HDAC were significantly reduced in the offspring of HFD mothers. The higher expression of *AgRP* was maintained but now also with reduced *POMC* expression compared to controls. This study provides additional evidence that prenatal HFD influences offspring's epigenetic processes at *POMC*. The changes were not associated with altered *POMC* expression until adulthood, though there were marked differences in body weight between the two groups.

As well as effects on offspring bodyweight and adiposity, maternal high energy diet has been shown to lead to impaired glucose homeostasis and greater insulin resistance in progeny. This has been evidenced by Ramamoorthy et al (32) who reported greater insulin resistance from 8 weeks of age that persisted into adulthood in rats exposed to HFD *in utero*. Furthermore, Zheng et al, demonstrated that *POMC* methylation was negatively associated with glucose response to a glucose load in a study in rats(34).

In summary, there is a strong evidence of an association between prenatal HFD and offspring hypermethylation in regulatory regions of *POMC* that persist until adulthood(32,33,35). A potential modulatory effect of postnatal FS diet on *POMC* methylation has been reported(34), however Ramamoorthy et al (32) demonstrated no additional influence from a postnatal high energy diet on the hypermethylation already seen in those with prenatal HFD exposure.

1.2 Maternal energy restriction

Maternal undernutrition was first demonstrated to be associated with offspring *POMC* hypomethylation and increased histone acetylation of H3K9 (associated with increased gene transcription), by Stevens et al using a sheep model with analysis of fetal hypothalamic tissue (39). However, these observed epigenetic changes were not associated with any change in *POMC* expression compared to controls (normally fed ewes). Importantly, their study found that alterations to the periconceptional nutritional environment, even when applied to a narrow window, can alter *POMC* gene methylation in later pregnancy. The study demonstrated that undernutrition exposure from as little as 2 days and up to 30 days before conception, was associated with significant *POMC* hypomethylation (64% lower than controls) in fetuses by day 133-135 of gestation. Pregnancies were terminated between day 131-135 of gestation, so an assessment of the postnatal phenotype was not made. Though 10-month old sheep from undernourished mothers (using the same protocol), had impaired glucose handling and increased body weight, it is not known if this was associated with alterations in *POMC* methylation or expression(40).

Begum et al (41), replicated the findings of *POMC* hypomethylation and increased H3K9 acetylation but also reported lower DNMT activity in the hypothalamus of both male and female offspring of energy restricted ewes (restricted between day -60 to +30 relative to conception). Similar to Stevens et al, the fetuses were euthanised between day 131-135 of gestation so there was no assessment of postnatal phenotype, however this study gives further evidence of an effect of periconceptional undernutrition on offspring's epigenetic marks at the *POMC* gene.

However, Coupe et al (42) reported no changes in DNA methylation in hypothalamic feeding-related genes including in the *POMC* promoter or enhancer regions, in the postnatal male rat hypothalamus in intrauterine growth restricted rats. Though the exact timing of calorific restriction was not clear, this would not support a modulatory effect of undernutrition on *POMC* methylation in rats.

Sex specific alterations in *POMC* expression are reported in offspring of energy restricted sheep, though the mechanism remains unclear. Begum et al (43) found that adult sheep exposed to undernutrition in early pregnancy demonstrated lower *POMC* expression in males (with increased fat mass in adulthood) but no difference in females(43). However, there was no assessment of *POMC* methylation as a mediator of these changes in gene expression. Ovilo et al (44) found female offspring (but not males) of late gestation nutritionally-restricted Iberian sows had a lower hypothalamic expression of *POMC*. There was increased body weight and fat but again there was no measurement of *POMC* methylation. Sexual dimorphism in hypothalamic circuits has been described and extensively reviewed (45). Deletion of ERRα from *POMC* neurons has no effect on male mice but leads to increased body weight, hyperphagia and increased lean mass in females(46). One possibility is that epigenetic changes involving the ERR are responsible for the sexual dimorphic response to maternal energy restriction. Alternatively, circulating estrogen could modulate the central actions of insulin or glucose on *POMC* neurons postnatally(47).

In summary, there is strong evidence from the studies of Stevens et al(39) and Begum et al (41) to suggest that periconceptional undernutrition has a modulatory effect on late gestational *POMC* epigenetic marks including hypomethylation, increased H3K9 acetylation and lower DNMT activity. There is limited evidence of an association between these changes and a subsequent postnatal phenotype. Evidence of phenotypic changes were highlighted by Begum et al(43), with higher fat mass and lower *POMC* expression in offspring of energy restricted ewes but it is not clear if this was mediated by epigenetic changes at *POMC*(43).

1.3 Mismatch in pre and postnatal vitamin and folate diet

A set of interlocking pathways, collectively known as one-carbon metabolism, provide methyl groups for methylation reactions including the methylation of cytosine bases and histone tails that in turn influence gene expression. These methylation reactions are controlled by methyltransferases that act on methyl groups produced by the conversion of S-adenosyl methionine (SAM) to S-adenosyl homocysteine (SAH). One-carbon metabolism is dependent on a number of nutrients and vitamins including folate, B12, B6, B2 and choline(48) (see Figure 3). It has been postulated that alterations in

folate or other one-carbon metabolite levels pre or postnatally could mediate epigenetic changes, putatively by altering the availability of methyl groups. For example in humans, maternal concentrations of folate in early pregnancy have been associated with offspring's DNA methylation and birth weight, illustrating the potential importance of prenatal exposure to one-carbon related nutrients and vitamins on DNA methylation and phenotype(49).

A mismatch between a high vitamin diet (HV) *in utero* and recommended vitamin diet (RV) in postnatal life has been shown to be associated with obesity in rats (50–52). Maintaining a similarly high vitamin intake in the postnatal diet can attenuate the development of the obese phenotype and can normalise *POMC* expression to that of controls. This was elegantly shown in a rat model by Cho et al (50) who showed that male offspring of rats fed HV, a 10-fold higher recommended intake of multivitamins ('AIN-93G' containing 13 vitamins(53) including Vitamin A, D, Riboflavin, Choline and Folic acid) in pregnancy, who were then weaned to RV, demonstrated higher body weight, food consumption and higher glucose response to a glucose load than those born to mothers fed the RV intake in pregnancy. Conversely, HV or HFol (high folate) postnatally prevented these associated phenotypic changes. Interestingly, postnatal HFol diet did not normalise *POMC* expression as the HV diet did. There was no difference in global hypothalamic DNA methylation between any of the dietary groups, though *POMC* methylation was not measured specifically.

Furthermore, it has been shown that HFol diet postnatally, alters *POMC* methylation in rats. This was evidenced by Cho et al(54), who gave male offspring of mothers fed HFol diet, either RV or HFol and compared them to controls fed RV pre and postnatally. HFol-RV rats demonstrated significantly higher food intake and body weight compared to controls and a 20% lower glucose response to an insulin load at 7 weeks of age (no differences at week 12 or 16 of life). HFol-HFol rats showed lower food intake and body weight compared to controls and demonstrated a significantly lower glucose response to a glucose load compared to HFol-RV and controls at week 10 (though no differences at week 14 or 18 of life). HFol in gestation reduced hypothalamic *POMC* expression compared to RV. The HFol – HFol groups had lower hypothalamic *POMC* methylation compared to the RV-RV and HFol-RV groups, providing evidence for ongoing postnatal plasticity in *POMC* DNA methylation mediated by increased dietary folate (even though global DNA methylation was not altered). It should be noted that although statistically significant, the differences in *POMC* methylation between the groups were small (HFol-HFol had 4% lower mean methylation compared to the other groups). This study supports an association between *POMC* methylation and a glucose response to a glucose load (glucose to *POMC* methylation coefficient of 0.7, p=0.03).

There is evidence that vitamin A can mediate epigenetic changes as it can increase glycine N-methyltransferase enzyme expression, which can lead to loss of methyl groups (55). In male rats exposed to a HV diet *in utero*, a postnatal diet high in vitamin A (HA, 10-fold higher vitamin A than recommended) can lower methylation at *POMC*, as demonstrated by Sánchez-Hernández et al (52). This study found that postnatal diets of either HV or HA, led to reduced post weaning weight gain and reduced food intake compared to male rats fed RV diets. HA postnatal diet was associated with significantly lower hypothalamic *POMC* promoter methylation compared to RV and HV, though the difference in mean methylation between the groups was small (3%). HV postnatal diet was associated with higher levels of *POMC* expression compared to RV but there was no difference in *POMC* expression between HV and HA diet. Though higher *POMC* gene expression was seen in the HV group compared to the control group, there was no difference in *POMC* promoter methylation, suggesting other epigenetic mechanisms (e.g. histone modifications) could account for the difference in expression. This study does support a modest effect of a postnatal HA diet in lowering mean *POMC* methylation and that a mismatch in vitamin diets between pre and postnatal period is associated with weight gain and increased food intake into adulthood.

In summary, there is strong evidence that a mismatch of HV diet prenatally and RV diet postnatally is associated with increased body weight and higher energy intake into adulthood(50–52,54). There is evidence that high concentrations of vitamin A(52) and folate(54) postnatally have a modest effect to lower *POMC* methylation in those exposed to HV diet prenatally.

1.4 Prenatal Stress

Prenatal stress (PNS) is associated with chronic disease, such as obesity, in adulthood (56). Epigenetic changes could be a conduit between PNS exposures and later phenotype. There is indeed evidence of epigenetic changes at *POMC* associated with PNS that predispose an individual to obesity when the animal is exposed to an obesogenic environment, such as HFD. Paternain et al showed that female PNS rats fed an HFD as adults demonstrated higher adiposity and greater insulin resistance when compared to non-PNS controls(57). HFD in adulthood led to increased hypothalamic *POMC* expression in both PNS rats and controls, but with an interaction between PNS, HFD and methylation at a single CpG site in the *POMC* promoter after adjustment for multiple testing. This study thus provides limited evidence that *POMC* methylation changes are influenced by PNS leading to increased susceptibility to HFD and the subsequent postnatal obese phenotype.

PNS has been shown to lead to greater weight loss in offspring when subjected to energy restriction. Boersma et al demonstrated that compared to controls, PNS male and female rats showed lower food intake and greater weight loss when subsequently exposed to activity-based anorexia (ABA); experimental conditions where food is limited to 1.5 hours per day and a running wheel in placed in their cage postnatally(58). ABA induced a reduction in *POMC* expression and an increased expression of *NPY* in all groups. There was no difference in *POMC* promoter methylation between groups. This study does not support the hypothesis that *POMC* methylation changes are associated with PNS. PNS does appear to influence postnatal energy balance under conditions of ABA, though this does not appear to be mediated by methylation changes at the *POMC* gene.

In summary, the study by Paternain et al(57), provides only limited evidence of an interaction between PNS and *POMC* methylation changes. However, both studies do provide evidence that PNS rats have greater adiposity when fed HFD as adults though conversely are more susceptible to weight loss under periods of nutritional scarcity. Whether these changes are mediated by *POMC* epigenetic alterations (other than DNA methylation) warrants further study.

1.5 Maternal drug exposure

Triclosan is a broad-spectrum antibiotic agent found in household and personal hygiene products (59). Exposure to the drug prenatally is thought to influence birth weight in humans(60), though the influence on epigenetic changes and later phenotype is little explored. Hua et al, showed a link between prenatal triclosan exposure and later obese phenotype in both male and female rats. Rats exposed to Triclosan in early/mid gestation (gestational day 6-14) had lower birth weight, but by day 30 of life showed increased mean *POMC* methylation (significantly higher in 6 CpGs across the *POMC* promoter), reduced *POMC* expression and subsequent early hyperphagic obesity and metabolic syndrome(61). This study gives support to the idea of an environmental agent acting prenatally to influence *POMC* methylation which is associated with the development of obesity and metabolic syndrome later in life.

2. Postnatal exposures, epigenetic alterations in the *POMC* gene and phenotype

There is also evidence that postnatal factors can influence *POMC* epigenetic marks and influence weight-related phenotype (see table 2). One possibility is that developmental plasticity continues into the postnatal period with an epigenetic adaptive response continuing beyond the 'classical' window of fetal programming. If this is the case, then alterations to *POMC* epigenetic marks in the postnatal period would allow the organism to more accurately *predict, adapt* and *respond* to future nutritional requirements.

Alternatively, postnatal changes may represent more widespread alterations to the epigenome brought about by environmental factors and may not target *POMC* specifically. Yet another explanation is that epigenetic changes are due to reverse causation where postnatal phenotype influences the epigenome.

2.1 Energy excess

Overnutrition during infancy, in animals and humans, is associated with the later development of obesity(62). There is evidence that *POMC* hypermethylation, induced by postnatal energy excess, may suppress the satiety response by impeding the action of peripheral signals such as leptin and insulin. This was demonstrated in a rodent model(63) whereby postnatal energy excess was induced by artificially creating 'small litters' (SL) of 3 pups per mother (as opposed to 10 pups per mother). Rats placed into SL from day 2 of life, demonstrated higher body weight, glucose, leptin and insulin by day 7 of life(63). Higher *POMC* promoter methylation, including in regions associated with nuclear factor K-B (NF-KB) and Sp1 binding on the *POMC* gene, was observed compared to controls (even after Bonferroni correction for multiple testing). Sp1 binding is crucial to mediate the effect of circulating leptin via STAT3-Sp1-complex formation. There was an inverse relationship between *POMC* methylation at the Sp1 binding site and *POMC* expression per unit of leptin and insulin, suggesting that methylation at this site impedes the anorectic effects of leptin and insulin resulting in lower *POMC* expression. This study supports a link between energy excess in the postnatal period and epigenetic changes at transcriptional regulatory sites in the *POMC* promoter.

Marco et al (64), used a rat over-feeding model and further demonstrated the interaction of *POMC* methylation, Sp1 binding and *POMC* expression. Male rats fed an HFD from the neonatal period into adulthood demonstrated increased weight, insulin and leptin with hypermethylation at 6 CpGs in the *POMC* promoter including at a Sp1 binding site. However, there was not the expected increase in hypothalamic *POMC* expression despite significantly higher levels of leptin, Sp1 and insulin (factors

associated with upregulation of *POMC* expression(4,6,23)). ChIP analysis demonstrated reduced Sp1 binding in the HFD group, despite higher levels of circulating Sp1, suggesting hypermethylation impeded the formation of the Sp1 complex. This study supports a link between energy excess in the postnatal period and methylation changes at CpGs near a key transcriptional regulatory site (Sp1) in the *POMC* promoter.

NF-KB is released in response to both acute and chronic inflammation. Though acute inflammation is associated with reduced food intake(65), obesity is associated with chronic low grade inflammation(66). RELA (v-rel reticuloendotheliosis viral oncogene homologue A (avian)) is a subunit of NF-KB. Shi et al(67), demonstrated that during acute inflammation, male mice showed reduced appetite and increased *POMC* expression. Using a STAT3 (Signal transducer and activator of transcription 3) knock out model they demonstrated that this upregulation of *POMC* was independent of the STAT3 pathway (mediated by leptin) but instead RELA upregulated *POMC* transcription by directly binding to the *POMC* promoter. In the same study, they showed that HFD male mice (modelling the chronic inflammation seen in obesity) had hypermethylation of the *POMC* promoter which subsequently impeded RELA-mediated *POMC* transcriptional activation. This study supports a link between energy excess and increased methylation in the *POMC* promoter but also proposes an alternative means of *POMC* transcriptional regulation by direct binding of RELA to the promoter region.

Mahmood et al (68), demonstrated a potential postnatal programming effect of high carbohydrate (HC) feeds on melanocortin system neuropeptides and adult body weight, though mediated predominantly through epigenetic changes in the *NPY* (neuropeptide-Y) gene. Female rats were fed with either HC or mother's milk (MM), and then weaned onto standard feed from day 24. By day 16, HC rats had higher *NPY* H3K9 acetylation, lower *NPY* gene methylation, lower *POMC* gene H3K9 gene acetylation but no difference in *POMC* promoter methylation. By day 100, only *NPY* methylation differences persisted (with increased expression of *NPY*) with significantly higher body weight in those fed HC diet compared to those fed MM in the postnatal period. This study suggests that a high energy diet can induce epigenetic changes at the *POMC* gene but these changes do not persist beyond the immediate postnatal period. Importantly, sustained epigenetic changes were only seen in the *NPY* gene, and not *POMC*.

It has been suggested that epigenetic changes could 'predispose' an individual to overconsumption in an obesogenic environment (69). Cifani et al, demonstrated there are epigenetic differences in

POMC and *NPY*, in those rats who are obesity-resistant or obesity-prone on exposure to an obesogenic environment (70). Male rats were fed a high energy feed for 5 weeks and classified as to whether they developed obesity or did not (diet induced obesity: DIO or diet resistant: DR). The DR rats consumed less, had similar weight to controls but also had decreased hypothalamic expression of *NPY* by 5 weeks (with hypermethylation at one CpG in *NPY*) and higher expression of *POMC* at 21 weeks (the mean *POMC* methylation was significantly lower though the difference was very small (~1%), however there were still significant differences at 4 CpG sites even after Bonferroni correction) compared to DIO rats. This study suggests that different neuropeptide gene methylation and expression could account for the divergent eating behaviour and weight gain, though overall differences in *POMC* methylation were small.

There is further evidence that high energy diets induce epigenetic changes that produce increased expression of orexigenic neuropeptides (*AgRP*/*NPY*), overriding any increased expression of the anorectic *POMC*, and leading overall to overconsumption and increased body weight. Lazzarino et al (71) fed female rats a CAF diet (Cafeteria diet, modelled on Western diet of high palatability and highly energy dense foods) from weaning and demonstrated an increased expression of *POMC* and reduced *AgRP* expression at the level of the ARC but increased expression of *AgRP*/*NPY* in the PVN. This was accompanied by lower methylation in the *POMC* promoter of the ARC and lower methylation in the *NPY* promoter in the PVN. CAF fed rats had higher body weight from week 10 and consistently consumed a higher energy intake compared to controls. Of note, the methylation of CpGs associated with Sp1 binding (thought to mediate anorectic effects of leptin) were not measured in this study. This study suggests that in response to CAF diet epigenetic and gene expression changes effect both *AgRP*/*NPY* and *POMC,* though the higher energy intake observed was mediated by higher *AgRP/NPY* expression in the PVN and not counteracted by increased *POMC* expression in the ARC.

In summary, the studies by Plagemann et al(63), Marco et al(64) and Zheng et al(34) provide the strongest evidence for nutritionally-driven postnatal changes to *POMC* epigenetic marks. However, it is not clear if the diet itself or the increased body weight leads to the epigenetic differences observed. Other studies suggest that *AgRP/NPY* genes are more susceptible to postnatal overnutrition than *POMC* and may be more implicated with the postnatal phenotype(68)(71).

2.2 Energy restriction

There are data from humans that suggest periods of extreme calorific restriction may lead to remission of type 2 diabetes mellitus(72) and a sustained reduction in BMI(73) long after the end of dietary intervention. Animal models suggest this 'metabolic memory' may be mediated via epigenetic changes in the appetite regulating neuropeptides. Liu et al (74), explored if periods of calorific restriction in the post weaning period could mitigate the development of obesity in adult life associated with SL during suckling. Male rats given moderate calorific restriction (24% reduction in daily calorific intake for 49 days) in the post weaning period, showed comparable food intake and slower weight gain compared to controls when subsequently fed ad libitum, such that they did not reach the weight of controls by day 140 of life. Furthermore, they had similar levels of insulin, leptin and expression of *NPY* and *POMC* to controls. In both control and calorific restriction groups there was increased methylation in the *NPY* promoter at a CpG important for E2F (E2 Factor) binding, compared to those from SL with no restriction. E2F is a transcription factor known to regulate *NPY* expression. There were no differences in *POMC* promoter methylation between the groups. This study demonstrates that a period of calorific restriction may attenuate the hyperphagia and obese phenotype associated with SL, though suggests that this does not appear to be mediated by alterations in *POMC* gene methylation.

However, Unnikrishnan et al, reported 65% lower hypothalamic *POMC* expression following short term dietary restriction in male mice (75). These alterations in *POMC* gene expression persisted even after 2 months of ad libitum feeding. There was no associated difference in DNA methylation at the *POMC* promoter, though DNA methylation was analysed from whole hypothalamic tissue and not specifically the ARC where *POMC* is expressed in the hypothalamus.

These two studies suggest that postnatal energy restriction does not lead to changes in *POMC* methylation but predominantly effects methylation marks at *NPY* (similar to studies of postnatal energy excess(68)(71)). Reported alterations in *POMC* gene expression could be related to other epigenetic processes not measured in these studies.

2.3 Fatty acids

Diets high in n–6 polyunsaturated fatty acids (PUFA) found in foods such as sunflower oil, and low in n–3 PUFA (found in fish oil for example) have been implicated in a number of chronic diseases including obesity (76). N-3 PUFA has been shown to reduce leptin expression (77), which is elevated in obesity due to the increased adipose tissue mass. Leptin regulates the anorectic actions of *POMC*

via the long form leptin receptor on *POMC* neurons, and it has been postulated that elevated leptin in obesity represents a form of leptin resistance(78). Fan et al, demonstrated that the addition of n-3 PUFA to the diet of obese male mice normalised *POMC* expression and leptin to comparable levels seen in controls (79). N-3 PUFA fed mice had lower weight compared to those fed n-6 PUFA. The changes in gene expression were not related to a difference in methylation in either *POMC* or the leptin promoter from adipose tissue.

Conjugated linoleic acid (CLA) is a fatty acid associated with the development of metabolic disease (80) and has been implicated in postnatal programming(81) in murine models. There is evidence from Zhang et al(82) to suggest that postnatal nutritional exposure to CLA causes *POMC* promoter hypermethylation, preventing Sp1 complex formation which in turn leads to reduced *POMC* expression and an increase in food intake. Lactating mice were fed a diet rich in either CLA or LA (linoleic acid) which altered milk composition such that CLA fed mice had increased milk glucose and insulin, lower milk lactose and triglycerides and altered fatty acid composition compared to LA fed mice. CLA fed pups had significantly lower *POMC* expression compared to LA fed pups with hypermethylation of two promoter CpGs. This region of the *POMC* promoter is at the location of a Sp1 binding site and ChIP assays demonstrated reduced Sp1-*POMC* promoter complex formation in CLA fed pups. Interestingly, CLA fed pups had low SAH (formed after a methyl group is removed from SAM in methylation reactions, see figure 3) and it was suggested this may explain the observed hypermethylation in this group. In the same study, CLA fed mice initially had lower body weight and restricted growth, but increased food intake. However, they were heavier as adults with impaired glucose homeostasis compared to controls. This study does support a modulatory effect of dietary CLAs on the *POMC* methylation and gene expression with an association with obesity in later life.

There is limited evidence that postnatal dietary fatty acids influence *POMC* gene methylation. However, Zhang et al (82) provide the strongest link with dietary CLAs associated with alterations at key regulatory regions of the *POMC* promoter.

2.4 Leptin

Due to leptin's known association with upregulation of *POMC*, Palou et al investigated if leptin treatment could influence postnatal programming of hypothalamic hormones including *POMC*(83). Leptin treatment in the suckling period led to lower body weight and food consumption in adulthood compared to controls. After the suckling period, male rats were fed either normal diet or HFD. At a CpG position close to the Sp1 binding site, leptin-treated rats demonstrated significantly

higher methylation when fed a normal diet and lower methylation when fed HFD (compared to non-leptin treated animals). However, there were no significant correlations between *POMC* methylation and *POMC* expression. The authors note these observations should be treated with caution until confirmed by other independent studies and significance was only reached when comparing individual groups (as overall there was a non-significant interaction using two-way ANOVA tests).

3. Other epigenetic regulatory factors and weight-related phenotypes

3.1 Methyl CpG Binding Protein (Mecp2)

Mecp2 was thought to be a repressor of *POMC* activity acting by recruitment of a repressor complex(84)(85) to methylated DNA. However, a study by Wang et al found that Mecp2 can act as a transcriptional activator in the hypothalamus(86). Male mice with a specific Mecp2 knock out (KO) in *POMC* neurons had higher body weight, energy intake, leptin and body fat percentage. The Mecp2 KO mice had higher *POMC* promoter methylation and lower *POMC* expression. Co-transfection of wild type Mecp2 and cAMP responsive element binding protein 1 (CREB1) led to increased *POMC* promoter activity, significantly more than when transfected with Mecp2 or CREB1 alone. This suggests that the activating properties of Mecp2 are dependent on an interaction with CREB1 to increase *POMC* expression. In contrast to previous studies (84)(85), this study specifically altered Mecp2 expression in *POMC* neurons and therefore better demonstrates the effect in this neuronal group.

Summary of evidence from Animal studies

From the evidence in animal models, it is clear there is a complex relationship between pre- and postnatal exposures that influence epigenetic processes associated with the *POMC* gene.

Epigenetic processes and marks are sensitive to prenatal nutritional and environmental exposures and have a consequence on later adult phenotype. There is evidence from these animal studies that supports the PAR hypothesis that mismatched *in utero* and *ex utero* environments are associated with obesity and poor metabolic health and can be to some degree explained by epigenetic changes in *POMC*. For example, there is considerable evidence that prenatal exposures to energy

excess(33)(35)(32) and periconceptional and early pregnancy energy restriction (39)(41) have different effects on *POMC* epigenetic marks and are associated with postnatal phenotype. A mismatch in pre and postnatal vitamin levels have been shown to be associated with higher body weight and increased food intake(50–52,54), and there is modest effects of vitamin A(52) and folate(54) to lower *POMC* methylation postnatally. Though limited to one study(61), the modulatory effect of Triclosan on *POMC* methylation appears significant and is strongly associated with obesity in adulthood.

The evidence of postnatal exposures influencing *POMC* epigenetic marks is more limited. Evidence from Plagemann et al(63) and Marco et al(64) provide the strongest link between postnatal energy excess and changes to DNA methylation at *POMC*, however other studies report that *AgRP*/*NPY* genes appear more sensitive than *POMC* to postnatal energy imbalance(68,71,74,75). The study by Zhang et al (82) suggests that dietary CLAs are the strongest specific nutritional factor to mediate methylation changes at *POMC* gene postnatally.

It is important to consider the potential for reverse causation when interpreting results from epigenetic studies. An approach to minimising the potential for reverse causation effects is to measure epigenetic marks(e.g. DNA methylation) prior to manifestation of the disease or phenotype. One advantage of animal studies is the ability to access functionally relevant tissue (such as ARC samples when examining *POMC*), however this may prevent assessment of epigenetic marks prior to the development of a particular phenotype as the animal is euthanised. To circumvent this, researchers often use subgroups of animals who are euthanised at different time points. Marco et al (33), euthanised a proportion of rodents on weaning and maintained a group into adulthood allowing comparisons of methylation before the adult phenotype developed. In this study, the same *POMC* promoter hypermethylation was seen at weaning and in adulthood suggesting that the methylation predated the development of the adult phenotype.

In all of the animal studies, samples harvested were either from the whole hypothalamus, ARC or PVN. One issue with analysing whole hypothalamus is that cell heterogeneity (i.e. the presence of cells that are not exclusively ARC or PVN cells) may dampen any positive signal and could risk type 2 errors ,or conversely increase the risk of type 1 errors in the case of confounding due to changing cell populations between treatment groups.

There are a number of studies which report no association between DNA methylation and *POMC* expression, suggesting that other epigenetic processes could drive differential expression. However,

studies that accounted for leptin or insulin levels when interpreting *POMC* gene expression more often found significant differences (63). Mechanistic studies are required to fully understand how alterations in *POMC* epigenetic pathways influence gene expression or phenotype, for example by assessing the effect of gene 'knock out' or 'knock down' models that target specific epigenetic processes. An exemplar of this approach is demonstrated by Wang et al (86)where the mechanism by which Mecp2 and CREB1 influence *POMC* methylation, gene expression and phenotype was elucidated using an MECP2 knock out in mouse POMC neurons.

It is important to note that whilst many studies did account for multiple comparisons(e.g., (32,33,35,41,57,58,63,64,70,74,75,83) some studies did not (38,61,67,68,71,79,82). The need for multiple testing corrections can be context dependent since methylation levels within genomic regions such as CpG islands may be correlated, suggesting that a correction isn't justified.

Human studies

The establishment and maintenance of cell-specific gene expression profiles is a key function of epigenetic processes, and epigenetic marks including DNA methylation are therefore generally tissue-specific. Thus a significant challenge in human epigenetic studies is selecting relevant accessible tissue to examine. Peripheral blood cell (PBC), buccal epithelial cells (BEC) or hair follicles are often used as they are easily accessed and minimally invasive. However, the degree of tissue discordance in DNA methylation can be a problematic and limit how much can be inferred from an accessible tissue (such as PBC) about the tissue of interest (i.e. the hypothalamus for *POMC*'s regulation of energy balance). In this respect, the use of human cadaveric samples can be useful to understand the correlation in methylation between the accessible tissue and the tissue of interest (87). An assessment of two or more samples originating from different germ layers e.g. PBC (mesoderm) and BEC (ectoderm) can be used to identify systemic methylation patterns.

4. Nutritional influences on the *POMC* gene – evidence in humans

As with animal studies (see Tables 1 and 2), there is evidence in humans that *POMC* gene methylation is associated with nutrition in early pregnancy and diet into adulthood.

4.1 Influence of periconceptional diet on the *POMC* gene

A body of work in mouse and humans has focussed on nutrition-associated changes in DNA methylation at metastable epialleles (MEs). MEs are epigenetic loci with characteristic methylation patterns that are established early in embryonic development(88) that are often associated with neighbouring transposable elements(89).

The most robust evidence for MEs comes from the Agouti viable yellow (A^{vy}) (89) and Axin-fused ($Axin^{Fu}$) mouse models(90). For example, isogenic A^{vy} mice show variable agouti expression dependent on DNA methylation at a cryptic promoter within an IAP (intracisternal A particle (IAP)) upstream of the *agouti* gene. A maternal diet rich in methyl donor nutrients folate, B12, choline and betaine gives rise to increased methylation at the IAP in A^{vy} offspring(89). The degree of methylation is associated with phenotype such that hypomethylation leads to increased ectopic agouti expression, yellow fur and an obese phenotype. In contrast, hypermethylation at the same region is associated with reduced expression and a lean mouse with brown or mottled fur.

A number of human studies have demonstrated associations between periconceptional diet and DNA methylation at putative MEs(91–93), and interestingly there are multiple lines of evidence, including sensitivity to maternal nutrition, that *POMC* is a putative human metastable epiallele.

Firstly, as with the *agouti* mouse, there is evidence that the methylation state of the *POMC* gene in humans is sensitive to maternal diet in early pregnancy. Mother-child paired blood samples from a Gambian cohort demonstrated an association between early pregnancy one-carbon metabolite concentrations in maternal plasma and offspring PBC *POMC* methylation(87). Specifically, there was a significant negative correlation for SAH and positive correlations with betaine and the ratio of SAM to SAH at a region spanning the intron2/exon3 boundary of the *POMC* gene. Offspring DNA methylation was also associated with Gambian season of conception, with lower DNA methylation at the *POMC* VMR (variably methylated region) in children conceived in the dry season compared to those conceived in the rainy season(87). In contrast to animal models of maternal overnutrition (32–34), neither maternal body weight at conception nor weight change in pregnancy correlated with offspring *POMC* methylation (87).

Secondly, *POMC* methylation appears to be set very early in embryonic development. Evidence for this comes from post mortem samples that demonstrate *POMC* DNA methylation is highly correlated

across tissues originating from different germ cell layers i.e. brain – ectoderm, and kidney or PBC – mesoderm(87). This suggests the methylation state was set prior to the separation of the germ layers at gastrulation; the canonical feature of an ME.

Thirdly, methylation at the *POMC* VMR is thought to be largely independent of genotype, at least in *cis*, with similar methylation patterns across genetically diverse cohorts(87). It is interesting to note that MEs were originally identified in isogenic mice and defined as being independent of genotype. While current evidence suggests this may be true at the *POMC* VMR, it has recently been proposed that MEs may be sensitive both to environment and local genotype(94).

Fourthly, *POMC* methylation is associated with the presence of neighbouring transposable elements. A considerable proportion of the human genome consists of transposable elements such as retrotransposons. The most common retrotransposons in the human genome are Alu elements(95). Kühnen et al, observed three Alu elements in intron 2(20) of the *POMC* gene (see Figure 1). These Alu elements are only found in humans and higher primates (e.g. chimpanzees) and are not found in more distant primates or mice. Those species with Alu elements, demonstrated the same hypermethylation pattern of intron 2, however those without Alu elements showed hypomethylation(27) suggesting that the presence of Alu elements drives hypermethylation in this region. This demonstrates a similar pattern seen in MEs where alteration in gene expression is driven by methylation at retrotransposons, similar to that seen with the IAP retrotransposon in the A^{vy} mouse(96).

Importantly, assessment of *POMC* methylation in new-born and adolescent blood samples demonstrated that the methylation pattern appears stable, suggesting that associations with postnatal phenotypes are not driven by reverse causation effects (87). This is further supported by Yoo et al (97), who assessed *POMC* methylation in a longitudinal birth cohort and showed that methylation was highly correlated from birth to childhood (r=0.80, p=0.0001).

4.2 *POMC* methylation and diet in adulthood

SFA (saturated fatty acid) diets are associated with increased visceral and hepatic steatosis and PUFA (polyunsaturated fatty acids) diets are associated with increased lean mass(98). There is evidence that these diets may influence DNA methylation. The LIPOGAIN study(99), a double-blind randomised controlled trial, gave healthy normal weight adults 7 weeks of a daily muffin either high in SFA (n=17) or PUFA (n=14). Adipose tissue samples were taken at baseline and at the end of intervention and used for methylation analysis. There was increased global methylation in adipose

tissue following both diets. However, *POMC* was one of a number of genes where mean methylation in adipose tissue increased only in response to PUFA and not SFA treatment.

5. *POMC* methylation and weight-related phenotypes

The majority of evidence related to *POMC* epigenetics and weight regulation has come from animal models. However, more recently, studies have explored the role of epigenetic regulation of *POMC* and weight-related phenotypes in humans (see Table 3).

5.1 Obesity

Kühnen et al (27) were first to examine the relationship between *POMC* DNA methylation and obesity in humans. In a case control study comparing 71 obese and 36 normal weight children, they reported a significant difference in average PBC *POMC* DNA methylation at a VMR overlapping the boundary of intron2/exon3 (average methylation 25% normal weight vs. 40% obese, $p<0.001$). This finding was replicated in a second case control study in children with comparable results. An association between *POMC* hypermethylation at the VMR and BMI was demonstrated in both PBC and cadaveric MSH neurons in adults (87). The effect size was largest in MSH neurons, where a 10% increase in methylation was associated with a 2.8kg/m^2 increase in BMI(87). Kühnen et al, also demonstrated that hypermethylation at the VMR decreased histone acetyltransferase P300 binding at the VMR (see Figure 1), leading to reduced expression of *POMC* from PBC. P300 is an enzyme that promotes transcription through histone acetylation (27). There was no association between BMI and DNA methylation in the CpG island in the *POMC* promoter region(27).

The association of BMI and *POMC* methylation appears dependent on the region of the gene studied. For example, a recent study by Acs et al(100) examining 82 obese children aged 3 to 18 years old did not show any correlation between PBC *POMC* methylation and BMI(100). This study examined DNA methylation at exon 1 of the *POMC* gene rather than the intron2/exon3 region studied by Kühnen et al (87). Another possible explanation for the lack of correlation between *POMC* methylation and BMI is that this study examined obese children with BMI >95th percentile, and there may not be a strong correlation with methylation and BMI at these extremes of weight compared to comparisons made with normal weight individuals or the general population. Furthermore, Yoo et al (97) reported no significant association between BMI or percentage body fat with *POMC* methylation in exon 3. A possible explanation is that the genomic region studied is

downstream of the region associated with P300 binding (27,87). In this study, PBC *POMC* methylation (at 4 CpG sites in exon 3) measured in cord blood was subdivided into categories of either high (>90th centile), mid or low (<10th centile) methylation. Birth weights (p=0.01) and ponderal indices (p=0.01) in the high *POMC* methylation group were significantly lower than in the mid-methylation group, but there were no differences in BMI z-score or percentage body fat between the groups by mid-childhood. It should be noted that there were only 10 children in each of the high and low groups and no clear justification was given for splitting the participants into methylation categories, rather than using methylation as a continuous variable. In a recent case control study of 79 controls and 41 overweight/obese children aged 7-9 years, Kwon et al (101) compared PBC *POMC* methylation (4 CpG sites in exon 3 and a similar region to Yoo et al(97)) and reported significantly lower methylation in the overweight/obese groups compared to normal weight individuals. This genomic region is downstream from intron2/exon3 border and again would not have included the region associated with P300 binding. Though statistically significant, mean differences in methylation were small, for example the methylation difference at *POMC* CpG site 2 was 50.3% (normal weight) vs 49.1% (overweight/obese), p<0.001. These differences highlight the need to study defined CpGs and it should be noted that there is limited coverage of the regions described in the studies above on methylation arrays such as the Illumina EPIC Array.

As mentioned above, there is evidence that *POMC* methylation is stable through infancy into adolescence and that hypermethylation at *POMC* may predate the onset of obesity(27,97). Kühnen et al, examined PBC DNA from individuals aged 5 or 13 years who later became obese and 8 of 21 individuals had the hypermethylation variant many years before the onset of obesity, implying that *POMC* hypermethylation is not merely a consequence of increased body mass (27).

In summary, the strongest evidence of an association between *POMC* methylation and BMI comes from the case control studies of Kuhnen et al(27,87). Other studies examine different regions of the gene where there is limited evidence for an association with body weight(97,101).

5.2 Metabolic outcomes

There are reported associations between *POMC* methylation (in exon 3) and metabolic outcomes that appear independent of BMI or adiposity. Yoo et al (97), reported significantly higher fasting triglycerides (TG) (though not total cholesterol) in children aged 7-9 years in the high and middle level *POMC* methylation groups compared to the low methylation group; high vs low (TG = 113.89mg/dl vs 57.97mg/dl, p=0.03) and middle vs low (TG= 67.29mg/dl vs 57.97mg/dl, p=0.01).

This is despite there being no difference in BMI or percentage body fat between the groups. Kwon et al (101), reported significant association between methylation at a CpG site in exon 3 and lower HDL (high density lipoprotein) cholesterol levels (β=-0.23, p=0.048) after adjusting for age, gender and BMI. Interestingly, polymorphisms in melanocortin signalling pathways have been linked to altered lipid metabolism, independent of body habitus suggesting a possible distinct causal mechanism(102).

There is evidence of an association between elevated insulin and increased *POMC* methylation. Yoo et al (97), demonstrated higher fasting insulin levels in the high and middle *POMC* methylation groups compared to low methylation group; high vs low (insulin 10.13 µIU/ml vs 7.1 µIU/ml vs p=0.05) and mid vs low (insulin = 7.64 µIU/ml vs. 7.1 µIU/ml, p=0.02). A non-significant trend to greater insulin resistance (homeostatic model assessment of insulin resistance (HOMA-IR)) between high (p=0.09) and mid (p=0.06) groups compared to low methylation group was also observed, and no significant difference in blood glucose was reported between the different methylation groups.

5.3 *POMC* methylation as a predictor of successful weight loss intervention

Crujeiras et al (103) explored the utility of *POMC* methylation as a potential biomarker for success in weight loss interventions. Eighteen men enrolled in a dietary intervention programme who successfully lost more than 5% of their body weight were reviewed at 32 weeks post intervention. Participants were divided into two groups: regainers (regained more than 10% body weight) and non-regainers (regained less than 10% body weight). Higher *POMC* promoter methylation was seen in regainers vs. non-regainers (p=0.02) with a percentage body weight regain to methylation correlation coefficient of 0.6. Interestingly, there was an opposite trend for *NPY* gene methylation.

5.4 *POMC* methylation and Anorexia Nervosa

Individuals with anorexia nervosa (AN) reportedly have low levels of folate(104) and elevated homocysteine(105); important components of one-carbon metabolism pathways (see Figure 3). Furthermore, lower global DNA methylation has been reported in AN compared to normal weight controls(106). This combined with *POMC*'s influence on appetite regulation makes the gene a key candidate to explore epigenetic influence on the development of AN. Ehrlich et al(107), explored the relationship of PBC *POMC* promoter DNA methylation and expression of *POMC* mRNA in both acutely admitted and weight recovered women with AN and normal weight female controls. Mean

POMC promoter methylation was neither different across nutritional states nor across disease groups. *POMC* mRNA expression was decreased in those with undernutrition and hypoleptinaemia. The study demonstrated that expression of *POMC* is linked to nutritional state rather than a distinct feature of AN.

Ehrlich et al(108) later confirmed observations from the earlier study that there was no effect in women of undernutrition or a diagnosis of AN on PBC *POMC* promoter DNA methylation. However, they did observe significant associations between cigarette smoking and PBC *POMC* DNA methylation. Overall PBC *POMC* promoter methylation was negatively associated with average number of cigarettes smoked per day (ρ=-0.287, p=0.002). Nicotine is known to induce hypophagia and this is thought to be mediated through *POMC* neuronal activation(109,110). Smokers have a lower body weight compared to non-smokers(111) and weight gain following smoking cessation has been observed(112). This is a cross sectional study and as such the direction of causality between smoking and *POMC* methylation is not known.

In summary, these studies do not suggest an association with *POMC* methylation and AN diagnosis.

Summary of evidence from Human studies

There is evidence in both children and adults that hypermethylation at the *POMC* VMR of intron2/exon3 is associated with obesity(27,87). There are a number of lines of evidence to suggest that this region of the *POMC* gene is a putative human ME and influenced by periconceptional nutritional status. An association between *POMC* gene methylation and body weight was not replicated in other studies that examined different regions of the gene(97,100,101).

Results from two studies in children show an association between *POMC* methylation in exon 3 and altered levels of lipids(97,101). There is also preliminary evidence in children that *POMC* methylation is associated with elevated fasting insulin(97). Crujeiras et al (103), report an association between hypermethylation in the *POMC* promoter and weight regain after a weight management intervention which demonstrates how *POMC* methylation measurement could inform clinical risk stratification and help guide tailored interventions.

Cell specific methylation patterns in heterogenous samples are a potential confounder in epigenetic studies and are generally accounted for in epigenome-wide association studies by adjusting for cell composition estimates derived from the DNA methylation data itself(113). The region of

intron2/exon3 described by Kuhnen et al is a putative ME, and as such demonstrated cross-tissue concordance in DNA methylation patterns in tissues derived from all 3 germ layers, notably between PBC and the hypothalamus(87). Therefore, in this instance correcting for cell composition should not be necessary as the methylation pattern is thought to be systemic. The only human study to correct for cell composition was Perfilyev et al(99) who used a reference-free method to correct for potential differences in adipose tissue cell type composition. Furthermore, a comparison of POMC methylation in the promoter or exon 1 between the hypothalamus and PBC has not been made before.

Periconception and gestation represent key windows where there is the potential for the prenatal environment to influence epigenetic reprogramming. However environmental and nutritional factors may influence DNA methylation throughout adult life(114). For example, methylation at a number of *POMC* CpGs was associated with PUFA diet in a study in adults as described above (99). Thus while there is evidence that the *POMC* ME region may be stable from periconception to late adolescence(87)(97), sensitivity to specific postnatal factors, and stability beyond this period warrants further study.

Further study is also needed to establish causative links between *POMC* methylation, altered gene expression and subsequent phenotype. Though Kuhnen et al(27), did see lower expression of *POMC* in PBC from hypermethylated individuals many studies do not examine *POMC* expression(97,99–101,103) or find no association between methylation and expression(107). Exploration of how *POMC* methylation could influence energy consumption or the satiety response has also not been studied before in humans.

6. Transgenerational inheritance of epigenetic marks at the *POMC* gene

Maternal exposures before and during pregnancy including environmental (e.g. toxins or stress) or nutritional factors can induce epigenetic changes in the offspring. In this case transmission of epigenetic changes is referred to as inter-generational epigenetic inheritance(115). In exposed mothers (F0), developing offspring (F1) and their germ cells (F2) are also exposed. Transmission of epigenetic changes resulting from maternal exposure in F0 that persists into the F3 generation or beyond (i.e. no direct exposure) is termed transgenerational inheritance. In contrast, where epigenetic changes result from an exposure in F0 males, effects in F1 are inter-generational whereas those apparent in F2 and beyond will be trans-generational(115). Evidence for trans-generational

epigenetic inheritance in humans is currently lacking and may be extremely rare due to epigenetic reprogramming at conception and during germ cell development(116).

Evidence for human inter-generational inheritance through the maternal line has been described above. There is provisional evidence from animal and human studies to suggest *POMC* epigenetic marks may be transmitted across generations and mediated via the paternal line. Firstly, evidence from animal models on the effect of fetal alcohol exposure on *POMC* epigenetic marks suggests the potential for transgenerational epigenetic inheritance via the male germline. Secondly, evidence from family trios in humans demonstrates a significant correlation between offspring PBC *POMC* methylation and paternal, but not maternal *POMC* methylation.

6.1 Fetal alcohol syndrome (FAS) and epigenetic inheritance

Recent studies have implicated a role for appetite regulating neuropeptides (including *POMC*) in alcohol dependence and craving(117)(118). *POMC* promoter methylation has been shown to be associated with craving in those with alcohol dependency(119,120). In human studies, paternal alcohol dependency has been associated with alterations in the hypothalamic-pituitary axis including changes to ACTH secretion (a derivative of *POMC*) in offspring (121,122). FAS is seen more frequently in the F2 generation of an alcohol abusing mother (i.e. her grandchildren) compared to F2 controls (123). Govorko et al (124) explored the possibility of intergenerational or transgenerational effects from alcohol exposure in a rat model by establishing two germlines: 1) breeding male fetal alcohol exposed rats and their male offspring with unexposed females and 2) breeding female fetal alcohol exposed rats and their female offspring with unexposed males. Hypermethylation of the *POMC* promoter and reduced *POMC* expression were seen in both female and male offspring in the F1 generation (of alcohol consuming mothers), but this pattern continued in male progeny in F2 and F3 from the male germline only. Thus a transgenerational effect was only seen via the male exposed germline. *POMC* promoter methylation was higher in sperm of male rats (F1-F3) from the male exposed germline suggesting a possible mechanism of epigenetic inheritance via methylation differences in the sperm. It is postulated that maternal fetal alcohol exposure to the developing progeny leads to transgenerational epigenetic transmission through the male germline thereafter (125).

Interestingly, maternal supplementation with choline (a one-carbon metabolite) altered the epigenome of the offspring such that *POMC* methylation and expression, and DNMT1 and Mecp2 levels were no different to non-alcohol exposed rats(126). It is known that alcohol interferes with

one-carbon metabolism(127) and therefore it seems intuitive to consider supplementation to normalise the levels. Additionally, giving the alcohol consuming pregnant rat either DNMT1 inhibitor or HDAC inhibitor reversed the methylation and expression changes in the offspring caused by prenatal alcohol exposure, suggesting a potential for future interventions(124). This demonstrates an interesting point of principle, albeit in rats, that changing the metabolome during gestation can alter the offspring epigenome in relation to *POMC* and mitigate the effects of alcohol. Were similar effects to be evident in humans, this could lead to a major new approach in preventative medicine with targeted maternal nutritional interventions to favourably influence the offspring's epigenome and break the intergenerational risk of diseases like obesity, as suggested by observed links between maternal concentrations of one-carbon metabolites, *POMC* methylation and obesity in humans(87).

6.2 Human family trios

Kühnen et al (87) examined 47 mother-father-offspring trios and demonstrated significant correlation with offspring PBC *POMC* methylation and the father's PBC *POMC* methylation but not the mother's, suggesting a potential intergenerational influence from the father. However, sperm methylation at this region was significantly lower than PBC, suggesting that the apparent paternal inheritance of epigenetic marks seen in the offspring was not mediated through sperm methylation. One potential mechanism is through modifications to sperm RNAs(128).

6.3 Y chromosome-linked patriline inheritance

Studies suggest a possible link between *POMC* expression, methylation and areas of the Y chromosome. The non-paring region of the Y chromosome (Y^{NPAR}) is exclusively transmitted between fathers and sons and includes functional genes such as SRY (sex determining region). Previous studies have suggested a possible interaction between SRY-androgen receptor binding and *POMC* methylation(119). A study in mice has demonstrated a significant Y^{NPAR} influence on brain β – endorphin (a derivative of *POMC*, see Figure 1) levels(129), suggesting a possible interaction with genetic polymorphisms in Y^{NPAR} on β – endorphin expression. Alternatively, it has been postulated that epigenetic changes on the Y^{NPAR} chromosome (for example caused by alcohol exposure) may influence *POMC* expression and/or methylation in the offspring and be a potential mechanism for epigenetic inheritance via the male line (130).

Conclusions

POMC is a key mediator of satiety and perturbations in the melanocortin system have been associated with dysregulation of energy balance. In animal models *POMC* gene methylation has been shown to be influenced by the prenatal and postnatal environment and associated with subsequent weight and appetite related phenotype in adulthood. In humans periconceptional nutrition has been associated with offspring methylation at *POMC*. Human studies have often demonstrated contradictory associations between *POMC* methylation and BMI and these appear to be dependent on the region of the gene studied. Therefore care should be exercised when selecting genomic regions for study. More prospective studies are needed to examine the influence of *POMC* DNA methylation on energy balance. Early studies suggest that *POMC* is an interesting candidate for exploring inter and transgenerational epigenetic inheritance in humans and future research should elucidate potential mechanisms for this.

There are potential clinical applications for using *POMC* epigenetic testing as a biomarker for early identification of obesity risk and as a predictor of response to obesity interventions. There are also potential pharmacological options with Setmelanotide, a MC4R agonist, demonstrating success in treating those with *POMC* deficiency(131), although it is yet to be established if this could prove an option for those with *POMC* hypermethylation. Looking to the future, a better understanding of nutritional factors influencing the epigenetic regulation of *POMC* could pave the way for maternal and paternal nutritional interventions that would provide a more favourable epigenotype, so reducing the risk of obesity in the next generation.

Funding statement

TC is funded by a Medical Research Council (UK) Clinical Research Training Fellowship. TC, MJS and AMP are supported by the MRC & the Department for International Development (DFID) under the MRC/DFID Concordat agreement (MRC Programme MC-A760-5QX00). This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

The images of the stomach and brain in figure 2 were made by Smashicons available from www.flaticon.com. The protein image in the graphical abstract was adapted from an icon from www.icons8.com. The rodent and human icons in the graphical abstract were made by koksikoks available from www.iconfinder.com.

Figures and tables

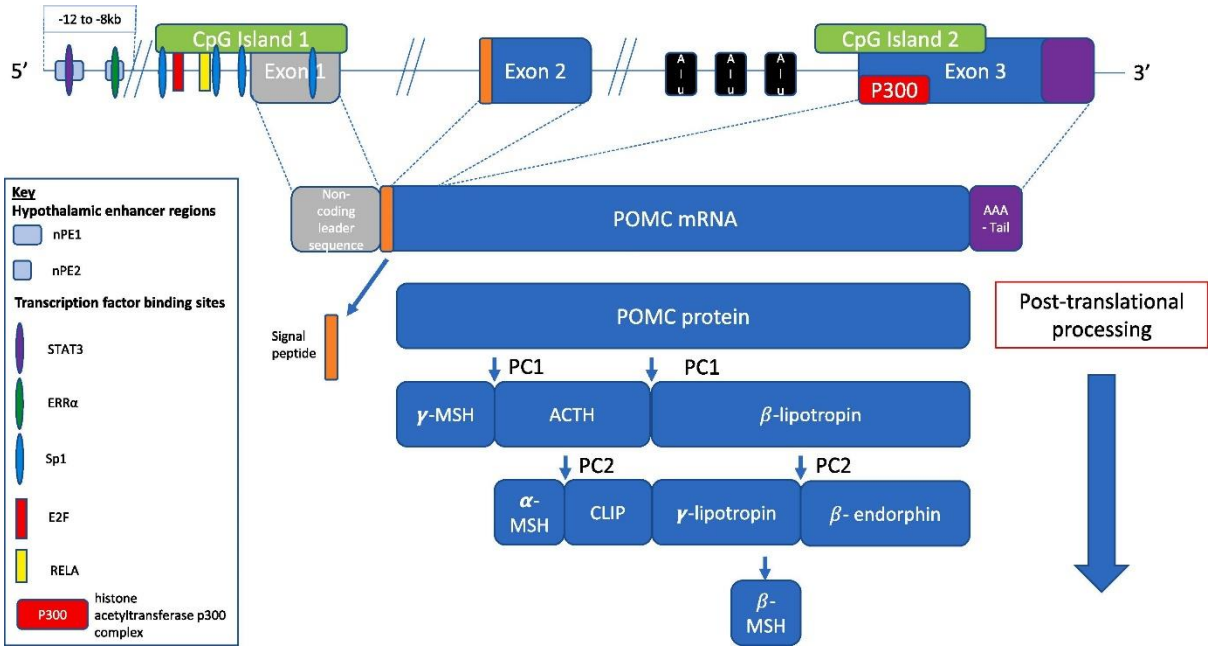


Figure 1. Human *POMC* Gene; transcription, translation and post-translational processes.

The human *POMC* gene consists of 3 exons and 2 large introns and is located between chromosome 2:25,383,722 to 25,391,722 (hg19, reverse strand). There are two CpG islands related to the *POMC* gene; the first in the promoter region and second over the boundary of intron2/exon3. Exon 1 (87 bp) contains a non-coding sequence and produces a short leader sequence that binds the ribosome at the start of translation. Exon 2 (152bp) gives rise to a small signal peptide and forms the N-terminal end of the *POMC* peptide. Exon 3 (835bp) produces the majority of the *POMC* peptide as well the signal for the addition of the poly-A tail. The figure provides schematic representation of the key transcriptional enhancers and binding sites related to leptin signalling and hypothalamic expression of neuropeptides (discussed in this article).

Key: CpG; cytosine-guanine dinucleotide, Alu; Alu element, P300; P300 complex binding domain, *POMC*; Proopiomelanocortin, PC1; Prohormone convertase 1, PC2; Prohormone convertase 2, -MSH; -melanocyte stimulating hormone, ACTH; Adrenocorticotropic hormone, CLIP; corticotropin-like intermediate peptide, AAA-tail; poly-A tail, nPE; neuro *POMC* enhancer, STAT3, Signal transducer and activator of transcription 3, ERRα; estrogen-related receptor alpha, Sp1; Specificity Protein 1, E2F; E2 Factor, RELA; v-rel reticuloendotheliosis viral oncogene homologue A (avian)

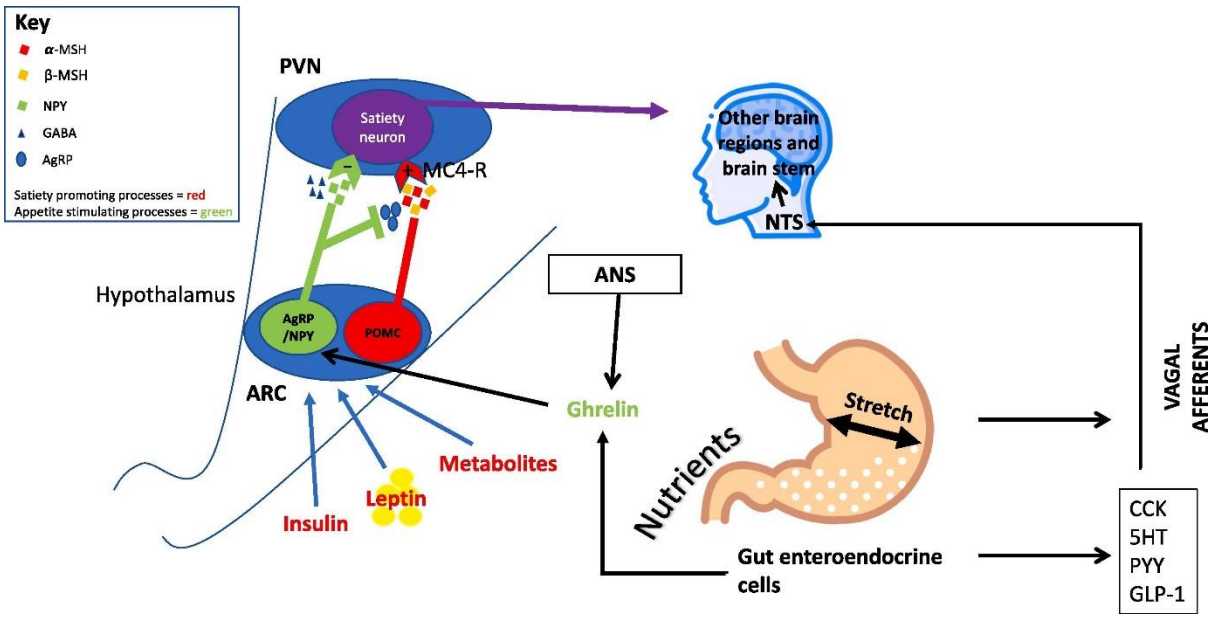


Figure 2. Schematic diagram of the melanocortin system and the control of appetite and satiety.

At the level of the hypothalamus, appetite and satiety regulating neurons (POMC, AgRP/NPY) of the arcuate nucleus (ARC) send projections to the paraventricular nucleus (PVN). The anorectic POMC expressing neurons are responsive to systemic signals including leptin, insulin and metabolites (such as glucose). α and β-MSH (derived from POMC) mediates the satiety signal via the action on MC4R. AgRP/NPY expressing neurons respond to ghrelin (which is predominantly under the control of the autonomic nervous system (ANS)), a hormone released by enteroendocrine cells that acts to increase appetite. AgRP antagonises the action of α-MSH at MC4R, whilst the neurotransmitters NPY and GABA convey an orexigenic signal via PVN neurons. Meal termination (satiation) is brought about via activation of vagal afferents from stomach stretch receptors and nutrient-induced release of enteroendocrine factors (CCK, 5HT, PYY, GLP-1). The vagal afferents send projections to the NTS (nucleus tractus solitarii) to bring about meal termination.

Abbreviations: POMC; Proopiomelanocortin, AgRP; Agouti-related peptide, NPY; neuropeptide-Y, α-MSH; alpha-melanocyte stimulating hormone, β-MSH; beta-melanocyte stimulating hormone, MC4R; melanocortin 4 receptors, GABA; gamma-aminobutyric acid, NTS; nucleus tractus solitarii, CCK; cholecystokinin, 5HT; 5-hydroxytryptamine, PYY; Peptide YY, GLP-1; glucagon-like peptide 1

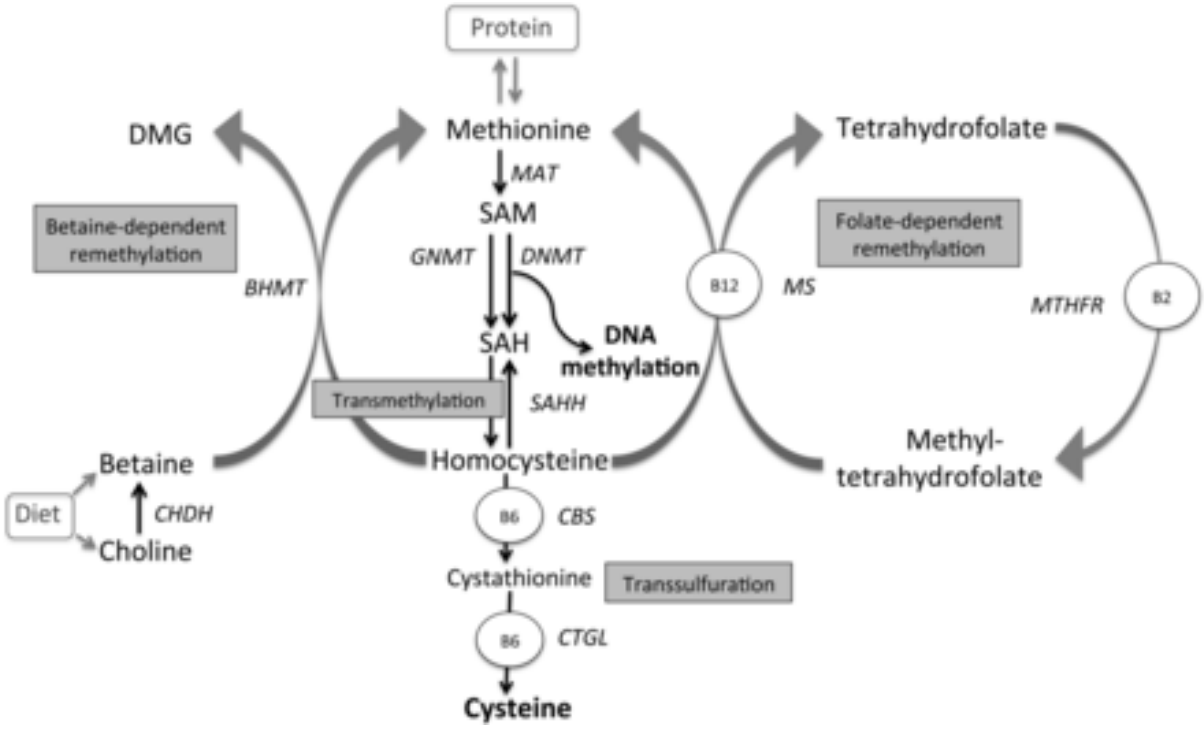


Figure 3. An overview of one-carbon metabolism.

Key: BHMT, betaine-homocysteine methyltransferase; B2, riboflavin; B6, vitamin B-6; B12, vitamin B-12; CBS, cystathionine-β-synthase; CHDH, choline dehydrogenase; CTGL, cystathionine-γ-lyase; DMG, dimethylglycine; DNMT, DNA methyltransferases; GNMT, glycine-N-methyltransferase; MAT, methionine adenosyltrasferase; MTHFR, methylenetetrahydrofolate reductase; MS, methionine synthase; SAH, S-adenosylhomocysteine; SAHH, S-adenosylhomocysteine hydrolase; SAM, S-adenosylmethionine.

Reproduced with permission from Paula Dominguez-Salas, Sophie E Moore, Darren Cole, Kerry-Ann da Costa, Sharon E Cox, Roger A Dyer, Anthony JC Fulford, Sheila M Innis, Robert A Waterland, Steven H Zeisel, Andrew M Prentice, Branwen J Hennig, DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women, *The American Journal of Clinical Nutrition*, Volume 97, Issue 6, June 2013, Pages 1217–1227, https://doi.org/10.3945/ajcn.112.048462.

Animal model	DNA Tissue source	Animal number*†	Prenatal Exposure	*POMC* epigenetic changes in offspring	Offspring phenotype	Reference
Rat	ARC	Offspring: n=7-11 per group (HFD-C & C-C)	Maternal Overnutrition	• ↑*POMC* promoter methylation persisted into adulthood	• Maternal HFD led to ↑birth weight, ↑calorific intake, ↑adult body weight, vulnerability to HFD	Marco et al, 2014(33)
Rat	Hypothalamus	Offspring: n=10 per group (C-C & FS-FS)	Maternal Overnutrition	• FS-FS diet led to ↓*POMC* methylation	• FS-FS led to ↑body weight into adulthood, impaired glucose handling and insulin resistance	Zheng et al, 2015(34)
Rat	ARC	Offspring: n=5-13 per group (C-C & HFD-C)	Maternal Overnutrition	• ↑*POMC* promoter methylation, ↑H3K9me2, ↑MBD1 & ↑SETDB1 • Lean controls had ↑5hme *POMC* promoter levels	• Maternal HFD led to ↑birth weight, ↑calorific intake, ↑adult body weight	Marco et al, 2016 (35)
Rat	ARC	Offspring: n=6 per group (Maternal HFD & C)	Maternal Overnutrition	• ↓DNMT1, ↓HDAC, ↓*POMC* expression in adulthood	• Maternal HFD led to ↑ adiposity, ↑ body weight	Desai et al, 2016 (38)
Rat	ARC and PVN	Offspring: n=6-22 per group (LFD & HFD)	Maternal Overnutrition	• ↑*POMC* promoter methylation persisted into adulthood	• Maternal HFD led to ↑ adiposity, ↑ body weight, vulnerability to HFD	Ramamoorthy et al, 2018 (32)
Sheep	Hypothalamus	Mothers: n=6-11 per group (CR & C)	Maternal Undernutrition	• ↓*POMC* promoter methylation, ↑histone acetylation of H3K9	• Not known	Stevens et al, 2010 (39)
Rat	Hypothalamus	Offspring: n = 40 per group (C-C, CR-C and CR-CR)	Maternal Undernutrition	• No changes in hypothalamic methylation in *POMC*	• ↓Birth weight and ↑ body weight by day 16 of life	Coupe et al, 2010 (42)
Sheep	ARC	Offspring: n=7-9 per group (C & underfed)	Maternal Undernutrition	• ↓*POMC* promoter methylation, ↑histone acetylation of H3K9, ↓DNMT1	• Not known	Begum et al, 2012 (41)
Rat	Hypothalamus	Offspring: n=10–14 per group (RV, HV, HA)	High Vitamin	• HA postnatal diet led to offspring ↓*POMC* promoter methylation	• Maternal HV diet led to ↑ body weight post weaning, ↑calorific intake – attenuated with HV or HA postnatal diet	Sánchez-Hernández et al, 2014 (52)
Rat	Hypothalamus	Offspring: n=8-14 per group (RV-RV, HV-RV, HV-HFol, HV-HV)	High Folate	• HFol-HFol diet led to ↓*POMC* promoter methylation, ↓*POMC* expression in adulthood	• Maternal HFol diet led to ↑calorific intake, ↑adult body weight, impaired glucose handling – attenuated with HFol postnatal diet	Cho et al, 2013 (54)
Rat	Hypothalamus	Offspring: n=6-8 per group (C & PNS)	Prenatal stress	• Interaction of methylation with PNS and postnatal HFD at CpG site in *POMC* promoter • HFD led to ↑*POMC* promoter methylation in non-stressed rats	• PNS led to ↑ adiposity, ↑insulin resistance	Paternain et al, 2012 (57)
Rat	ARC and PVN	Offspring: n=4-8 per group (BWMC, SEDC, RWC, ABA)	Prenatal stress	• No difference in *POMC* methylation between groups	• PNS rats showed ↓ food intake and ↑ weight loss when exposed to ABA	Boersma et al, 2017 (58)
Rat	ARC and PVN	Offspring: n=10 per group (Triclosan 4 or 8mg/kg/day & C)	Drug exposure (Triclosan)	• ↑*POMC* promoter methylation, ↓*POMC* expression	• Triclosan exposure associated with ↑Calorific intake, ↑adult body weight, metabolic syndrome in adulthood	Hua et al, 2018 (61)

Table 1. Animal models of prenatal exposures, epigenetic changes in *POMC* gene and associated phenotype

Key: ARC; arcuate nucleus of the hypothalamus, PVN; paraventricular nucleus of the hypothalamus, HFD; High Fat Diet, C;Control diet, FS; high fat, high sucrose diet, LFD; low fat diet, CR; calorie restricted diet, RV; recommended vitamin diet, HA; High Vitamin A diet, HV; High Vitamin diet, HFol High folate, BWMC; Body Weight Matched Controls, SEDC; Sedentary Controls, RWC; Running Wheel Controls, ABA; Activity Based Anorexia, PNS; Pre Natal Stress, MBD1; methyl binding protein 1, SETDB1; SET domain binding 1, 5hme; 5-hydroxymethylcytosine, , DNMT1; DNA methyltransferase 1, HDAC; histone deacetylase. † diets separated by a hyphen refer to diet pre and post natally * range of numbers in each experimental group reflects the different numbers used for different aspects of the experimental design

Animal model	DNA Tissue source	Animal number*	Postnatal Exposure	*POMC* epigenetic changes	Phenotype	Reference
Rat	Hypothalamus	n=9 per group (SL & C)	Energy excess	• SL ↑ *POMC* promoter methylation compared to controls • Methylation impeded the anorectic effects of leptin and insulin resulting in ↓ *POMC* expression	• SL rats demonstrated ↑ body weight, glucose, leptin and insulin by week of age	Plagemann et al, 2009 (63)
Rat	ARC	n=16-18 per group (HFD & C)	Energy excess	• HFD ↑ methylation at a Sp1 binding site which impeded Sp1 binding	• HFD fed rats from the neonatal period into adulthood demonstrated ↑ weight, ↑ insulin and ↑ leptin	Marco et al, 2013 (64)
Mouse	ARC	n=20 per group (LFD & HFD)	Energy excess	• HFD ↑ methylation at/near to RELA and Sp1 binding sites • HFD-induced chronic inflammation inhibits activation of *POMC* transcription	• Acute inflammation suppresses food intake	Shi et al, 2013 (67)
Rat	Hypothalamus	n=8 per group (HC & C)	Energy excess	• By day 16, HC rats ↓ *POMC* gene H3K9 gene acetylation and no difference in *POMC* promoter methylation compared to controls	• HC diet group had ↑ body weight, ↑ insulin and ↑ leptin by day 100.	Mahmood et al 2013 (68)
Rat	Hypothalamus	C group (n=6) HFD (n=50)	Energy excess	• DR rats had an overall ↓ *POMC* methylation compared to DIO rats	• DR rats consumed less than DIO rats, had similar weight to controls	Cifani et al, 2015 (70)
Rat	ARC and PVN	n=16 per group (CAF & C)	Energy excess	• CAF associated with ↓ *POMC* promoter methylation and ↑ expression of *POMC*	• CAF fed rats ↑body weight and ↑energy intake compared to controls	Lazzarino et al, 2017 (71)
Rat	Hypothalamus	n=9-12 per group (C, CR, SL, AL)	Energy restriction	• No differences in *POMC* methylation between the groups	• Post weaning CR group had comparable food intake and ↓ weight gain compared to controls when fed AL	Liu et al, 2013 (74)
Mouse	Hypothalamus	n=5 per group (CR, AL, CR-AL)	Energy restriction	• No differences in *POMC* methylation between the groups	• No comment	Unnikrishnan et al, 2017 (75)
Mouse	Adipose tissue and Hypothalamus	n=15 per group (n3 PUFA deficient, Soy Oil, Soy/Fish Oil, Fish Oil, C)	Fatty acid	• n-3 PUFA normalised *POMC* expression in obese mice • No differences in *POMC* methylation between the groups	• n-3 PUFA fed mice had ↓ body weight compared n-6 PUFA fed mice	Fan et al, 2011 (79)
Mouse	Hypothalamus	n=16 per group (LA & CLA)	Fatty acid	• CLA diet led to ↑*POMC* promoter methylation and ↓ *POMC* expression compared to LA fed group	• CLA fed mice ↑adult body weight and impaired glucose homeostasis compared to controls	Zhang et al, 2014 (82)
Rat	Hypothalamus	n=6-8 per group (Control: HFD & NF, Leptin: HFD & NF)	Leptin	• Leptin treated rats demonstrated ↑ *POMC* methylation when fed a normal diet and ↓ *POMC* methylation when fed HFD	• Leptin treatment led to ↓ body weight and ↓ energy intake in adulthood compared to controls	Palou et al, 2011 (83)

Table 2. Animal models of postnatal exposures, epigenetic changes in *POMC* gene and associated phenotype

Key: ARC; arcuate nucleus of the hypothalamus, PVN; paraventricular nucleus of the hypothalamus, SL; small litters, C;Control diet HFD; High Fat Diet, LFD; Low Fat Diet, HC; high carbohydrate, DIO; diet induced obesity, DR; diet resistant, CAF; Cafeteria Diet, CR; Calorie restricted diet, AL; ad libitum, PUFA; polyunsaturated fatty acids, LA; linoleic acid, CLA; Conjugated linoleic acid, NF; normal feeds, RELA; v-rel reticuloendotheliosis viral oncogene homologue A, Sp1; specificity protein 1, STAT3; Signal transducer and activator of transcription 3.

* range of numbers in each experimental group reflects the different numbers used for different aspects of the experimental design

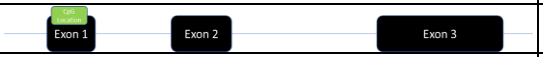
Study type and participants	DNA Tissue Source	Outcomes and summary of findings	Illustration of region of interest in *POMC* gene †	Reference
Case control Anorexia Nervosa; recovered (n=30) & acutely underweight (n=31) vs. normal weight adult women (n=30)	PBC	• Mean promotor *POMC* methylation was not different across nutritional states nor across disease groups • ↓*POMC* mRNA expression in those with malnutrition and hypoleptinaemia		Ehrlich et al, 2010 (107)
Case control Anorexia Nervosa; recovered (n=21) & acutely underweight (n=40) vs. normal weight adult women n=54	PBC	• Mean promotor *POMC* methylation was not different across nutritional states nor across disease groups but negatively associated with smoking		Ehrlich et al, 2012 (108)
Two Case control studies Obese vs. normal weight children (n=71 vs. n=36 & n=100 vs. n=54)	PBC	• Higher *POMC* methylation in obese vs. normal weight • *POMC* methylation stable in childhood • *POMC* methylation set early in development • *POMC* hypermethylation associated with reduced *POMC* expression		Kuhnen et al, 2012 (27)
Longitudinal cohort study Children from birth to mid-childhood (n=90)	PBC	• Higher triglycerides and insulin in the high *POMC* methylation group compared to mid and low *POMC* methylation groups • No difference in BMI or adiposity between groups • *POMC* methylation highly correlated from birth to mid childhood		Yoo et al, 2013 (97)
Case control study Adults following a weight loss intervention; 'regainers (n=7) vs non-regainers (n=11)'	PBC	• Higher methylation in regainers vs. non regainers • % weight regain correlated with % *POMC* methylation		Crujeiras et al, 2013 (103)
Case control study Adults obese (n=103) vs. normal weight (n=125) **Cross sectional study of cadaveric ARC samples and BMI** (n=41) **Brain and Kidney samples to assess *POMC* methylation across different germ layers (**n=16) **Longitudinal cohort** Methylation assessed from birth to adolescence (n=52) **Mother-Child pairs from Gambia** n=144 mother-child pairs	PBC ARC Brain & Kidney PBC PBC	• Positive correlation between BMI and *POMC* methylation in both PBC and ARC samples • Non-tissue specificity of *POMC* methylation • *POMC* methylation stable through childhood • *POMC* associated with one-carbon metabolites around conception		Kuhnen et al, 2016 (87)
Cross sectional Obese children (n=82)	PBC	• No correlation between *POMC* methylation and BMI		Acs et al, 2017 (100)
Randomised controlled trial Healthy normal weight adults, SFA (n=17) vs PUFA (n=14) diets	Adipose tissue	• PUFA but not SFA diet altered mean methylation at *POMC*	Illumina Human Methylation 450K array (26 Individual CpGs related to *POMC* gene, 6 significant CpGs †)	Perfilyev et al, 2017 (99)
Case control study Overweight/obese (n=41) vs. normal weight (n=79) children	PBC	• Lower methylation in overweight/obese groups compared to normal weight • Association between hypermethylation in CpG site 2 & lower HDL-cholesterol levels		Kwon et al, 2018 (101)

Table 3. Summary table of human studies of *POMC* gene methylation in relation to energy balance and metabolic disease outcomes

Key: ARC; arcuate nucleus of the hypothalamus, PBC; peripheral blood cells, *POMC*; Proopiomelanocortin, BMI; Body Mass Index, ARC; arcuate nucleus, CpG; cytosine-guanine dinucleotide, PUFA; polyunsaturated fatty acids SFA; saturated fatty acid, HDL; high density lipoprotein.

† Where reported, the genomic coordinates refer to hg19 genome build; **Ehrlich et al 2010 and 2012**: 68 CpGs between chr2:25,392,258 – 25,391,492 **, Kuhnen et al, 2012**: 10 CpGs between chr2:25,384,508-25,384,832, **Yoo et al, 2013**: 4 CpGs between chr2:25383999-25384108 **et al, 2013**: 2 CpGs demonstrated significance between chr2:25,391,046-25,391,545 (total of 52 CpGs in this genomic region), **Kuhnen 2016**: 9 CpGs between chr2:25,384,508-25,384,832, **Acs et al, 2017**: CpG Island in exon 1, **Perfilyev et al, 2017**: 6 CpGs at chr2:25,391,670, chr2:25,384,809, chr2:25,389,989, chr2:25,384,293, chr2:25,384,762, chr2:25,391,505 **Kwon et al, 2018**: 4 CpGs between chr2: 25383999-25384108

References

1. Toda C, Santoro A, Kim JD, Diano S. POMC Neurons: From Birth to Death. Annu Rev Physiol. 2017;79(1):209–36.

2. Cone RD. Anatomy and regulation of the central melanocortin system. Nat Neurosci. 2005;8(5):571–8.

3. Mountjoy KG. Pro-Opiomelanocortin (POMC) Neurones, POMC-Derived Peptides, Melanocortin Receptors and Obesity: How Understanding of this System has Changed Over the Last Decade. J Neuroendocrinol. 2015;27(6):406–18.

4. Balthasar N, Coppari R, McMinn J, Liu SM, Lee CE, Tang V, et al. Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. Neuron. 2004;42(6):983–91.

5. Parton LE, Ye CP, Coppari R, Enriori PJ, Choi B, Zhang CY, et al. Glucose sensing by POMC neurons regulates glucose homeostasis and is impaired in obesity. Nature. 2007;449(7159):228–32.

6. Qiu J, Zhang C, Borgquist A, Nestor CC, Smith AW, Bosch MA, et al. Insulin excites anorexigenic proopiomelanocortin neurons via activation of canonical transient receptor potential channels. Cell Metab. 2014;19(4):682–93.

7. Schioth HB, Bouifrouri AA, Rudzish R, Muceniece R, Watanobe H, Wikberg JES, et al. Pharmacological comparison of rat and human melanocortin 3 and 4 receptors in vitro. Regul Pept. 2002 Jun;106(1–3):7–12.

8. Biebermann H, Castañeda TR, van Landeghem F, von Deimling A, Escher F, Brabant G, et al. A role for β-melanocyte-stimulating hormone in human body-weight regulation. Cell Metab. 2006;3(2):141–6.

9. Andermann ML, Lowell BB. Toward a Wiring Diagram Understanding of Appetite Control. Neuron. 2017;95(4):757–78.

10. Krude H, Biebermann H, Luck W, Horn R, Brabant G, Gruters A. Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. Nat Genet. 1998 Jun;19(2):155–7.

11. Farooqi IS, Drop S, Clements A, Keogh JM, Biernacka J, Lowenbein S, et al. Heterozygosity for a POMC-null mutation and increased obesity risk in humans. Diabetes. 2006;55(9):2549–53.

12. Krude H, Biebermann H, Schnabel D, Tansek MZ, Theunissen P, Mullis PE, et al. Obesity Due

to Proopiomelanocortin Deficiency: Three New Cases and Treatment Trials with Thyroid Hormone and ACTH4-10. J Clin Endocrinol Metab. 2003;88(10):4633–40.

13. Gibney ER, Nolan CM. Epigenetics and gene expression. Heredity (Edinb). 2010 May 12;105(1):4–13.

14. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012 May 29;13(7):484–92.

15. Riggs AD, Porter TN. Overview of Epigenetic Mechanisms. In: Epigenetic Mechanisms of Gene Regulation. 1996. p. 29–45.

16. Hannon E, Knox O, Sugden K, Burrage J, Wong CCY, Belsky DW, et al. Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. PLoS Genet. 2018;14(8):e1007544.

17. Jaenisch R, Bird A. Epigenetic regulation of gene expression : how the genome integrates intrinsic and environmental signals. 2003;33(march):245–54.

18. James P, Sajjadi S, Tomar AS, Saffari A, Fall CHD, Prentice AM, et al. Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism. Int J Epidemiol. 2018;47(6):1910–37.

19. Takahashi H, Hakamata Y, Watanabe Y, Kikuno R, Miyata T, Numa S. Complete nucleotide sequence of the human corticotropin-beta-lipotropin precursor gene. Nucleic Acids Res. 1983 Oct 11;11(19):6847–58.

20. Kuehnen P, Krude H. Alu elements and human common diseases like obesity. Mob Genet Elements. 2012;2:4:197–201.

21. de Souza FSJ, Santangelo AM, Bumaschny V, Avale ME, Smart JL, Low MJ, et al. Identification of Neuronal Enhancers of the Proopiomelanocortin Gene by Transgenic Mouse Analysis and Phylogenetic Footprinting. Mol Cell Biol. 2005;25(8):3076–86.

22. Ladyman SR, Grattan DR. JAK-STAT and feeding. JAK-STAT. 2013;2(2):e23675.

23. Yang G, Lim CY, Li C, Radda GK, Li C, Cao X, et al. FoxO1 inhibits leptin regulation of pro-opiomelanocortin promoter activity by blocking STAT3 interaction with specificity protein 1. J Biol Chem. 2009;284(6):3719–27.

24. Drouin J. 60 YEARS OF POMC: Transcriptional and epigenetic regulation of POMC gene

expression. J Mol Endocrinol. 2016;56(4):T99–112.

25. Hill JW. Gene Expression and the Control of Food Intake by Hypothalamic POMC/CART Neurons. Open Neuroendocrinol J. 2010;3:21–7.

26. Bumaschny VF, de Souza FSJ, López Leal RA, Santangelo AM, Baetscher M, Levi DH, et al. Transcriptional Regulation of Pituitary POMC Is Conserved at the Vertebrate Extremes Despite Great Promoter Sequence Divergence. Mol Endocrinol. 2007;21(11):2738–49.

27. Kuehnen P, Mischke M, Wiegand S, Sers C, Horsthemke B, Lau S, et al. An alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. PLoS Genet. 2012;8(3):1–12.

28. Lacaze-Masmonteil T, de Keyzer Y, Luton JP, Kahn A, Bertagna X. Characterization of proopiomelanocortin transcripts in human nonpituitary tissues. Proc Natl Acad Sci. 1987;84(20):7261–5.

29. Gluckman PD, Hanson MA. Living with the past: Evolution, development, and patterns of disease. Science (80-). 2004;305(5691):1733–6.

30. Gluckman PD, Hanson MA, Spencer HG. Predictive adaptive responses and human evolution. Trends Ecol Evol. 2005;20(10):527–33.

31. Guerrero-Bosagna C. Evolution with No Reason: A Neutral View on Epigenetic Changes, Genomic Variability, and Evolutionary Novelty. Bioscience. 2017 May 1;67(5):469–76.

32. Gali Ramamoorthy T, Allen TJ, Davies A, Harno E, Sefton C, Murgatroyd C, et al. Maternal overnutrition programs epigenetic changes in the regulatory regions of hypothalamic Pomc in the offspring of rats. Int J Obes. 2018;42(8):1431–44.

33. Marco A, Kisliouk T, Tabachnik T, Meiri N, Weller A. Overweight and CpG methylation of the Pomc promoter in offspring of high-fat-diet-fed dams are not “reprogrammed” by regular chow diet in rats. FASEB J Off Publ Fed Am Soc Exp Biol. 2014 Sep;28(9):4148–57.

34. Zheng J, Xiao X, Zhang Q, Yu M, Xu J, Wang Z, et al. Maternal and post-weaning high-fat, high-sucrose diet modulates glucose homeostasis and hypothalamic POMC promoter methylation in mouse offspring. Metab Brain Dis. 2015;30(5):1129–37.

35. Marco A, Kisliouk T, Tabachnik T, Weller A, Meiri N. DNA CpG Methylation (5-Methylcytosine) and Its Derivative (5-Hydroxymethylcytosine) Alter Histone Posttranslational Modifications at the Pomc Promoter, Affecting the Impact of Perinatal Diet on Leanness and Obesity of the

Offspring. Diabetes. 2016 Aug;65(8):2258–67.

36. Shi D-Q, Ali I, Tang J, Yang W-C. New Insights into 5hmC DNA Modification: Generation, Distribution and Function. Front Genet. 2017 Jul 19;8:100.

37. Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell. 2012 Dec;151(7):1417–30.

38. Desai M, Han G, Ross MG. Programmed hyperphagia in offspring of obese dams: Altered expression of hypothalamic nutrient sensors, neurogenic factors and epigenetic modulators. Appetite. 2016;99:193–9.

39. Stevens A, Begum G, Cook A, Connor K, Rumball C, Oliver M, et al. Epigenetic changes in the hypothalamic proopiomelanocortin and glucocorticoid receptor genes in the ovine fetus after periconceptional undernutrition. Endocrinology. 2010;151(8):3652–64.

40. Todd SE, Oliver MH, Jaquiery AL, Bloomfield FH, Harding JE. Periconceptional undernutrition of ewes impairs glucose tolerance in their adult offspring. Pediatr Res. 2009;65(4):409–13.

41. Begum G, Stevens A, Smith EB, Connor K, Challis JRG, Bloomfield F, et al. Epigenetic changes in fetal hypothalamic energy regulating pathways are associated with maternal undernutrition and twinning. FASEB J. 2012;26(4):1694–703.

42. Coupé B, Amarger V, Grit I, Benani A, Parnet P. Nutritional programming affects hypothalamic organization and early response to leptin. Endocrinology. 2010;151(2):702–13.

43. Begum G, Davies A, Stevens A, Oliver M, Jaquiery A, Challis J, et al. Maternal undernutrition programs tissue-specific epigenetic changes in the glucocorticoid receptor in adult offspring. Endocrinology. 2013;154(12):4560–9.

44. Óvilo C, González-Bulnes A, Benítez R, Ayuso M, Barbero A, Pérez-Solana ML, et al. Prenatal programming in an obese swine model: Sex-related effects of maternal energy restriction on morphology, metabolism and hypothalamic gene expression. Br J Nutr. 2014;111(4):735–46.

45. Chowen JA, Freire-Regatillo A, Argente J. Neurobiological characteristics underlying metabolic differences between males and females. Prog Neurobiol. 2019;(179):18–32.

46. Xu Y, Nedungadi TP, Zhu L, Sobhani N, Irani BG, Davis KE, et al. Distinct hypothalamic neurons mediate estrogenic effects on energy homeostasis and reproduction. Cell Metab. 2011;14(4):453–65.

47. Ramamoorthy TG, Begum G, Harno E, White A. Developmental programming of hypothalamic

neuronal circuits: Impact on energy balance control. Front Neurosci. 2015;9:126.

48. Friso S, Udali S, De Santis D, Choi S-W. One-carbon metabolism and epigenetics. Mol Aspects Med. 2017 Apr;54:28–36.

49. Hoyo C, Daltveit AK, Iversen E, Benjamin-Neelon SE, Fuemmeler B, Schildkraut J, et al. Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. Epigenetics. 2014 Aug;9(8):1120–30.

50. Cho CE, Sanchez-Hernandez D, Reza-Lopez SA, Huot PSP, Kim Y-I, Anderson GH, et al. Obesogenic phenotype of offspring of dams fed a high multivitamin diet is prevented by a post-weaning high multivitamin or high folate diet. Int J Obes. 2013;37(9):1177–82.

51. Sanchez-Hernandez D, Poon AN, Kubant R, Kim H, Huot PSP, Cho CE, et al. A gestational diet high in fat-soluble vitamins alters expression of genes in brain pathways and reduces sucrose preference, but not food intake, in Wistar male rat offspring. Appl Physiol Nutr Metab. 2015;40(4):424–31.

52. Sánchez-Hernández D, Cho CE, Kubant R, Reza-López SA, Poon AN, Wang J, et al. Increasing vitamin A in post-weaning diets reduces food intake and body weight and modifies gene expression in brains of male rats born to dams fed a high multivitamin diet. J Nutr Biochem. 2014;25(10):991–6.

53. Reeves PG. Components of the AIN-93 diets as improvements in the AIN-76A diet. J Nutr. 1997 May;127(5 Suppl):838S-841S.

54. Cho CE, Sánchez-Hernández D, Reza-López SA, Huot PSP, Kim YI, Anderson GH. High folate gestational and post-weaning diets alter hypothalamic feeding pathways by DNA methylation in Wistar rat offspring. Epigenetics. 2013;8(7):710–9.

55. Rowling MJ, McMullen MH, Schalinske KL. Vitamin A and its derivatives induce hepatic glycine N-methyltransferase and hypomethylation of DNA in rats. JNutr. 2002;132(3):365–9.

56. Tamashiro KLK, Moran TH. Perinatal environment and its influences on metabolic programming of offspring. Physiol Behav. 2010;100(5):560–6.

57. Paternain L, Batlle MA, De La Garza AL, Milagro FI, Martínez JA, Campión J. Transcriptomic and epigenetic changes in the hypothalamus are involved in an increased susceptibility to a high-fat-sucrose diet in prenatally stressed female rats. Neuroendocrinology. 2012;96(3):249–60.

58. Boersma GJ, Liang N, Lee RS, Albertz JD, Kastelein A, Moody LA, et al. Failure to upregulate Agrp and Orexin in response to activity based anorexia in weight loss vulnerable rats characterized by passive stress coping and prenatal stress experience. Psychoneuroendocrinology. 2016;67:171–81.

59. US Food and Drug Administration. 5 Things to Know About Triclosan [Internet]. [cited 2018 Dec 5]. Available from: https://www.fda.gov/ForConsumers/ConsumerUpdates/ucm205999.htm

60. Etzel TM, Calafat AM, Ye X, Chen A, Lanphear BP, Savitz DA, et al. Urinary triclosan concentrations during pregnancy and birth outcomes. Environ Res. 2017;156:505–11.

61. Hua X, Xiong J-W, Zhang Y-J, Cao X-Y, Sun P, Wu J, et al. Exposure of pregnant mice to triclosan causes hyperphagic obesity of offspring via the hypermethylation of proopiomelanocortin promoter. Arch Toxicol. 2019 Feb;93(2):547–58.

62. Ojha S, Saroha V, Symonds ME, Budge H. Excess nutrient supply in early life and its later metabolic consequences. Clin Exp Pharmacol Physiol. 2013 Nov;40(11):817–23.

63. Plagemann A, Harder T, Brunn M, Harder A, Roepke K, Wittrock-Staar M, et al. Hypothalamic proopiomelanocortin promoter methylation becomes altered by early overfeeding: An epigenetic model of obesity and the metabolic syndrome. J Physiol. 2009;587(20):4963–76.

64. Marco A, Kisliouk T, Weller A, Meiri N. High fat diet induces hypermethylation of the hypothalamic Pomc promoter and obesity in post-weaning rats. Psychoneuroendocrinology. 2013;38(12):2844–53.

65. Grossberg AJ, Scarlett JM, Marks DL. Hypothalamic mechanisms in cachexia. Physiol Behav. 2010 Jul;100(5):478–89.

66. Ellulu MS, Patimah I, Khaza'ai H, Rahmat A, Abed Y. Obesity and inflammation: the linking mechanism and the complications. Arch Med Sci. 2016/03/31. 2017 Jun;13(4):851–63.

67. Shi X, Wang X, Li Q, Su M, Chew E, Wong ET, et al. Nuclear factor κb (NF-κB) suppresses food intake and energy expenditure in mice by directly activating the Pomc promoter. Diabetologia. 2013;56(4):925–36.

68. Mahmood S, Smiraglia DJ, Srinivasan M, Patel MS. Epigenetic changes in hypothalamic appetite regulatory genes may underlie the developmental programming for obesity in rat neonates subjected to a high-carbohydrate dietary modification. J Dev Orig Health Dis. 2013;4(6):479–90.

69. Stöger R. Epigenetics and obesity. Pharmacogenomics. 2008 Dec;9(12):1851–60.

70. Cifani C, Micioni Di Bonaventura M V., Pucci M, Giusepponi ME, Romano A, Di Francesco A, et al. Regulation of hypothalamic neuropeptides gene expression in diet induced obesity resistant rats: possible targets for obesity prediction? Front Neurosci. 2015;9(June):1–9.

71. Lazzarino GP, Andreoli MF, Rossetti MF, Stoker C, Tschopp MV, Luque EH, et al. Cafeteria diet differentially alters the expression of feeding-related genes through DNA methylation mechanisms in individual hypothalamic nuclei. Mol Cell Endocrinol. 2017;450:113–25.

72. Lean MEJ, Leslie WS, Barnes AC, Brosnahan N, Thom G, McCombie L, et al. Primary care-led weight management for remission of type 2 diabetes (DiRECT): An open-label, cluster-randomised trial. Lancet. 2018;391(10120):541–51.

73. Willi SM, Martin K, Datko FM, Brant BP. Treatment of type 2 diabetes in childhood using a very-low-calorie diet. Diabetes Care. 2004 Feb;27(2):348–53.

74. Liu H-W, Srinivasan M, Mahmood S, Smiraglia DJ, Patel MS. Adult-onset obesity induced by early life overnutrition could be reversed by moderate caloric restriction. AJP Endocrinol Metab. 2013;305(7):E785–94.

75. Unnikrishnan A, Jackson J, Matyi SA, Hadad N, Wronowski B, Georgescu C, et al. Role of DNA methylation in the dietary restriction mediated cellular memory. GeroScience. 2017;39(3):331–45.

76. Ailhaud G, Massiera F, Weill P, Legrand P, Alessandri J-M, Guesnet P. Temporal changes in dietary fats: role of n-6 polyunsaturated fatty acids in excessive adipose tissue development and relationship to obesity. Prog Lipid Res. 2006 May;45(3):203–36.

77. Ukropec J, Reseland JE, Gasperikova D, Demcakova E, Madsen L, Berge RK, et al. The hypotriglyceridemic effect of dietary n-3 FA is associated with increased beta-oxidation and reduced leptin expression. Lipids. 2003 Oct;38(10):1023–9.

78. Myers MG, Heymsfield SB, Haft C, Kahn BB, Laughlin M, Leibel RL, et al. Challenges and opportunities of defining clinical leptin resistance. Cell Metab. 2012;15(2):150–6.

79. Fan C, Liu X, Shen W, Deckelbaum RJ, Qi K. The regulation of leptin, leptin receptor and pro-opiomelanocortin expression by N-3 PUFAs in diet-induced obese mice is not related to the methylation of their promoters. Nutr Metab. 2011;8:1–9.

80. Tsuboyama-Kasaoka N, Takahashi M, Tanemura K, Kim HJ, Tange T, Okuyama H, et al.

Conjugated linoleic acid supplementation reduces adipose tissue by apoptosis and develops lipodystrophy in mice. Diabetes. 2000;49(9):1534–42.

81. Poulos SP, Sisk M, Hausman DB, Azain MJ, Hausman GJ. Pre- and postnatal dietary conjugated linoleic acid alters adipose development, body weight gain and body composition in Sprague-Dawley rats. J Nutr. 2001;131(10):2722–31.

82. Zhang X, Yang R, Jia Y, Cai D, Zhou B, Qu X, et al. Hypermethylation of Sp1 binding site suppresses hypothalamic POMC in neonates and may contribute to metabolic disorders in adults: impact of maternal dietary CLAs. Diabetes. 2014 May;63(5):1475–87.

83. Palou M, Picó C, McKay JA, Sánchez J, Priego T, Mathers JC, et al. Protective effects of leptin during the suckling period against later obesity may be associated with changes in promoter methylation of the hypothalamic pro-opiomelanocortin gene. Br J Nutr. 2011;106(5):769–78.

84. Wu Y, Patchev A V., Daniel G, Almeida OFX, Spengler D. Early-Life stress reduces dna methylation of the pomc gene in male mice. Endocrinology. 2014;155(5):1751–62.

85. Gangisetty O, Bekdash R, Maglakelidze G, Sarkar DK. Fetal Alcohol Exposure Alters Proopiomelanocortin Gene Expression and Hypothalamic-Pituitary-Adrenal Axis Function via Increasing MeCP2 Expression in the Hypothalamus. PLoS One. 2014;9(11):e113228.

86. Wang X, Lacza Z, Sun YE, Han W. Leptin resistance and obesity in mice with deletion of methyl-CpG-binding protein 2 (MeCP2) in hypothalamic pro-opiomelanocortin (POMC) neurons. Diabetologia. 2014;57(1):236–45.

87. Kühnen P, Handke D, Waterland RA, Hennig BJ, Silver M, Fulford AJ, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. Cell Metab. 2016;24(3):502–9.

88. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. Trends Genet. 2002;18(7):348–51.

89. Waterland R a, Jirtle RL. Transposable Elements : Targets for Early Nutritional Effects on Epigenetic Gene Regulation. Mol Cell Biol. 2003;23(15):5293–300.

90. Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG. Maternal methyl supplements increase offspring DNA methylation at Axin fused. Genesis. 2006;44(9):401–6.

91. Finer S, Iqbal MS, Lowe R, Ogunkolade BW, Pervin S, Mathews C, et al. 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. 2016 Nov 1;6(11):e011768.

92. Silver MJ, Kessler NJ, Hennig BJ, Dominguez-Salas P, Laritsky E, Baker MS, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. Genome Biol. 2015;16:118.

93. Dominguez-Salas P, Moore SE, Baker MS, Bergen AW, Cox SE, Dyer RA, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. Nat Commun. 2014;5:3746.

94. Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. Sci Adv. 2018;4(7):eaat2624.

95. Price AL, Eskin E, Pevzner PA. Whole-genome analysis of Alu repeat elements reveals complex evolutionary history. Genome Res. 2004;14(11):2245–52.

96. Morgan HD, Sutherland HGE, Martin DIK, Whitelaw E. Epigenetic inheritance at the agouti locus in the mouse. Nat Genet. 1999;23(3):314–8.

97. Yoo JY, Lee S, Lee HA, Park H, Park YJ, Ha EH, et al. Can proopiomelanocortin methylation be used as an early predictor of metabolic syndrome? Diabetes Care. 2014;37(3):734–9.

98. Rosqvist F, Iggman D, Kullberg J, Cedernaes J, Johansson HE, Larsson A, et al. Overfeeding polyunsaturated and saturated fat causes distinct effects on liver and visceral fat accumulation in humans. Diabetes. 2014;63(7):2356–68.

99. Perfilyev A, Dahlman I, Gillberg L, Rosqvist F, Iggman D, Volkov P, et al. Impact of polyunsaturated and saturated fat overfeeding on the DNA-methylation pattern in human adipose tissue: A randomized controlled trial. Am J Clin Nutr. 2017;105(4):991–1000.

100. Ács O, Péterfia B, Hollósi P, Luczay A, Török D, Szabó A. Methylation Status of CYP27B1 and IGF2 Correlate to BMI SDS in Children with Obesity. Obes Facts. 2017;10(4):353–62.

101. Kwon EJ, You YA, Park B, Ha EH, Kim HS, Park H, et al. Association between the DNA methylations of POMC, MC4R, and HNF4A and metabolic profiles in the blood of children aged 7-9 years. BMC Pediatr. 2018;18(1):1–8.

102. Perez-Martinez P, Garcia-Rios A, Delgado-Lista J, Delgado-Casado N, Malagon MM, Marin C, et al. A variant near the melanocortin-4 receptor gene regulates postprandial lipid metabolism in a healthy Caucasian population. Br J Nutr. 2011;106(4):468–71.

103. Crujeiras AB, Campion J, Díaz-Lagares A, Milagro FI, Goyenechea E, Abete I, et al. Association of weight regain with specific methylation levels in the NPY and POMC promoters in leukocytes of obese men: A translational study. Regul Pept. 2013;186:1–6.

104. Castro J, Deulofeu R, Gila A, Puig J, Toro J. Persistence of Nutritional Deficiencies after Short-term Weight Recovery in Adolescents with Anorexia Nervosa. Int J Eat Disord. 2004;35(2):169–78.

105. Frieling H, Römer K, Röschke B, Bönsch D, Wilhelm J, Fiszer R, et al. Homocysteine plasma levels are elevated in females with anorexia nervosa. J Neural Transm. 2005;112(7):979–85.

106. Frieling H, Gozner A, Römer KD, Lenz B, Bönsch D, Wilhelm J, et al. Global DNA hypomethylation and DNA hypermethylation of the alpha synuclein promoter in females with anorexia nervosa. Mol Psychiatry. 2007 Feb 27;12:229.

107. Ehrlich S, Weiss D, Burghardt R, Infante-Duarte C, Brockhaus S, Muschler MA, et al. Promoter specific DNA methylation and gene expression of POMC in acutely underweight and recovered patients with anorexia nervosa. J Psychiatr Res. 2010;44(13):827–33.

108. Ehrlich S, Walton E, Roffman JL, Weiss D, Puls I, Doehler N, et al. Smoking, but not malnutrition, influences promoter-specific DNA methylation of the proopiomelanocortin gene in patients with and without anorexia nervosa. Can J Psychiatry. 2012;57(3):168–76.

109. Mineur YS, Abizaid A, Rao Y, Salas R, DiLeone RJ, Gündisch D, et al. Nicotine decreases food intake through activation of POMC neurons. Science (80-). 2011;332(6035):1330–2.

110. Huang H, Xu Y, van den Pol AN. Nicotine excites hypothalamic arcuate anorexigenic proopiomelanocortin neurons and orexigenic neuropeptide Y neurons: similarities and differences. J Neurophysiol. 2011;106(3):1191–202.

111. Chiolero A, Faeh D, Paccaud F, Cornuz J. Consequences of smoking for body weight, body fat distribution, and insulin resistance. Am J Clin Nutr. 2008;87(4):801–9.

112. Filozof C, Fernández Pinilla MC, Fernández-Cruz A. Smoking cessation and weight gain. Obes Rev. 2004;5(2):95–103.

113. Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: A review and recommendations. Epigenomics. 2017;9(5):757–68.

114. Martin EM, Fry RC. Environmental Influences on the Epigenome: Exposure- Associated DNA Methylation in Human Populations. Annu Rev Public Heal. 2018;39:309–33.

115. Heard E, Martienssen RA. Transgenerational Epigenetic Inheritance : Myths and Mechanisms. Cell. 2014;157(1):95–109.

116. Horsthemke B. A critical view on transgenerational epigenetic inheritance in humans. Nat Commun. 2018;9(1):2973.

117. Hillemacher T, Kahl KG, Heberlein A, Muschler MA, Eberlein C, Frieling H, et al. Appetite- and volume-regulating neuropeptides: Role in treating alcohol dependence. Curr Opin Investig Drugs. 2010 Oct;11(10):1097–106.

118. Fortuna JL. Sweet preference, sugar addiction and the familial history of alcohol dependence: shared neural pathways and genes. J Psychoactive Drugs. 2010 Jun;42(2):147–51.

119. Muschler MAN, Lenz B, Hillemacher T, Kraus C, Kornhuber J, Frieling H, et al. CAGn repeat of the androgen receptor is linked to proopiomelanocortin promoter methylation - Relevance for craving of male alcohol-dependent patients? Psychopharmacology (Berl). 2014;231(10):2059–66.

120. Muschler MAN, Hillemacher T, Kraus C, Kornhuber J, Bleich S, Frieling H. DNA methylation of the POMC gene promoter is associated with craving in alcohol dependence. J Neural Transm. 2010;117(4):513–9.

121. Hernandez-Avila CA, Oncken C, Van Kirk J, Wand G, Kranzler HR. Adrenocorticotropin and cortisol responses to a naloxone challenge and risk of alcoholism. Biol Psychiatry. 2002 Apr;51(8):652–8.

122. Zimmermann U, Spring K, Wittchen H-U, Himmerich H, Landgraf R, Uhr M, et al. Arginine vasopressin and adrenocorticotropin secretion in response to psychosocial stress is attenuated by ethanol in sons of alcohol-dependent fathers. J Psychiatr Res. 2004;38(4):385–93.

123. Kvigne VL, Leonardson GR, Borzelleca J, Welty TK. Characteristics of grandmothers who have grandchildren with fetal alcohol syndrome or incomplete fetal alcohol syndrome. Matern Child Health J. 2008 Nov;12(6):760–5.

124. Govorko D, Bekdash RA, Zhang C, Sarkar DK. Male germline transmits fetal alcohol adverse effect on hypothalamic proopiomelanocortin gene across generations. Biol Psychiatry. 2012;72(5):378–88.

125. Mead EA, Sarkar DK. Fetal alcohol spectrum disorders and their transmission through genetic and epigenetic mechanisms. Front Genet. 2014;5(JUN):1–10.

126. Bekdash RA, Zhang C, Sarkar DK. Gestational choline supplementation normalized fetal alcohol-induced alterations in histone modifications, DNA methylation, and proopiomelanocortin (POMC) gene expression in β-endorphin-producing POMC neurons of the hypothalamus. Alcohol Clin Exp Res. 2013;37(7):1133–42.

127. Fowler AK, Hewetson A, Agrawal RG, Dagda M, Dagda R, Moaddel R, et al. Alcohol-induced one-carbon metabolism impairment promotes dysfunction of DNA base excision repair in adult brain. J Biol Chem. 2012;287(52):43533–42.

128. Chen Q, Yan W, Duan E. Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. Nat Rev Genet. 2016;17(12):733–43.

129. Botbol M, Roubertoux PL, Carlier M, Trabado S, Brailly-Tabard S, Perez-Diaz F, et al. Modulation of brain β-endorphin concentration by the specific part of the Y chromosome in mice. PLoS One. 2011;6(3):e16704.

130. Sarkar DK. Male germline transmits fetal alcohol epigenetic marks for multiple generations: A review. Addict Biol. 2016;21(1):23–34.

131. Kühnen P, Clément K, Wiegand S, Blankenstein O, Gottesdiener K, Martini LL, et al. Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. N Engl J Med. 2016;375(3):240–6.

The literature review relating to periconceptional influences on *PAX8* methylation and phenotype is much more limited and therefore the pertinent studies are summarised in Chapter 1 (section 1.1.7 and 1.3) and detailed further in Chapter 7.

Chapter 4 POMC methylation influence on weight and adiposity

Summary of the chapter

In this chapter, I report the seasonal change in body weight and fat mass index (FMI) in both women and children. Using cosinor modelling, I model the effect of *POMC* methylation on seasonal changes in weight and fat mass. From the cosinor model for both body weight and FMI, I derived the mean value over the year, the magnitude of change over the year and the timing of the peak of change. Furthermore, secondary analysis explored the effect of *POMC* methylation of leptin and DXA-derived fat mass index.

4.1 Introduction

As described in earlier chapters (see chapters 1,2 and 3), POMC sits as a central component of the melanocortin system and regulates appetite and satiety[1]. Homozygous gene mutations in POMC are associated with early onset obesity in both animals[2] and humans[3–5]. Those with heterozygous mutations have an intermediate overweight phenotype suggesting a gene dosage effect[6]. In humans, *POMC* variably methylated region (VMR, referring to the genomic region studied in this work) hypermethylation is associated with lower peripheral blood cell (PBC) POMC expression[7]. POMC mediates the satiety response via the actions of α- and β-MSH on melanocortin 4 receptors (MC4R) in the paraventricular nucleus (PVN) of the hypothalamus[8]. In humans, the POMC VMR has been identified as a putative ME with systemic methylation i.e. consistent methylation pattern across [7,9]. Therefore assessment of methylation in peripheral bloods cells (PBC) can be a proxy for methylation at the hypothalamus[7].

Hypermethylation of the *POMC* VMR is associated with obesity in children[7] and adults[9] from Germany. This hypermethylated variant is thought to be stable through childhood and predates the development of obesity during adolescence (suggesting that the methylation variant is not caused by obesity in these individuals). Hypermethylation in this region is thought to impede p300 (part of acetyltransferase complex) binding and thus influence chromatin modelling and gene expression[7].

Most cases of obesity develop over many years[10]. Small imbalances of energy intake and expenditure are associated with a gradual increase in weight e.g. for 1kg per year increase in weight would need only a +71kj/day energy imbalance[11]. High weight gain has been reported to be in excess of a 2.5% annual increase in body weight[12]. In most settings, to prospectively examine the effect of *POMC* methylation of body weight changes over time would take many years. The seasonal changes in energy balance witnessed in the Gambia provide an exciting experimental model. The rainy season (July-October) is associated with increase agricultural workload, depleted food supplies and higher prevalence of infectious disease. The dry season (October-June) coincides with the agricultural harvest and as such is associated with increased food availability. These seasonal changes are associated with well documented fluctuations in energy balance and weight (see Figure 1.8). These seasonal changes in energy balance provides an exciting opportunistic experimental model to interrogate how *POMC* methylation influences weight gain and loss brought about by the seasons. This model tests the central hypothesis that increasing *POMC* methylation, by reducing *POMC* expression (and reducing satiety signal), will be associated more weight/fat gain in the

harvest season and protect against weight/fat loss in the hungry season. By examining both children and adults, potential different effects across the life course can be assessed.

In addition to studying possible weight effects, this section examines *POMC* methylation effects specifically related to fat mass changes over the year. As well as regulating appetite and body weight, there is increasing evidence of an interaction with POMC neurons and adipose tissue. Adipose tissue consists of both white adipose tissue (WAT) and brown adipose tissue (BAT). WAT primarily stores excess energy whereas BAT stores lower levels of fat and can be utilised to produce heat through the oxidation of fatty acids[13]. Hypothalamic circuits and importantly POMC neurons have been shown to be important factors related to white adipose tissue lipolysis (via upregulation of α-MSH)[14] and stimulate BAT[15,16] and implicated in WAT 'browning' (active thermogenesis in WAT)[17]. Leptin is an adipocyte derived hormone that helps regulate long energy balance by acting as a biological signal of fat stores. Leptin interacts with *POMC* neurons in the ARC of the hypothalamus[18] to increase POMC expression and thus satiety signal[19]. The relationship between seasonal changes in circulating leptin levels and *POMC* methylation is explored to further describe any relationship between leptin (as a *long* term measure of energy balance and fat) and *POMC* methylation.

4.2 Methods

4.2.1 Subject recruitment

A total of 493 children were recruited from a potential 572 from the ENID trial[20]. A total of 513 mothers were recruited from a potential of 691 mothers from the ENID trial. A group of these mothers (n=118) and children (n=118) were recruited into a study subset. The study subset were recruited from the villages of Keneba, Jali, Kantong Kunda, Manduar and Tankular. Village sensitisation and consenting occurred between February-April 2018.

Inclusion criteria:

- Enrolled in ENID trial (children) or parent of enrolled child (adults)
- Resides predominantly in West Kiang, The Gambia

Exclusion criteria:

- Unwell on day of study participation
- Pregnancy (at any time)

4.2.2 Study timepoints and activity

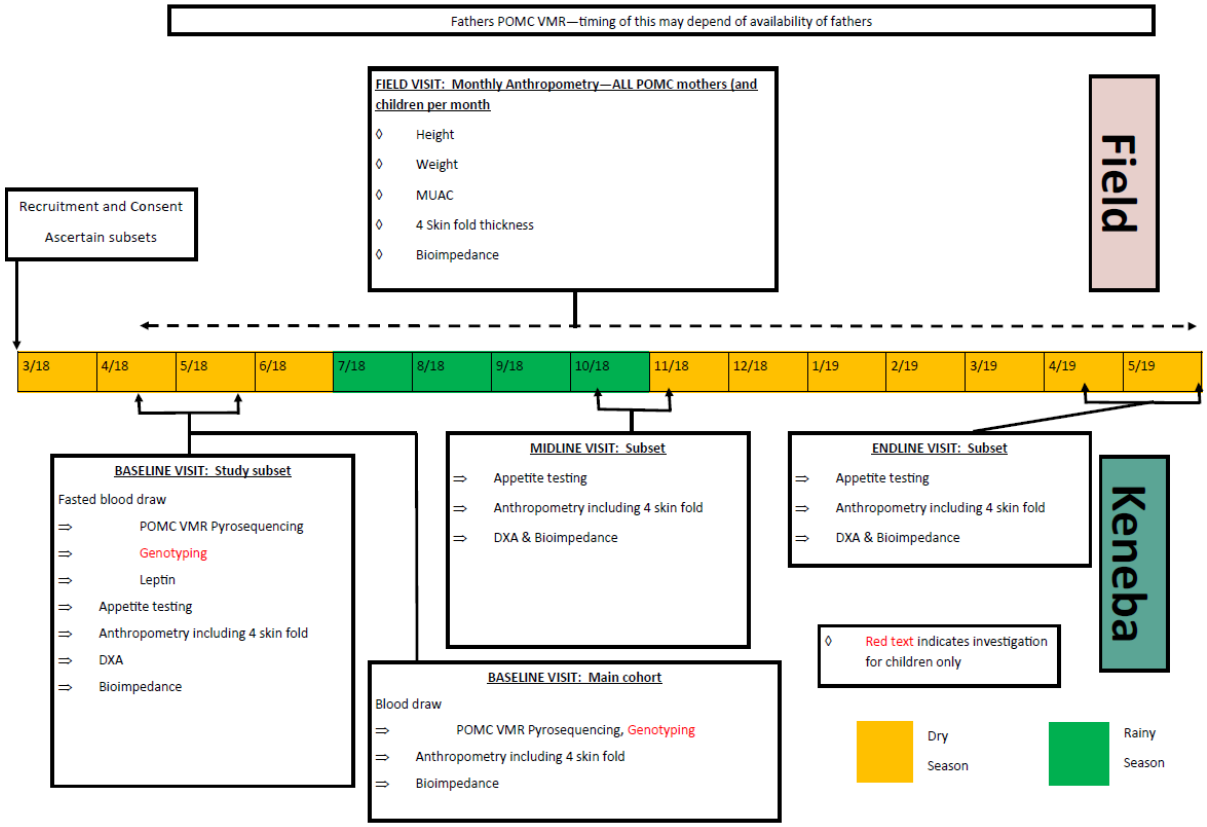


Figure 4.1. Overview of the POMC study including study timepoints and planned activity. Key: MUAC: mid-upper arm circumference, DXA=dual energy x-ray absorptiometry. The dry and rainy season is demarcated by orange and green colour to reflect changes in seasonal rainfall.

Three time-points were termed **baseline**, **midline** and **endline** and reflected peaks of seasonal weight loss/gain i.e. dry-harvest: April-May 2018; rainy-hungry: October-November 2018; and dry-harvest April-May 2019 (see Figure 4.1). The timing of peaks of season was an estimate of peaks of weight loss and gain informed by previous data in both women and children (see Chapter 1, Figure 1.8). I would have preferred to use May and October as the timepoints, but the study started in Mid-April to avoid Ramadan.

Study baseline (main cohort and subset): Participants attended Keneba field station from 16th April until 15th May 2018 (before the start of Ramadan 16th May until 14th June 2018). Five mls of fasted blood was taken. All children were measured in triplicate for weight, height, MUAC, skin fold thickness (as per MRCG@LSTHM SOP:NUTP.SOP.2009) and had a measurement of bioimpedance using TANITA BC-418 MA body composition analyser (as per MRCG@LSHTM SOP:NUTP.SOP.2018).

Monthly Field visits (main cohort and subset): All participants had monthly (12 consecutive months from May 16th onwards) scheduled field visits for anthropometry including weight, height, middle upper arm circumference (MUAC), skin fold thickness and a measurement of bioimpedance. Two field teams visited participants in a fixed order to standardise the period between each measurement. 'Mopping' days were included to attempt to visit participants again if they were unavailable during a planned visit.

During each encounter, the mother was asked if she was pregnant and offered antenatal care accordingly. If pregnant, she would exit the study.

Subset activity: At baseline, midline and endline time-points participants were scheduled a whole body DXA (Dual energy X-ray Absorptiometry) scan. Mothers had a urinary pregnancy test prior to scanning.

A measurement of appetite and satiety (see chapter 5 for more detail of methods and results) was taken at baseline, midline and endline. A further 5 ml fasted blood draw was taken at midline and endline for leptin (as a biochemical assessment of adiposity).

To replace participants who exited the study from the study subset (and thus maintain numbers in the subset), field workers approached mothers and children from the main cohort in the month prior to midline and endline to ascertain if they would consent to subset activity at these timepoints. Twenty subset mothers exited the study by midline and an additional 14 mothers were recruited from the main cohort in the subset for midline activity. A further 16 subset mothers exited the study by endline and an additional 14 mothers were recruited from the main cohort in to the subset at endline. One subset child exited the study by midline and 2 children (recruited simultaneously by separate field workers) were recruited from the main cohort in to the subset at midline. An additional child exited the study by endline but was not replaced due to over-recruitment before midline activity.

A total of 513 mothers (n=395 main study and n=118 study subset) and 493 children (n=375 main study and n=118 study subset) were recruited.

A summary of attendance at study activity at each timepoint is shown in Table 4.1. A total of 126 (24.6% of recruited mothers) mothers dropped out of the study with majority due to pregnancy (103/127). Other reasons for mothers leaving the study included consent withdrawn (n=8), moved out of the study area (n=6) and other reason not specified (n=9), see table 4.2. Drop out in children was far less common with only 16 leaving the study (3.2% of recruited children), see Table 4.2.

Children withdrew from study activity due to either consent being withdrawn (n=8) or they moved outside of the study area (n=8).

4.2.3 Anthropometry

Standing height was calculated as the mean of measures taken in triplicate to the nearest millimetre using a portable stadiometer (Seca 213). Weight was similarly calculated from measures in triplicate to the nearest 0.1kg using electronic scales (Seca 803), with participants clothed, but with shoes and coat removed. BMI was calculated as weight (kg) divided by height^2 (m^2). WAZ (weight for age) and BMI standard deviation score (SDS) for each child were calculated using WHO reference ranges[21]. BMI categories for adults were defined as <18.5kg/m2(underweight), >18.5 and <25 (normal weight), >25kg/m2 (overweight), >30kg/m2(obesity)[22]. BMI categories were calculated as per WHO standards and defined for children as <-2 SD (underweight), >-2 SD and <+1 SD (normal weight), >+1SD (overweight) and >+2 SD (obese)[23].

4.2.4 Bioimpedance

Body composition was measured by TANITA body composition analyser BC-418. Sex, height, and age were inputted into the analyser when prompted. The participant stepped on to the weighing platform with bare feet and ensured the heels were in direct contact with the posterior electrode and the front part of the feet were in contact with the anterior electrodes. The grip electrode were grasped with both hands. The impedance measurement was made and the result printed out. The body composition analyser provided an estimation of fat free mass and fat mass (two-compartment model of body composition) by passing an imperceivable electrical current through the body and measuring the impedance to the current. In general terms, constituents of fat free mass (e.g. muscle, bone, organs) is generally well conducted compared the poorly conductive fat mass. The analyser converts the impedance measurements into a measures of body composition with inbuilt computer prediction models which incorporate height, sex, and age into the models. The technology also allows a portable 'field' based measure of body composition.

4.2.5 Dual Energy x-ray Absorptiometry (DXA)

A whole body DXA scan was performed using the GE-Lunar Prodigy scanner (GE Medical, Waltham, MA; software version 13.60.033). DXA is a widely used non-invasive painless tool for assessment of body composition and provides a three-compartment model of body composition including measures of fat mass, lean mass, and bone mineral content. DXA scan administers a very low radiation dose (children and adults were exposed to a total of 2.88 μSv over the study) far less than one day of background radiation (6-7 μSv). As opposed to bioelectrical measurements of body composition such as bioimpedance, the measurements are independent of sample based prediction equations. Furthermore, the advantage of three (DXA) vs two compartment (bioimpedance) methods is that bone mineral content is specifically measured where it is included in fat free mass estimations in a two compartment model. This allows for a more accurate measure of lean mass. DXA is therefore considered the reference method for fat mass assessment in clinical research[24,25] though there has been high concordance between bioimpedance and DXA measurements reported[25].

For both bioimpedance and DXA results, fat mass index (FMI) was calculated and used as the metric of adiposity in analysis. FMI was calculated by fat mass (kg)/metres^2. Comparisons between bioimpedance measurements were made to assess concordance in FMI between the two measures.

Study Activity	Month 0 Apr-May BASELINE	Month 1 May-Jun	Month 2 Jun-Jul	Month 3 Jul-Aug	Month 4 Aug-Sept	Month 5 Sept-Oct	Month 6 Oct-Nov MIDLINE	Month 7 Nov-Dec	Month 8 Dec-Jan	Month 9 Jan-Feb	Month 10 Feb-Mar	Month 11 Mar-Apr	Month 12 Apr-May ENDLINE
Study subset. Number of participants with study activity – Mothers (enrolled 118)													
Anthropometry	115	118	117	116	115	105	105	102	102	98	99	81	94
Bioimpedance	114	118	105	116	115	105	103	99	101	98	96	80	91
DXA	100						95						88
Main study. Number of participants with study activity – Mothers (enrolled 395)													
Anthropometry	271	316	317	316	311	307	288	285	283	260	263	246	258
Bioimpedance	266	312	306	312	306	301	283	279	278	257	260	240	255
DXA[£]	0						14						12
Study subset. Number of participants with study activity – Children (enrolled 118)													
Anthropometry	115	115	117	115	114	112	116	112	115	114	113	109	113
Bioimpedance	115	114	106	114	113	112	110	112	114	113	111	109	112
DXA	108						113						110
Main study. Number of participants with study activity – Children (enrolled 375)													
Anthropometry	299	343	344	343	332	332	332	333	342	340	329	330	324
Bioimpedance	293	339	340	337	320	322	321	326	340	337	318	326	318
DXA[£]	0						2						79

Table 4.1 Overview of attendance of mothers and children at study activity by study group. [£] main study participants with DXA activity represents those who were additionally recruited into subset activity due to subset participant exit or additional study activity from the PAX8 study

Total number of participants with study activity - Mothers (total enrolled in study = 513)													
Anthropometry	386	434	434	432	426	412	393	387	385	358	362	327	352
Bioimpedance	380	430	411	428	421	406	386	378	379	355	356	320	346
Total exited study (cumulative n) and reason[$]	n= 3 (3) P=3	n=32 (35) P=22,C=3, O=7	n=10 (45) P=8, C=2	n=9 (54) P=8, O=1	n=15 (69) P=14, O=1	n =4 (73) P=4	n=11 (84) P=9, M=2	n=10 (94) P=8, C=1, M=1	n=8 (102) P=6, C=1, M=1	n=4 (106) P=3, C=1	n=5 (111) P=3, M=2	n=7 (118) P=7	n=8 (126) P=8
Of which were subset exits (cumulative n)	0	8 (8)	2 (10)	2 (12)	4 (16)	0	4 (20)	4 (24)	2 (26)	0 (26)	2 (28)	5 (33)	3 (36)
DNA anthropometry	124	43	33	26	17	27	35	31	25	48	39	67	34
DNA anthropometry	130	47	56	30	22	33	42	40	31	51	45	74	40
Total number of participants with study activity - Children (total enrolled in study = 493)													
Anthropometry	414	458	461	458	446	444	448	445	457	454	442	439	437
Bioimpedance measurement	408	453	446	451	433	434	431	438	454	450	429	435	430
Total exited study (cumulative n) and reason[$]	n=0 (0)	n= 5 (5) C=5	n=1 (6) C=1	n=1 (7) M=1	n=0 (7)	n=0 (7)	n=1 (8) M=1	n=2 (10) M=2	n=2 (12) M=2	n=1(13) M=1	n=1(14) M=1	n=0 (14)	n=2 (16) C=2
Of which were subset exits (cumulative n)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1)	0 (1)	0 (1)	0 (1)	1 (2)	0 (2)	0 (2)
DNA anthropometry	79	30	26	28	40	42	38	39	25	27	39	42	42
DNA bioimpedance	85	35	41	35	53	52	55	46	28	31	52	46	49

Table 4.2 Overall attendance at study activity with details of study drops and reasons. Key: DNA = Did not attend study activity. [$] Study exit key: P=pregnant, C= consent withdrew, M=moved outside study area, O=Other

4.2.6 Laboratory processes

Blood sampling overview

Five mls of fasted venous blood was taken at each blood draw (4ml EDTA, 1ml serum). The schedule and planned laboratory analysis is shown in Table 4.3.

Study participant	Genotype	*POMC* DNA Methylation	Leptin	Plasma storage	Serum storage
Mothers (**baseline** subset and main cohort)		✓	✓ (subset only)	✓	✓
Children (**baseline** subset and main cohort)	✓	✓	✓ (subset only)	✓	✓
Subset mothers and children only (**midline**)			✓	✓	✓
Subset mothers and children only (**endline**)			✓	✓	✓

Table 4.3 Summary of scheduled study blood draws and tests performed

Sample processing and tests (see Table 4.4)

EDTA sample collection tubes are spun at 1800g (rcf) 10min at 4°C and then aliquoted as per Table 4.4. Aliquot tubes were frozen at -70°C. The remaining cellular fraction (~1.5ml) was used for DNA extraction using the Chemagic360 (MRCG@LSHTM SOP:SOP-NUT-007). DNA samples were quantified and the purity checked using the NanoDrop protocol (MRCG@LSHTM SOP:NUT.SOP.4007) prior to transfer to the genomics laboratory in Fajara.

Serum sample collection tubes were spun at 1800g (rcf) 10min at 4°C and aliquoted into 0.5ml microtubes and frozen at -70°C. The cellular fraction was discarded.

Collection tube	Vol collect	Aliquot name	Aliquot type	Aliquot vol (ul)	Test performed
EDTA	4ml	EDTA_AQ1	Plasma	500	Stored
		EDTA_AQ2	Plasma	500	Stored
		EDTA_AQ3	Plasma	500	Stored
		DNA_AQ1	DNA	500-1000 (~160ug)	1. *POMC* methylation in all participants 2.Additionally genotype in children
Serum 1.2ml	1ml	Serum_AQ1	Serum	500	1. Leptin in subset participants only 2. Samples stored other participants

Table 4.4 Blood sample aliquot protocol.

Leptin

Human Leptin Immunoassay (Quantikine ELISE, R&D systems) was used to measure leptin from serum samples as per manufacturer's instructions. Leptin measurements were conducted in Keneba Field Station Laboratory, MRCG at LSHTM by me and Ebrima Bah (scientific officer).

In summary:

1. 100 µL of Assay Diluent RD1-19 was added to each well.
2. 100 µL of standard, control, or sample was put in each per well (all standard, control, or samples were performed in duplicate). The 96 well plate layout included Human leptin standard (varying concentrations of 1000 pg/mL, 500 pg/mL, 250 pg/mL, ,125 pg/mL, 62.5 pg/mL, 31.3 pg/mL, 15.6 pg/mL n=16 wells), control leptin solution of high, normal, and low leptin (n=6 wells), and duplicate samples (n=84 wells, n=42 samples)
3. The plate was covered with the adhesive strip provided and incubated for 2 hours at room temperature.
4. Each well was aspirated and washed (by filling each well with 400 µL of wash buffer) and the process repeated for a total of four washes. After the last wash the plate was inverted and blotted it against clean paper towels.

5. 200 µL of Human Leptin Conjugate was added to each well. The plate was covered with the adhesive strip provided and incubated for 1 hours at room temperature.
6. The aspiration and wash was repeated as per step 4.
7. 200 µL of Substrate Solution to was added to each well and incubated for 30 minutes at room temperature. Whilst protecting the plate from light with tin foil.
8. 50 µL of Stop Solution was added to each well - the colour in the wells changed from blue to yellow.
9. The optical density of each well was assessed within 30 minutes using a microplate reader set to 450 nm.

The covariation of the means (COV) was calculated for low and high control by taking the mean/standard deviation x 100. The mean high control was 484.4 (SD=53.7) with COV of 11.1% and the mean low control was 90.6 (SD=10.3) with COV of 10.3%. The inter-assay COV calculated by taking the mean of the high and low COV and was 10.7%.

4.2.7 *POMC* methylation measurement

POMC VMR methylation was measured by pyrosequencing

To obtain a measurement of DNA methylation requires three key processes: bisulfite conversion, followed by polymerase chain reaction (PCR) of bisulfite converted DNA and then pyrosequencing of the PCR product. Bisulfite conversion converts unmethylated cytosine to uracil whereas methylated cytosine remains unchanged. PCR converts the uracil to a thymine, whereas methylated cytosine remains a cytosine. Pyrosequencing is then performed to discriminate between C and T single nuclear polymorphisms in the PCR product. The pyrosequencer detects luminescence from the release of pyrophosphate on incorporation of a nucleotide into the complementary strand.

For example at the CpG site below, we can see the difference in the nucleotides of PCR product dependent on the methylation status of the cytosine in the original strand.

	Unmethylated cytosine	Methylated cytosine
Original top strand	CG	^{m}CG
Bisulfite converted	UG	CG
PCR	TG	CG

The bisulphite conversion and PCR steps were conducted at the MRC The Gambia. The pyrosequencing was performed by Peter Kuhnen's (collaborator and advisory committee member) team in Berlin.

I worked closely with the genomics team in an attempt to perform the pyrosequencing in The Gambia on the Pyromark Q48 machine. Like many centres internationally, we had much difficulty in getting accurate results on this machine.

The main issues were concerning low peak height (low luminescence), significant baseline drift and potentially miss priming such that the pyrosequencing runs consistently failed. An example pyrosequencing run is shown in Figure 4.2.

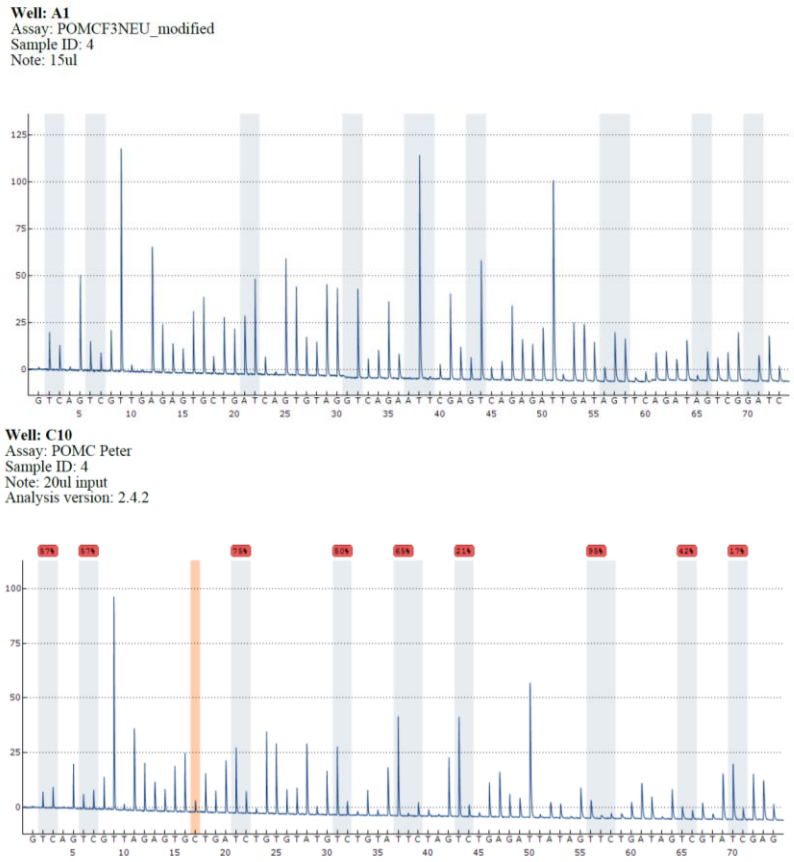


Figure 4.2. Pyrograms demonstrating failed POMC pyrosequencing runs. Both images demonstrate baseline drift and low peak height thus resulting in failed runs. Top image using 15 ul PCR input, bottom image using 20ul PCR input.

Low peak height:

To mitigate this problem I tried increasing the PCR input amount and performed a gradient of input between 10-20ul of PCR product. However, beyond 20ul the wells were full so this prevented increasing the product further. There was minimal and unsatisfactory improvement in peak height with the increasing input volumes.

Baseline drift:

As you can see from the pyrograms (Figure 4.2), there was a pattern of the baseline to drift downwards. A common cause of this phenomenon is temperature instability during the pyrosequencing run. After consulting with the machine manual and with colleagues internationally we tried a number of approaches including i) turning the machine on 30-45min before a run ii) turning it on and using straight away iii) placing thermometers around the room to find the coolest position in the room. Furthermore, we ensured air conditioning units in the rooms were functioning and remained on throughout all runs. The pyrosequencing reagents were taken out of the fridge and allow to warm to room temperature. There was no improvement in baseline drift on the pyrograms with any of these strategies.

Miss priming:

There were also runs where we were getting some peaks recorded where we would expect a negative response. We reordered primers from another company to see if there is an issue with our sequencing primers and tried different primer concentrations, with no difference noted.

To test if there was issues with the pyrosequencer at MRC The Gambia, a set of 4 samples were sent to Germany to see if successful runs on the Q24 could be achieved. Samples were sent as either PCR product produced by me in MRC The Gambia or extracted DNA (without bisulfite conversion). This allowed a comparison of methylation outputs from PCR prepared in MRC The Gambia and the PCR product produced in Germany from extracted unbisulfite converted DNA (sent from Gambia). There were no issues with low peak height, baseline drift, or miss priming on the runs performed in Germany on the Pyromark Q24. The methylation results were highly similar between PCR produced from DNA in the Germany and PCR product from MRC The Gambia suggesting that the PCR product from The Gambia was of sufficient quality (Table 4.5).

Study ID												
		Pos. 1	Pos. 2	Pos. 3	Pos. 4	Pos. 5	Pos. 6	Pos. 7	Pos. 8	Pos. 9	Mean	Peakheight (Dispensation 11 T-Peak)
PM-C536E		Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)		
	PCR Product (produced in Germany from DNA sent from Gambia)	82.59	81.72	80.79	73.85	70.98	74.2	66.5	57.18	47.18	**70.5544444**	60
	PCR Product (pro duced in MRC Gambia)	87.5	84.28	77.34	75.41	67.52	67.2	67.84	56.66	41.04	**69.4211111**	650
		Pos. 1	Pos. 2	Pos. 3	Pos. 4	Pos. 5	Pos. 6	Pos. 7	Pos. 8	Pos. 9	Mean	Peakheight (Dispensation 11 T-Peak)
PM-C537J		Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)		
	PCR Product (produced in Germany from DNA sent from Gambia)	56.92	51.51	45.31	33.69	35.76	35.76	31.81	26.46	21.12	**37.5933333**	100
	PCR Product (pro duced in MRC Gambia)	58.77	53.77	44.88	36.94	35.98	17.19	36.02	24.35	18.78	**36.2977778**	650
		Pos. 1	Pos. 2	Pos. 3	Pos. 4	Pos. 5	Pos. 6	Pos. 7	Pos. 8	Pos. 9	Mean	Peakheight (Dispensation 11 T-Peak)
PM-C538G		Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)		
	PCR Product (produced in Germany from DNA sent from Gambia)	67.19	64.14	48.35	33.52	35.67	31.6	30.34	26.41	19.01	**39.5811111**	400
	PCR Product (pro duced in MRC Gambia)	63.64	60.17	44.84	31.75	32.2	28.78	30.24	23.23	16.51	**36.8177778**	650
		Pos. 1	Pos. 2	Pos. 3	Pos. 4	Pos. 5	Pos. 6	Pos. 7	Pos. 8	Pos. 9	Mean	Peakheight (Dispensation 11 T-Peak)
PM-C539D		Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)	Meth. (%)		
	PCR Product (produced in Germany from DNA sent from Gambia)	37.01	31.75	17.36	11	10.5	9.63	10.52	7.62	6.74	**15.7922222**	250
	PCR Product (pro duced in MRC Gambia)	35.38	31.53	19.28	11.13	11.16	8.07	10.94	6.51	3.68	**15.2977778**	500

Table 4.5. Comparisons in methylation output on PCR product derived either from The Gambia or Germany. Note peak height much higher that the pyrograms in Figure 4.2.

Therefore, PCR product was produced in MRC The Gambia and exported on dry ice -20 degrees to Germany for pyrosequencing. The methodology of the bisulfite conversion, PCR process and pyrosequencing used in the samples in the study is outlined below.

Plate randomisation

Participants were randomised to 12 x 96 well PCR plates using the OSAT package in R[26] which aims to distribute biological groups and confounding factors across sample batches. Sex, BMI quartile, participant category and season of conception were variables assigned to be balanced across the plates. The distribution of these variables across the 12 plates is shown in Annex 4.1.

Bisulphite conversion

The sample bisulphite conversion and PCR was tested and developed by me and Abdoulie Kanteh (Scientific officer, Genomics team, MRCG The Gambia At LSHTM). Bisulphite conversion and PCR reaction was completed on all samples in Fajara Genomics facilities, MRCG The Gambia At LSHTM by me, Abdoulie Kanteh and Ebrima Bah (Trainee Scientific Officer, Nutrition theme, MRCG The Gambia At LSHTM).

DNA aliquots were produced with a concentration of 500ng DNA in 20ul. DNA underwent bisulphite conversion using the EZ DNA methylation-gold kit (Zymo).

The process of bisulphite conversion is summarised below:

1. 130 µl of CT conversion reagent was added to 20 µl of DNA in a PCR tube. The reagent and DNA was mixed by pipetting up and down and then centrifuged to bring the fluid to the bottom of the tube. Note that the CT conversion reagent was light sensitive so this reaction was conducted in a darkened room and the PCR tubes covered in foil
2. The tube was placed in a thermocycler with the following settings
    a. 98^{o}C for 10 minutes
    b. 64^{o}C for 150 minutes
    c. Reduced to 4^{o}C until retrieved for onward processing

These first two steps denatures the DNA with heat (to create a single strand) and sulphonation of cytosine bases. Bisulphite reacts with cytosine to produce cytosine sulphonate, whereas 5'methyl-cytosine does not react.

3. 600 µl of M-binding buffer is placed in the Zymo-spin IC column and is mounted on a collection plate
4. The samples from steps 1 and 2 above were added to a Zymo-spin IC column with the M-binding buffer and pipetted up and down to mix. The plate was centrifuged for 5 minutes at 3,000g. The flow through was discarded.
5. 100 µl of M-wash buffer was added to the column. The plate was centrifuged for 5 minutes at 3,000g.
6. 200 µl of M-desulphonation buffer was added to the column and left to stand for 20 minutes at room temperature. After this time, the plate was centrifuged for 5 minutes at 3,000g. The flow through was discarded.
7. 200 µl of M-wash buffer was added to the column. The plate was centrifuged for 5 minutes at 3,000g. A further 200 µl of M-wash buffer was added to the column and the plate then centrifuged for 10 minutes at 3,000g.
8. The column was placed on the elution plate. 12 µl of pre-warmed (60 ^{o}C) M-elution buffer was added to the column. The plate was left for 5 minutes at room temperature after which time the plate was centrifuged for 3 minutes at 3000g to elute the DNA
9. The eluted DNA was extracted from the elution plate to use in the PCR reaction.

These next steps convert cytosine sulphonate to uracil sulphonate by hydrolytic deamination and then convert uracil sulphonate to uracil by desulphonation. The bisulphite converted DNA is then cleaned up, eluted and ready for use in the PCR reaction.

PCR reaction

Bisulphite converted DNA was used in a PCR reaction to produce an amplicon length of 325bp. A 96 well plate was prepared as per the quantities reported in Table 4.6.

POMC PCR	**1x Sample**	**X100**
Distilled water	39.1	3910
Coral load	6	600
MgCl2 (25mM)	6	600
dNTPs (10 mM)	1.2	120
forward primer (10µM)	0.5	50
reverse primer (10µM)	0.5	50
Qiagen HotStarTaq DNA Polymerase	0.7	70
DNA template: Bis-DNA	6	
Totals	60	6000

Table 4.6. PCR reaction reagent list and quantity

The PCR primers used are reported in Annex 4.2. The PCR plate was placed in the thermocycler with the following protocol:

1. 95 ^{o}C for 15 minutes
2. 95 ^{o}C for 30 seconds
3. 54.5 ^{o}C for 30 seconds
4. 72 ^{o}C for 45 seconds
5. Steps 2,3 and 4 were repeated 43 times
6. 72 ^{o}C for 8 minutes
7. Held at 4 ^{o}C until retrieved

The PCR product was tested using QIAxel Advanced Instrument (automated gel electrophoresis machine), see Figure 4.3 for an example of the output. The PCR product was then frozen to -20 ^{o}C and transported on dry ice to Germany.

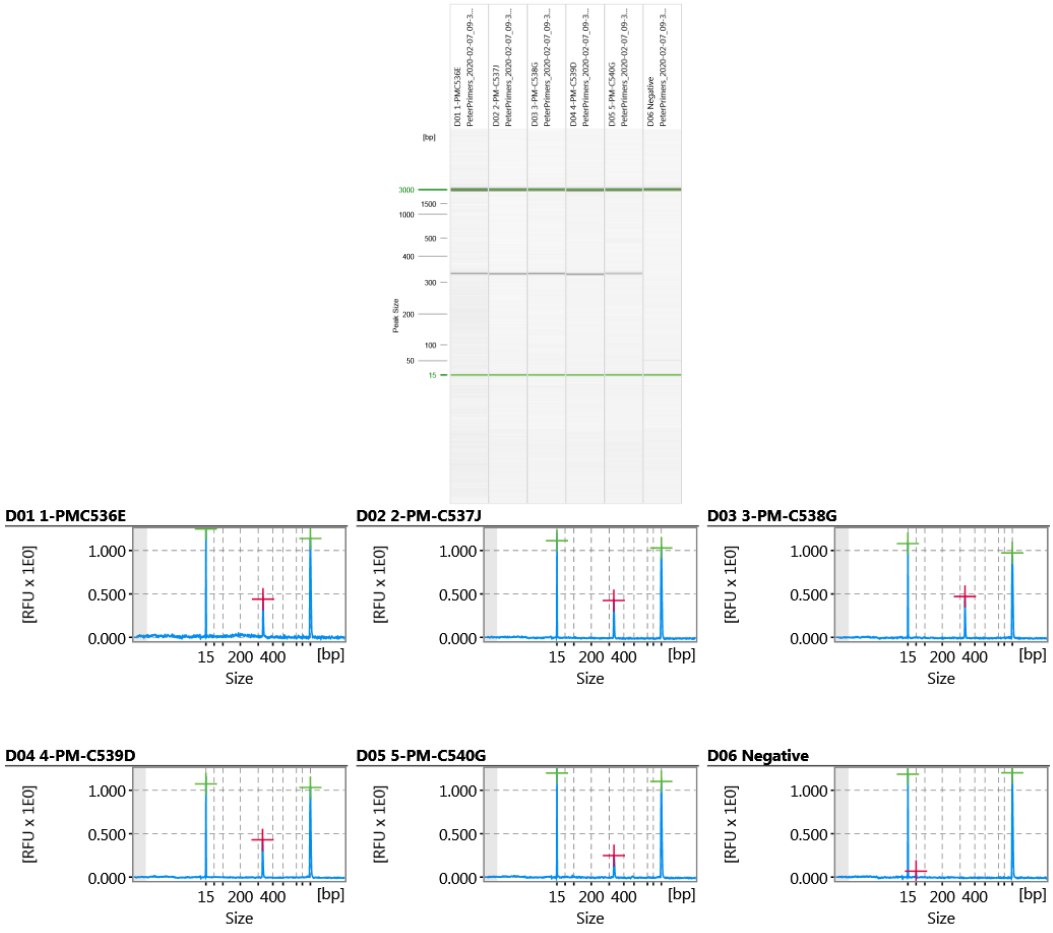


Figure 4.3. Qiaxel report (automated gel electrophoresis) of 5 samples and one negative control demonstrating consistent PCR product with an amplicon length of 325 bp with protocol.

Pyrosequencing

The pyrosequencing was conducted by Lara Lechner (Medical Student (MD PhD candidate), Kuhnen Laboratory, Charite University, Berlin, Germany) on the Qiagen Q24 pyrosequencer using the protocol summarised below.

1. PCR product purification
    a. 44µl of SPRI-Beads (0.8 ratio) was added to 55µl PCR-template.
    b. The PCR-bead mixture was then incubated in the dark for 10 min at room temperature
    c. The templates were placed on a magnet stand and until all beads are aggregated.
    d. The fluid was removed by pipette
    e. 200µl of 80% Ethanol (freshly mixed) was then added to the beads

f. Templates were then moved from one side to another on the magnet stand and then the beads were allowed to aggregate. This process was repeated twice.
g. The ethanol was then removed by pipette
h. A further 200µl of 80% Ethanol was added to the beads and step f. was repeated. The beads were allowed to stand for 3-5 minutes.
i. The ethanol was then removed by pipette
j. The beads were allowed to dry and then eluted with 41µl of Milli Q Water.
k. The beads were allowed to aggregate and then the eluate to transferred to new tubes.

2. Pyrosequencing
    a. A well vortexed mastermix of 1080 µl binding buffer and 27µl sepharose beads for our 24-Well Pyrosequencing was produced.
    b. 39µl of the DNA-purification-eluate and 40µl binding buffer with 1µl sepharose beads is put in each well of the primer plate
    c. A mastermix of 673 µl annealing buffer and 2,025µl 100µM primer (don´t vortex, just shake) was produced and then 25µl 0,3µM primer with annealing buffer was put on each well of the primer plate.
3. The primer plate was placed in the pyrosequencer and sequencing commenced.

POMC methylation measure

DNA methylation data from CpGs -2 to +7 (referring to their position in relation to the intron 2 and exon 3 boundary of the *POMC* gene, see Figure 4.4 and 4.5) were available (440=mothers, 444=children, 99=fathers). A single methylation measure was produced to summarise the methylation pattern across key CpGs across the region. This single *POMC* methylation measure was used in all the subsequent analysis. The methylation pattern across the 9 CpGs is shown in Figure 4.6, and showed a higher percentage methylation in the CpGs in intron2 and with a trend for lower methylation in exon 3 as reported previously[7,27].

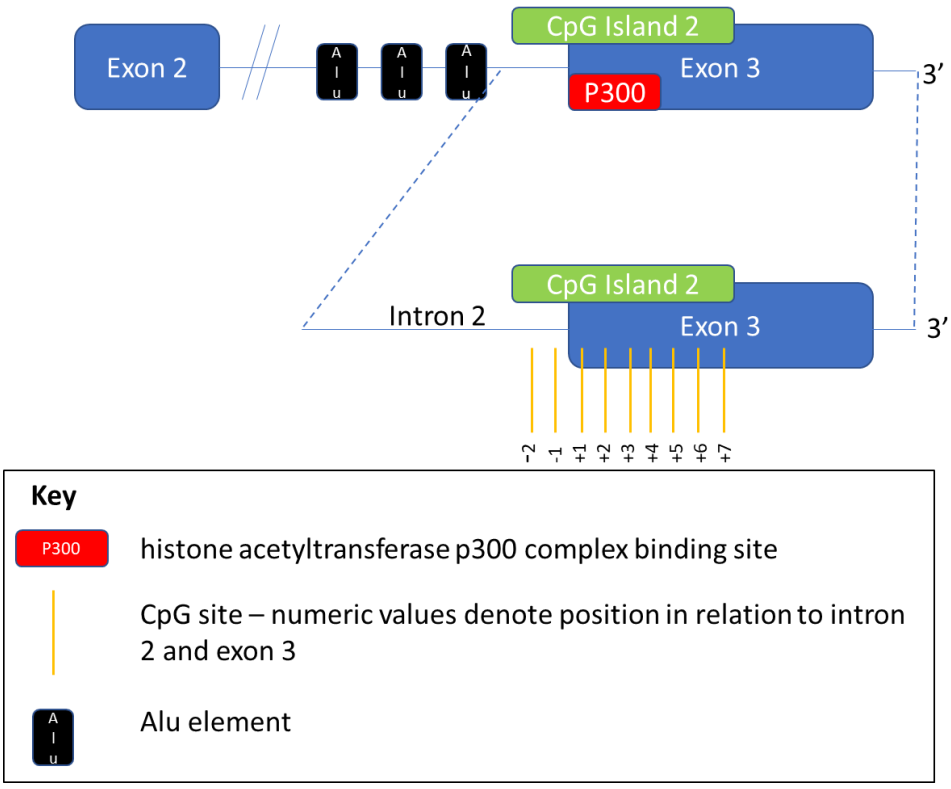


Figure 4.4 Schematic representation of the location of POMC VMR CpGs (not to scale).

```
 25,384,560       570       580       590       600       610       620       630       |hg19 (+)
          |        |         |         |         |         |         |         |         |
25,161,690       700       710       720       730       740       750       760       |hg38 (+)
         |        |         |         |         |         |         |         |         |
     CGTCGCCATTTCCCGGGAACATGGGAGTCTCGGCCGAGAGGTCGGGCTTGCAGGCCCGGATGCACTCCTGGGGGAAGACGCG |
     +7 +6        +5               +4  +3       +2             +1                   -1-2 |CpGs
     exon <<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<< exon][intron         |POMC gene


CpG coordinates:

           hg38        hg19
-2    25,161,767  25,384,636
-1    25,161,765  25,384,634
+1    25,161,743  25,384,612
+2    25,161,729  25,384,598
+3    25,161,721  25,384,690
+4    25,161,717  25,384,686
+5    25,161,700  25,384,668
+6    25,161,690  25,384,658
+7    25,161,687  25,384,655
```

Figure 4.5. CpG coordinates for CpGs (-2 to +7) for both hg38 and hg19 genomic builds

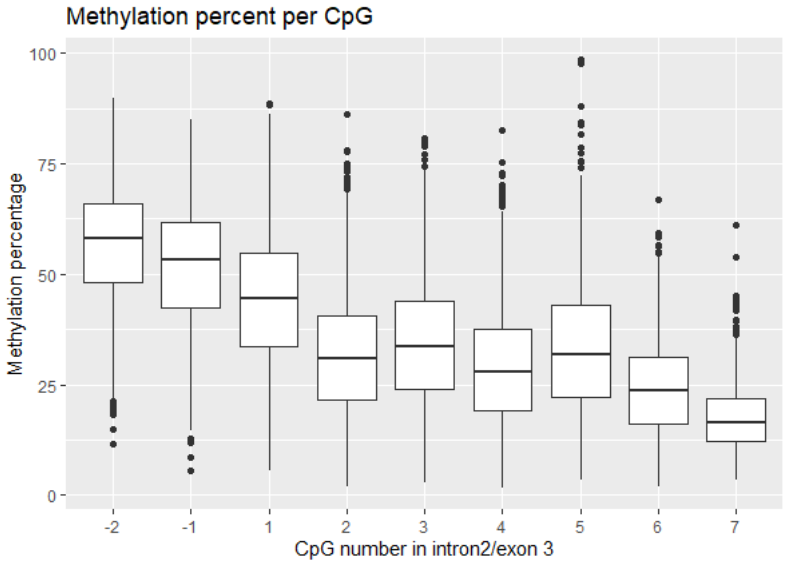


Figure 4.6. Boxplots of percentage methylation at each of the CpGs between -2 and +7 (relative to intron2/exon3 boundary)

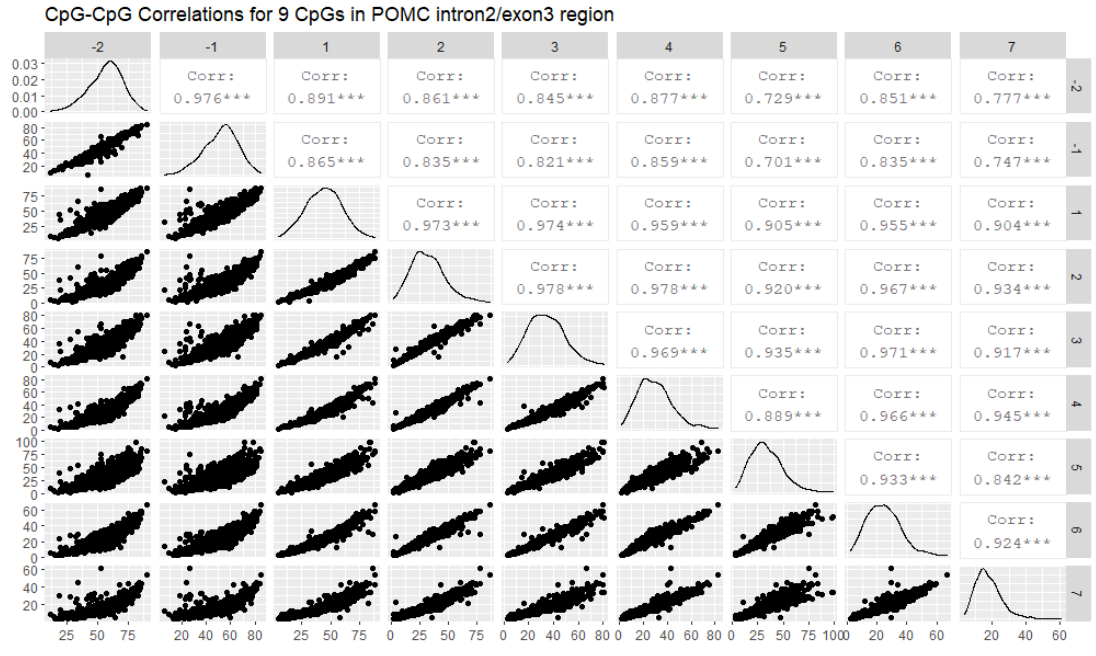
DNA methylation data from CpGs +1 to +5 inclusive were selected to form the single methylation measure as methylation at these CpGs was i) highly correlated (see Figure 4.7), ii) previously associated with obesity[7,9] and iii) had been reported to be sensitive to periconceptional nutrition[9] (see Annex 4.3 for evidential summary table).

i) CpG-CpG correlations demonstrated that the CpGs within exon 3 (+1 to +7) were highly correlated (Pearson R = 0.842 – 0.978). There was lower correlation with the CpGs in intron 2 and exon 3 (e.g. -1 and +5 demonstrated a Pearson R of 0.701(see Figure 4.7)).

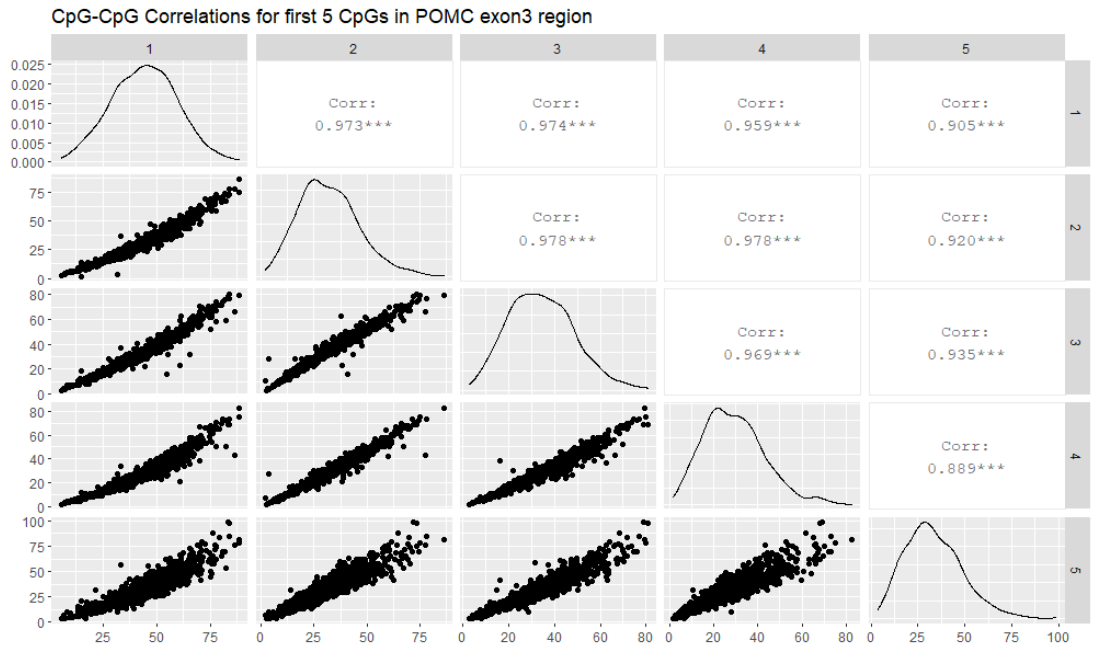
ii) Previous studies by Kuhnen et al[7,9], had identified significant associations between DNA methylation at CpGs +1, +2, +3 and +5 (in exon 3) and obesity in adults and children[7,27].

iii) Previous studies by Kuhnen et al[7,9], had reported a significant relationship between maternal periconceptional circulating levels of key one-carbon metabolites and DNA methylation in CpGs in exon 3; SAH (all CpGs between +1 and +5), SAM:SAH ratio (CpGs +1,+3 and +5) and betaine (CpGs +3, +4 and +5).

To create a single measure of *POMC* methylation, a methylation z score was created for each CpG with the mean of the z scores calculated to give a mean *POMC* methylation z score across the 5 CpGs. Z scores were produced within participant groups (i.e. calculated separately for mothers, fathers and children) and calculated separately for each sex for the children. The resultant methylation z score was normally distributed.



a)



b)

Figure 4.7. CpG-CpG correlations between CpGs in intron 2 and exon 3 of POMC gene (a) and the first 5 CpGs in exon 3 (b). ****= P ≤ 0.001.*

Genotyping

Extracted DNA was shipped on dry ice (-20) to University of Bristol Genomics laboratory. Genotype was determined using the H3 Africa array from baseline children's samples. The H3 Africa array was developed by the H3 Africa consortium to identify common and rare genetic associations with common and rare traits among African populations. The array contains 2,271,503 markers with a single bead chip able to process 8 samples and 200ng of DNA required for input[28].

4.3 Statistical analysis

All statistical analysis were performed using R version 3.6.2[29].

4.3.1 Primary analysis: Assessing the association between weight and fat trajectories and *POMC* methylation

Cosinor modelling of seasonal weight and fat mass index change

Maternal weight, child weight for age z score, maternal FMI and child FMI were plotted as a function of time to assess the presence of seasonal rhythmicity.

Mean population cosinor models were fitted.

Cosinor models have three parameters: MESOR (midline estimating statistic of rhythm), Amplitude and Acrophase, see Figure 4.8.

- The MESOR refers to the rhythm adjusted mean i.e. the mean value across the year.
- The Amplitude is defined as half of the extent of predictable change within the cycle i.e. the degree of excursion from the mean.
- The Acrophase relates to the timing of the peaks within the cycle (NB the acrophase is expressed in negative degrees in relation to a reference time set to 0°, with 360° equal to the 'period' (duration of one cycle or the unit circle) [30].

  For example, in a 12 month cycle, if the peak was recorded at month 8 then the acrophase would be offset by 4 months. The acrophase could be described as being offset by a 1/3 (i.e. peak is 4 months before the end of the cycle of 12 month i.e. 4 divided by 12 = 1/3). There are 360 degrees in a circle therefore the acrophase is ***phase advanced*** (-360 / 3 =) -120 degrees before the defined start of the unit cycle.

Where a significant association with acrophase was made, the coefficient was converted to time by the following process:

The cosinor package expresses the acrophase coefficients as radians.

Degrees were calculated by multiplying radians by $180^{o}/\pi$.

Time in relation to cycle = (Degrees/360) x period

The mean population cosinor fits sin/cosine regression models to data from multiple individuals at multiple time points. This statistical procedure fits individual cosinor models for an individual's time series data then averages individual cosinor parameters to produce a population mean for each of the cosinor parameter (see Figure 4.8).

The cosinor2 package in R was used to generate the cosinor models[31] and allows the calculation of confidence intervals using previous described methods[32]. The period (the duration of one cycle) was set at 12 (months) to reflect the experimental model occurring over a 12 month period and the duration of the sinusoidal curve of weight (and weight for age z score) and FMI change. A rhythm detection test was performed on mean population cosinor models to ascertain model fit and acts as a global test of significance for the estimated model. cosinor2 calculates a 'percent rhythm' (r^2) which is the proportion of variance explained by the rhythm, calculated from the correlation between observed and estimate data (r).

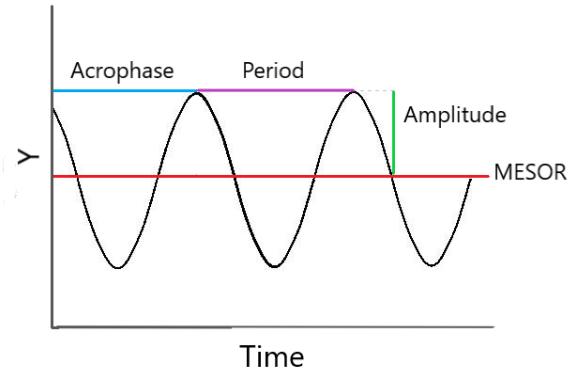


Figure 4.8. Rhythm characteristics and schematic depiction of COSINOR parameters

Linear regression modelling of individual consinor parameters and POMC methylation

Individual cosinor parameters were extracted from the mean population cosinor model and were transformed to normal distributions where required. These outcomes effectively parameterises (or summarises) the individual trajectories in terms of MESOR, amplitude and acrophase.

The association between weight/fat mass index trajectories and *POMC* methylation was assessed in linear regression models. The cosinor model parameters (MESOR, amplitude and acrophase) were modelled separately. In the linear regression models, one of MESOR, amplitude and acrophase was the dependent variable (*outcome*), and mean *POMC* methylation z score the independent variable (*predictor*) to assess if mean *POMC* methylation z score influenced each cosinor parameter. Each model was adjusted for relevant covariates. Where reported, coefficients (β) associated with log transformed dependent variables were back transformed using (exp(β) - 1) x 100, to represent percentage change in dependent variable per unit increase in the corresponding predictor.

4.3.2 Secondary analysis: Assessing the association between leptin and DXA-derived FMI and *POMC* methylation

For these analyses, data from the study subset were used. Additional subset measures of adiposity included leptin and measure of DXA-derived fat mass index measurements taken at baseline, midline and endline.

All outcome variables (leptin and DXA-derived fat mass index) were normally distributed or transformed to normal distribution where required. Change in leptin and DXA-derived fat mass index between baseline and midline was calculated by subtracting midline value from baseline value and between midline and endline by subtracting endline value from midline value. The change in leptin and FMI were made the dependent variables (*outcome*) in linear regression models with mean *POMC* methylation z score (*predictor*) and adjusted for other relevant covariates.

4.3.3 Assessing performance of different fat measures against DXA-calculated fat mass

DXA-derived fat mass was correlated with TANITA-derived fat mass and Spearman correlation coefficient and p-value reported. DXA-derived fat mass was correlated with leptin and Spearman correlation coefficient and p-value reported.

4.3.4 Sample size

The effect of *POMC* methylation on seasonal weight or fat change is unknown. However, a sample size of n=500 would provide 80% power to detect a relatively small Cohen's effect size (f^2) of 0.03 in a multiple regression adjusted for age and sex, with a multiple testing correction for 5 independent outcome variables if required.

4.4 Results

4.4.1 Maternal results – subset and main cohort comparisons

There were no differences in mothers baseline characteristics between main study and study subset participants with regards to age, weight, BMI, BMI category or mean *POMC* methylation z score. Baseline characteristics of the mothers is shown in Table 4.7.

Mothers	**Overall**	**Main**	**Subset**	**p-value**
Mean age (years) (SD, range)	37.87 (6.49, 22.78 - 53.78)	37.92 (6.76)	39.20 (5.40)	0.07
Weight (kg) (SD)	61.10 (11.94)	60.72 (11.90)	61.98 (12.05)	0.34
BMI (kg/m2) (SD)	23.48 (4.48)	23.40 (4.54)	23.67 (4.33)	0.59
BMI category (%)				0.95
Underweight	8.8	8.9	8.7	
Normal Weight	64.8	65.3	63.5	
Overweight	18.1	18.1	18.3	
Obese	8.3	7.7	9.6	
Mean *POMC* methylation z score (SD)	0.20 (0.97)	0.18 (0.97)	0.27(0.97)	0.38

Table 4.7. Maternal baseline characteristics. Comparisons between study groups for age, weight, BMI and mean POMC methylation z score were assessed using student t test. Differences in BMI category between study groups was made using chi squared test (using the numbers in each category however % in each category shown in table). **Key:** SD=standard deviation, BMI=body mass index.

4.4.2 Maternal weight – summary statistics

Both weight and BMI was plotted over time from baseline to month 12 (see Figure 4.9). There were similar patterns of weight and BMI change across the year with a peak at baseline (April-May) and a nadir in month 5 during the rainy season (September-October). Weight was used for all subsequent analysis.

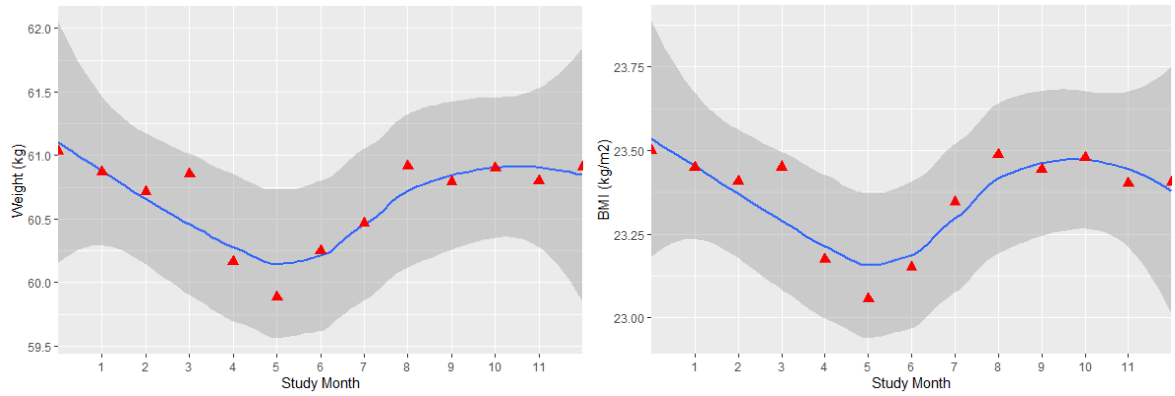


Figure 4.9. Maternal weight (left) and BMI (right) plotted against study month 0-12. Red triangles indicate mean value at each time point. Loess regression line (blue) with 95% confidence intervals

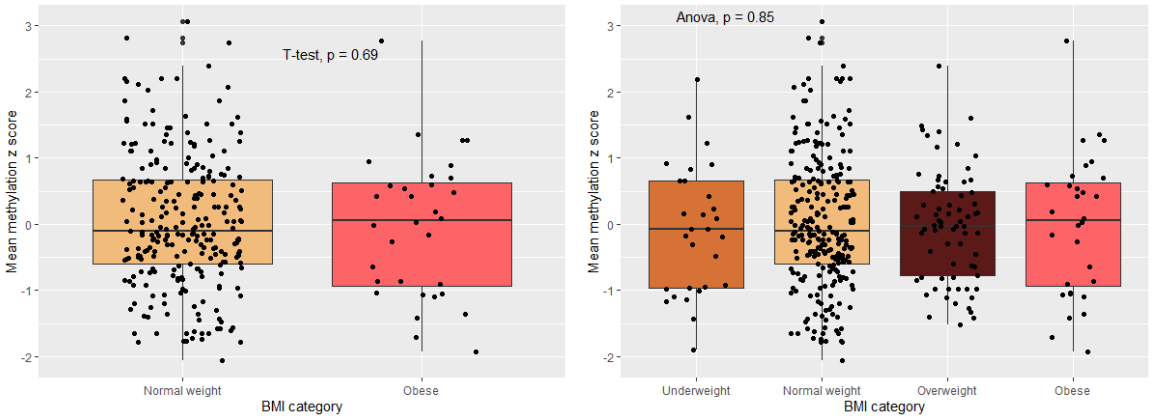


Figure 4.10. Boxplots of maternal BMI category at baseline and mean POMC methylation z score. Left plot: Normal weight vs. Obese. Right plot: Comparison across all BMI categories.

There was no difference in mean *POMC* methylation z score by BMI category at baseline (Figure 4.10).

4.4.3 Maternal weight trajectories - cosinor modelling

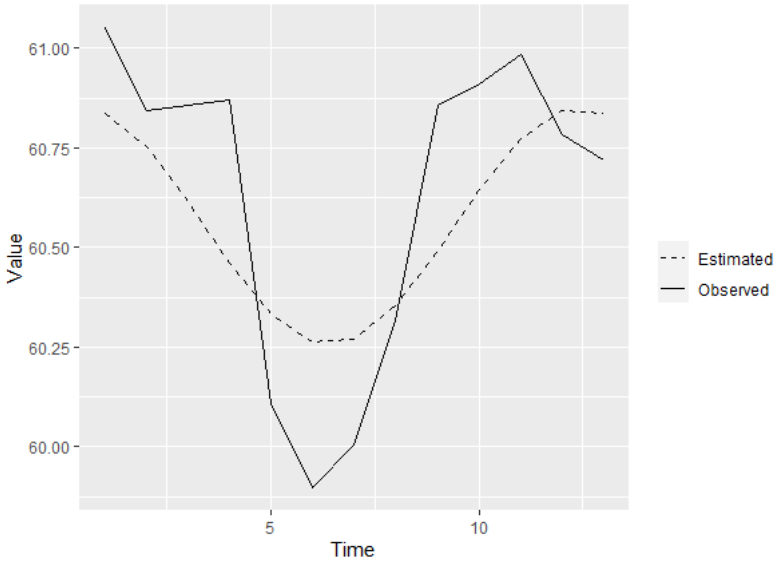


Figure 4.11. Plot of fitted vs observed values from mean population cosinor model of maternal weight.

Maternal individual weight trajectories were modelled using the cosinor package in R, as described in Section 4.3.1. The mean population cosinor model was fitted for maternal weight change over the study (see Figure 4.11). This model included data from 371 women. Those with more than 5 missing values were excluded from the model. This cut off was selected to provide a good model fit and yet maximising study participants in the model. The rhythm detection test was highly significant suggesting a good model fit (F=19.67 , p= 7.60×10^{-9}) and confirming a cyclical pattern of weight change across the year. The percent rhythm was 0.67 (p=0.0006), meaning that 67% of variance was explained by the rhythm. The population cosinor parameters were; MESOR = 60.55, (95%CI, 59.34 - 61.77), amplitude = 0.30 (0.20 - 0.39) and, acrophase= -0.21 (-0.68 – 0.20) (See Figure 4.8 for an explanation of these terms) .

4.4.4 Maternal weight results - association of maternal modelled weight trajectories with *POMC* methylation

There was no significant correlation between maternal weight MESOR, amplitude or acrophase and mean *POMC* methylation z score (see Figure 4.12).

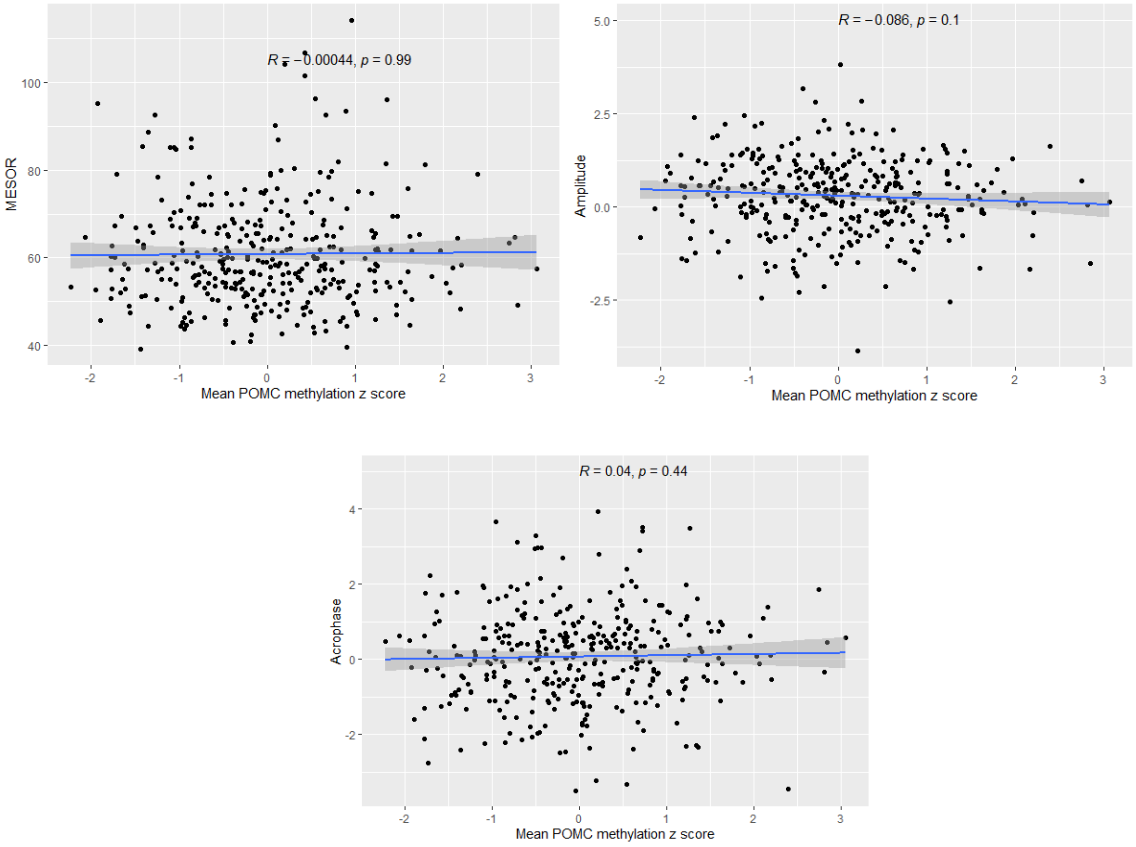


Figure 4.12 Scatterplot of maternal weight MESOR, amplitude and acrophase plotted against mean POMC methylation z score. Linear regression line fitted with 95%CI.

In linear regression models adjusted for age, height and height^2 there was no significant association between mean *POMC* methylation z score and neither maternal weight MESOR, amplitude nor acrophase (see Table 4.8). There was a significant association with age and all cosinor parameters. There was a significant association between age and 1/MESOR (β=--0.00005, p=0.05) suggesting that the rhythm adjusted mean is positively associated with age i.e. weight rises with increasing age (see Figure 4.13). There was a significant association between age and amplitude (β=0.024, p=0.003), meaning for every year increase in age there was a 0.024kg increase in weight change across the year. There was a significant association between age and acrophase (β=-0.020, p=0.04), meaning for every year increase in age the timing of the weight change peak was 1.1 days earlier.

Mother's weight Model predictors	**Dependent variable:** 1/MESOR (inverse transformation)	**Dependent variable:** Amplitude	**Dependent variable:** Acrophase
Mean *POMC* methylation z score	-0.00004 (0.0002)	-0.061 (0.053)	0.037 (0.067)
Age (years)	-0.00005* (0.00002)	0.024*** (0.008)	-0.020** (0.010)
Height (metres)	-0.136 (0.120)	-15.403 (41.663)	-94.314 (52.064)
Height^2 (metres^2)	0.037 (0.037)	5.407 (12.900)	27.942 (16.120)
Observations	347	347	347
R^2	0.107	0.041	0.051
Note:	*p<0.1 **p<0.05 ***p<0.01		

Table 4.8. Linear regression model of maternal weight MESOR, amplitude and acrophase regressed against mean POMC methylation z score adjusted for relevant covariates. Regression coefficient and (standard error) shown in brackets.

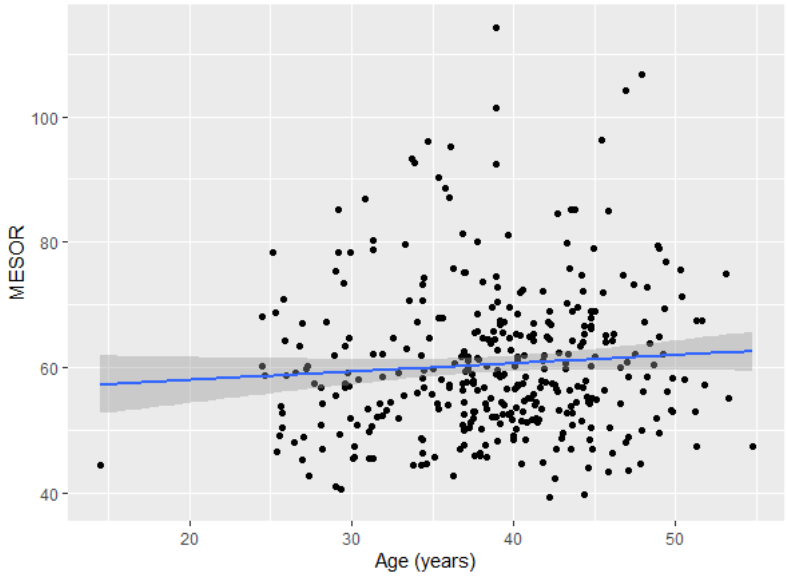


Figure 4.13 The relationship between maternal MESOR and age.

4.4.5 Maternal FMI – summary statistics

The plot of maternal FMI change over the year is shown in Figure 4.14. In a similar finding to maternal weight, study month 5 (September-October) had the lowest mean FMI. Month 8 (December-January) showed the peak of FMI which was earlier than the peak for weight (April-May). Maternal FMI per timepoint was calculated from the fat mass measurement derived from bioimpedance.

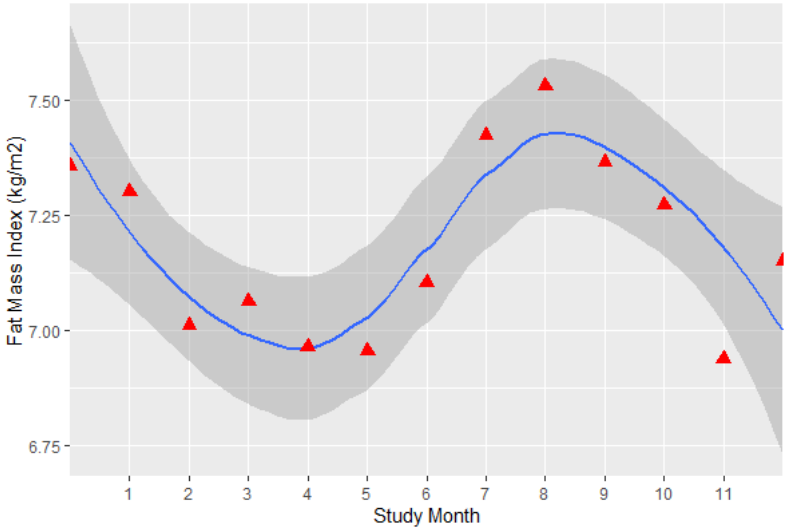


Figure 4.14. Maternal FMI plotted against study month 0-12. Red triangles indicate mean value at each time point. Loess regression line (blue) with 95% confidence intervals

To assess the validity of bioimpedance fat mass measurement, comparisons were made with DXA-derived measure of body fat. Bioimpedance fat mass and DXA-derived fat mass from 277 paired DXA-bioimpedance measurements showed the two measures were highly correlated (Spearman R=0.97, $p=2.2x10^{-16,}$ see Figure 4.15).

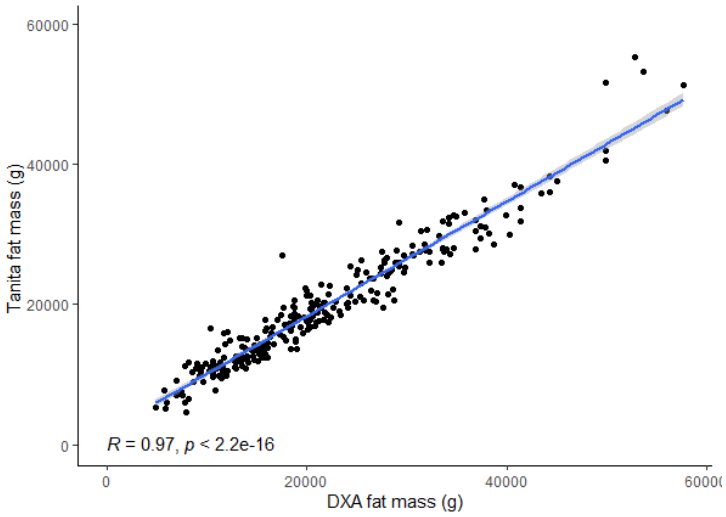


Figure 4.15. Scatterplot of maternal bioimpedence (TANITA) derived fat mass and DXA-derived fat mass. Linear regression line fitted with 95%CI

4.4.6 Maternal FMI trajectories- cosinor modelling

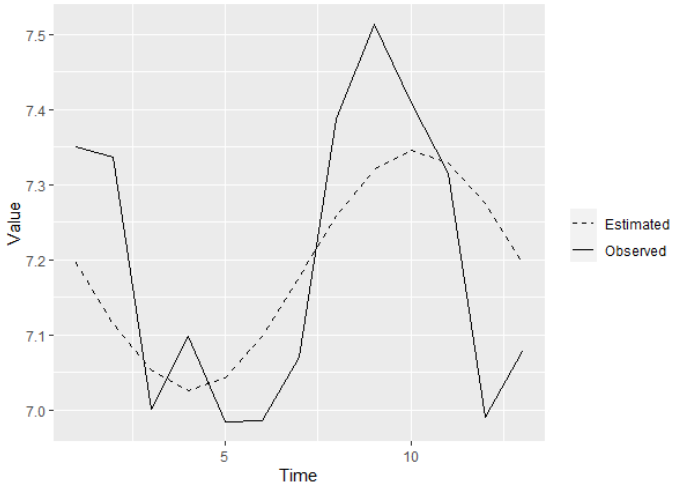


Figure 4.16. Plot of fitted vs observed values from mean population cosinor model of maternal FMI

The mean population cosinor model was fitted for maternal FMI change over the study (see Figure 4.16). This model included data from 367 women. As for maternal weight, those with greater than 5 missing values were excluded from the model. This cut off was used as it provided a balance between model fit and maximising study participants in the model. The rhythm detection test was highly significant suggesting a good model fit (F=27.20 , p=9.7x10^{-12}) which confirms a seasonal rhythm in FMI change. The percent rhythm was 0.42 (p=0.016), meaning that 42% of variance was

explained by the rhythm. The population cosinor parameters were; MESOR = 7.19 (95%CI, 6.86 – 7.51), amplitude = 0.16 (0.12 - 0.39), acrophase= -5.30 (-5.09 to -5.51).

4.4.7 Maternal FMI results - association of maternal modelled FMI trajectories with *POMC* methylation

There was a significant negative correlation between maternal FMI amplitude and mean *POMC* methylation z score (Pearson's R=-0.16, p=0.0026). There was no significant correlation between neither maternal FMI MESOR nor acrophase and mean *POMC* methylation z score, see Figure 4.17.

In linear regression models adjusted mother's age, mean *POMC* methylation z score was negatively associated with maternal FMI amplitude (β=-0.045, p=0.011) meaning that for every SD increase in methylation z score the amplitude of FMI change reduced by -0.045 kg/m^2 (see Table 4.9).

There was no significant association with maternal FMI MESOR nor acrophase and mean *POMC* methylation z score (see Table 4.9). Age was significantly associated with MESOR (β=0.009, p=0.01) such that for every year increase in age the rhythm adjusted (year round mean) FMI increased by 0.9%. Age was significantly associated with amplitude (β=0.008, p=0.004) such that for every year increase in age the amplitude of FMI change increased by 0.008kg.

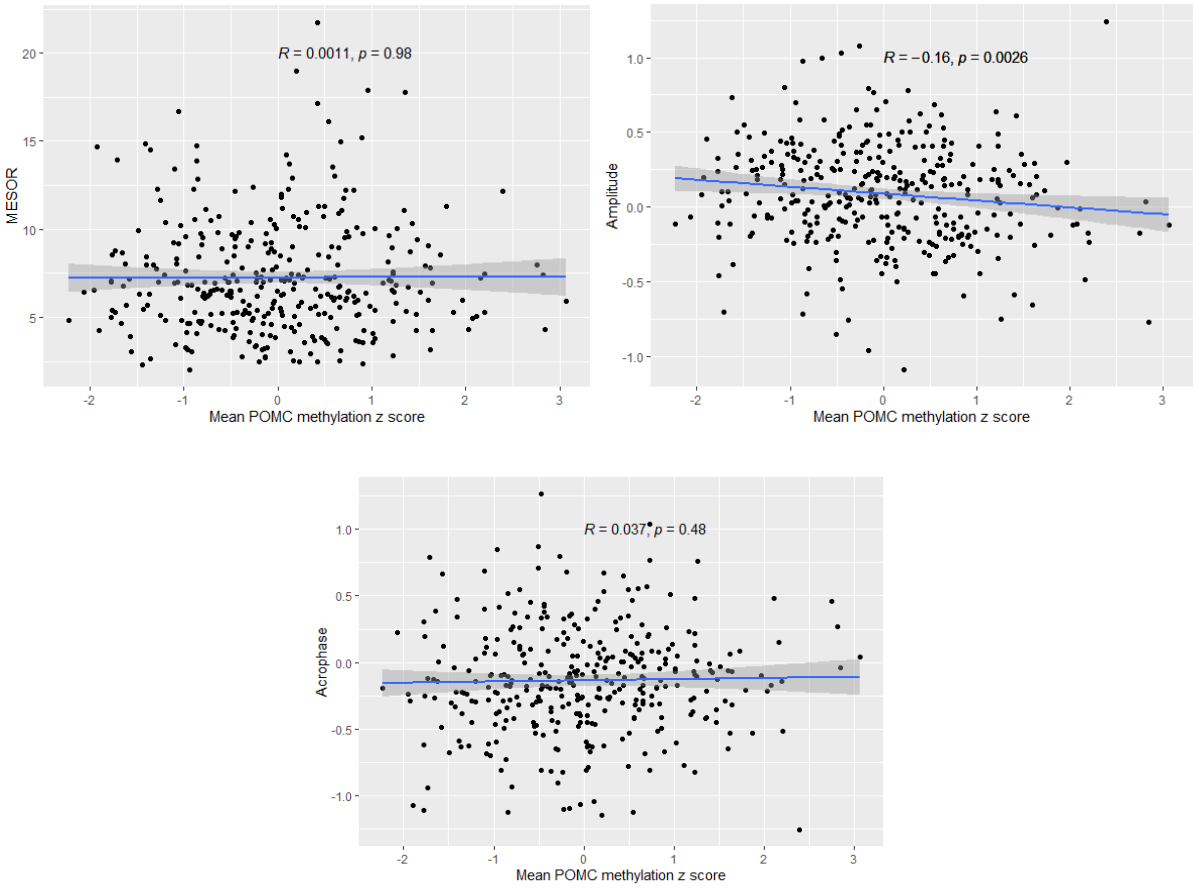


Figure 4.17. Scatterplot of maternal FMI MESOR, amplitude and acrophase plotted against mean POMC methylation z score. Linear regression line fitted with 95%CI

Maternal FMI Model predictors	**Dependent variable:** logMESOR	**Dependent variable:** Amplitude	**Dependent variable:** Acrophase
Mean *POMC* methylation z score	0.003 (0.023)	-0.045** (0.018)	0.008 (0.021)
Age (years)	0.009** (0.003)	0.008*** (0.003)	-0.006 (0.003)
Observations	366	366	366
R^2	0.018	0.040	0.009
Note:	* p<0.1 ** p<0.05 *** p<0.01		

Table 4.9 Linear regression model of maternal FMI logMESOR, amplitude and acrophase regressed against mean POMC methylation z score adjusted mother's age. Regression coefficient and (standard error) shown in brackets.

4.4.8 Maternal results –association between maternal leptin and DXA-derived FMI and *POMC* methylation

Additional measures of adiposity included leptin and DXA-derived fat mass index taken at baseline, midline and endline in *subset* mothers. These additional measures were taken for a number of reasons i) DXA is considered the reference method for fat mass assessment in clinical research and therefore provided a highly accurate measurement of fat mass ii) leptin is a biochemical measure of fat mass and considered a *long* term measure of energy balance and fat iii) leptin interacts with *POMC* neurons in the ARC of the hypothalamus[18] to increase POMC expression and thus satiety signal[19] and therefore the relationship between circulating leptin levels and *POMC* methylation was explored.

Seasonal change in maternal leptin and DXA-derived fat mass

Measurements of leptin at baseline, midline and endline were highly correlated (Spearman R between 0.73 and 0.85) with DXA measurements of total fat mass (see Figure 4.18). Sixty-four mothers had 3 leptin measurements and 65 had 3 DXA measurements. There was no change in DXA-derived FMI across the three timepoints and overall there was no significant difference between leptin levels at baseline, midline and endline period (p=0.99, Kruskall-Wallis test) despite a trend for leptin increasing across the study (Figure 4.19).

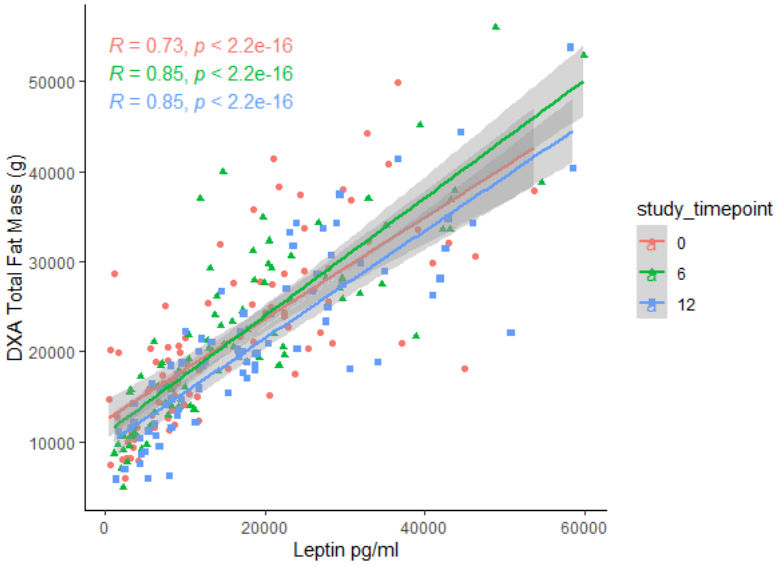


Figure 4.18 Correlation between paired maternal baseline, midline, endline leptin and DXA fat mass. Spearman correlation coefficient showed with P value. Linear regression line shown with 95% CI for each

timepoint. study time point 0=baseline, 6=midline and 12=endline

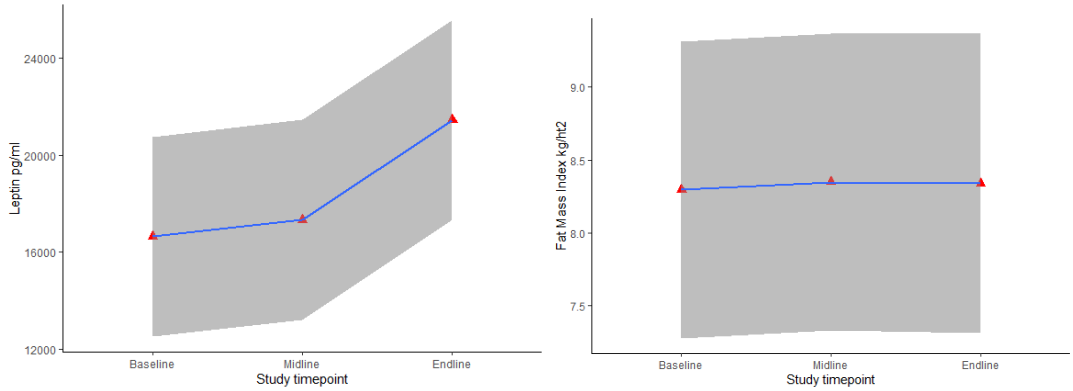


Figure 4.19. Trend of maternal leptin levels (left, n-64 women with data from all three timepoints) and FMI (right, n=65 women with data from all three timepoints) across baseline, midline and endline with measurements at each timepoint. Regression line with 95%CI shown.

Maternal leptin and DXA-derived fat mass and *POMC* methylation

There was no significant correlation between leptin levels at baseline and mean *POMC* methylation z score (R=-0.095, p=0.38, n=95).

The change in FMI and leptin between i) baseline and midline ii) midline and endline was compared in multiple linear regression models adjusted for relevant covariables (see Table 4.10). Higher *POMC* methylation was associated with a smaller difference in leptin between midline and endline. Such that for every unit increase in mean *POMC* methylation z score there was an associated 3,488 pg/ml smaller change in leptin between study midline and endline. This was consistent with the finding of a reduced amplitude of FMI change (or more stable fat mass) demonstrated in the mothers FMI consinor models (see Table 4.9).

Model Predictors	*Dependent variable:*			
	DXA-derived FMI change		**Leptin change**	
	Baseline minus midline (kg/m^2)	Midline minus Endline (kg/m^2)	Baseline minus midline (pg/ml)	Midline minus Endline (pg/ml)
Mean *POMC* methylation z score	0.041 (0.19)	-0.052 (0.15)	497.43 (1,306.65)	**-3,488.31** ** **(1,668.12)**
Age (years)	0.007 (0.022)	-0.020 (0.025)	222.90 (248.97)	-169.99 (299.36)
Height (metres)			41,909.60(24,374.74)	15,709.600(28,266.88)
Observations	66	59	65	76
R^2	0.003	0.012	0.056	0.062
Note:	** p<0.05 *** p<0.01			

Table 4.10 Linear regression models examining effect of mean POMC methylation z score and change in FMI and leptin between study timepoints. Regression coefficient and (standard error) shown in brackets

4.4.9 Child results – subset and main cohort comparisons

For children, those in the subset weighed more (18.23 vs 17.62kg, p=0.03), had a higher BMI z score (-0.76 vs -.1.18, p<0.001) with significant differences in BMI category (p=0.016) compared to those in the main study children (see Table 4.11). There was no difference in age or mean *POMC* methylation z score.

Children	**Overall**	**Main**	**Subset**	**p-value**
Mean age (SD, range)	6.15 (0.91, 4.18 - 7.69)	6.19 (0.88)	6.18(0.97)	0.82
Sex	Male 54% Female 46%	Male 53% Female 47%	Male 56% Female 44%	0.63
Weight (kg) (SD)	17.79 (2.59)	17.62 (2.57)	18.23 (2.58)	**0.03**
Mean BMI (kg/m2) z score (SD)	-1.06 (0.86)	-1.18 (0.82)	-0.76 (0.90)	**<0.001**
BMI category (%)				**0.016**
Underweight	13.3	15.4	7.8	
Normal Weight	85.7	84.2	89.6	
Overweight	1.0	0.3	2.6	
Obese	0	0	0	
Mean methylation z score (SD)	-0.13 (0.94)	-0.13(0.92)	-0.14(0.99)	**0.86**

Table 4.11 Baseline characteristics of POMC study children. Comparisons between study groups for age, weight, BMI and mean POMC methylation z score were assessed using student t test. Differences in BMI category between study groups was made using chi square test (using the numbers in each category however % in each category shown in table). **Key:** SD=standard deviation, BMI=body mass index.

4.4.10 Child weight – summary statistics

Both weight for age z score and BMI z score were plotted over time from baseline to month 12 (see Figure 4.20). There were similar patterns of weight for age and BMI change across the year with peaks at baseline (April-May) and a nadir in month 5 (September-October). Weight for age was used as the weight outcome for subsequent analysis.

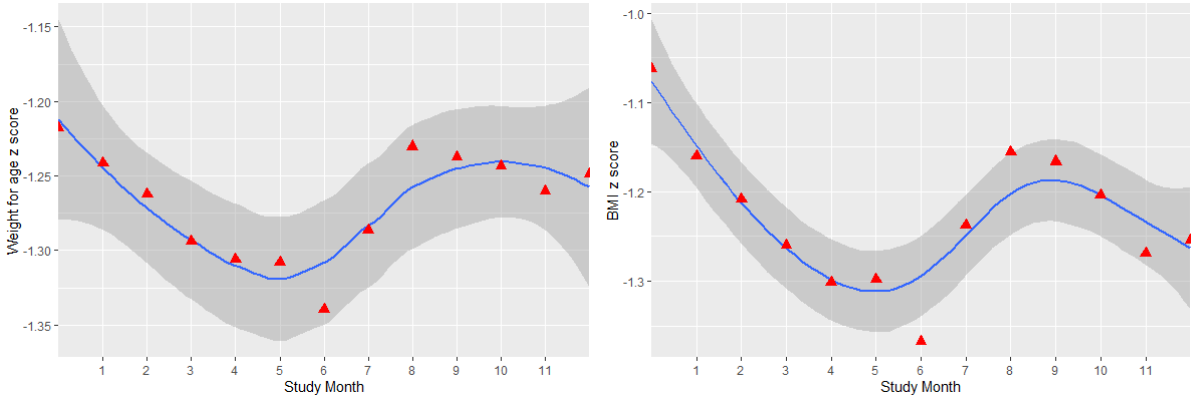


Figure 4.20. Child weight (left) and BMI (right) plotted against study month 0-12. Red triangles indicate mean value at each time point. Loess regression line (blue) with 95% confidence intervals

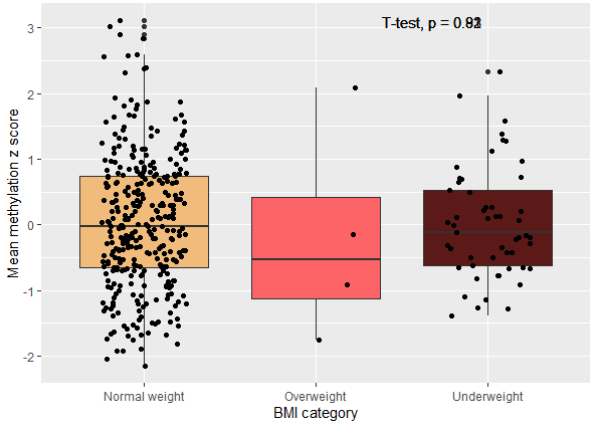


Figure 4.21. Boxplots of child BMI category at baseline and mean POMC methylation z score.

There was no difference in mean POMC methylation z score and BMI category at baseline (see Figure 4.21). Note no child was classified as obese.

4.4.11 Child weight trajectories- cosinor modelling

The mean population cosinor model was fitted for child weight for age z scores across the study period (see Figure 4.22). This model included data from 420 children. As for maternal consignor models, those with greater than 5 missing values were excluded from the model. The rhythm detection test was highly significant suggesting a good model fit (F= 49.5, $p=3.9x10^{-20}$) confirming the presence of a seasonal rhythm of weight change. The percent rhythm was 0.77 ($p=7.7x\ 10^{-5}$), meaning that 77% of variance was explained by the rhythm. The mean population cosinor parameters were MESOR =-1.28 (95% CI -1.35 to -1.20), amplitude = 0.042 (0.033 to 0.050) and acrophase = -6.25 (-6.45 to -6.06).

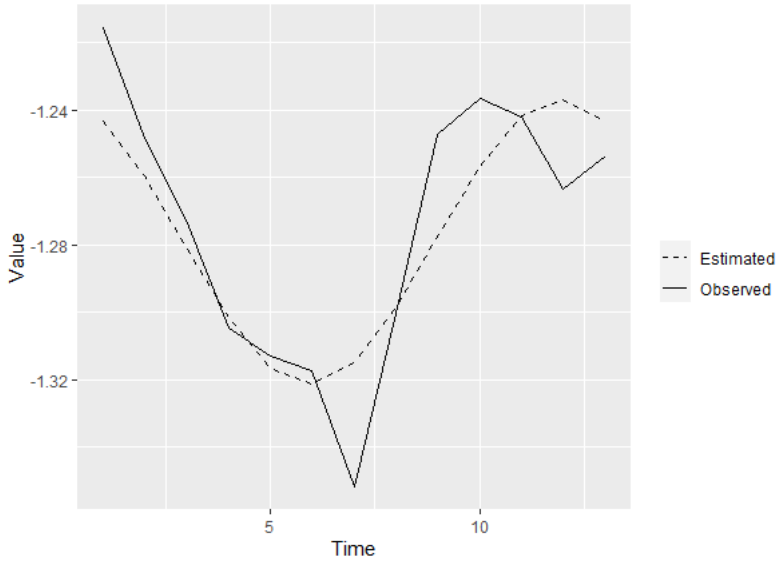


Figure 4.22. Plot of fitted vs observed values from mean population cosinor model of child weight for age z scores.

4.4.12 Child weight results - association of child modelled weight trajectories with *POMC* methylation

There was no significant correlation between child weight for age z score MESOR, amplitude nor acrophase and mean *POMC* methylation z score (see Figure 4.23).

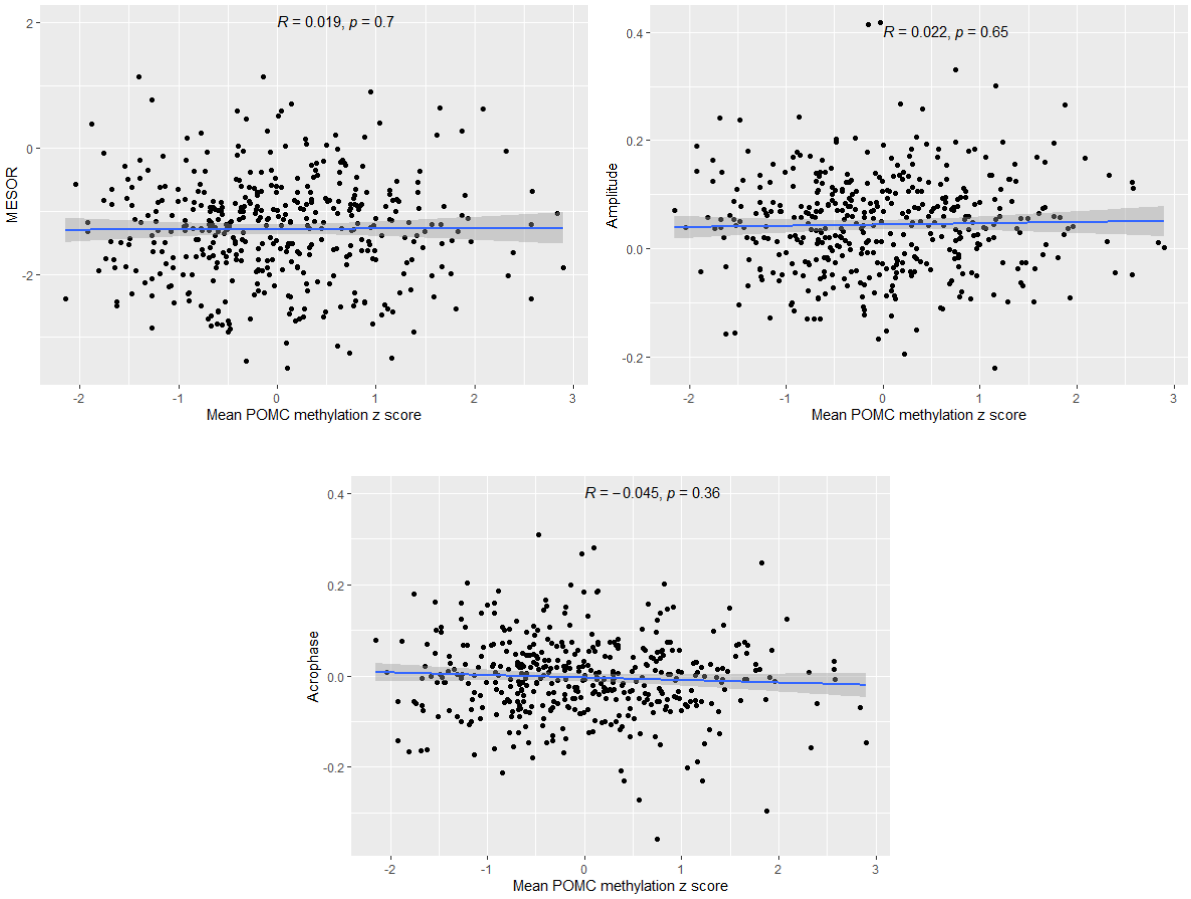


Figure 4.23. Scatterplot of child weight for age z score MESOR, amplitude and acrophase plotted against mean POMC methylation z score. Linear regression line fitted with 95%CI

In linear regression models adjusted for sex, height and height^2 there was no significant association between mean *POMC* methylation z score and child weight for age MESOR, amplitude nor acrophase (see Table 4.12). There was a significant association with sex and weight such that girls had 0.21 kg higher weight for age z score (p=0.003).

Child weight for age Model predictors	**Dependent variable:** MESOR	**Dependent variable:** Amplitude	**Dependent variable:** Acrophase
Mean *POMC* methylation z score	0.001 (0.049)	0.004 (0.006)	-0.001 (0.006)
Sex (Female)	0.210*** (0.070)	0.005 (0.009)	-0.011 (0.009)
Height (metres)	-12.003 (12.796)	1.967 (1.631)	-1.246 (1.566)
Height^2 (metres^2)	7.704 (5.534)	-0.845 (0.705)	0.545 (0.677)
Mean *POMC* methylation z score*Sex (interaction)	-0.073 (0.073)	-0.004 (0.009)	-0.010 (0.009)
Observations	420	420	420
R^2	0.252	0.005	0.012
Note:	*p<0.1 **p<0.05 ***p<0.01		

Table 4.12 Linear regression model of children's weight MESOR, amplitude and acrophase regressed against mean POMC methylation z score adjusted for relevant covariates. Regression coefficient and (standard error) shown in brackets.

4.4.13 Child FMI – summary statistics

FMI was plotted over time from baseline to month 12 (see Figure 4.24). The peak of FMI was seen at baseline (April-May) and a nadir in month 3-4 (July to September).

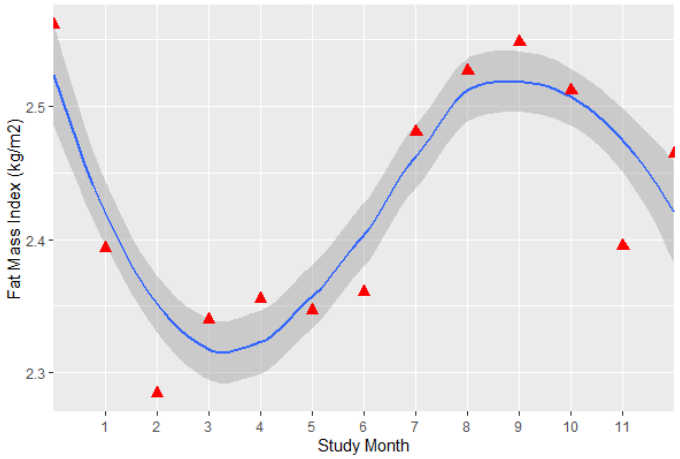


Figure 4.24. Child FMI plotted against study month 0-12. Red triangles indicate mean value at each time point. Loess regression line (blue) with 95% confidence intervals

To assess the validity of bioimpedance measure of fat mass comparisons were made with DXA-derived fat mass. Bioimpedance fat mass and DXA-derived fat mass from 397 paired DXA-bioimpedance measurements showed the two measures were well correlated (Spearman R=0.75, $p=2.2x10^{-16,}$ see figure 4.25).

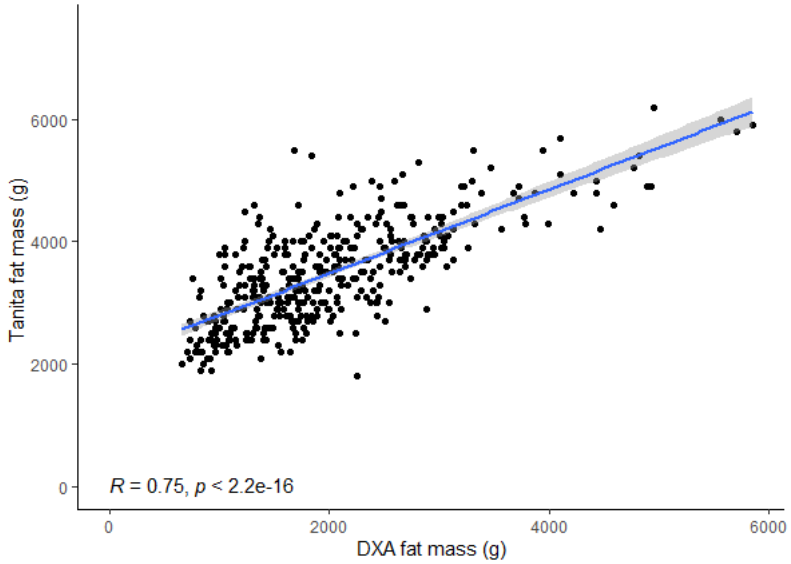


Figure 4.25. Scatterplot of child bioimpedence (TANITA) derived fat mass and DXA-derived fat mass. Linear regression line fitted with 95%CI

4.4.14 Child FMI - cosinor modelling

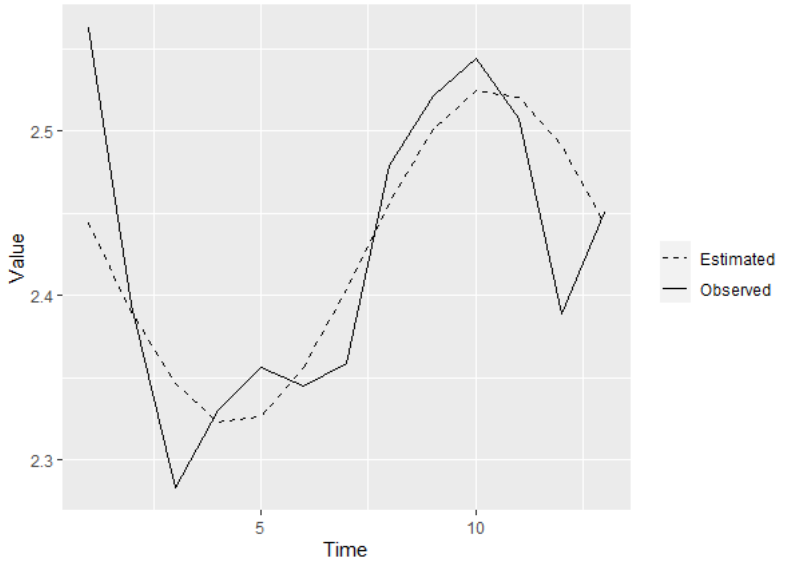


Figure 4.26. Plot of fitted vs observed values from mean population cosinor model of child FMI.

The mean population cosinor model was fitted for child FMI across the study period (Figure 4.26). This model included data from 413 children. As for mother's consinor models, those with greater than 5 missing values were excluded from the model. The rhythm detection test was highly significant suggesting a good model fit (F= 178.3, p=1.4x10^{-57}) and the presence of a seasonal rhythm

for FMI change. The percent rhythm was 0.67 (p=0.0007), meaning that 67% of variance was explained by the rhythm. The mean population cosinor parameters were MESOR =2.4 (95% CI 2.38 to 2.46), amplitude = 0.10 (0.09 to 0.11) and acrophase = -5.43 (-5.35 to -5.52).

4.4.15 Child results - association of child modelled FMI trajectories with POMC methylation

There was no significant correlation between child FMI MESOR, amplitude or acrophase and mean *POMC* methylation z score (see Figure 4.27).

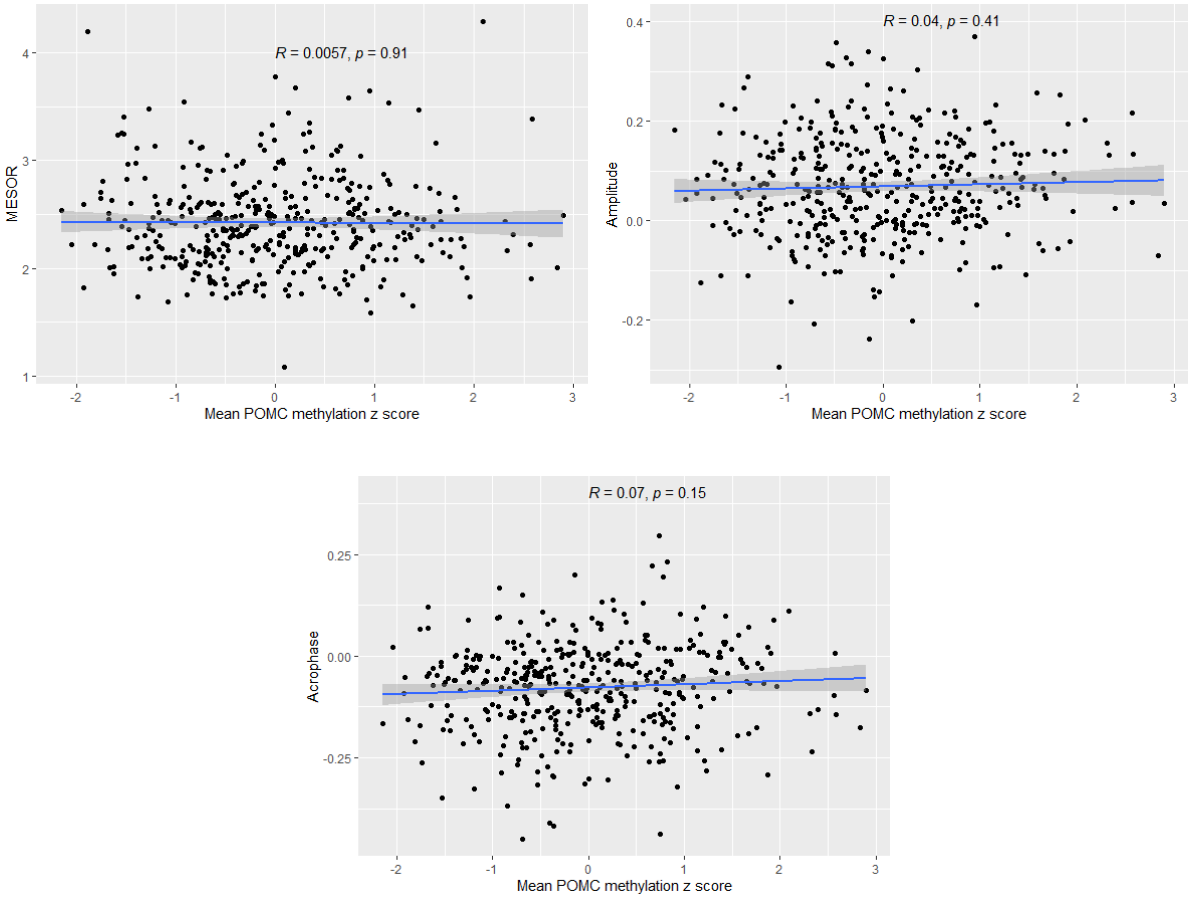


Figure 4.27. Scatterplot of child FMI MESOR, amplitude and acrophase plotted against mean POMC methylation z score. Linear regression line fitted with 95%CI

In linear regression models adjusted for sex and age there was no significant association between mean *POMC* methylation z score and child weight for age MESOR, amplitude or acrophase (see Table 4.13). There was a significant association with sex and FMI such that girls a higher rhythm adjusted mean logFMI ($p=1.93 \times 10^{-14}$). When the log coefficient was back extrapolated it meant that girls had a 13.3% higher FMI compared to boys.

There was weak evidence of an interaction between mean *POMC* methylation z score and sex (see Table 4.13 and Figure 4.28) on logMESOR FMI. For girls there was a negative association between mean *POMC* methylation z score and logMESOR FMI but for boys there was the opposite relationship (positive association). The results should be viewed with caution due to the number of associations considered here and there was no adjustment for multiple testing. This analysis was not part of the primary hypothesis testing.

Child FMI Model predictors	**Dependent variable:** logMESOR	**Dependent variable:** Amplitude	**Dependent variable:** Acrophase
Mean *POMC* methylation z score	0.016 (0.011)	0.001 (0.007)	0.012 (0.008)
Sex (Female)	0.125*** (0.016)	-0.005 (0.010)	-0.016 (0.011)
Age (years)	-0.010 (0.008)	0.016*** (0.005)	-0.019*** (0.006)
Mean *POMC* methylation z score *Sex (interaction)	-0.037** (0.016)	0.002 (0.010)	-0.004 (0.011)
Observations	413	413	413
R^2	0.149	0.026	0.037
Note:	* p<0.1 ** p<0.05 *** p<0.01		

Table 4.13. Linear regression model of child FMI logMESOR, amplitude and acrophase regressed against mean POMC methylation z score adjusted for relevant covariates. Regression coefficient and (standard error) shown in brackets.

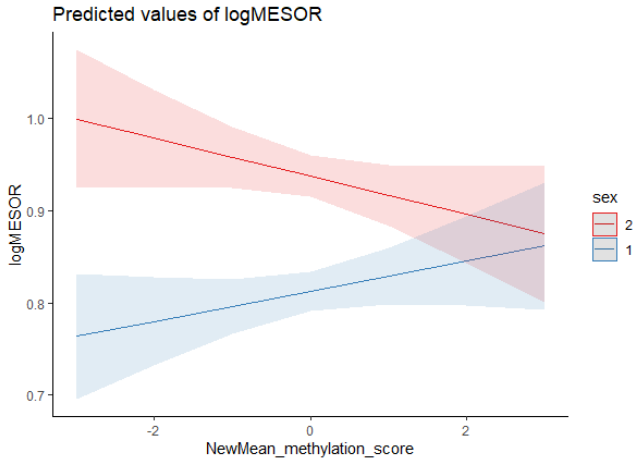


Figure 4.28. Plot of the interaction between Sex and mean POMC methylation z score and log MESOR (from FMI model from table 4.12). Sex: 2=Female, 1=Male

4.4.16 Child results - association between child leptin and DXA-derived FMI and *POMC* methylation

Additional measures of adiposity included leptin and measure of DXA-derived fat mass index measurements taken at baseline, midline and endline in *subset* children. The reason for assessing the relationship between these additional measures of fat and energy balance is the same as for the mothers (see section 4.4.8).

Seasonal change in child leptin and DXA-derived fat mass

Measurements of leptin at baseline, midline and endline were well correlated with DXA measurements of total fat mass (Spearman R between 0.63 and 0.76, see Figure 4.29). One hundred and four children had 3 leptin measurements with a trend for a reduction in leptin at midline and returning to baseline levels by endline (Figure 4.30), though overall there was no significant difference between leptin levels at baseline, midline and endline (p=0.9, Kruskall-Wallis test). One hundred children had 3 DXA scan measurements. Overall there was no significant difference between DXA-derived FMI at baseline, midline and endline (p=0.5, Kruskall-Wallis test).

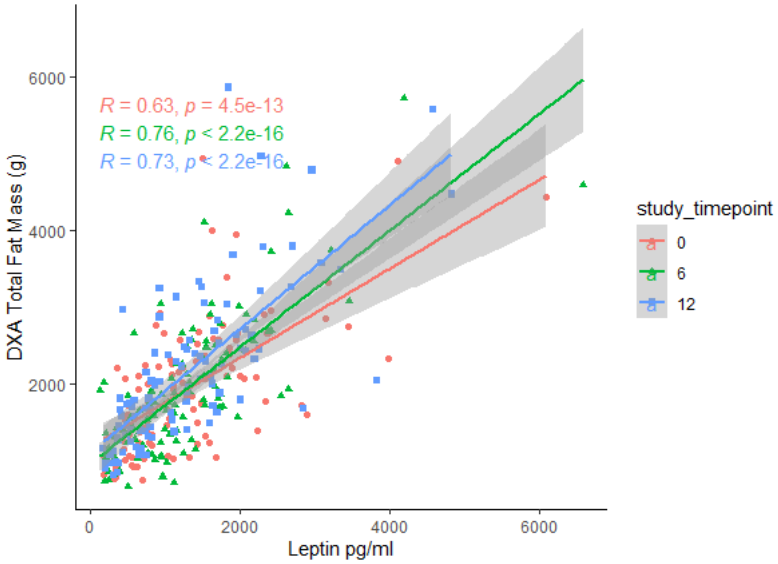


Figure 4.29. Correlation between paired child baseline, midline, endline measurements of leptin and DXA fat mass. Spearman correlation coefficient showed with P value. Linear regression line shown with 95% CI for each timepoint. Study time point 0=baseline, 6=midline and 12=endline

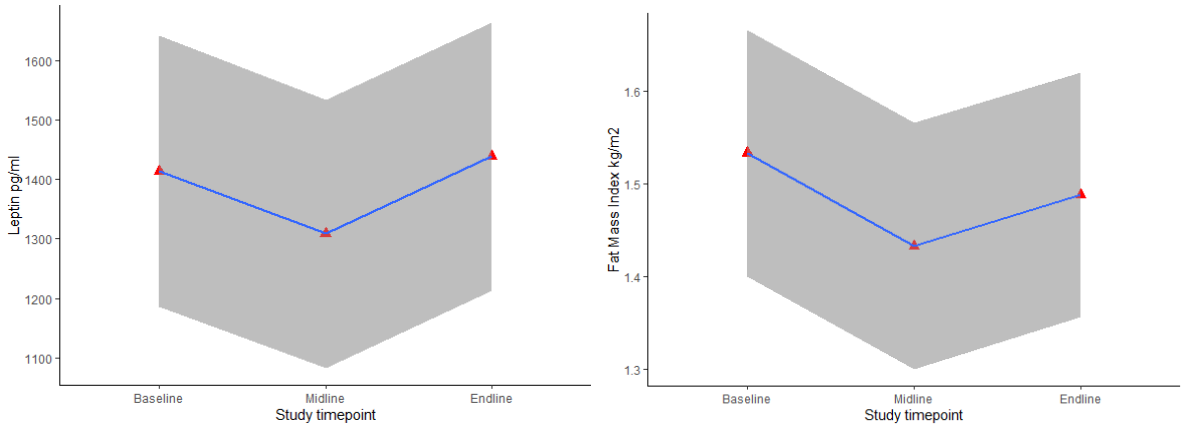


Figure 4.30. Trend of child leptin levels (left, n-104 children with data from all three timepoints) and FMI (right, n=100 children with data from all three timepoints) across baseline, midline and endline with measurements at each timepoint. Regression line with 95%CI shown.

Child leptin and DXA-derived fat mass and *POMC* methylation

There was no significant correlation between leptin levels at baseline and mean *POMC* methylation z score (R=-0.086, p=0.38, n=104).

Change in FMI and leptin between i) baseline and midline ii) midline and endline were compared in multiple linear regression models adjusted for relevant covariables (see Table 4.14). There was no significant association between mean *POMC* methylation z score and any change in leptin or FMI between any of the study points.

Model Predictors	*Dependent variable:*			
	DXA-derived FMI change		**Leptin change**	
	Baseline minus midline (kg/m^2)	Midline minus Endline (kg/m^2)	Baseline minus midline (pg/ml)	Midline minus Endline (pg/ml)
Mean *POMC* methylation z score	0.043 (0.04)	-0.032 (0.04)	-15.92 (106.08)	-92.79 (104.19)
Sex (female)	-0.10* (0.06)	0.048 (0.05)	20.60 (154.91)	-94.04 (151.50)
Age (years)	-0.071** (0.027)	-0.062** (0.027)	-76.87 (136.15)	-105.53 (132.32)
Height (metres)			-622.12 (2,031.63)	-25.25 (1,955.88)
Mean *POMC* methylation z score *Sex (interaction)	-0.051 (0.06)	0.066 (0.06)	50.587 (160.30)	-50.76 (156.72)
Observations	94	94	95	96
R^2	0.099	0.095	0.028	0.057
Note:	*p<0.1 **p<0.05 ***p<0.01			

Table 4.14. Linear regression models examining effect of mean POMC methylation z score and change in FMI and leptin between study timepoints. Regression coefficient and (standard error) shown in brackets.

4.5 Discussion

The summary of the important results are outlined below.

Maternal results:

- *POMC* methylation was not associated with seasonal weight change
- *POMC* methylation was negatively associated with maternal FMI amplitude meaning that for every SD increase in *POMC* methylation z score the amplitude of FMI change reduced by 0.045 kg/m^2
- Higher *POMC* methylation was associated with a smaller change in leptin between midline and endline such that for every unit increase in mean *POMC* methylation z score there was an associated 3,488 pg/ml smaller change in leptin between study midline and endline.

Child results:

- *POMC* methylation was not associated with seasonal weight change
- *POMC* methylation was not associated with seasonal FMI change
- There was no significant association between mean *POMC* methylation z score and any change in leptin or FMI between any of the study points.
- There was weak evidence of an interaction between mean *POMC* methylation z score and sex on logMESOR FMI; for girls there was a negative association between mean *POMC* methylation z score and logMESOR FMI but for boys there was the opposite relationship (positive association).

These results suggest that *POMC* methylation can predict the degree of *change* in fat mass across the year in Gambian women. The population mean FMI change across the year was 0.2 kg/m^2 (i.e. twice the mean amplitude). For every SD increase in *POMC* methylation the change in FMI was reduced by 0.045 (nearly half the mean amplitude). It is important to contextualise this with what we know about POMC's role in energy balance. POMC is a central component of the melanocortin system and regulates energy balance by inducing satiety (via α-MSH activation of MC4R receptors) and increasing energy expenditure[33]. Higher *POMC* methylation at the VMR has been associated with lower POMC expression potentially due to the DNA methylation disrupting binding of histone acetyltransferase P300 complex (involved in chromatin acetylation and gene activation) [7]. Higher methylation could be driving a lower satiety signal and lower energy expenditure which could lead to less fluctuation in fat mass in response to seasonal nutritional pressures. However, the potential

influence of *POMC* methylation on satiety in the study population is described in depth in the next chapter and may not be a key driver for the observed differences in fat mass change across the year.

It should be noted that there was no reported association with body weight model parameters and mean *POMC* methylation z score. The mean seasonal population weight change in mothers in this study was 0.6kg (see Figure 4.9) which is much lower than the previously reported ~3.5kg change in 2009-10[34]. Body weight is a function of both fat, lean mass, and bone and therefore body weight may have been maintained by changes in muscle mass during the agricultural harvest in the rainy season. The change in muscle mass would not be thought to be influenced by the actions of POMC.

Differential methylation in adipose tissue at numerous genes (though not *POMC* specifically) has been associated with BMI, fat mass and fat distribution in an adult European cohort[35]. A recent study in a cohort of 230 African-American adults identified methylation at CpGs (analysed from adipose tissue) from the *POMC* VMR that were strongly associated with POMC expression. Furthermore, methylation at the *POMC* VMR was causally linked to BMI via its role in epigenetic regulation of the expression of the *POMC* transcript[36]. The effect of *POMC* methylation on fat mass index change (but not weight) may therefore be linked to an indirect or direct effect of POMC on adipose tissue.

In animal studies, MC4R receptors (a target of POMC derived α-MSH) were expressed in sympathetic nervous system (SNS) neurons that innervate both brown and white adipose tissue[37,38]. A key role of the SNS is to regulate thermogenesis, leptin mobilisation and lipolysis in adipose tissue. This suggests that methylation-driven alterations in POMC expression could influence lipolysis and fat mass in adipose tissue. POMC deficient mice have been shown to have reduced sympathetic output to adipose tissue resulting in unexpectedly high adiponectin levels in obese mice which normalised once POMC expression was restored in the hypothalamus[39], suggesting a role for POMC neurons in controlling adipose tissue derived adiponectin[40].

The finding that change in leptin between midline and endline was reduced in those with higher methylation is congruent with the finding of a smaller change in FMI. The effect on leptin may simply reflect a difference in fat mass but could also point to a direct effect of POMC on leptin mobilisation from adipose tissue. The observation of a significant *POMC* methylation effect on leptin was only seen between rainy (midline) and dry (endline) seasons. This may suggest that the effect on fat mass/leptin was most apparent during the period of weight gain when women were relatively nutritional surfeit. There is an homeostatic interplay between circulating leptin, hunger and fat [41–43]. *POMC*-methylation driven altered signalling between leptin and POMC neurons could explain the difference different seasonally driven FMI responses seen. Furthermore, obesity has been described

as condition where the defended level of fat mass is increased[41] and therefore the more stable fat mass seen in those with higher *POMC* methylation may reflect a more robust defence of fat mass.

For children's (rhythm-adjusted) mean FMI, there was evidence of a weakly significant interaction between mean *POMC* methylation z score and sex, although results from this secondary analysis should be treated with caution. Overall girls had higher FMI than boys which has been reported before in pre-pubertal children of both European[44] and African ancestry[45]. Girls demonstrated an inverse relationship between FMI and *POMC* methylation but the reverse was true for boys.

Sexual dimorphism in hypothalamic circuits and specifically relating to POMC neurons has been consistently reported in animal models[46]. Sex specific DNA methylation patterns in *POMC* have been seen in offspring of rats fed a high fat diet in pregnancy[47]. In a murine model, lack of POMC expression was associated with increased appetite, reduced energy expenditure and greater adiposity with a greater effect seen in female mice[48]. Lower POMC expression has been associated with energy restricted male sheep but not female[49]. Furthermore, sex specific effects on body weight have been seen in mice lacking GLUT2[50] or GABA type B[51] receptors on POMC neurons with a male mice having an increased effect on body weight compared to females. One possible explanation in these studies of the observed sex specific effects of POMC on body weight and adiposity could be an effect of oestrogen effecting POMC expression via oestrogen-receptors on POMC neurons[52,53]. Children in this study were prepubertal (aged between 5-8 years of age) and therefore the interaction of sex and *POMC* methylation on FMI is very unlikely to mediated by differences in oestrogen signalling. Another potential explanation could lie with differences in sex steroid exposure in early life. In rat models, differences in synaptic organisation in neural feeding circuits has been postulated to be linked to the early sex steroid exposure[54,55]. The apparent opposite effect on fat mass is unexpected and certainly warrants further research to elucidate the underlying physiological mechanism.

Children's FMI change across the year was not associated with *POMC* methylation as it was for adults. It is widely reported that the genetic influence on body weight and BMI changes over the life course. For example, Kera and colleagues found that the polygenic effect on weight emerges in infancy but increases in adulthood[56]. It is unclear how epigenetic modifications could differentially influence body weight, fat mass or BMI across the life course but it may help explain the significant effect seen in adults but not children in my study. However, the association between *POMC* VMR hypermethylation and obesity was seen in both children and adults in a German cohort[7,9]. In my study there was no association between *POMC* methylation and BMI category in both women and children. The effect of POMC on body weight and fat may be situational. For example in rural

Gambia, *POMC* hypermethylation may have limited scope to influence body weight due to nutritional scarcity across the population. However, in more obesogenic environments (a sum of numerous environmental and conditions of life that promotes obesity in an individual or a population[57]) such as found in Germany[58], the opportunity to consume excess calories and have a longstanding positive energy balance (leading to obesity) is far greater. An advantage to the *POMC* study (over the cross-sectional study designs by Kuhnen and colleagues[7,9]) is that I could observe prospectively the influence of *POMC* methylation on the annual dynamic changes in body weight and fat and thus gain a greater understanding of its role in energy balance. To the best of my knowledge there is one other human *prospective* study related to *POMC* epigenetics. This reported that methylation in the *POMC* promotor region predicted the success of weight loss intervention[59]. In this study, lower methylation levels were associated with a greater chance of success with weight loss interventions. Though in a different region of the gene this suggests that *POMC* methylation may be associated with subsequent weight change.

Obesity is known to be more prevalent with age. In the *POMC* study, a positive association with fat mass index was reported in women. It is widely reported that fat mass increases and muscle mass reduces with age[60] in keeping with what was observed in Gambian women. Furthermore, increasing age was associated with greater amplitude of weight and FMI change across the year. The cause of this maybe multifaceted and was outside the remit of investigation for this study. For example, it is unclear if age is related to the amount of agricultural work load, family food allocation or an underlying physiological process.

4.5.1 Limitations

Though significant findings were presented all had a p value >0.01. Though I have attempted to rationalise outcomes measures for example I did not analyse anthropometric measures of fatness such as skin fold thickness or MUAC. However, there were still multiple tests performed in the analysis. There was no correction for multiple testing in the analysis. It is important to replicate these findings in this population and explore similar findings in disparate populations.

There was significant drop out by mothers in this study; mostly due to pregnancy which may have influenced power to detect weight change differences. This taken together with the lower than expected seasonal effect size, may have influenced power to detect significant associations between *POMC* methylation and weight changes. The interaction of *POMC* methylation, sex and FMI could not be explored in adults as there were no adult males recruited to the study. Exploring if a *POMC* methylation-sex interaction with FMI persists into adulthood should be considered in future study.

Methylation was measured in leucocytes and was a proxy for the proposed effect in the hypothalamus. MEs (such as *POMC*) are defined by systemic (i.e. not tissue specific) methylation and therefore it would be appropriate not to adjust for cell composition[9,61]. A previous study has demonstrated *POMC* methylation measured in cells from the arcuate nucleus of the hypothalamus and leucocytes are correlated[9].

The relationship between POMC methylation and seasonal driven fat mass index changes may not be unidirectional. An advantage of the study design is that *POMC* methylation measure was taken prior to the start of the weight and fat mass monitoring i.e. before the phenotype is observed. However, it could be that repeated patterns of weight and fat changes in previous years had influenced *POMC* methylation levels. For example, if large swings in FMI in previous years altered *POMC* methylation by decreasing methylation then the observed association between *POMC* methylation and women's fat mass index observed in this study could be the *result* of previous body composition fluxes. Though DNA was obtained at midline and endline (to explore seasonally driven POMC methylation changes) the samples were not processed due to time and logistical constraints related to the COVID-19 pandemic. Analysis of these samples would be beneficial to explore methylation stability and the potential for reverse causation effects.

4.5.2 Conclusions

This study identifies a significant association between *POMC* methylation on maternal FMI change across the year. Those with higher *POMC* methylation demonstrated a smaller change in fat mass across the year. This association was not capitulated in children however a significant sex and *POMC* methylation interaction with FMI was observed.

These findings suggest a potentially important role of *POMC* methylation in regulating seasonally driven changes in FMI in Gambian women.

4.6 References

1. Andermann ML, Lowell BB. Toward a Wiring Diagram Understanding of Appetite Control. *Neuron*. 2017;95(4):757-778. doi:10.1016/j.neuron.2017.06.014

2. Raffan E, Dennis RJ, O'Donovan CJ, et al. A Deletion in the Canine POMC Gene Is Associated with Weight and Appetite in Obesity-Prone Labrador Retriever Dogs. *Cell Metab*. 2016;23(5):893-900. doi:10.1016/j.cmet.2016.04.012

3. Coll AP, Farooqi IS, Challis BG, Yeo GSH, O'Rahilly S. Proopiomelanocortin and Energy Balance: Insights from Human and Murine Genetics. *J Clin Endocrinol Metab*. 2004;89(6):2557-2562. doi:10.1210/jc.2004-0428

4. Farooqi IS, Keogh JM, Yeo GSH, Lank EJ, Cheetham T, O'Rahilly S. Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene. *N Engl J Med*. 2003;348(12):1085-1095. doi:10.1056/NEJMoa022050

5. Krude H, Biebermann H, Luck W, Horn R, Brabant G, Grüters a. Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat Genet*. 1998;19(2):155-157. doi:10.1177/000992289903800416

6. Farooqi IS, Drop S, Clements A, et al. Heterozygosity for a POMC-null mutation and increased obesity risk in humans. *Diabetes*. 2006;55(9):2549-2553. doi:10.2337/db06-0214

7. Kuehnen P, Mischke M, Wiegand S, et al. An alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet*. 2012;8(3):1-12. doi:10.1371/journal.pgen.1002543

8. Landeghem F Van, Deimling A Von, Escher F, et al. A role for b -melanocyte-stimulating hormone in human body-weight regulation. 2006;(February):141-146. doi:10.1016/j.cmet.2006.01.007

9. Kühnen P, Handke D, Waterland R a, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016:1-8. doi:10.1016/j.cmet.2016.08.001

10. Hall KD, Sacks G, Chandramohan D, et al. Quantification of the effect of energy imbalance on bodyweight. *Lancet*. 2011. doi:10.1016/S0140-6736(11)60812-X

11. Christiansen E, Garby L, Sørensen TIA. Quantitative analysis of the energy requirements for

development of obesity. *J Theor Biol*. 2005;234(1):99-106. doi:https://doi.org/10.1016/j.jtbi.2004.11.012

12. Brown WJ, Flores TR, Keating SE, Mielke GI. Trajectories and determinants of weight gain in two cohorts of young adult women born 16 years apart. *Int J Obes*. 2021;45(7):1553-1564. doi:10.1038/s41366-021-00819-0

13. Saely CH, Geiger K, Drexel H. Brown versus white adipose tissue: a mini-review. *Gerontology*. 2012;58(1):15-23. doi:10.1159/000321319

14. Gómez-Valadés AG, Pozo M, Varela L, et al. Mitochondrial cristae-remodeling protein OPA1 in POMC neurons couples Ca2+ homeostasis with adipose tissue lipolysis. *Cell Metab*. 2021;33(9):1820-1835.e9. doi:https://doi.org/10.1016/j.cmet.2021.07.008

15. Rodrigues KC da C, Pereira RM, Campos TDP de, et al. The Role of Physical Exercise to Improve the Browning of White Adipose Tissue via POMC Neurons . *Front Cell Neurosci* . 2018;12:88.

16. Zheng H, Patterson LM, Phifer CB, Berthoud H-R. Brain stem melanocortinergic modulation of meal size and identification of hypothalamic POMC projections. *Am J Physiol Integr Comp Physiol*. 2005;289(1):R247-R258. doi:10.1152/ajpregu.00869.2004

17. Dodd GT, Decherf S, Loh K, et al. Leptin and Insulin Act on POMC Neurons to Promote the Browning of White Fat. *Cell*. 2015;160(1):88-104. doi:https://doi.org/10.1016/j.cell.2014.12.022

18. Balthasar N, Coppari R, McMinn J, et al. Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. *Neuron*. 2004;42(6):983-991. doi:10.1016/j.neuron.2004.06.004

19. Cowley MA, Smart JL, Rubinstein M, et al. Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature*. 2001;411(6836):480-484. doi:10.1038/35078085

20. Moore SE, Fulford AJC, Darboe MK, Jobarteh ML, Jarjou LM, Prentice AM. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: The ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth*. 2012. doi:10.1186/1471-2393-12-107

21. WHO (World Health Organization). WHO Anthro for personal computers. Software for assessing growth and development of the world's children. 2011.

22. WHO. WHO obesity and overweight definitions. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight.

23. WHO GMGRS. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. *World Heal Organ*. 2006. doi:10.4067/S0370-41062009000400012

24. Andreoli A, Scalzo G, Masala S, Tarantino U, Guglielmi G. Body composition assessment by dual-energy X-ray absorptiometry (DXA). *Radiol Med*. 2009;114(2):286-300. doi:10.1007/s11547-009-0369-7

25. Achamrah N, Colange G, Delay J, et al. Comparison of body composition assessment by DXA and BIA according to the body mass index: A retrospective study on 3655 measures. *PLoS One*. 2018;13(7):e0200465-e0200465. doi:10.1371/journal.pone.0200465

26. Yan L, Ma C, Wang D, et al. OSAT: a tool for sample-to-batch allocations in genomics experiments. *BMC Genomics*. 2012;13(1):689. doi:10.1186/1471-2164-13-689

27. Kühnen P, Handke D, Waterland RA, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016;24(3):502-509. doi:10.1016/j.cmet.2016.08.001

28. Illumina. *Infinium™ H3Africa Consortium Array V2*.; 2020.

29. R Core Team. R: A Language and Environment for Statistical Computing. *R Found Stat Comput Vienna, Austria*. 2013;0:{ISBN} 3-900051-07-0. doi:http://www.R-project.org/

30. Cornelissen G. Cosinor-based rhythmometry. *Theor Biol Med Model*. 2014;11:16. doi:10.1186/1742-4682-11-16

31. Mutak A. Extended Tools for Cosinor Analysis of Rhythms: Cosinor2. 2018.

32. Bingham C, Arbogast B, Guillaume GC, Lee JK, Halberg F. Inferential statistical methods for estimating and comparing cosinor parameters. *Chronobiologia*. 1982;9(4):397-439.

33. Toda C, Santoro A, Kim JD, Diano S. POMC Neurons: From Birth to Death. *Annu Rev Physiol*. 2017;79(1):209-236. doi:10.1146/annurev-physiol-022516-034110

34. Dominguez-Salas P, Moore SE, Cole D, et al. DNA methylation potential: Dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am J Clin Nutr*. 2013;97(6):1217-1227. doi:10.3945/ajcn.112.048462

35. Agha G, Houseman EA, Kelsey KT, Eaton CB, Buka SL, Loucks EB. Adiposity is associated with DNA methylation profile in adipose tissue. *Int J Epidemiol*. 2015;44(4):1277-1287. doi:10.1093/ije/dyu236

36. Sharma NK, Comeau ME, Montoya D, et al. Integrative Analysis of Glucometabolic Traits, Adipose Tissue DNA Methylation, and Gene Expression Identifies Epigenetic Regulatory Mechanisms of Insulin Resistance and Obesity in African Americans. *Diabetes*. 2020;69(12):2779 LP - 2793. doi:10.2337/db20-0117

37. Song CK, Jackson RM, Harris RBS, Richard D, Bartness TJ. Melanocortin-4 receptor mRNA is expressed in sympathetic nervous system outflow neurons to white adipose tissue. *Am J Physiol Regul Integr Comp Physiol*. 2005;289(5):R1467-76. doi:10.1152/ajpregu.00348.2005

38. Song CK, Vaughan CH, Keen-Rhinehart E, Harris RBS, Richard D, Bartness TJ. Melanocortin-4 receptor mRNA expressed in sympathetic outflow neurons to brown adipose tissue: neuroanatomical and functional evidence. *Am J Physiol Regul Integr Comp Physiol*. 2008;295(2):R417-28. doi:10.1152/ajpregu.00174.2008

39. Yu H, Chhabra KH, Thompson Z, et al. Hypothalamic POMC deficiency increases circulating adiponectin despite obesity. *Mol Metab*. 2020;35:100957. doi:https://doi.org/10.1016/j.molmet.2020.01.021

40. Caron A, Lee S, Elmquist JK, Gautron L. Leptin and brain–adipose crosstalks. *Nat Rev Neurosci*. 2018;19(3):153-165. doi:10.1038/nrn.2018.7

41. Guyenet SJ, Schwartz MW. Regulation of Food Intake, Energy Balance, and Body Fat Mass: Implications for the Pathogenesis and Treatment of Obesity. *J Clin Endocrinol Metab*. 2012;97(3):745-755. doi:10.1210/jc.2011-2525

42. Rosenbaum M, Leibel RL. Role of leptin in energy homeostasis in humans. *J Endocrinol*. 2014;223(1):T83-T96. doi:10.1530/JOE-14-0358

43. Rosenbaum M, Goldsmith R, Bloomfield D, et al. Low-dose leptin reverses skeletal muscle, autonomic, and neuroendocrine adaptations to maintenance of reduced weight. *J Clin Invest*. 2005;115(12):3579-3586. doi:10.1172/JCI25977

44. Taylor RW, Gold E, Manning P, Goulding A. Gender differences in body fat content are present well before puberty. *Int J Obes Relat Metab Disord J Int Assoc Study Obes*. 1997;21(11):1082-1084. doi:10.1038/sj.ijo.0800522

45. Guy Ikambo Wanghi, Leslie Lytle ARB and EKS. Body Mass Index Correlates with Body Fat

Percentage in Children and Adolescents in the Democratic Republic of the Congo: A cross sectional study. *Afr J Health Sci*. 2021;34(1).

46. Chowen JA, Freire-Regatillo A, Argente J. Neurobiological characteristics underlying metabolic differences between males and females. *Prog Neurobiol*. 2019;(179):18-32. doi:10.1016/j.pneurobio.2018.09.001

47. Schellong K, Melchior K, Ziska T, Henrich W, Rancourt RC, Plagemann A. Sex-specific epigenetic alterations of the hypothalamic Agrp-Pomc system do not explain "diabesity" in the offspring of high-fat diet (HFD) overfed maternal rats. *J Nutr Biochem*. 2020;75:108257. doi:10.1016/j.jnutbio.2019.108257

48. Burke LK, Doslikova B, D'Agostino G, et al. Sex difference in physical activity, energy expenditure and obesity driven by a subpopulation of hypothalamic POMC neurons. *Mol Metab*. 2016;5(3):245-252. doi:10.1016/j.molmet.2016.01.005

49. Begum G, Davies A, Stevens A, et al. Maternal undernutrition programs tissue-specific epigenetic changes in the glucocorticoid receptor in adult offspring. *Endocrinology*. 2013;154(12):4560-4569. doi:10.1210/en.2013-1693

50. Dennison CS, King CM, Dicken MS, Hentges ST. Age-dependent changes in amino acid phenotype and the role of glutamate release from hypothalamic proopiomelanocortin neurons. *J Comp Neurol*. 2016;524(6):1222-1235. doi:https://doi.org/10.1002/cne.23900

51. Ito Y, Banno R, Shibata M, et al. GABA type B receptor signaling in proopiomelanocortin neurons protects against obesity, insulin resistance, and hypothalamic inflammation in male mice on a high-fat diet. *J Neurosci*. 2013;33(43):17166-17173. doi:10.1523/JNEUROSCI.0897-13.2013

52. Chowen JA, Argente J, Vician L, Clifton DK, Steiner RA. Pro-opiomelanocortin messenger RNA in hypothalamic neurons is increased by testosterone through aromatization to estradiol. *Neuroendocrinology*. 1990;52(6):581-588. doi:10.1159/000125647

53. Xu Y, Nedungadi TP, Zhu L, et al. Distinct hypothalamic neurons mediate estrogenic effects on energy homeostasis and reproduction. *Cell Metab*. 2011;14(4):453-465. doi:10.1016/j.cmet.2011.08.009

54. Matsumoto A, Arai Y. Sexual dimorphism in 'wiring pattern' in the hypothalamic arcuate nucleus and its modification by neonatal hormonal environment. *Brain Res*. 1980;190(1):238-242. doi:https://doi.org/10.1016/0006-8993(80)91173-7

55. Madrid JA, Lopez-Bote C, Martín E. Effect of neonatal androgenization on the circadian rhythm of feeding behavior in rats. *Physiol Behav*. 1993;53(2):329-335. doi:https://doi.org/10.1016/0031-9384(93)90213-Y

56. Khera A V., Chaffin M, Wade KH, et al. Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. *Cell*. 2019;177(3):587-596.e9. doi:10.1016/j.cell.2019.03.028

57. Swinburn B, Egger G, Raza F. Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity. *Prev Med (Baltim)*. 1999;29(6):563-570. doi:https://doi.org/10.1006/pmed.1999.0585

58. Präger M, Kurz C, Böhm J, Laxy M, Maier W. Using data from online geocoding services for the assessment of environmental obesogenic factors: a feasibility study. *Int J Health Geogr*. 2019;18(1):13. doi:10.1186/s12942-019-0177-9

59. Crujeiras AB, Campion J, Díaz-Lagares A, et al. Association of weight regain with specific methylation levels in the NPY and POMC promoters in leukocytes of obese men: A translational study. *Regul Pept*. 2013;186:1-6. doi:10.1016/j.regpep.2013.06.012

60. St-Onge M-P, Gallagher D. Body composition changes with aging: the cause or the result of alterations in metabolic rate and macronutrient oxidation? *Nutrition*. 2010;26(2):152-155. doi:10.1016/j.nut.2009.07.004

61. Candler T, Kühnen P, Prentice AM, Silver M. Epigenetic regulation of POMC; implications for nutritional programming, obesity and metabolic disease. *Front Neuroendocrinol*. 2019;54:100773. doi:10.1016/j.yfrne.2019.100773

Chapter 5 POMC methylation and appetite and satiety

Summary of the chapter

In this chapter, I summarise the relevance of exploring appetite with POMC's role in appetite regulation and the satiety cascade. I describe the background methodology that informed the study design. I report the results of the appetite test used for POMC study subset mothers and children. The relationship between *POMC* methylation and subjective measures of satiety, amount consumed and speed of eating during test meals is examined.

5.1 Introduction

The melanocortin system integrates systemic signals and neural pathways to regulate food intake and energy expenditure[1]. Energy balance is controlled by a delicate interplay between anorexic (appetite suppressing) and orexigenic (appetite promoting) pathways. The dominant pathway is determined by the status of the individual as they oscillate between a 'fed' and 'starved' state dependent on their energy balance. In the 'fed' state, peripheral signals such as leptin[2], (released from adipocytes), insulin[3] (produced by β-cells of the pancreas) and metabolites (such as glucose[4]) cross the blood brain barrier and bind to receptors on POMC (proopiomelanocortin) neurons in the arcuate nucleus (AVN) of the hypothalamus. POMC, a prohormone, is then cleaved to produce α- and β-MSH which interact with MC4R receptors in the paraventricular nucleus of the hypothalamus (PVN) to convey a satiety signal[5]. In the 'starved' state, increased expression of AgRP (agouti related peptide) and neuropeptide Y (NPY) is driven by a reduction in leptin and insulin and an increase in circulating ghrelin. Ghrelin, a hormone derived from enteroendocrine cells, is released in response to food deprivation and binds to a receptor on AgRP/NPY and acts to increase food intake and appetite[6]. AgRP/NPY neurons project from the AVN to the PVN. Here they convey an orexigenic response by antagonising MC4R and by directly inhibiting satiety neurons in the PVN by the actions of NPY and GABA[1,5] (see Figure 1.4).

Dysregulation of the melanocortin system can lead to disorders of energy balance and satiety. The most common monogenic cause of obesity is due to mutations in the *MC4R* gene[7]. Furthermore, individuals with *MC4R* mutations report a history of increased appetite and children with *MC4R* mutations have been shown to eat three times the amount eaten by unaffected siblings during an ad libitum meal[8]. Conversely, gain of function mutations in MC4R are associated with low BMI and are protective against obesity[9].

POMC is a key mediator of satiety within the melanocortin system[10,11]. Mutations in *POMC* among flat coat retriever dogs are associated with increased appetite and food motivation[12]. Administration of synthetic MSH in POMC deficient rats leads to a reduction in weight and normalisation of food intake[13]. POMC deficiency caused by biallelic mutations in the *POMC* gene are associated with a triad of: early onset *hyperphagic* obesity, red hair, and adrenal insufficiency[14,15]. Early onset hyperphagia appears a near universal feature from reported cases. Setmelanotide, an MC4R agonist, has led to significant weight loss in those with leptin receptor and POMC deficiency[16,17]. Importantly, Setmelanotide was associated with a significant reduction in hunger scores in these individuals.

Childhood eating behaviour is influenced by both genetic and environmental factors. Evidence from twin and family studies suggests that appetitive traits demonstrate between 50-75% heritability[18–21]. Alterations in the *FTO (fat mass and obesity associated)* gene have been associated with higher energy intake during a test meal[22] and increased consumption of palatable food following a meal[23], so called 'eating in the absence of hunger'. Epigenetic influence on energy intake and appetite in humans has had limited study.

As discussed in earlier chapters, *POMC* hypermethylation is associated with obesity in children and adults[24,25]. *POMC* hypermethylation is associated with lower POMC expression[25]. Understanding that increased appetite and disordered satiety is a key feature of those with *POMC* deficiency and other disorders of the melanocortin system, I hypothesised that *POMC* hypermethylation, resulting in lower POMC expression and a weaker satiety signal would be associated with a lower measure of reported satiety following a meal and increased food intake.

5.2 Appetite testing

5.2.1 A background to appetite testing

Accurately assessing human appetite, satiety and food intake can be challenging with numerous methodological considerations for the researcher[26,27]. Controlling the multi-faceted components of the satiety cascade (see Figure 5.1) to test the variable of interest poses a particular challenge.

It is important to firstly define the key processes relating to human food intake. The definitions below were agreed by the Appetite Regulation Task Force[26] and will be used when describing these terms in this chapter.

- Appetite: (i) Covers the whole field of food intake, selection, motivation, and preference and (ii) Refers specifically to qualitative aspects of eating, sensory aspects or responsiveness to environmental stimulation that can be contrasted with the homeostatic view based on eating in response to physiological stimuli such as energy deficit.

- Hunger: (i) Construct or intervening variable that connotes the drive to eat. Not directly measurable but can be inferred from objective conditions and (ii) Conscious sensation reflecting a mental urge to eat.

- Satiation: process that leads to the termination of eating; therefore controls meal size. Also known as intra-meal satiety.

- Satiety: process that leads to inhibition of further eating, decline in hunger, increase in fullness after a meal has finished. Also known as post-ingestive satiety or inter-meal satiety.

The satiety cascade (see Figure 5.1), first described by John Blundell in the 1990s[28], describes the cognitive and physiological processes that occur following a meal and synergise to control meal termination and the time to the next eating episode. Our knowledge of many of the physiological processes that govern this pathway has expanded in recent years and the complexity of the cascade has become more apparent. Key processes in the satiety cascade relate to i) Meal quality (reward, pleasure, meal palatability, individual cognitive associations and expectations) ii) Meal quantity (gastric stretch, osmotic load) iii) Post-ingestive gastro-intestinal peptides promoting satiation (CCK (cholycystokinin), GLP-1 (Glucagon-like peptide), PYY (peptide YY)) iv) Nutrient status (e.g. glucose metabolism, insulin release) v) Energy balance such as resting energy expenditure and fat free mass vi) post-absorptive processes and the melanocortin system.

Experimental approaches to measuring components of human food intake are often subject to compromise. Designing an experimental protocol is a trade-off between laboratory-centred study with greater experimental control, precision and accuracy and free-living naturalistic study with less control and precision but with greater ecological relevance[36]. Methods used to test satiation and satiety are distinct, though the physiological processes may overlap[26].

Satiation relates to the control of meal size. Overconsumption during an eating episode is an important consideration in the development of obesity. An *ad libitum* meal following an overnight fast provides researchers with a quantitative measure of meal size. The choice of test meal is important as numerous factors are shown to influence meal size including calorific content, macronutrient context , texture and palatability[26,27,36,37].

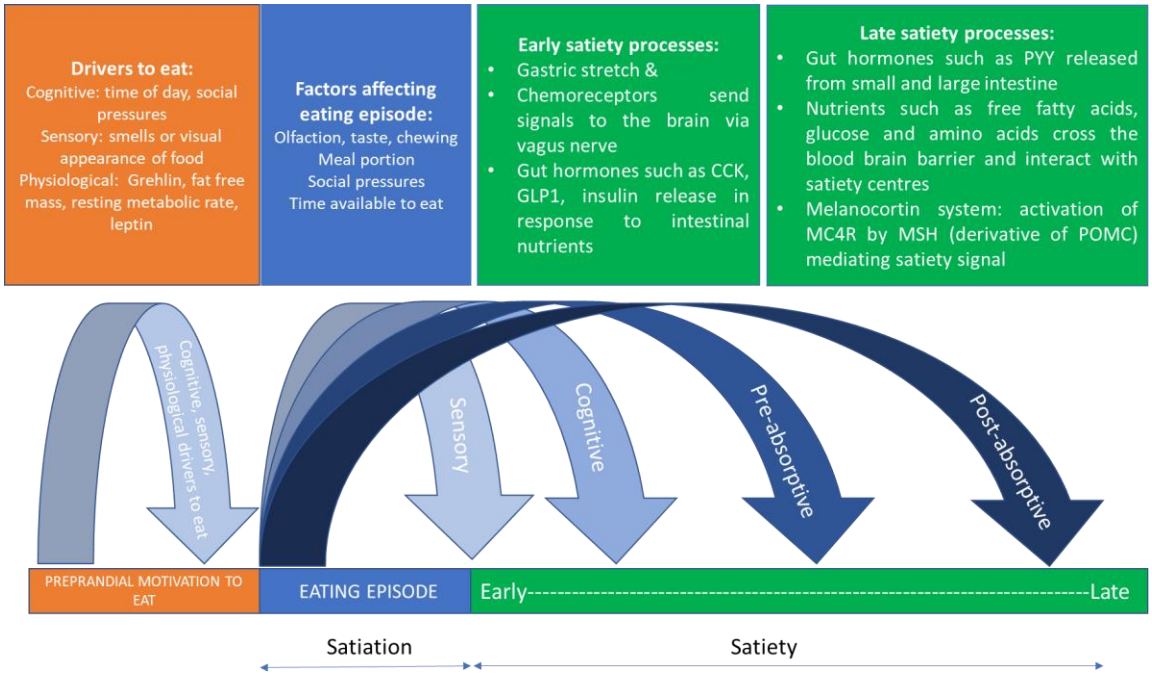


Figure 5.1 The Satiety Cascade. Adapted from Blundell et al, 1991[29]. The drivers to eat are varied and multifaceted. Societal and psychological factors such as time of day, social occasions, and individual cognitive associations and expectations about a meal contribute to our drive to eat. Sensory factors such as smell and sight of food are also important processes. The enteroendocrine hormone ghrelin is a key hormone linked to promotion of appetite, acting in the hypothalamus to increase the orexigenic signal. Ghrelin is released by the gut with concentrations increasing in relation to fasting and falling after eating[30]. Tonic and long term regulation of eating is regulated by factors such as leptin[31], fat-free mass[32] and resting metabolic rate[33,34]. Sensory feedback during and immediately after a meal provide the main control over satiation e.g. afferent signals from mouth to the brain tend to act as positive feedback whereas signals from the stomach and gut tend to provide negative feedback[35]. As described in the introduction to this chapter, pre and post absorptive processes interact with the melanocortin system to contribute to appetite regulation. Peripheral signals such as gut hormones (CCK (cholecystokinin), GLP1 (glucagon-like peptide), PYY (peptide YY), leptin, metabolites (such as glucose, amino acids, free fatty acids) and insulin promote the release of POMC in the arcuate nucleus (AVN). Post translational cleavage of POMC produces melanocortin stimulating hormone (MSH) which binds to melanocortin receptor 4 (MC4R) in the paraventricular nucleus (PVN) to convey a satiety response.

Satiety, the suppression of eating following a meal, can be measured as either the degree of suppression of hunger following a meal or the amount eaten at the subsequent eating event. Combining a measure of satiety into experimental design is often achieved by using a either a fixed portion meal or preload meal before an ad libitum meal[27,38]. The degree of hunger suppression is commonly measured by inter-meal appetite ratings provided by the participant at fixed intervals[26].

5.2.2 A background to Visual Analogue Scales of Hunger

A visual analogue scale (VAS) is an instrument used to measure a subjective characteristic that cannot be measured *directly*. VASs are used to ask respondents to score a characteristic on a scale between two end points denoting the extremes of the subjective feeling in question. VASs have been used widely in health settings and for research. The VAS for pain being a notable example used widely in children and adults[39,40].

The most commonly used VAS in appetite research uses a 100mm continuous line with subjects asked to indicate where they lie in relation to the characteristic in question e.g. hunger, fullness, desire to eat. The VAS has taken different forms in appetite research[41]. In summary, the VAS can be either i) a continuous line with anchored statements at the extremes e.g. how hungry do you feel? – not at all or extremely hungry[42] ii) a line punctuated by numbers 1 to 10[41] iii) a line punctuated with descriptions of hunger e.g. greatest hunger imaginable, extremely hungry, moderately hungry, slightly hungry[43] iv) a pictorial representations of hunger[44,45]. Increasingly VAS are recorded via electronic data capture[46].

Importantly, the use of VAS to measure appetite and satiety has been shown to be reliable, valid and reproducible[47–49].

It is important to select the most appropriate VAS tool for the study population. Non-literate, non-numerate and children are groups that need careful consideration of the methodology used. Pictorial appetite-related VAS have been used before for both children[44,45,50] and adults[51]. A summary of pictorial VASs used in children from high income countries are summarised in Table 5.1 below.

To my knowledge, the use of a VAS for appetite has not been utilised in Gambian children or adults and reflects a wider lack of development of these tools in low and middle income countries (LMIC).

5.2.3 Testing appetite and satiety in LMIC

Measuring appetite in LMIC can be an important component of health research. A recent systematic review of methods used to measure appetite in children less than 5 years old from LMIC identified 23 studies since 1995[52]. The majority of studies used caregiver assessment or quantification of feeding or appetite. Only 6 studies reported validation processes, such as comparing the amount consumed in an observed test meal with daily energy intake. Previous studies have performed

relatively rudimentary appetite tests in children and adults in LMIC income countries involving a single measurement of consumption of an ad libitum breakfast/snack[53–58] (see Table 5.2). For both children and adults from LMIC, where appetite testing has been used, the test centres around the assessment of ad libitum consumption of a porridge based breakfast following an overnight fast. . Detailed dietary assessments and dietary intakes of children and mothers from The Gambia has been well documented previously[59].

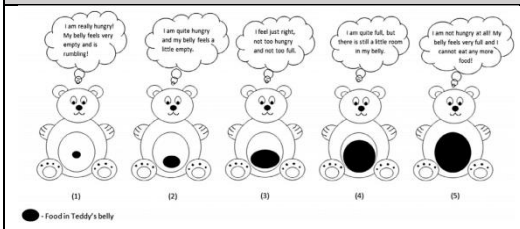
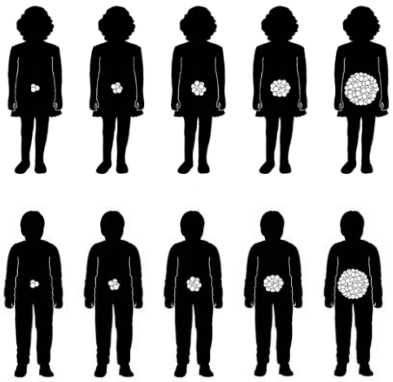
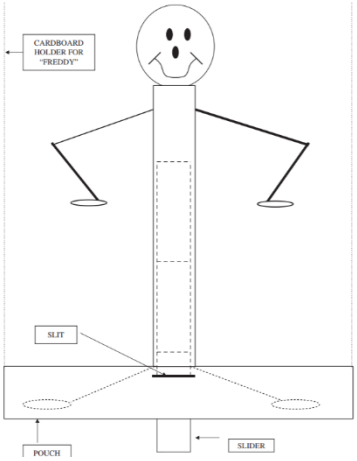
Visual Analogue Scale used	Age and Ethnicity	Methods	Conclusion	Reference
	5-9 years 54 Majority Caucasian British	Use VAS in story and in context of eating episode	Significant negative correlation between pre-snack hunger rating and ad libitum snack intake	Bennett et al, 2014[45]
	4-6 years USA 9 Caucasian 2 African American 3 Asian 4 Hispanic 2 other	Used VAS in context of imagined food stimuli Asked to identify where hunger felt	Children have the capacity to report, quantitatively, internal feelings of satiety by means of a five level scale Did not test reported satiety levels against actual food intake Hunger identified in the stomach	Faith et al, 2002[44]
	4-5 years USA 8 Caucasian 1 African-American 1 Hispanic 1 other	Tested the potential of VAS for quantifying sensations such as fullness	Children can be trained to use VAS to quantify differences in portion sizes of foods Not real eating situations – only simulated	Keller et al, 2006[50]

Table 5.1. Examples of pictorial appetite-related VAS. VAS = visual analogue score

Study Location	Age	Methods	Conclusion	Reference
Kenya	'School aged'	Quantitatively measured the ad libitum porridge (Uji) consumption (midmorning)	Hookworm treatment led to increased intake of mid-morning snack	Latham et al, 1990[57]
Kenya	6-11 years	Quantitatively measured the ad libitum porridge (Uji) consumption (midmorning) Question to participants "Do you consider your appetite in the last few days to be: very poor, poor, average, good, or very good?"	Iron supplements resulted in improved appetite (in terms of both energy intake of the snack and child report of appetite) as compared with children receiving the placebo	Lawless et al, 1994[56]
Benin	18-30 months	Quantitatively measured the consumption of ad libitum breakfast (either porridge or rice based) Compared amount eaten ad libitum with diet diary over 3 days of observed weighed record When finished eating, five minute pause then offered to eat again (to assess 'eating in absence of hunger')	Porridge based breakfast had better correlation with observed dietary intake Appetite measurement has a better reproducibility for the total intake from 2 or 3 eating episodes than from a unique eating episode	Dossa et al, 2001 [54] and 2002 [53]
South Africa	2months – 2 years	Quantitatively measured the consumption of ad libitum breakfast (Nestle Nestum No2 Cereal) for 3 days over 2 week period Five minute break and then offered to continue eating (to assess 'eating in absence of hunger') Reduced appetite among HIV infected children.	Reduced appetite among HIV infected children	Mda et al, 2010[60] and 2011[61]

Table 5.2. Summary of studies from LMIC that use an appetite test. HIV= human immunodeficiency virus

5.3 Pilot appetite test

A pilot appetite test was developed to assess the feasibility of using an ad libitum meal test for *POMC* subset children and their mothers. The pilot test also trialled the use of a 5 point pictorial VAS in the *POMC* subset children. The 'POMC subset' were 118 mothers and child pairs recruited from the villages of Keneba, Jali, Kantong Kunda, Manduar and Tankular. Appetite testing was a key part of 'subset' activity. The pilot The design of the pilot test was developed using similar methodology employed in Africa by Dossa et al (Benin) and Mda et al (South Africa) [53,54,60,61].

5.3.1 Methodology

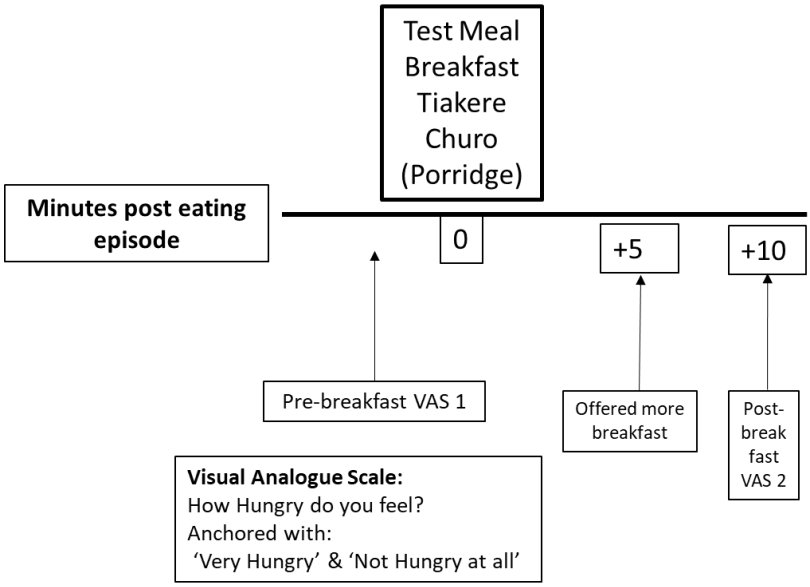


Figure 5.2. An overview of the pilot appetite test. Key: VAS = Visual Analogue Score. Note VAS only offered to children (not their mothers). See annex 5.1 for more details.

Participants were fasted overnight and attended Keneba field station for the test meal at 8am. The test meal was given in an outside covered recreational area. Mothers and children were sat together though with 5 other mother-child pairs distributed within the outside covered recreational area. Participants declaring an allergy to any of the ingredients in the test meal were excluded. Children were shown a five-point VAS (see Figure 5.3, taken from Faith et al[44]) before eating (corresponding to VAS 1 in Figure 5.2). Children were asked "how hungry are you?" and to select an image from the VAS (the field workers anchored the statements with picture on far left 1 ("very hungry") to picture on far right 5 ("not at all")). Children and Mothers were given an ad libitum

breakfast of local Gambian porridge (Tiakere Churo, energy density 3.4 KJ/g). The Tiakere Churo was prepared by the study cook to an agreed recipe (see Table 5.3 below).

Breakfast: Tiakere Churo recipe

Ingredient	Amount (g)
Long grain white rice (raw)	705
Groundnut (fresh and raw)	665
Sugar	555
Milk (sour)	865

Estimated ingredient to provide enough breakfast for 10 adults. Ingredients proportionally increased to cover the number of participants attending each day

Macronutrient composition

Energy density of breakfast: 3.4 kJ/g	
Macronutrient component	Percentage (%) of energy
Protein	10
Fat	30
Carbohydrate	60

Table 5.3. Ingredients in Tiakere Churo (top) with macronutrient composition of the meal (bottom). Metabolisable energy density was calculated using raw ingredient weights per 100 g and 17 kJ, 37 kJ, and 16 kJ per g for protein, fat, and carbohydrate, respectively[62].

Tiakere Churo was filled up to a line 1 cm from the bowl rim. Participants were invited to eat as much as they wish and they were informed that if they finish the bowl it will be refilled until they have eaten the amount they would like. The eating episode was timed using an electronic timer (SLS Timer Interval Stopwatch, model number TIM0250). When the participants had finished eating they were asked to inform the field worker who recorded the time taken to eat to the nearest second. Children were again asked to score the VAS again (corresponding to VAS 2 in Figure 5.2) "how hungry are you?" and to select an image from the VAS (the field workers anchored the statements with picture on far left 1 ("very hungry") to picture on far right 5 ("not at all")). The amount consumed (g) was calculated by the weight of bowl before eating minus weight of bowl after eating. The bowl (and contents) were measured by Salter electronic kitchen scales (model

number 1035 SSBKDR) to the nearest gram. Participants were offered more porridge 5 minutes after stopping eating.

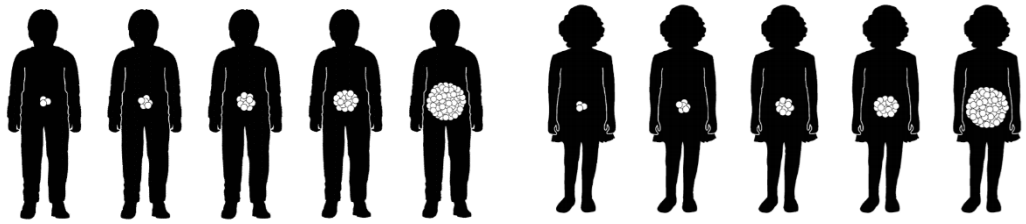


Figure 5.3. Visual Analogue Scale for hunger used for children in baseline appetite test. Source: Faith et al, 2002[44]

Assessment of body composition

For all study timepoints, participants (subset mothers and children) had a whole body DXA scan performed using the GE-Lunar Prodigy scanner (HE Medical, Waltham, MA,; software version 13.60.033). The relationship between both total lean mass and total fat mass and appetite test outcomes speed of eating and amount eaten was explored using Spearman correlation coefficient and used as adjustment covariates where significant relationships were found. Previously studies have reported a positive association between appetite measures and lean or fat free mass[32,63].

5.3.2 Statistical analysis

The purpose of the pilot test was to see if the test was to i) assess if the porridge meal was acceptable for participants, ii) assess the amount eaten and time to eat (may help inform further future appetite testing) iii) assess if children could perform a VAS for hunger iv) assess any relationship between VAS and amount eaten in the ad libitum meal v) explore the relationship between the outcome measures and potential adjustment covariates.

Outcomes recorded in the test meal were amount eaten (g) and time to eat (seconds). Differences in i) pre and post meal VAS scores (testing for a change in VAS after eating) ii) pre meal low (VAS score 1,2 combined) vs high (VAS score 3,4 and 5 combined) VAS were assessed by Mann-Whitney U test. Kruskall-Wallis test was used to test multiple pre meal VAS score and amount eaten in test meal.

Spearman rank correlation coefficients were assessed between amount eaten and age, DXA-derived fat mass and lean mass and weight. The associations between amount eaten and sex were explored using Mann-Whitney U test.

Child's age (years), amount eaten (g) , time to eat (s), total fat mass (g), VAS scores were not normally distributed and median and interquartile range (IQR) were reported. Measures of lean mass (kg), mothers age (years) and weight (kg) were normally distributed and mean and standard deviation reported.

5.3.3 Results

Children

	Children
Number of participants	Total attended for appetite test n= 113 n=108 participated in test meal n=4 refused Tiakere Churo n=1 allergic to Tiakere Churo
Median Age [IQR, range]	6.02 years [1.81, 4.26 to 7.68]
Sex	Male n=64 Female n= 49
Median amount of Tiakere Churo eaten (g) [IQR, range]	175g [140.25, 13 to 614]
Completed VAS	1st Hunger rating (VAS 1) = 113 2nd Hunger rating (VAS 2) = 102
Median time to eat (seconds)	Median=274 seconds [IQR 165]

Table 5.4. Summary of pilot appetite test for children. VAS = visual analogue score, IQR = interquartile range, SD=standard deviation, s= seconds

113 children attended the pilot appetite test and 108 (95.6%) children participated in the test (see Table 5.4). 4.4% (n=5) children did not proceed with the appetite test (n=4 refused the breakfast, n=1 reported an allergy to the Tiakere Churo). The median amount eaten was 175g [IQR 140.25]

taking on average 274 seconds to complete the meal. Boys ate more than girls (195g vs 140g, p=0.0074, see Figure 5.4). There were significant positive correlations between amount eaten and age (R=0.2, p=0.038) and DXA-derived total lean mass (R=0.25, p=0.01, see Figure 5.5). There was no significant correlation between either amount eaten and weight or DXA-derived Total Fat Mass.

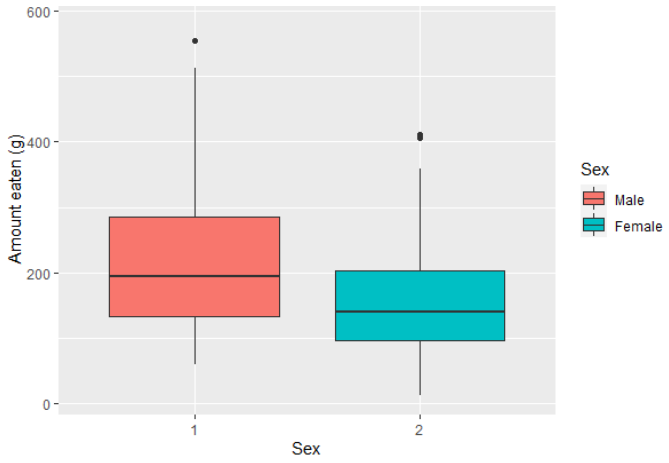


Figure 5.4. Boxplot of amount eaten by Sex. P-value = 0.0074. Mann Whitney U test used to test the null hypothesis.

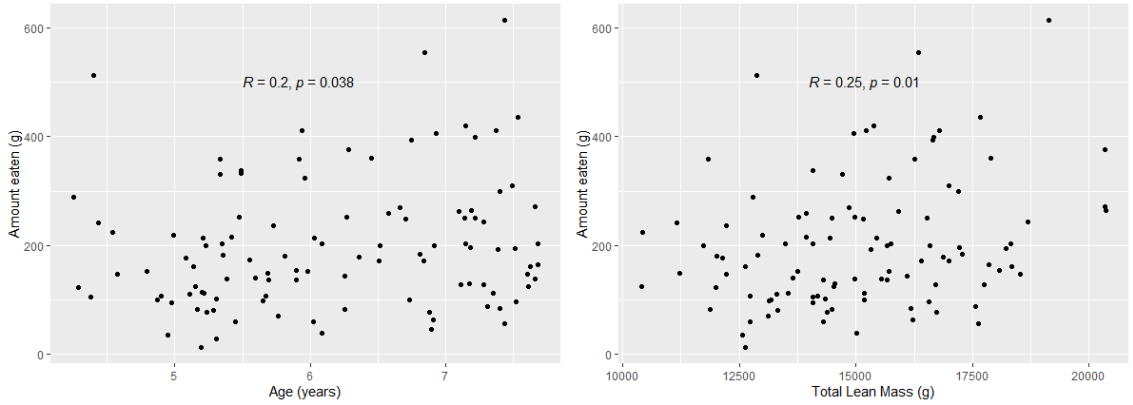


Figure 5.5. Scatterplot between amount eaten (g) and age (years) and DXA-derived Total Lean Mass (g). R is Spearman correlation coefficient with p-value

VAS

102/113 (90%) of children were able to score a pre and post meal VAS. 94/102 (92%) participants scored higher post meal compared to pre meal i.e. reported feeling less hungry after eating. 7/102 (6.9%) scored the same pre and post meal i.e. reported the same degree of hunger and 1/102 (1%) scored lower post meal compared to pre meal i.e. felt more hungry after eating. Overall pre and post meal VAS differed significantly (2 vs 5, p<0.001, see Figure 5.6).

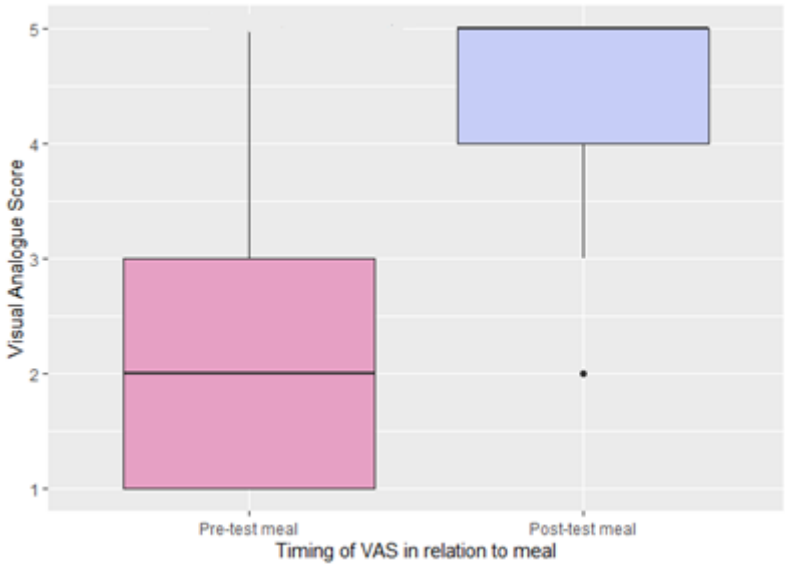


Figure 5.6. Boxplot of pre and post meal VAS scores. P-value <2.2x10^{-16}. Mann Whitney U test used to test the null hypothesis.

The relationship between pre meal VAS scores and amount eaten during the test meal is shown in Figure 5.7. Overall, there was limited evidence of a link between amount eaten and pre-meal VAS with amount eaten during the test meal and pre meal VAS scores (p=0.069). However, when the pre meal VAS were dichotomised into low (score of 1 or 2) and high (3,4 or 5) scores there was a significant difference in amount eaten during the meal (p=0.022, see Figure 5.8).

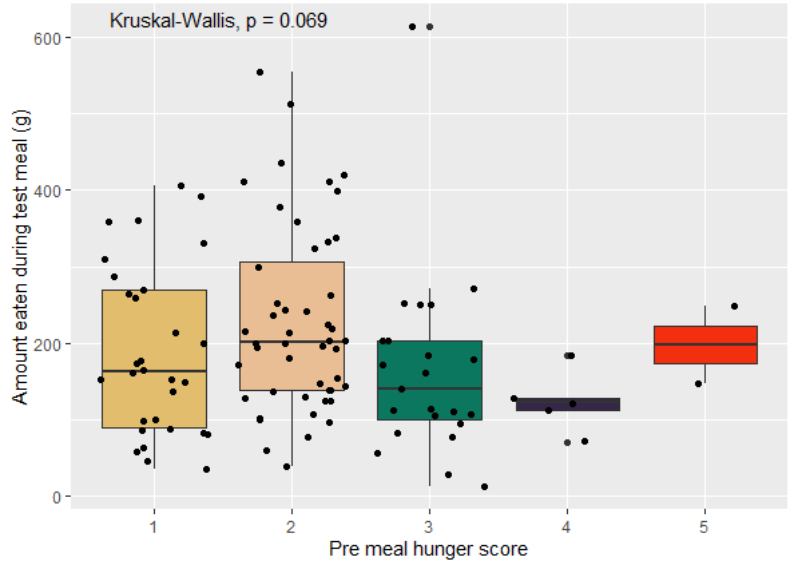


Figure 5.7. Boxplot of pre meal VAS score and amount eaten during test meal. Kruskall-Wallis test used to test the null hypothesis.

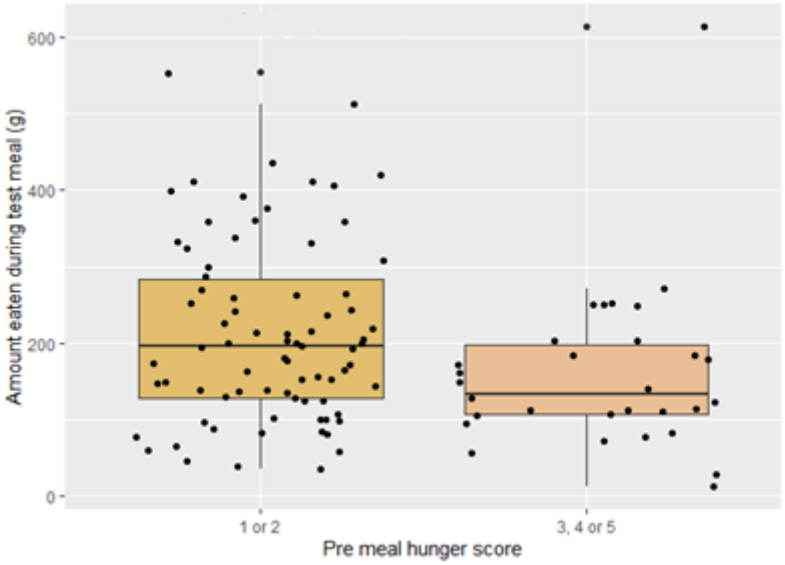


Figure 5.8. Boxplot of Low vs High pre meal VAS score and amount eaten during test meal. P-value = 0.022. Mann Whitney U test used to test the null hypothesis.

Mothers

99 mothers attended for the appetite test with 89 (89.9%) who participated in the test. 10.1% (n=10) mothers did not proceed with the appetite test (n=4 refused the breakfast, n=6 reported an allergy to the Tiakere Churo). The median amount eaten was 290g [IQR 310] taking on average 217 seconds to finish eating. In contrast to the children, there were no significant correlations between amount eaten and age, weight or any DXA measure of lean or fat mass.

	Mothers
Number of participants	Total attended for appetite test n= 99 n=89 participated in test meal n=4 refused Tiakere Churo n=6 allergic to Tiakere Churo
Mean age [SD, range]	39.3 years [5.25, 25.4 to 49.8]
Median amount of Tiakere Churo eaten (g) [IQR, range]	290g [310, 8 to 766]
Median time to eat (seconds)	Median=217 seconds [IQR 212]

Table 5.5. Summary of pilot appetite test for mothers. IQR = interquartile range, SD=standard deviation, s= seconds

5.3.4 Summary and learning points

In summary, the purpose of the pilot test was to see if the test was to:

i. assess if the porridge meal was acceptable for participants

The pilot test showed that the vast majority of mothers (89.9%) and children (95.6%) participated in the test. No participant ate again after the 5 minute pause and therefore this element of the pilot test was not able to discriminate those who might be susceptible to eat in the absence of hunger. This may also reflect the propensity to develop sensory specific satiety with homogenously textured meals such as porridge[37].

ii. assess the amount eaten and time to eat

The field workers were able to measure the amount eaten and record the time take to eat. There was a wide range seen with the amount eaten in both mothers (range = 8 to 766g) and children (range = 13 to 614g)

iii. assess if children could perform a VAS for hunger

The majority of children could use a VAS of hunger. 90% of children were able to score a pre and post meal VAS. Of those, 92% off participants scored higher post meal compared to pre meal i.e. recorded feeling less hungry after eating. Overall pre and post meal VAS differed significantly.

iv. assess any relationship between VAS and amount eaten in the ad libitum meal

Overall there was not a strong relationship between VAS and amount eaten. Though when the VAS scores were dichotomised there was a significant relationship. This suggested that the VAS could be improved. There was not a test used for mothers and therefore this proposed an area for development.

v. explore the relationship between the outcomes measures and potential adjustment covariates

In children, there was a significant positive correlation between amount eaten and age and total lean mass (DXA) and has been described previously [33]. Boys ate significantly more than girls.

In mothers, there was no association between amount eaten and age nor any measure of body composition

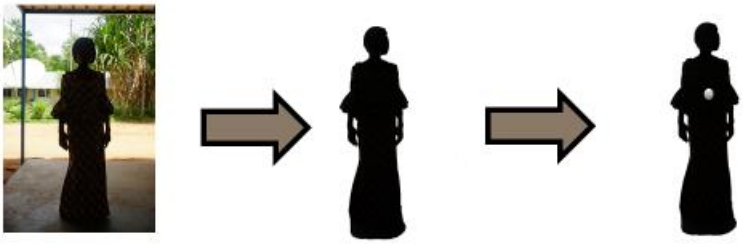
5.4 Revised appetite test

The appetite test was revised for midline and endline study activity. Two key areas were developed and included in the new design i) the development of a Gambian VAS for hunger to improve performance from the pilot and to include a VAS for mothers ii) the appetite test was revised to follow the model of a pre-load meal followed by an ad libitum meal (see Figure 5.10) as used in previous studies and widely utilised in appetite research[26,27,38].

A preload-test meal study design is widely used to study short term eating behaviour[26]. The preload is provided to participants to consume before an ad libitum 'test' meal. Depending on the nature of the research question, typically 'preloads' can either act as a standardised meal prior to an ad libitum lunch or can be manipulated (e.g. varying protein content) to assess the effect of dietary interventions on later eating behaviour. The revised test used Tiakere Churo as the preload breakfast followed by an ad libitum lunch. The subjective satiety response after the preload, measured by multiple VAS, can measure the decay in satiety after this standardised eating episode.

5.4.1 Developing Gambian VAS for hunger

A Gambian specific VAS was developed with an aim of improving the performance of the tool. Normal BMI volunteer mothers and children (within +/- 1SD of BMI as defined by WHO[64]) were used to develop a culturally typical silhouette (see Figure 5.9) and a scale of 1-8 (mothers) and 1-5 (children) with each increment having an increase in the 'bubbles' (similar to Faith et al[44]) in the abdomen signifying reducing hunger or increasing satiety.



a)

How Hungry do you feel?
Anchored with: 'Very Hungry' & 'Not Hungry at all

b)

How Hungry do you feel?
Anchored with: 'Very Hungry' & 'Not Hungry at all

c)

How Hungry do you feel?
Anchored with: 'Very Hungry' & 'Not Hungry at all

d)

Figure 5.9. Developing the Gambian VAS for hunger. a) The original back lit photograph of the volunteer mother is shown in the first image. The silhouette was captured using Microsoft Paint 3D[65] using the 'magic select' function, shown in the second image. A circular 'bubble' was then added to the upper abdomen to start the first step on the VAS, shown in the third image. b) Increasing numbers of 'bubbles' were added to the abdomen at each step of the VAS to denote being more sated and to create an 8 step VAS for mothers c) The 5 step VAS for boys (created by the same process as for mothers in described in a)) d) The 5 step VAS for girls (created by the same process as for mothers in described in a))

5.4.2 New appetite test methodology

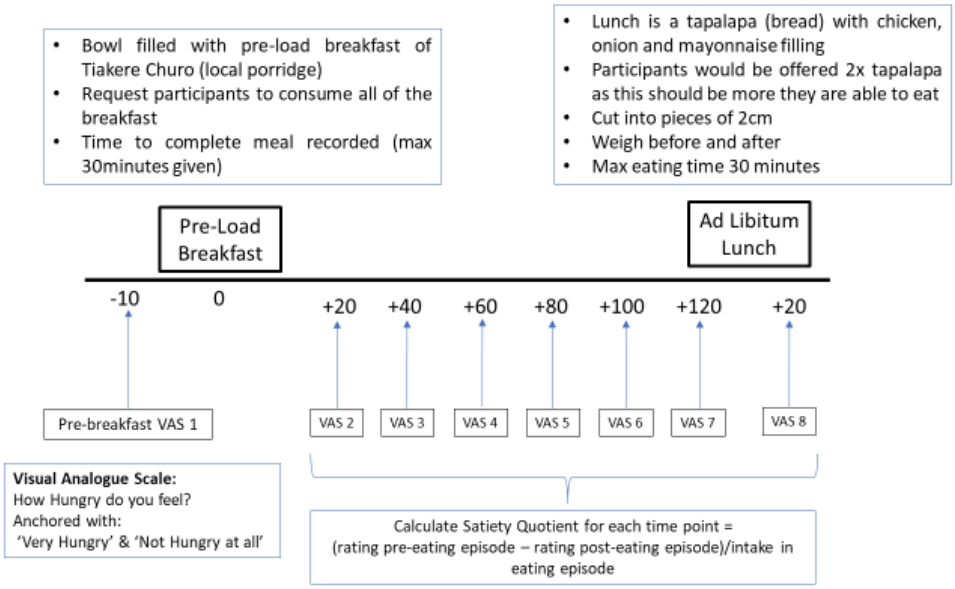


Figure 5.10. Overview of revised appetite test with pre-load breakfast and ad libitum lunch. See annex 5.2 for full details of protocol

Preload breakfast

Participants were fasted overnight and attended for the test meal at 8am. Participants declaring an allergy to any of the ingredients in the test meal were excluded. The appetite test was conducted indoors in one of two study rooms. A mother and child pairs were sat together at a table (see figure 5.11a). Each study room had four tables in total. The participants' weight and height were inputted into tablets to calculate the amount of preload breakfast to give. The preload breakfast was personalised to account for estimated resting metabolic rate (RMR), considered a key driver of appetite[33]. RMR for mothers, boys and girls were estimated using the following equations:

i) Mother(kj/day)= 616.93 – 14.9(Age in years) + 35.12(weight in Kg) + 19.83(height in cm) – 271.88(Ethnicity*)[66] *African decent = 1

ii) Boys(kcal/d)= (16.6 x weight in Kg) + (77 x height in metres) + 572[67]

iii) Girls(kcal/d)= (7.4 x weight in Kg) + (482 x height in metres) + 217[67]

The Tiakere Churo recipe gave an energy density of 3.4 kJ/g (NB 1 Kcal= 4.184 kj) as used in previous research in Keneba and the pilot test[68]. Tiakere Churo was prepared by the study cook to an agreed recipe (see Table 5.3). The preload portion was calculated as 20% of the estimated RMR for mothers and 15% of the estimated RMR for children as used in previous studies[33,63,69,70]. The electronic tablet calculated the amount of Tiakere Churo to put in the bowl and given to the participant based on the equations above. Children were shown a five-point VAS (see Figure 5.9c+d) before eating. Children were asked "how hungry are you?" and to select an image from the VAS (the field workers anchored the statements with picture on far left 1 ("very hungry") to picture on far right 5 ("not hungry at all")). Mothers were shown an eight-point VAS (see Figure 5.9b) before eating. Mothers were asked "how hungry are you?" and to select an image from the VAS (the field workers anchored the statements with picture on far left 1 ("very hungry") to picture on far right 8 ("not hungry at all")). The eating episode was timed using an electronic timer (SLS Timer Interval Stopwatch, model number TIM0250). When the participants had finished eating they were asked to inform the field worker who would record the time taken to eat to the nearest second. The amount of preload consumed (g) was calculated by the weight of the bowl before eating minus weight of the bowl after eating. The bowl (and contents) were measured by Salter electronic kitchen scales (model number 1035 SSBKDR) to the nearest gram.

Inter-meal VAS

After eating the preload breakfast, VAS scores were recorded every 20 minutes until 120 minutes (see Figure 5.10).

Ad libitum lunch

An ad libitum lunch of a filled "tapalapa" sandwich was given 120 minutes after the breakfast. The ad libitum lunch included two standard sized tapalapa (local bread) with 100g of filling (energy density = 10.4kJ/g, see table 5.6 below for recipe and energy content).

The tapalapa was cut into 2 cm pieces, put in a bowl and weighed prior to giving to the participant (see Figure 5.11b). The participant was invited to eat as much as they wish. The participants were told they have up to 30 minutes to eat, after which the remaining food will be removed and weighed. The field worker weighed the bowl and any remaining food by Salter electronic kitchen scales (model number 1035 SSBKDR) to the nearest gram. The eating episode was timed using an electronic timer (SLS Timer Interval Stopwatch, model number TIM0250). When the participants had finished eating they were asked to inform the field worker who would record the time taken to eat to the nearest second.

Lunch: Filled tapalapa sandwich

Ingredient	Amount (g)
Filling	
12 x tins of chicken	2376
Chopped onion	768
Mayonaise	1020
Bread	
Tapalapa (white bread)	153 per loaf

Estimated ingredients to provide enough breakfast for 40 fillings for Tapalapa i.e. 20 people. 100g of filling put in each tapalapa i.e. one filled tapalapa = 253g

Macronutrient composition

Energy density of filling = 9.5 kJ/g	
Macronutrient component	Percentage (%) of energy
Protein	21
Fat	74
Carbohydrate	5
Energy density of tapalapa = 11kJ/g	
Protein	14
Fat	6
Carbohydrate	80
Energy density of 1 filled tapalapa = 10.4kJ/g	

Table 5.6. Ingredients in the filled tapalapa (top) with macronutrient composition of the meal(bottom). Metabolisable energy density was calculated using raw ingredient weights per 100 g and 17 kJ, 37 kJ, and 16 kJ per g for protein, fat, and carbohydrate, respectively[62].

a)

b)

Figure 5.11. Appetite test. a) Mother and child sat at a table in the study room. b) Example of ad libitum lunch preparation and bowl presented to participants.

5.4.3 Statistical analysis

Hypotheses

- Increased POMC methylation, associated with reduced *POMC* expression, will be associated with a more rapid fall in satiety following a meal (see Figure 5.12)
- Increased POMC methylation will be associated with greater consumption during an ad libitum meal
- Increased POMC methylation will be associated with faster speed of eating

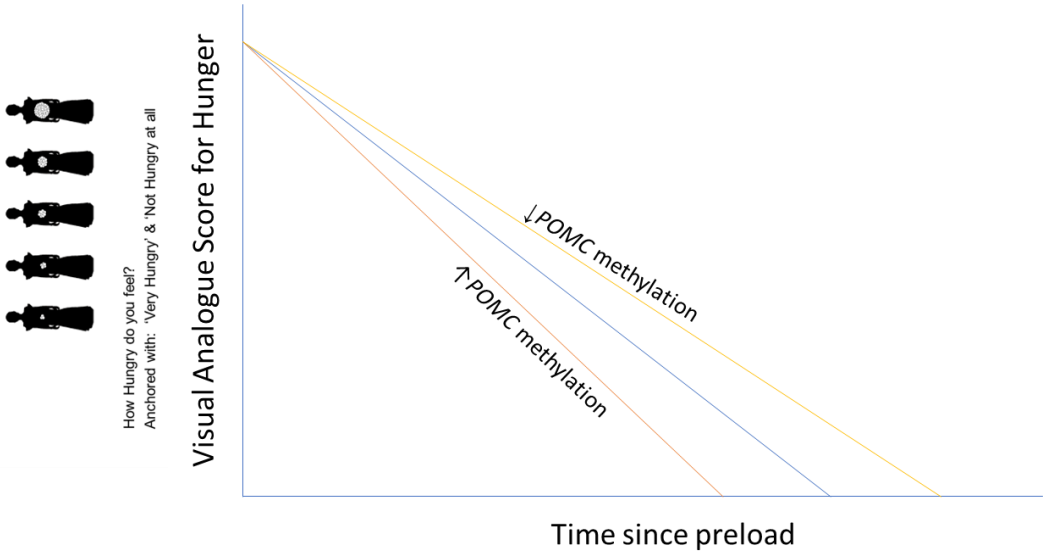


Figure 5.12 Schematic representation of hypothesised effect of altered POMC methylation on satiety response measured by VAS. Red line = higher POMC methylation. Yellow line = lower POMC methylation

Descriptive statistics

For children, age (years), amount eaten during ad libitum lunch (g), time to eat (seconds), and eating speed (g/second) were not normal normally distributed and therefore median, interquartile range (IQR) and range were reported. Speed of eating (g/second) was calculated by dividing amount eaten during the ad libitum lunch and breakfast (g) by time taken to eat (seconds). Amount eaten and speed of eating were normally distributed after square root transformation.

For mothers, time to eat (seconds) and eating speed (g/second) were not normally distributed and therefore median, interquartile range (IQR) and range were reported. Amount eaten during ad libitum lunch (g) and age (years) were normally distributed and mean and standard deviation (SD) reported. Speed of eating were normally distributed after square root transformation.

The relationship between outcome variables and covariates age, sex (children) and body composition were assessed by Spearman correlation coefficient.

***POMC* methylation measure**

A single measure of POMC methylation was taken for the 5 CpGs in exon 3 of the POMC gene (see chapter 4.2.7) measured from a blood sample taken at study baseline.

Outcome measures

The baseline test was a feasibility study only and therefore the appetite test measurements taken in midline and endline were used for hypothesis testing.

Inter-meal VAS scores reported every 20 minutes between preload and ad libitum meal, were used to assess the subjective reports of hunger. Scores 1-5 were used for children and 1-8 for mothers.

The amount eaten (g) and speed of eating (g/second) of during ad libitum meal were the outcomes from the ad libitum meal.

A single measure for amount eaten during ad libitum meal and speed of eating during ad libitum were produced by combining results across the two study timepoints (midline and endline). The reason for this was to take data from two eating episodes rather than examine a single meal, and to help alleviate issues around multiple testing.

In children, amount eaten and speed of eating were pre-adjusted for sex, age and lean mass (significant covariates identified in pilot study) in linear regression models at midline and endline. Z scores of the residuals from these models were produced for midline and endline and the mean z score taken for analysis. In mothers, amount eaten and speed of eating were unadjusted (no significant relationship was identified with age or body composition measures) and converted to a z score for midline and endline. The mean z score was taken for analysis. Using z scores accounted for potential seasonal differences in eating behaviour due to potential seasonal pressures on eating behaviour and energy balance.

Statistical modelling

Linear mixed effect models

To test the hypothesis that increased *POMC* methylation will be associated with a more rapid fall in satiety following a meal, linear mixed effect models were produced to assess how *POMC* methylation influenced the slope of VAS scores over time. Random slope models were produced so to allow the intercept and the slope to vary for each person (2 level model) and for each person *and* each study time point (3 level model). Two linear mixed effect models were produced; 2-level (level 1: repeated measure inter-meal VAS scores after preload breakfast, level 2: participant)and 3-level

random slope model (level 3: study timepoint). Model fit was determined by a combination of lowest AIC (Akaike Information Criteria) and BIC (Bayesian information criterion) and a log likelihood nearest zero.

Mean *POMC* methylation z score was included in the models as a fixed effect with an interaction with time to allow for an assessment of effect on both the intercept and slope. Sex was fitted into the models for children as a fixed effect but was not significant associated with intercept or slope and did not improve model fit as determined by loglihood test. Therefore sex was not adjusted in final models. The Satterthwaite approximation was used to calculate degrees of freedom and p-values. Models were fitted using the full information maximum likelihood (FIML). Models were fitted in R using lmer function of the lme4 package[71].

POMC methylation and ad libitum meal outcomes

For both mothers and children, the mean amount eaten z score and the mean speed of eating z score was then correlated with mean *POMC* methylation z score with the R calculated by Pearson correlation coefficient.

All model covariates were assessed for multicollinearity, and standard tests were performed to ensure that linear modelling assumptions were met.

5.5 Child Appetite testing results

5.5.1 Descriptive statistics

A summary of midline and endline results is shown in Table 5.7.

	Midline	**Endline**
Number recruited to subset	n=118	n=116
Number participated in test	Breakfast n=114 (male n= 65, female n= 49) Lunch n=111 (male n= 64, female n= 47)	Breakfast n=111 (male n= 61, female n= 50) Lunch n=111 (male n= 61, female n= 50)
Number not participating	*Breakfast* 1 allergy - excluded 3 refused breakfast - excluded *Lunch* 1 allergy – already excluded 3 refused breakfast - excluded 3 refused lunch	*Breakfast* 1 allergy - excluded 4 refused breakfast - excluded *Lunch* 1 allergy - already excluded 4 refused breakfast - excluded
Median age (years) [IQR, range]	6.5 years [1.8, 4.8 to 8.2]	7.0 [1.8, 5.2 to 8.7]
Median amount of preload breakfast eaten (g) [IQR, range]	168.5g [19, 72 to 210]	174g [26.5, 27 to 280]
Median time to eat (seconds) Preload breakfast [IQR, range]	281.5 seconds [129.8, 111 to 963]	255.5 seconds [123.5, 97 to 607]
Median amount of ad libitum lunch eaten (g) [IQR, range]	258g [163, 90 to 692]	280g [154.8, 53 to 643]
Median time to eat (seconds) of ad libitum lunch [IQR, range]	1020.5 seconds [531.5, 26 to 1438]	1024.5 seconds [512.8, 25 to 1435]

Table 5.7. Summary of midline and endline appetite tests for children. IQR = interquartile range, SD=standard deviation.

5.5.2 Ad libitum meal outcomes association with covariates

There was a significant positive correlation between age and amount eaten during ad libitum lunch (Spearman Rho = 0.21, p=0.025). There was a highly significant positive correlation between amount eaten during ad libitum lunch and total lean mass (Spearman Rho = 0.35, p=00023). There was no significant relationship between amount eaten and DXA-derived fat mass. There was a positive correlation between eating speed and DXA-derived total lean mass (Spearman Rho = 0.31, p=0.0014) and age (Spearman Rho = 0.22, p=0.024).

5.5.3 VAS summary

The individual VAS scores are plotted in Figure 5.13 below. The average trend (calculated by loess method) and mean score for each time point demonstrate a fall in VAS score after the preload (i.e. becoming less hungry or more sated) and a gradual decay in the satiety signal and return to baseline by +120 minutes.

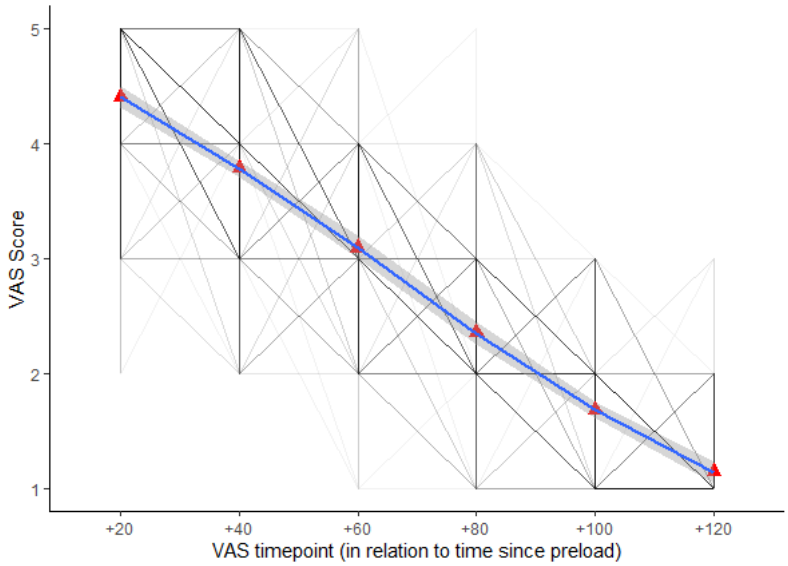


Figure 5.13. VAS score over time. Combined midline and endline plots of VAS score over time. Blue line shows trend over the time calculated by loess method with 95% CI. Colour density of individual line plots graded in grey.

5.5.4 *POMC* methylation and measures of eating behaviour and satiety

Methylation data were available for 105/111 (94.6%) children at midline and 106/111 (95.5%) at endline.

Subjective measures of satiety and *POMC* methylation

Inter-meal VAS scores (+20 to +120 post preload) were modelled using different linear mixed effect models (see methods 5.4.3). The 3-level random intercept model (model 2 in Table 5.8) provided the best model fit.

	Measure of model fitness		
Model	**AIC**	**BIC**	**loglikelihood**
1. 3-level, random slope model	2608.0	2664.9	-1293.0
2. 2-level, random slope model	2805.6	2847.0	-1394.8

Table 5.8. Comparison of model fit. AIC= Akaike Information Criteria, BIC= Bayesian information criterion.

The parameter estimates derived from the 3-level random slope model are shown in Table 5.9 below. The intercept and slope estimates demonstrate an average +20 minute post preload score of 5.07 [95% CI 4.77 – 5.37] and an average fall in VAS score of -0.67 [95% CI -0.77 - -0.56, p<0.001] every 20 minutes i.e. as expected there was a significant reduction in satiety over time.

There was no evidence of an association between mean *POMC* methylation z score and VAS slope intercept or slope i.e. *POMC* methylation was neither associated with subjective feelings of hunger after eating nor the fall in satiety between meals.

Amount eaten and *POMC* methylation

Amount eaten was pre-adjusted for sex, age, and lean mass in linear regression models at midline and endline. Z scores of the residuals from these models were produced for midline and endline and the mean z score taken for analysis. There was no evidence of an association between the amount eaten during an ad libitum lunch and mean *POMC* methylation z score (Pearson R = -0.068, p=0.48, see Figure 5.14).

	VAS Score		
Predictors	**Estimates**	**95% CI**	**p-value**
(Intercept)	5.07	4.77 to 5.37	<0.001
VAS timepoint	-0.67	-0.77 to -0.56	0.0019
POMCm Z-score	0.14	-0.0015 to 0.27	0.052
VAS timepoint * POMCm Z-score (interaction)	-0.02	-0.046 to 0.006	0.13
Model information			
Number of participants	214 tests over 2 timepoints		
Observations	1302		

Table 5.9. Linear mixed effect model for VAS scores over time. VAS = visual analogue scale, CI = confidence interval. POMCm Z-score = Mean POMC methylation z score

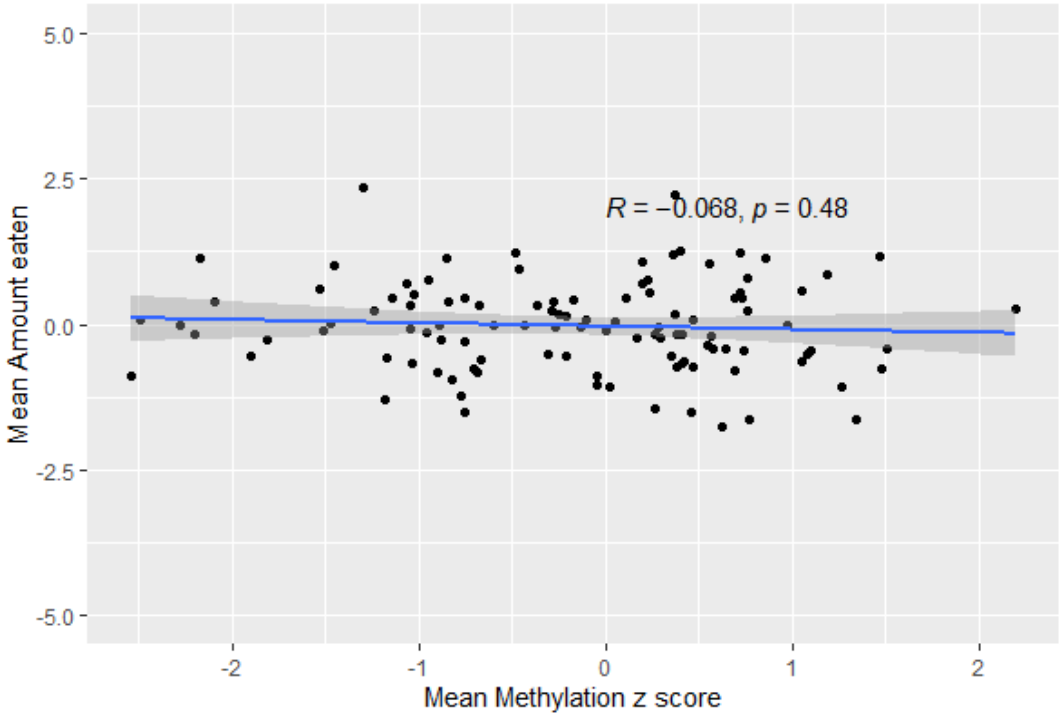


Figure 5.14. Scatterplot demonstrating the relationship between the amount eaten during ad libitum lunch and mean POMC methylation z score. Note the mean amount eaten is calculated as the mean of the residual z scores from models adjusting for sex, age and lean mass at midline and endline. Linear regression line shown in blue with 95% CI.

Eating speed and *POMC* methylation

Eating speed was pre-adjusted for sex, age, and lean mass in linear regression models at midline and endline. Z scores of the residuals from these models were produced for midline and endline and the mean z score taken for analysis. There was no association between the eating speed during the ad libitum meal and mean *POMC* methylation z score (Pearson R = -0.042, p=0.67, see Figure 5.15).

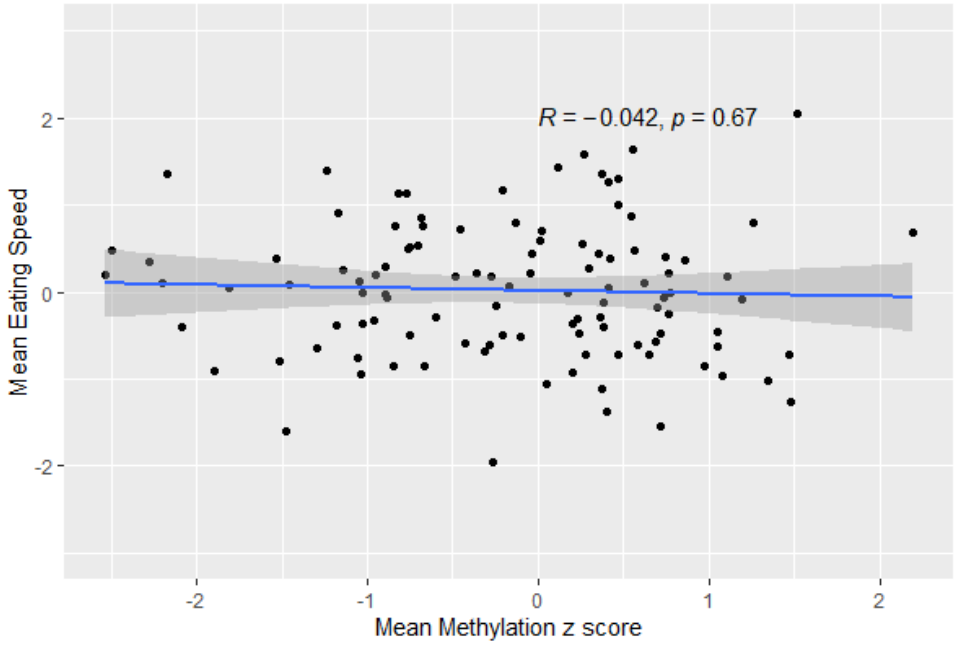


Figure 5.15. Scatterplot demonstrating the relationship between the eating speed during ad libitum lunch and mean POMC methylation z score. Note the mean eating speed is calculated as the mean of the residual z scores from models adjusting for sex, age and lean mass at midline and endline. Linear regression line shown in blue with 95% confidence interval.

5.5.5 Summary of child appetite test results

The results provide no evidence of an association between *POMC* methylation and any outcomes measures of appetite or satiety.

5.6 Maternal Appetite testing results

5.6.1 Descriptive statistics

A summary of midline and endline results is shown in Table 5.10. Compared to the children, there was a substantial number of mothers who did not participate in the test e.g. for those who ate lunch midline children 94% vs midline mothers 72% and endline children 96% vs endline mothers 67%.

	Midline	**Endline**
Number recruited to subset	n=114	n=98
Number participated in test	Breakfast n=85 Lunch n=82	Breakfast n=68 Lunch n=66
Mean Age [SD, range]	39.5 years [5.7, 25.2 to 50.4]	40.3 years [5.5, 26.4 to 50.8]
Number not participating	*Breakfast* 18 allergy – excluded 11 refused breakfast – excluded *Lunch* 18 allergy – already excluded 11 refused breakfast – excluded 3 refused lunch	*Breakfast* 22 allergy – excluded 8 refused breakfast – excluded *Lunch* 22 allergy – already excluded 8 refused breakfast -excluded 2 refused lunch
Median amount of preload breakfast eaten (g) [IQR, range]	292g [54, 32 to 563]	283g [68.8, 44 to 522]
Median time to eat (seconds) Preload breakfast [IQR, range]	323 seconds [120, 39 to 1062]	321.5 seconds [122.5, 188 to 1311]
Mean amount of ad libitum lunch eaten (g) [SD, range]	320.7 [102.3, 90 to 583]	399.5g [133.9, 15 to 641]
Median time to eat (seconds) of ad libitum lunch	944.5 seconds [465.8, 15 to 1439]	1006 seconds [357.5, 36 to 1433]

Table 5.10. Summary of midline and endline appetite tests for mothers. IQR = interquartile range, SD=standard deviation.

5.6.2 Ad libitum meal outcomes association with covariates

There were no significant correlations between amount eaten and speed of eating during the ad libitum lunch, and age or any measure of body composition derived from DXA scans or body weight.

5.6.3 VAS summary

The individual VAS scores were plotted from +20 to +120 and shown in Figure 5.16. The average trend (calculated by loess method) and mean score for each time point demonstrated a fall in VAS score after the preload demonstrating a decay in satiety with time and return to baseline by +120 minutes.

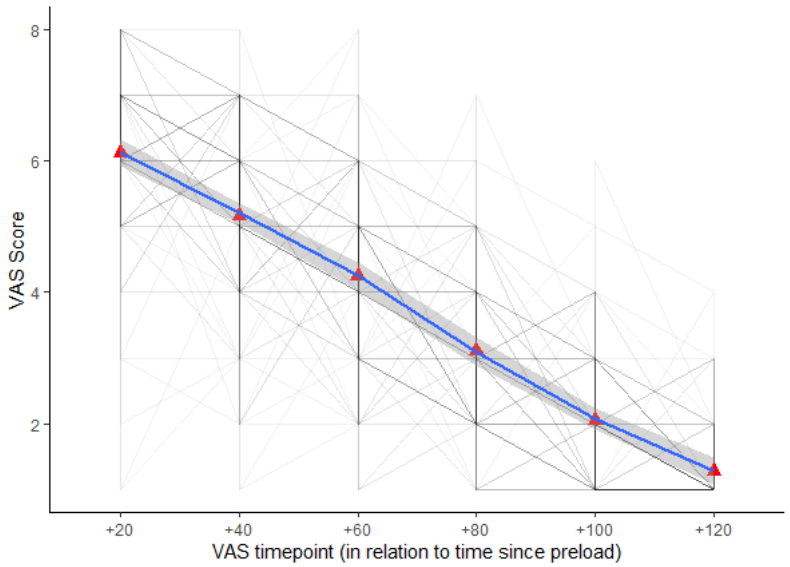


Figure 5.16. VAS score over time. Combined midline and endline plots of VAS score over time. Blue line shows trend over the timepoints calculated by loess method with 95% confidence interval. Colour density of individual line plots graded with in grey.

5.6.4 *POMC* methylation and measures of eating behaviour and satiety

POMC methylation data were available for 85/89 (95.5%) at baseline, 78/82 (95.1%) at midline and 64/66 (97.0%) at endline.

Subjective measures of satiety and *POMC* methylation

Different linear mixed effect models were produced with a 3-level random slope model (model 1 in Table 5.11) providing the best model fit. However, this model produced a singular fit suggesting the random effect structure was too complex for the data available. To improve model fit with this regard it is recommended to avoid fitting overly complex models[72], select a model that balances predictive accuracy against overfitting[72,73] and remove random terms to produce a non-singular fit[74]. With this in mind, the second model (2-level random slope model) was selected as it produced a non-singular fit.

	Measure of model fitness		
Model	**AIC**	**BIC**	**loglikelihood**
1. 3-level, random slope model	2679.6	2717.6	-1331.8
2. 2-level, random slope model	2443.9	2496.2	-1211.0

Table 5.11. Comparison of model fit. AIC= Akaike Information Criteria, BIC= Bayesian information criterion,

The coefficient estimates from the random intercept model are shown in Table 5.12 below. The average intercept and slope estimates demonstrate an average +20 minute post preload score of 7.14 [95% CI 6.80 – 7.48] and an average fall in VAS score of -0.99 [95% CI -1.05 - -0.93, p<0.001] every 20 minutes i.e. as expected time after eating is significantly associated with a reduction in satiety.

A similar lack of a *POMC* methylation effect was seen in mothers as for children. There was no evidence of an association between mean *POMC* methylation z score and intercept or slope i.e. *POMC* methylation was neither associated with subjective feelings of hunger after eating nor the fall in satiety between meals.

	VAS Score		
Predictors	**Estimates**	**95% CI**	**p-value**
(Intercept)	7.14	6.80 to 7.48	<0.001
VAS timepoint	-0.99	-1.05 to -0.93	<0.001
Mean methylation z score	0.24	-0.092 to 0.56	0.15
VAS timepoint * Mean methylation z score	-0.043	-0.10 to 0.015	0.14
Model information			
Number of participants	84		
Number of time points	2		
Observations	856		

Table 5.12. Linear mixed effect models for VAS over time. VAS = visual analogue scale, CI = confidence interval

Amount eaten and *POMC* methylation

There was no significant association between mean POMC methylation z score and amount eaten during ad libitum lunch (Pearson R = -0.041, p=0.71, see Figure 5.17).

Speed of eating and *POMC* methylation

There was no significant association between mean POMC methylation z score and amount eaten during ad libitum lunch (Pearson R = -0.072, p=0.53, see Figure 5.18).

5.6.5 Summary of maternal appetite test results

The results provide no evidence of an association between maternal *POMC* methylation and any outcomes measures of appetite or satiety.

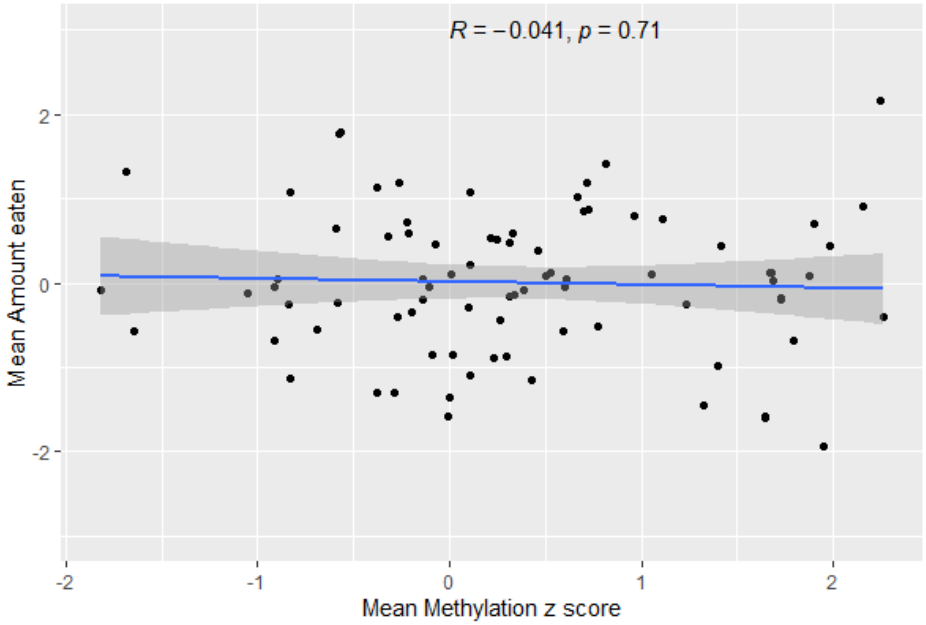


Figure 5.17. Scatterplot demonstrating the relationship between the amount eaten during ad libitum lunch and mean POMC methylation z score. Note the mean amount eaten is calculated as a mean z scores of the amount eaten at midline and endline. Linear regression line shown in blue with 95% CI.

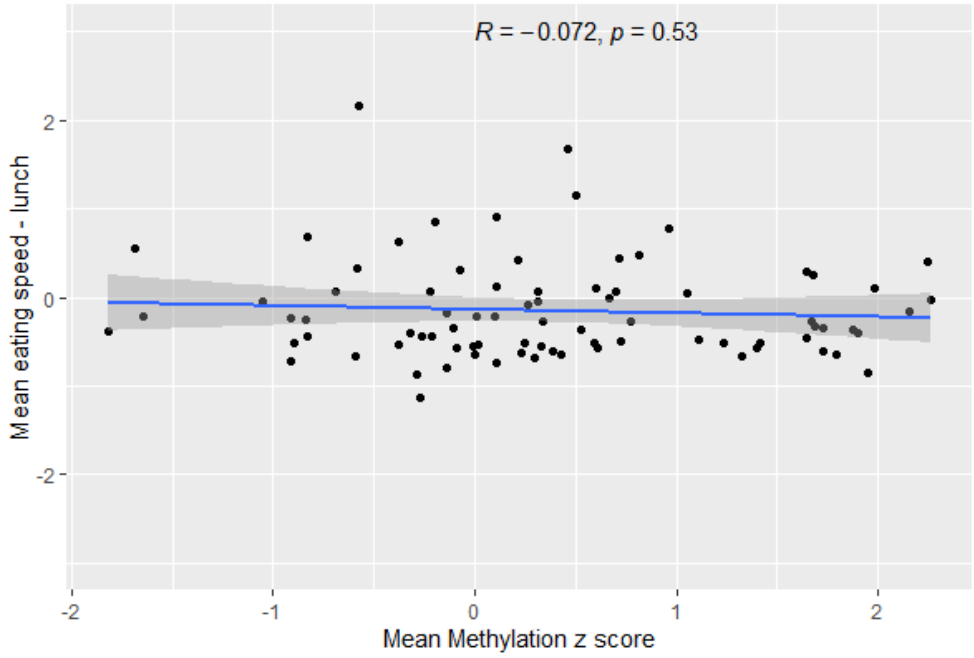


Figure 5.18. Scatterplot demonstrating the relationship between eating speed during ad libitum lunch and mean POMC methylation z score. Note the mean amount eaten is calculated as a mean z scores of the amount eaten at midline and endline. Linear regression line shown in blue with 95% CI.

5.7 Discussion

The results from the appetite test demonstrate no effect of *POMC* methylation on inter-meal satiety, amount consumed or speed of eating during an ad libitum meal. POMC is a mediator of satiety with evidence from animal studies that identified POMC neurons in the ARC of the hypothalamus and the nucleus tractus solitarius (NTS) of the brainstem play different roles in the satiety cascade[75]. POMC NTS neuron stimulation results in acute termination of eating, whereas only chronic stimulation of the ARC neurons are implicated in reduced food intake[76,77]. Interestingly ablation of POMC ARC neurons has been shown to lead to increased food intake and higher body weight however this is not so after ablation of POMC NTS neurons[76]. These studies suggest that the POMC neurons in the ARC are central to late satiety signals and long term energy balance. Previous studies have shown that human *POMC* methylation in the neurons from ARC of the hypothalamus are correlated with BMI[78] and that higher methylation is associated with lower POMC expression[25]. However, there is no evidence from this study to support a *POMC* methylation effect on satiety. There was however a *POMC* methylation interaction on the VAS slope that was directionally consistent in both children

and mothers (children = -0.02, p value 0.13 and mothers= -0.042, p value 0.14). This did not reach statistical significance and this may be as there is indeed no effect of *POMC* methylation on satiety signal in this population. Alernatively, the study could have been underpowered to detect an association with *POMC* methylation or the VAS tool may lack the precision to detect very small differences in satiety signal driven by *POMC* methylation. There are tools for measuring subjective feelings of hunger and satiety that use 100mm continuous line[41] and thus may enable more accurate measurements beyond to 1-5 and 1-8 scale used in this study. The reason for using a pictorial scale was many individuals are non-literate and/or non-numerate. Therefore, further appetite testing in different populations using more precise tools warrants further investigation.

One could postulate that a steeper fall in satiety could be associated with a shorter inter-meal interval or a compensatory increase in intake in a subsequent meal. However, food intake was only measured over 2 tests days and therefore the association between the satiety test scores and longer term eating behaviour remains untested. For obesity to develop there need on only be a very small imbalance between energy intake and expenditure[79,80]. By modelling the energy dynamics associated with the rise in average weight gain seen in the USA over recent decades, Hall et al stated that a persistent imbalance of +30 kJ per day underlies the weight gain observed[80]. Longer term follow up of eating behaviours or a more naturalistic study may help gain insight in the effect of differential methylation on longer term eating behaviours and energy intake.

POMC methylation was neither associated with amount eaten in the follow-up ad libitum meal nor eating speed in this study. Differences in amount eaten during an ad libitum meal have been seen in disorders of the melanocortin system. For example, individuals with *MC4R* mutations have been shown to eat three times the amount eaten by unaffected siblings during an ad libitum meal[8]. Similarly, alterations in the *FTO* gene have been associated with higher energy intake during a test meal[22,23]. Interestingly, *POMC* neuronal knock out rats eat larger meals and eat faster rather than more frequent eating episodes. Again, the *POMC* study looked at only two eating episodes over a year and therefore may not have captured differences in eating behaviours or maybe underpowered to detect the small differences in the amount eaten. Differential *POMC* methylation in humans could be influencing eating behaviour via the later post-absorptive satiety signal or tonic satiety signals rather than influencing satiation, meal size or eating speed.

This study demonstrated that a controlled appetite test can be performed in children and women in this setting. Importantly, non-numerate women and children can provide meaningful insights into subjective feelings of hunger. This has implications for appetite testing in other LMICs and potentially a clinical utility to formally assess appetite for conditions such as childhood malnutrition.

In the pilot test, the correlation between pre meal appetite scores and the amount eaten in an ad libitum meal demonstrated that these subjective feelings relayed by children relate in some part to meal size. VASs for hunger have been shown to reflect subsequent intake in a test meal in adults[82,83] and pictorial VAS for hunger has correlated with food intake in primary aged children[45]. This study suggests that children in mid-childhood are able to relate subjective feelings of hunger in a similar way to children from high income countries[44,50].

It is important to control for physiologic confounders of appetite such as age, sex and body size or composition[81]. The association between total lean mass and amount eaten and speed of eating in children is consistent with previous studies in children and adolescents from Europe[84–86]. Lean mass is a major contributor of resting metabolic rate and it is this that is thought to drive appetite and energy intake[33]. There was no association between appetite measures and any measure of body composition in adult women. Previously, studies in obese adults have identified a positive association between lean mass and daily and test meal energy intake[32,34,63]. A minority of women in the study were obese and therefore lean mass may not be associated with energy intake in this population of women.

The lack of an effect of *POMC* methylation on parameters in the appetite test are aligned with the observation of no association between maternal and child seasonal change in weight and *POMC* methylation (see chapter 4). An association between maternal FMI change and POMC methylation was observed. Energy imbalance has been associated with change in weight and fat mass[87], and an altered satiety response could contribute to changes in energy balance by effecting energy intake though not observed in this study. Fat loss is associated with a reduction in circulating leptin and provides a reduction in the leptin-driven positive feedback to POMC neurons which in turn leads to a reduction in POMC-mediated satiety signal, increased calorific intake and restoration or body fat stores[88,89]. Though from this study, there is no evidence from this appetite test that alterations in maternal satiety signal explain the change in fat mass observed across the year.

5.7.1 Limitations

The VAS used for mothers and children were limited to an eight and five point scale respectively. This can be considered a 'blunt' tool to express feelings of hunger compared to more elaborate VASs or the continuous 100mm line used in other settings. Only 2 test meals were measured at 2 time points across the year. More frequent testing may be needed to get more accurate assessment of

food intake and satiety. Naturalistic studies (e.g. observed dietary intake) may provide more ecologically relevant data though are time consuming, expensive and lack the control and precision of laboratory based studies. In The Gambia, shared bowl communal eating is commonplace and therefore the presentation of the food provided in the study may have been deemed unusual to some participants. Participants arrived fasted though the activity levels the day before could not be controlled or standardised. Due to infrastructure restraints there was only 2 rooms available to conduct the appetite testing. The presence of other people (particularly family and friends) during a meal has shown to increase the duration and energy intake of a meal in adults in the USA[90,91]. Therefore, where possible, study design should limit interaction with other participants or be done alone[81].

5.8 Conclusion

This study has provided no evidence that *POMC* methylation is associated with subjective reports of inter-meal satiety in both women and children.

5.9 References

1. Baldini G, Phelan KD, Rock L, Sciences D, Rock L. The melanocortin pathway and control of appetite- progress and therapeutic implications. *J Endocrinol*. 2019;241(1):1-54. doi:10.1530/JOE-18-0596.The

2. Balthasar N, Coppari R, McMinn J, et al. Leptin receptor signaling in POMC neurons is required for normal body weight homeostasis. *Neuron*. 2004;42(6):983-991. doi:10.1016/j.neuron.2004.06.004

3. Qiu J, Zhang C, Borgquist A, et al. Insulin excites anorexigenic proopiomelanocortin neurons via activation of canonical transient receptor potential channels. *Cell Metab*. 2014;19(4):682-693. doi:10.1016/j.cmet.2014.03.004

4. Parton LE, Ye CP, Coppari R, et al. Glucose sensing by POMC neurons regulates glucose homeostasis and is impaired in obesity. *Nature*. 2007;449(7159):228-232.

doi:10.1038/nature06098

5. Andermann ML, Lowell BB. Toward a Wiring Diagram Understanding of Appetite Control. *Neuron*. 2017;95(4):757-778. doi:10.1016/j.neuron.2017.06.014

6. Pradhan G, Samson SL, Sun Y. Ghrelin: much more than a hunger hormone. *Curr Opin Clin Nutr Metab Care*. 2013;16(6):619-624. doi:10.1097/MCO.0b013e328365b9be

7. Stutzmann F, Tan K, Vatin V, et al. Prevalence of melanocortin-4 receptor deficiency in Europeans and their age-dependent penetrance in multigenerational pedigrees. *Diabetes*. 2008;57(9):2511-2518. doi:10.2337/db08-0153

8. Farooqi IS, Keogh JM, Yeo GSH, Lank EJ, Cheetham T, O'Rahilly S. Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene. *N Engl J Med*. 2003;348(12):1085-1095. doi:10.1056/NEJMoa022050

9. Lotta LA, Mokrosiński J, Mendes de Oliveira E, et al. Human Gain-of-Function MC4R Variants Show Signaling Bias and Protect against Obesity. *Cell*. 2019;177(3):597-607.e9. doi:10.1016/j.cell.2019.03.044

10. Coll AP, Farooqi IS, Challis BG, Yeo GSH, O'Rahilly S. Proopiomelanocortin and Energy Balance: Insights from Human and Murine Genetics. *J Clin Endocrinol Metab*. 2004;89(6):2557-2562. doi:10.1210/jc.2004-0428

11. Millington GW. The role of proopiomelanocortin (POMC) neurones in feeding behaviour. *Nutr Metab (Lond)*. 2007;4:18. doi:10.1186/1743-7075-4-18

12. Raffan E, Dennis RJ, O'Donovan CJ, et al. A Deletion in the Canine POMC Gene Is Associated with Weight and Appetite in Obesity-Prone Labrador Retriever Dogs. *Cell Metab*. 2016;23(5):893-900. doi:10.1016/j.cmet.2016.04.012

13. Yaswen L, Diehl N, Brennan MB, Hochgeschwender U. Obesity in the mouse model of pro-opiomelanocortin deficiency responds to peripheral melanocortin. *Nat Med*. 1999;5(9):1066-1070. doi:10.1038/12506

14. Krude H, Biebermann H, Schnabel D, et al. Obesity Due to Proopiomelanocortin Deficiency: Three New Cases and Treatment Trials with Thyroid Hormone and ACTH4-10. *J Clin Endocrinol Metab*. 2003;88(10):4633-4640. doi:10.1210/jc.2003-030502

15. Çetinkaya S, Güran T, Kurnaz E, et al. A Patient with Proopiomelanocortin Deficiency: An Increasingly Important Diagnosis to Make. *J Clin Res Pediatr Endocrinol*. 2018;10(1):68-73.

doi:10.4274/jcrpe.4638

16. Clément K, van den Akker E, Argente J, et al. Efficacy and safety of setmelanotide, an MC4R agonist, in individuals with severe obesity due to LEPR or POMC deficiency: single-arm, open-label, multicentre, phase 3 trials. *lancet Diabetes Endocrinol*. 2020;8(12):960-970. doi:10.1016/S2213-8587(20)30364-8

17. Kühnen P, Clément K, Wiegand S, et al. Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. *N Engl J Med*. 2016;375(3):240-246. doi:10.1056/NEJMoa1512693

18. Llewellyn CH, van Jaarsveld CHM, Plomin R, Fisher A, Wardle J. Inherited behavioral susceptibility to adiposity in infancy: a multivariate genetic analysis of appetite and weight in the Gemini birth cohort. *Am J Clin Nutr*. 2012;95(3):633-639. doi:10.3945/ajcn.111.023671

19. Fisher JO, Cai G, Jaramillo SJ, Cole SA, Comuzzie AG, Butte NF. Heritability of hyperphagic eating behavior and appetite-related hormones among Hispanic children. *Obesity (Silver Spring)*. 2007;15(6):1484-1495. doi:10.1038/oby.2007.177

20. Carnell S, Haworth CMA, Plomin R, Wardle J. Genetic influence on appetite in children. *Int J Obes (Lond)*. 2008;32(10):1468-1473. doi:10.1038/ijo.2008.127

21. Wood AC. Gene-Environment Interplay in Child Eating Behaviors: What the Role of "Nature" Means for the Effects of "Nurture." *Curr Nutr Rep*. 2018;7(4):294-302. doi:10.1007/s13668-018-0254-x

22. Jonsson A, Franks PW. Obesity, FTO gene variant, and energy intake in children. *N Engl J Med*. 2009;360(15):1571-1572; author reply 1572. doi:10.1056/NEJMc090017

23. Wardle J, Llewellyn C, Sanderson S, Plomin R. The FTO gene and measured food intake in children. *Int J Obes*. 2009;33(1):42-45. doi:10.1038/ijo.2008.174

24. Kühnen P, Handke D, Waterland RA, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016;24(3):502-509. doi:10.1016/j.cmet.2016.08.001

25. Kuehnen P, Mischke M, Wiegand S, et al. An alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet*. 2012;8(3):1-12. doi:10.1371/journal.pgen.1002543

26. Blundell J, De Graaf C, Hulshof T, et al. Appetite control: Methodological aspects of the

evaluation of foods. *Obes Rev*. 2010;11(3):251-270. doi:10.1111/j.1467-789X.2010.00714.x

27. Livingstone MBE, Robson PJ, Welch RW, Burns AA, Burrows MS, McCormack C. Methodological issues in the assessment of satiety. *Näringsforskning*. 2000;44(1):98-103. doi:10.3402/fnr.v44i0.1776

28. Blundell JE, Lawton CL, Cotton JR, Macdiarmid JI. Control of human appetite: implications for the intake of dietary fat. *Annu Rev Nutr*. 1996;16:285-319. doi:10.1146/annurev.nu.16.070196.001441

29. Blundell J. Pharmacological approaches to appetite suppression. *Trends Pharmacol Sci*. 1991;12(4):147-157. doi:10.1016/0165-6147(91)90532-w

30. Cummings DE. Ghrelin and the short- and long-term regulation of appetite and body weight. *Physiol Behav*. 2006;89(1):71-84. doi:10.1016/j.physbeh.2006.05.022

31. Maffei M, Halaas J, Ravussin E, et al. Leptin levels in human and rodent: measurement of plasma leptin and ob RNA in obese and weight-reduced subjects. *Nat Med*. 1995;1(11):1155-1161. doi:10.1038/nm1195-1155

32. Blundell JE, Caudwell P, Gibbons C, et al. Body composition and appetite: fat-free mass (but not fat mass or BMI) is positively associated with self-determined meal size and daily energy intake in humans. *Br J Nutr*. 2012;107(3):445-449. doi:10.1017/S0007114511003138

33. Blundell JE, Caudwell P, Gibbons C, et al. Role of resting metabolic rate and energy expenditure in hunger and appetite control: a new formulation. *Dis Model Mech*. 2012. doi:10.1242/dmm.009837

34. Hopkins M, Finlayson G, Duarte C, et al. Modelling the associations between fat-free mass, resting metabolic rate and energy intake in the context of total energy balance. *Int J Obes (Lond)*. 2016;40(2):312-318. doi:10.1038/ijo.2015.155

35. Bray GA. Afferent signals regulating food intake. *Proc Nutr Soc*. 2000;59(3):373-384. doi:10.1017/s0029665100000422

36. Gibbons C, Finlayson G, Dalton M, Caudwell P, Blundell JE. METABOLIC PHENOTYPING GUIDELINES: Studying eating behaviour in humans. *J Endocrinol*. 2014;222(2):G1-G12. doi:10.1530/joe-14-0020

37. Rolls BJ, Hetherington M, Burley VJ. The specificity of satiety: The influence of foods of different macronutrient content on the development of satiety. *Physiol Behav*.

1988;43(2):145-153. doi:10.1016/0031-9384(88)90230-2

38. Rouhani MH, Surkan PJ, Azadbakht L. The effect of preload/meal energy density on energy intake in a subsequent meal: A systematic review and meta-analysis. *Eat Behav*. 2017;26:6-15. doi:10.1016/j.eatbeh.2016.12.011

39. von Baeyer CL. Children's self-report of pain intensity: what we know, where we are headed. *Pain Res Manag*. 2009;14(1):39-45. doi:10.1155/2009/259759

40. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage*. 2011;41(6):1073-1093. doi:10.1016/j.jpainsymman.2010.08.016

41. Forde CG. Chapter 7 Measuring Satiation and Satiety. In: ; 2021:151-182. doi:10.1016/b978-0-08-101743-2.00007-8

42. Hill AJ, Blundell JE. Nutrients and behaviour: research strategies for the investigation of taste characteristics, food preferences, hunger sensations and eating patterns in man. *J Psychiatr Res*. 17(2):203-212. doi:10.1016/0022-3956(82)90023-1

43. Solah VA, Meng X, Wood S, et al. Effect of training on the reliability of satiety evaluation and use of trained panellists to determine the satiety effect of dietary fibre: a randomised controlled trial. *PLoS One*. 2015;10(5):e0126202. doi:10.1371/journal.pone.0126202

44. Faith MS, Kermanshah M, Kissileff HR. Development and preliminary validation of a silhouette satiety scale for children. *Physiol Behav*. 2002. doi:10.1016/S0031-9384(02)00702-3

45. Bennett C, Blissett J. Measuring hunger and satiety in primary school children. Validation of a new picture rating scale. *Appetite*. 2014. doi:10.1016/j.appet.2014.03.011

46. Gibbons C, Caudwell P, Finlayson G, King N, Blundell J. Validation of a new hand-held electronic data capture method for continuous monitoring of subjective appetite sensations. *Int J Behav Nutr Phys Act*. 2011;8(1):57. doi:10.1186/1479-5868-8-57

47. Whybrow S, Stephen JR, Stubbs RJ. The evaluation of an electronic visual analogue scale system for appetite and mood. *Eur J Clin Nutr*. 2006;60(4):558-560. doi:10.1038/sj.ejcn.1602342

48. Stubbs RJ, Hughes DA, Johnstone AM, et al. The use of visual analogue scales to assess

motivation to eat in human subjects: a review of their reliability and validity with an evaluation of new hand-held computerized systems for temporal tracking of appetite ratings. *Br J Nutr*. 2000;84(4):405-415. doi:DOI: 10.1017/S0007114500001719

49. de Graaf C. The validity of appetite ratings. *Appetite*. 1993;21(2):156-160. doi:10.1016/0195-6663(93)90008-8

50. Keller KL, Assur SA, Torres M, et al. Potential of an analog scaling device for measuring fullness in children: development and preliminary testing. *Appetite*. 2006;47(2):233-243. doi:10.1016/j.appet.2006.04.004

51. Sadoul BC, Schuring EAH, Symersky T, Mela DJ, Masclee AAM, Peters HPF. Measuring satiety with pictures compared to visual analogue scales. An exploratory study. *Appetite*. 2012;58(1):414-417. doi:https://doi.org/10.1016/j.appet.2011.10.016

52. Ickes SB, Hossain M, Ritter G, et al. Systematic Review of Tools and Methods to Measure Appetite in Undernourished Children in the Context of Low- and Middle-Income Countries. *Adv Nutr*. 2018;9(6):789-812. doi:10.1093/advances/nmy042

53. Dossa RAM, Ategbo EAD, Van Raaij JMA, De Graaf C, Hautvast JGAJ. An appropriate tool for appetite testing and evaluation in young children in Benin. *Appetite*. 2002. doi:10.1006/appe.2001.0457

54. Dossa RAM, Ategbo ED, Raaij JMA Van, Graaf C De, Hautvast JGAJ. Multivitamin-Multimineral and Iron Supplementation Did Not Improve Appetite of Young Stunted and Anemic Beninese Children. *J Nutr*. 2001.

55. Mda S, van Raaij JMA, Macintyre UE, de Villiers FPR, Kok FJ. Duration of hospitalization and appetite of HIV-infected South African children. *Matern Child Nutr*. 2011;7(2):175-187. doi:10.1111/j.1740-8709.2009.00228.x

56. J.W. L, M.C. L, L.S. S, S.N. K. Iron supplementation improves appetite and growth in anemic Kenyan primary school children. *J Nutr*. 1994.

57. Latham MC, Stephenson LS, Kurz KM, Kinoti SN. Metrifonate or praziquantel treatment improves physical fitness and appetite of Kenyan schoolboys with Schistosoma haematobium and hookworm infections. *Am J Trop Med Hyg*. 1990. doi:10.4269/ajtmh.1990.43.170

58. Rehman AM, Woodd S, Chisenga M, et al. Appetite testing in HIV-infected African adults recovering from malnutrition and given antiretroviral therapy. *Public Health Nutr*. 2015;18(4):742-751. doi:10.1017/S1368980014000718

59. Prentice AM, Paul AA. Fat and energy needs of children in developing countries. *Am J Clin Nutr*. 2000;72(5):1253s-1265s. doi:10.1093/ajcn/72.5.1253s

60. Mda S, van Raaij JMA, MacIntyre UE, de Villiers FPR, Kok FJ. Improved appetite after multi-micronutrient supplementation for six months in HIV-infected South African children. *Appetite*. 2010. doi:10.1016/j.appet.2009.09.021

61. Mda S, van Raaij JMA, Macintyre UE, de Villiers FPR, Kok FJ. Duration of hospitalization and appetite of HIV-infected South African children. *Matern Child Nutr*. 2011. doi:10.1111/j.1740-8709.2009.00228.x

62. NATIONS FAAOOTU. *Food Energy - Methods of Analysis and Conversion Factors*. Rome; 2002.

63. Blundell JE, Finlayson G, Gibbons C, Caudwell P, Hopkins M. The biology of appetite control: Do resting metabolic rate and fat-free mass drive energy intake? *Physiol Behav*. 2015;152:473-478. doi:https://doi.org/10.1016/j.physbeh.2015.05.031

64. WHO GMGRS. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. *World Heal Organ*. 2006. doi:10.4067/S0370-41062009000400012

65. Microsoft 365. Paint 3D. 2016.

66. Vander Weg MW, Watson JM, Klesges RC, Eck Clemens LH, Slawson DL, McClanahan BS. Development and cross-validation of a prediction equation for estimating resting energy expenditure in healthy African-American and European-American women. *Eur J Clin Nutr*. 2004;58(3):474-480. doi:10.1038/sj.ejcn.1601833

67. Joint WHO/FAO/UNU Expert Consultation. *World Health Organ Tech Rep Ser*.; 1985.

68. Jenny Foley. The potential role of protein in determining total energy intake in rural Gambian men: the Protein Leverage Hypothesis. 2006.

69. Gibson MJ, Dawson JA, Wijayatunga NN, et al. A randomized cross-over trial to determine the effect of a protein vs. carbohydrate preload on energy balance in ad libitum settings. *Nutr J*. 2019;18(1):69. doi:10.1186/s12937-019-0497-4

70. Caudwell P, Finlayson G, Gibbons C, et al. Resting metabolic rate is associated with hunger, self-determined meal size, and daily energy intake and may represent a marker for appetite. *Am J Clin Nutr*. 2013;97(1):7-14. doi:10.3945/ajcn.111.029975

71. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J*

Stat Software; Vol 1, Issue 1 . October 2015.

72. Matuschek H, Kliegl R, Vasishth S, Baayen H, Bates D. Balancing Type I error and power in linear mixed models. *J Mem Lang*. 2017;94:305-315. doi:https://doi.org/10.1016/j.jml.2017.01.001

73. Bates D, Kliegl R, Vasishth S, Baayen H. Parsimonious Mixed Models. *arXiv*. 2015;1506.

74. Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang*. 2013;68(3):255-278. doi:https://doi.org/10.1016/j.jml.2012.11.001

75. Cassidy RM, Tong Q. Hunger and Satiety Gauge Reward Sensitivity . *Front Endocrinol* . 2017;8:104.

76. Zhan C, Zhou J, Feng Q, et al. Acute and long-term suppression of feeding behavior by POMC neurons in the brainstem and hypothalamus, respectively. *J Neurosci*. 2013;33(8):3624-3632. doi:10.1523/JNEUROSCI.2742-12.2013

77. Aponte Y, Atasoy D, Sternson SM. AGRP neurons are sufficient to orchestrate feeding behavior rapidly and without training. *Nat Neurosci*. 2011;14(3):351-355. doi:10.1038/nn.2739

78. Kühnen P, Handke D, Waterland R a, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016:1-8. doi:10.1016/j.cmet.2016.08.001

79. Hill JO, Wyatt HR, Peters JC. Energy balance and obesity. *Circulation*. 2012;126(1):126-132. doi:10.1161/CIRCULATIONAHA.111.087213

80. Hall KD, Sacks G, Chandramohan D, et al. Quantification of the effect of energy imbalance on bodyweight. *Lancet*. 2011. doi:10.1016/S0140-6736(11)60812-X

81. Benelam B. Satiation, satiety and their effects on eating behaviour. *Nutr Bull*. 2009;34(2):126-173. doi:https://doi.org/10.1111/j.1467-3010.2009.01753.x

82. Porrini M, Crovetti R, Testolin G, Silva S. Evaluation of satiety sensations and food intake after different preloads. *Appetite*. 1995;25(1):17-30. doi:10.1006/appe.1995.0038

83. Barkeling B, Rössner S, Sjöberg A. Methodological studies on single meal food intake characteristics in normal weight and obese men and women. *Int J Obes Relat Metab Disord J Int Assoc Study Obes*. 1995;19(4):284-290.

84. Vainik U, Konstabel K, Lätt E, Mäestu J, Purge P, Jürimäe J. Diet misreporting can be corrected: confirmation of the association between energy intake and fat-free mass in adolescents. *Br J Nutr*. 2016;116(8):1425-1436. doi:10.1017/S0007114516003317

85. Cameron JD, Sigal RJ, Kenny GP, et al. Body composition and energy intake - skeletal muscle mass is the strongest predictor of food intake in obese adolescents: The HEARTY trial. *Appl Physiol Nutr Metab = Physiol Appl Nutr Metab*. 2016;41(6):611-617. doi:10.1139/apnm-2015-0479

86. Steinsbekk S, Llewellyn CH, Fildes A, Wichstrøm L. Body composition impacts appetite regulation in middle childhood. A prospective study of Norwegian community children. *Int J Behav Nutr Phys Act*. 2017;14(1):70. doi:10.1186/s12966-017-0528-5

87. Strasser B, Spreitzer A, Haber P. Fat loss depends on energy deficit only, independently of the method for weight loss. *Ann Nutr Metab*. 2007;51(5):428-432. doi:10.1159/000111162

88. Guyenet SJ, Schwartz MW. Regulation of Food Intake, Energy Balance, and Body Fat Mass: Implications for the Pathogenesis and Treatment of Obesity. *J Clin Endocrinol Metab*. 2012;97(3):745-755. doi:10.1210/jc.2011-2525

89. Morton GJ, Cummings DE, Baskin DG, Barsh GS, Schwartz MW. Central nervous system control of food intake and body weight. *Nature*. 2006;443(7109):289-295. doi:10.1038/nature05026

90. de Castro JM, de Castro ES. Spontaneous meal patterns of humans: influence of the presence of other people. *Am J Clin Nutr*. 1989;50(2):237-247. doi:10.1093/ajcn/50.2.237

91. de Castro JM. Family and friends produce greater social facilitation of food intake than other companions. *Physiol Behav*. 1994;56(3):445. doi:10.1016/0031-9384(94)90286-0

Chapter 6 Periconceptional, maternal and genetic predictors of *POMC* methylation

Summary of the chapter

In this chapter, I explore periconceptional factors including season of conception, maternal nutritional biomarkers, mother's BMI, pregnancy supplementation that potentially influence offspring's *POMC* methylation. The influence of child genotype and parent-child *POMC* methylation correlations are also reported.

6.1 Introduction

Epidemiological evidence points to a prenatal period of embryonic plasticity whereby early environmental exposures can set a developmental course associated with postnatal and adult health outcomes[1–3]. Epigenetic modifications are believed to act as biological conduit between environmental exposures and later disease[4–6]. There is a burgeoning body of evidence in humans that environmental exposures to mother (and father) can influence offspring's epigenetics landscape and disease susceptibility.

Maternal prenatal environmental exposures, such as maternal diabetes [7–11], maternal BMI [12–15], maternal famine exposure [16–22], maternal tobacco use [23–26], maternal nutrition [27–38], are associated with epigenetic modifications in their offspring. There is emerging evidence to suggest that paternal factors, such as obesity may also influence sperm and offspring DNA methylation[39–44] . The periconceptional (or early gestational) period is an important time for epigenetic reprogramming and has been identified as a key window for environmental exposures to influence offspring's epigenome [16,19,36]. Establishing causal relationships between prenatal exposures, alterations to the epigenome and subsequent postnatal phenotype is challenging. However using techniques such as Mendelian Randomisation, mediation analysis and prospective longitudinal studies have aimed to established causal links[22,45,46].

Offspring's *POMC* methylation has been shown to be sensitive to mother's periconceptional milieu in animal[47] and humans studies[34].

Data from a Gambian cohort of 144 mother-child dyads demonstrated a link between the season of conception and offspring DNA methylation with higher mean methylation (across 6 MEs) in those conceived in the rainy season[28]. Increased DNA methylation at a variably methylated region (VMR) at the intron2/exon3 border of the *POMC* gene has been associated with those conceived in the rainy season[34].

One-carbon metabolites provide methyl groups for DNA methylation reactions (see Figure 1.3). One-carbon metabolite concentrations in maternal plasma taken in very early pregnancy (mean (SD) 8.6±4.0 weeks gestation) has been associated with offspring *POMC VMR* methylation taken in infancy (mean (SD) 3.6±0.9 months of age)[34]. Specifically, DNA methylation at the POMC VMR was negatively correlated with S-adenosyl homocysteine (SAH) and positively correlated with betaine and S-adenosyl methionine (SAM):SAH ratio[34]. Furthermore, *POMC* methylation was significantly correlated between the new-born period and childhood/early adolescence, suggesting stability of

methylation through childhood. In a study of 47 family trios (mother-father-child trios) from Germany, a significant correlation between offspring and paternal but not maternal *POMC* methylation was found[34]. A pattern of apparent patriline inheritance of epigenetic marks at *POMC* has also been reported in animal studies[48,49] whereby persistence of DNA methylation patterns associated with foetal alcohol exposure were seen in male progeny in F2 and F3 generations from the male germline only[50].

This chapter explores how the maternal periconceptional nutritional environment influences *POMC* methylation in the offspring. Furthermore, the influence of factors such as age and sex are explored. Intergenerational patterns of *POMC* methylation are also examined within family trios. The effect of proximal genotype on *POMC* methylation is reported.

6.2 Methods

The methodology outlined below describes how predictors of *POMC* VMR methylation were explored. Both 'within child' and maternal factors were explored. The predictors explored included child's sex, child's age, ENID supplementation group, maternal BMI, season of conception, maternal circulating early pregnancy nutritional biomarkers and genotype. *POMC* methylation was explored at age 2 years and in mid childhood. The reasons for this were two fold i) to look for longitudinal differences in the effects ii) there was bigger sample size with mid childhood measures compared to 2 year measurements.

6.2.1 ENID (Early Nutrition and Immune Development) Trial and maternal supplementation

The *POMC* study followed up children born to mothers who participated in the ENID trial, carried out in 2010-2014 ([51], see section 2.2.1).

The ENID trial[51] and cohort is described in section 2.2.1. In brief, women were recruited in early pregnancy (10-20 weeks) and randomised to receive either i) Iron-Folate (standard care) ii) multiple micronutrient (MMN) iii) Energy, protein, and lipid with Iron-Folate; or iv) energy, protein, and lipid with MMN supplements for the remainder of their pregnancy. There were no differences in maternal (BMI, age, parity) or infant characteristics (birthweight, birth length, sex or gestational age)

across the study arms[52]. From 6 to 18 months of age, infants were further randomized to a lipid-based nutritional supplement, with or without additional MMN.

6.2.2 MDEG 2 (Methyl Donors and Epigenetics 2) study and maternal nutritional biomarker data

Women recruited to the ENID trial had provided a 10ml venous blood sample at the time of the first study visit after their first missed menses and pregnancy was confirmed by a positive urinary hCG test. The women's first blood sample was therefore taken in early pregnancy (<20 weeks gestation) and *before* they commenced any trial supplements. Blood samples were centrifuged at 1800g (rcf), and the plasma was drawn into 2ml microtubes and then frozen and stored at -70^{o}c.

The MDEG2 study utilised these stored samples from 350 women from the ENID trial. From these women, stored first study visit blood samples were analysed for levels of nutritional biomarker analysis[53] (see section 2.2.2, MDEG2 study). The mean age of the women in MDEG2 was 28.8 years (95% CI, 28.1-29.5) with a mean BMI of 20.8 kg/m^{2}(20.5-21.1). The mean gestational age (assessed by ultrasound) when the first blood sample (i.e. sample used for nutritional biomarker analysis) was taken was 11.6 weeks gestation (95% CI 11.4-11.9). The MDEG2 women were selected to give an even distribution by month of booking and by the earliest gestational age i.e. nearest to time of conception.

Biomarkers were analysed at Child and Family Research Institute at the University of British Columbia using liquid chromatography-tandem mass spectrometry (one-carbon metabolites: choline, betaine, dimethylglycine (DMG), homocysteine, B6 vitamers (4-pyridoxic acid (PA), pyridoxal (PL) and pyridoxal-5'-phosphate (PLP), vitamin B2 (riboflavin), uracil, uridine), Abbott AxSYM autoanalyzer (folate, vitamin B12) and Hitachi L-8900 amino acid analyser (amino acids: cysteine, methionine, serine, glycine, alanine, arginine, aspartic acid, glutamic acid, histidine, isoleucine, leucine, lysine, phenylalanine, proline, threonine, tryptophan, tyrosine and valine). The inflammatory marker AGP was measured using the Cobas Integra 400 plus autoanalyser at MRC The Gambia, Keneba field station.

6.2.3 Season of conception definitions

Season of conception (SoC) was defined as 'dry' (January-June) and 'rainy' (July-December) as previously described[54] with the conception date calculated from a gestational age estimation obtained from antenatal ultrasound at ENID trial recruitment.

6.2.4 *POMC* methylation measures

For the analysis in this chapter, two measures of POMC methylation were used:

i) At age 2 years (only described in this chapter in relation to association with predictors)
ii) In mid-childhood (same measure of *POMC* methylation reported in chapters 4 and 5)

This allowed for assessment of associations between methylation predictors at different time points and to see if any association change over time.

The 2 year methylation measure used the Agilient SureSelect Methyl-seq target enrichment system with target-enriched DNA, including the *POMC* region of interest, sequenced using the Illumina NovaSeq platform. Compared to pyrosequencing (as used for the mid-childhood measure) this technology has an advantage of being able to efficiently measure DNA methylation across larger parts of the genome.

6.2.4.1 POMC Methylation measurement at 2 years of age

Children from the original ENID trial had genomic DNA isolated from their blood at age 2 years. DNA was enriched and bisulfite-converted using the custom Agilent SureSelect Methyl-seq target enrichment system on a subset of these children (n=521)[55,56]. The Methyl-seq targeted capture included probes for CpGs +1 to+5 (relative to exon 3, see figure 4.5 for relative genomic coordinates) in the *POMC* VMR. Target-enriched DNA was then sequenced using the Illumina NovaSeq platform at the Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA. Reads were mapped to the human genome (hg38) by Noah Kessler (collaborator bioinformatician) using Bismark v0.20.0[57] with default options, which was also used to extract methylation values after mapping. Methylation calls from opposite strands of the same CpG site were combined. Within each

individual, CpG sites were considered 'covered' if they had a read depth of at least 20x; un-covered sites were excluded from analyses leaving 283 ENID participants with methylation data for CpGs between +1 and +5 (relative to exon 3 of the POMC VMR, see figure 4.5).

6.2.4.2 POMC Methylation measurement in mid-childhood

Methylation for the mid-childhood (*POMC* study) participants were measured by pyrosequencing (as described in chapter 4.2.7 methods section).

For both ENID 2-year samples and POMC mid-childhood samples, a single measure of POMC methylation was taken for 5 CpGs in exon 3 of the POMC gene (see chapter 4.2.7). Methylation z scores were calculated for each CpG and a mean z score calculated across the 5 CpGs.

6.2.5 Genotype

450 of 493 enrolled in the study children from the POMC study had genotype data available derived from the Illumina H3Africa genotyping array[58].

Noah Kessler, a bioinformatician previously with the nutrition theme at MRC The Gambia at LHSTHM now at the University of Cambridge, prepared the genotype data to provide me with genotype for each child. His methodology is outlined below.

Initially, all 22 SNPs on the H3Africa genotyping array within 10kb of the *POMC* gene were considered (hg38: chr2:25150853-25178690). Publicly available data from the Gambian Genome Variation Project (encompassing genomic data from the Fula, Wollof, Jola and Mandinka ethnic groups) were accessed via Ensembl[59]. Linkage disequilibrium, defined as the non-random association of alleles at different loci[60] was measured by r^2. The r^2 is the square of the correlation coefficient of two alleles i.e. the correlation between presence or absence of a particular allele at one locus and the presence or absence of a particular allele at another locus[61]. Figure 6.1 below demonstrates the LD r^2 for the 22 SNPs within 10kb of the POMC gene (the blue arrows show the location of the *POMC* variably methylated region (VMR)).

SNPs affecting methylation in the POMC VMR were then identified as follows:
For each CpG-SNP pair, the median CpG methylation for all genotypes with ≥20 individuals was

calculated, and range of the medians reviewed. SNPs with an effect size of ≥10% (difference between genotypes) were carried forward (N=7). These were then filtered by removing any SNP which was in LD (r^2≥0.8) with a higher-effect-size SNP (see Figure 6.1), leaving four SNPs for review. The selected SNPs were chr2:25156150 (hg38) or rs11892647 (rs SNP ID), chr2:25158279 or rs6751851, chr2:25161964 or rs6713532 and chr2:25171781 or rs6545976.

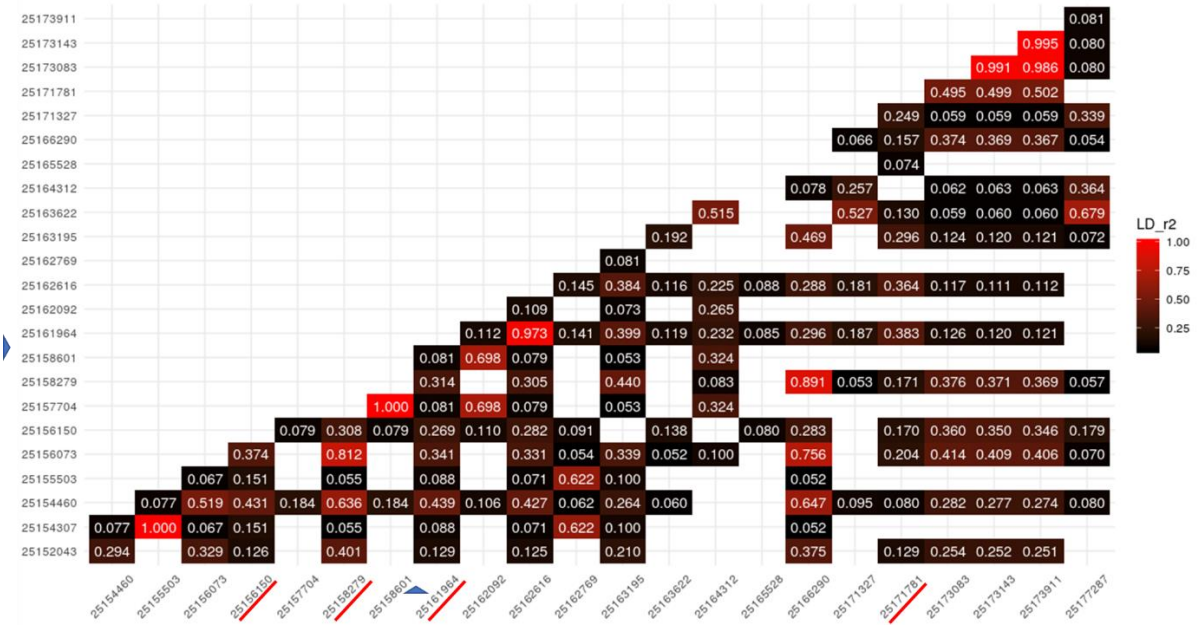


Figure 6.1 Pairwise linkage disequilibrium (LD) plot of 22 SNPs located within 10kb from the POMC gene region of interest (shown with blue arrows). Genomic regions are mapped to hg38 on chromosome 2 and data taken from Gambian Genome Variation Project was accessed via Ensembl[59]. Pearson R^2 reported. SNPs selected for further analysis underlined in red on the x axis.

6.2.6 Statistical analysis

Both 2-year and mid childhood mean *POMC* methylation z score were normally distributed. The relationship between *POMC* methylation z score and predictors were assessed using appropriate statistical tests as outlined below. To examine for any change in periconceptional effect (e.g. pregnancy supplementation, SoC, periconceptional nutritional biomarker) through childhood, *POMC* methylation measurements at both 2 years and mid-childhood were used in analysis.

Sex

Differences between male and female percentage methylation across CpGs in the *POMC* region of interest (CpGs -2 to +7 in the *POMC* VMR) was assessed by the student t test.

Child' s age

The relationship between age at *POMC* methylation measurement and *POMC* mean methylation z score was reported using Spearman Correlation coefficient.

ENID supplementation group

Comparisons of mean *POMC* methylation z score across ENID pregnancy supplementation groups (FeFol, MMN, PE, PE and MMN) were assessed by ANOVA (Analysis of variance) test. Comparisons of mean *POMC* methylation z score between ENID infant supplementation groups (no MMN vs MMN) were assessed by ANOVA (Analysis of variance) test.

Maternal BMI

The correlation between periconceptional maternal BMI and mean *POMC* methylation z score was reported using Spearman Correlation coefficient.

SoC and *POMC* methylation

Previous analyses have considered season as a dichotomised variable for ease of interpretation. The dichotomised comparisons across each CpG were made as part of initial exploratory analysis though SoC effect was formally assessed using Fourier analysis outlined below.

Previously, methylation across metastable epialleles has been associated with SoC (see chapter 1.1.5).

- Seasonal oscillations around a mean can be modelled using Fourier terms in linear regression models[62].

- A likelihood ratio test was used to compare these two models i.e. comparing if having Fourier terms in the model (and therefore seasonality) was significantly associated with *POMC* methylation. Seasonal trends were visualised by plotting Fourier terms against mean *POMC* methylation.

The Fourier terms introduce cosine and sine waves to model SoC differences in *POMC* methylation.

- Firstly, date of conception (calculated from a gestational age estimation obtained from antenatal ultrasound at ENID trial recruitment) was converted to radians so time represented a cyclical variable (i.e. 31st December (365th day of year) is adjacent to 1st January (1st day of year).
- To convert date of conception into radians; time was expressed a number between 0 – 1 by dividing conception day of year by 365 e.g. conception date of 25th October is the 298th day of year and can be expressed as 298/365 = 0.816. This number was then multiplied by 2π to produce time in radians (θ) i.e. 0.816 x 2π = 5.12.
- The Fourier terms were created by taking the sin and cosine of the radians (θ) and inputted as predictors in a linear regression model (with mean *POMC* methylation z score as outcome variable) to assess seasonal patterns of offspring *POMC* methylation.
- Linear regression models (adjusted for age in children in mid-childhood) are produced with and without Fourier terms.

Biomarker analysis

All maternal biomarkers were preadjusted for gestational age, maternal BMI, maternal age and inflammation (AGP) and then back extrapolated to date of conception using previously described methods[28]. Those biomarkers not normally distributed were log-transformed and all biomarkers were scaled and centred to enable comparison of standardised coefficients. In multiple linear regression models, *POMC* mean methylation z-score was the dependent variable with each biomarker fitted as a predictor and adjusted for sex. Models were fitted for each biomarker individually. Both 2 year and mid-childhood methylation measurement were assessed. All model covariates were assessed for multicollinearity, and standard tests were performed to ensure that linear modelling assumptions were met.

Genotype

The association between genotype and mean POMC methylation z score was explored. ANOVA test was performed to assess the influence of genotype on mean *POMC* methylation z score. Dunn's multiple comparison test, a post hoc test, using Bonferroni correction was used to determine the significant differences between individual genotypes.

To test if SoC was influenced by genotype, 2x3 contingency tables were produced as SoCxgenotype. Any associations were then assessed by chi squared test.

Intergenerational POMC methylation correlations

Associations between parent and offspring mean *POMC* methylation z score were reported using Pearson's correlation coefficient. 95 mother-father-child trios had methylation data available.

Stability of POMC methylation from mid-childhood to after puberty/adolescence

Stability of *POMC* methylation with age was assessed by Pearson's correlation coefficient by correlating paired ENID 2 year *POMC* methylation with *POMC* mid-childhood methylation (n=265).

6.3 Results

Factors predicting children's *POMC* methylation

6.3.1 Sex

Across all CpGs girls had significantly higher percentage methylation (measured in mid-childhood) across all CpGs compared to boys (see figure 6.2). Given this finding, mean *POMC* methylation z scores were calculated separately for boys and girls to allow adjustment and comparison in later analyses (see section 4.2.7).

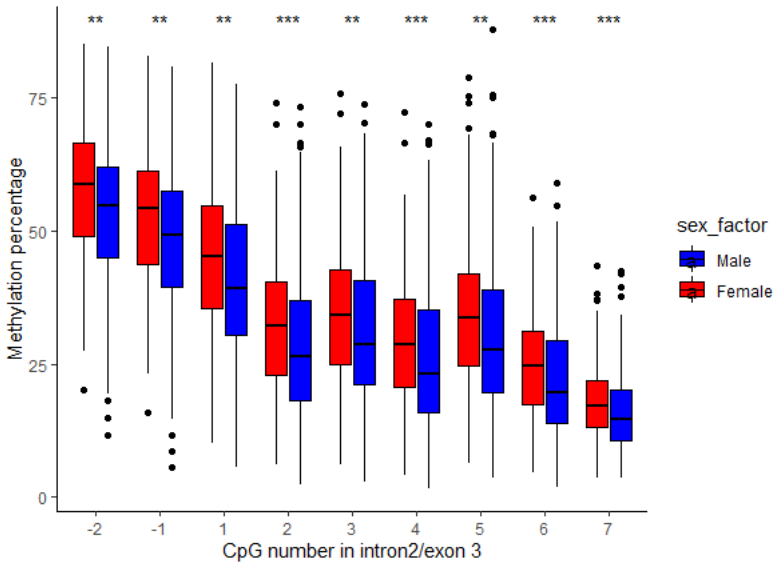


Figure 6.2 Percentage methylation across the 9 CpGs in the intron2/exon3 region of the POMC gene by Sex in mid childhood. P-value calculated by student t test. Key: p-value = **<0.01 ***<0.001

6.3.2 Age

Only the methylation measurement in mid-childhood (age between 5-8 years) was assessed as there was no difference in age at measurement for the 2 year methylation measure. Children's age was positively correlated with mean methylation z score (Spearman R = 0.093, p=0.05, see Figure 6.3). There was no association between mean *POMC* methylation z score and height (Spearman R = 0.058, p=0.26, n=442) suggesting that the age association was not driven by height.

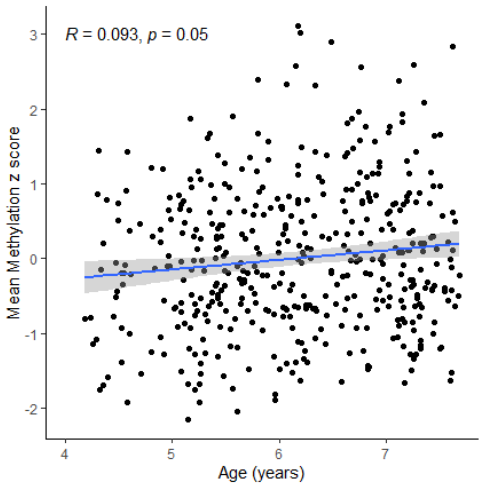


Figure 6.3. Scatterplot demonstrating the relationship between age and mean POMC methylation z score. Linear regression line with 95% CI shown. Spearman Correlation coefficient reported with p-value. N=442.

6.3.3 ENID supplementation

There was no difference in 2-year or mid-childhood mean *POMC* methylation z score between the pregnancy supplementation groups, see Figure 6.4.

There was no effect of infant multiple micronutrient supplementation on mean *POMC* methylation at age 2 (p=0.42) or in mid-childhood (p=0.088).

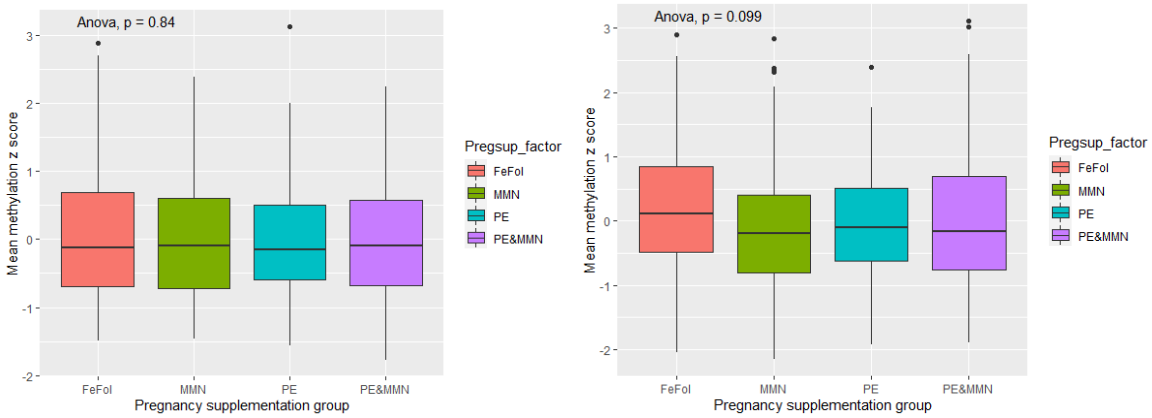


Figure 6.4 Boxplots comparing mean POMC methylation z score in 2 year (left, n=283) and mid-childhood (right, n=442) and mother's ENID pregnancy supplementation group. Group differences assessed by ANOVA.

6.3.4 Maternal BMI

There was no significant correlation between maternal periconceptional BMI and offspring's mean *POMC* methylation z score (see Figure 6.5).

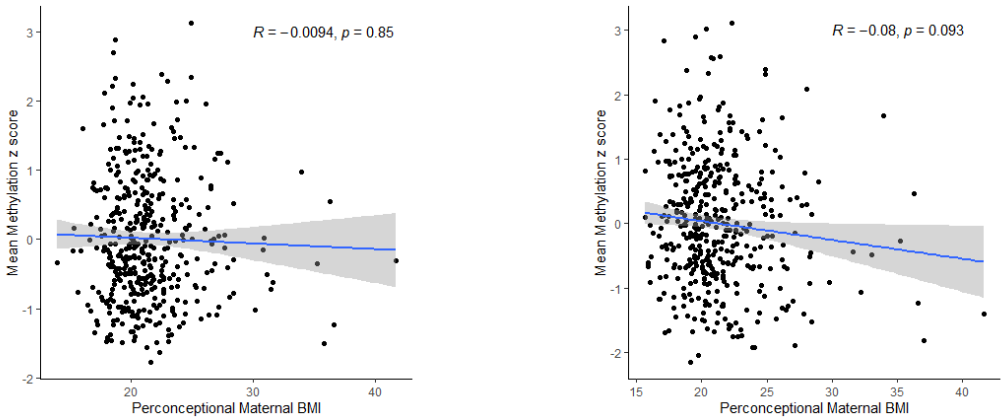


Figure 6.5 Scatterplot demonstrating the relationship between periconceptional maternal BMI and mean POMC methylation z score (left plot – methylation at age 2, n=283, right plot – methylation in mid-childhood, n=442). Linear regression line with 95% CI shown. Spearman correlation coefficient reported with p-value.

6.3.5 Season of conception

Firstly, to explore the effect of SoC on *POMC* methylation, SoC was dichotomised into rainy and dry season (see methods section). Percentage methylation at each CpG in the VMR was plotted by SoC, see Figure 6.6. There was a pattern of higher methylation in those conceived in the rainy season consistent with findings from previous [28,29,31,32,63].

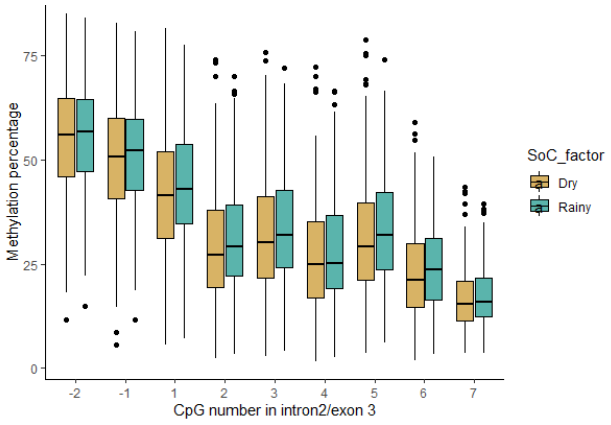


Figure 6.6 Percentage mid childhood POMC methylation by CpG across region by Season of Conception (SoC). SoC was defined as 'rainy' (January-June) and 'dry' (July-December) as previously described[54] with the conception date calculated from a gestational age estimation obtained from antenatal ultrasound at ENID trial recruitment

Mean *POMC* methylation z score was plotted by estimated date of conception. The loess curve plotted across the year demonstrated a trend for higher methylation in those conceived towards the end of the rainy season and lower methylation in those conceived towards the end of the dry season (see Figure 6.7).

To further explore the potential influence of SoC, regression models were fitted with Fourier terms to capture the seasonal fluctuations in *POMC* methylation. Models were fitted separately using the 2 year methylation measure and a mid-childhood methylation measure (adjusted for age) to assess the effect of SoC at two different points in childhood. Models with and without the Fourier terms were compared using a likelihood ratio test to assess if the model fitted better with Fourier included.

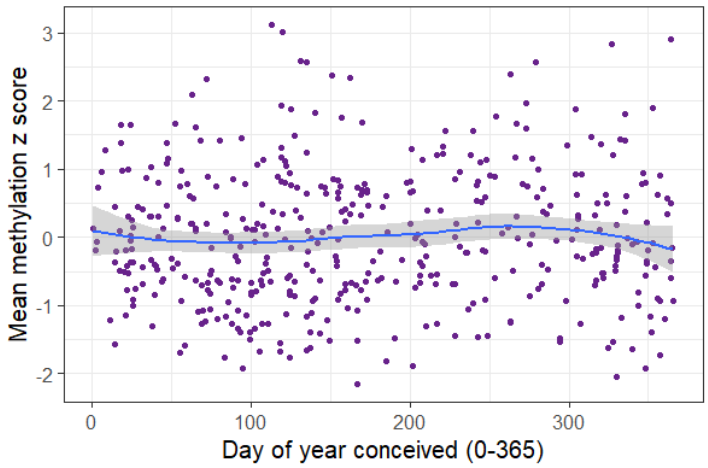


Figure 6.7 Mean POMC methylation z score plotted against day of year conceived. Loess curve plotted with 95% CI. Seasons are defined as conceptions 'dry' (January-June) and 'rainy' (July-December).

SoC association with 2 year methylation measure

There was evidence of a SoC effect (LRT p=0.019, see Table 6.1), driven by the sin Fourier term (sin coefficient = -0.18 (p=0.016). There was a peak of *POMC* methylation for conceptions in August to September (see Figure 6.8).

SoC association with mid-childhood methylation measure

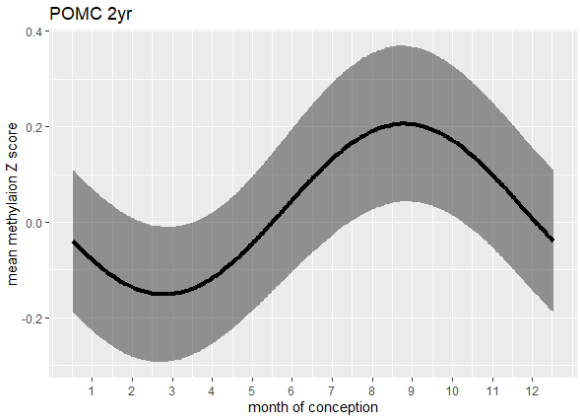
To assess if the seasonal effect on *POMC* methylation persisted beyond 2 years of age, the effect of SoC was examined again in mid-childhood. In regression models, adjusted for age, the seasonal effect was attenuated (see Table 6.2). The sin coefficient fell to -0.13 (compared to -0.18 at age 2 years of age) and there was no evidence of a significant SoC effect (LRT = 0.14).

	Dependent variable: **Age 2 years mean *POMC* methylation z score**	
	Model 1 with seasonal Fourier terms	**Model 2 without seasonal Fourier terms**
sin(doc.theta)	-0.18** (0.076)	
cos(doc.theta)	-0.080 (0.080)	
Observations	283	283
R^2	0.023	0.000
Note:	*p<0.1 **p<0.05 ***p<0.01	
Likelihood ratio test : **p-value=0.019.**		

Table 6.1 Linear regression model assessing SoC effect on POMC methylation at 2 years of age. Model 1: Model with Fourier terms. Model 2: Random intercept model with Fourier terms removed. Models compared using likelihood ratio test.

	Dependent variable: **Mid-childhood mean *POMC* methylation z score**	
	Model 1 with seasonal Fourier terms	**Model 2 without seasonal Fourier terms**
Age (years)	0.12** (0.051)	0.13** (0.051)
sin(doc.theta)	-0.13* (0.066)	
cos(doc.theta)	-0.026 (0.066)	
Observations	442	442
R^2	0.022	0.014
Note:	*p<0.1 **p<0.05 ***p<0.01	
Likelihood ratio test : **p-value = 0.14**		

Table 6.2 Linear regression model assessing SoC effect on POMC methylation in mid-childhood. Model 1: Model with Fourier terms. Model 2: Model with Fourier terms removed. Models compared using likelihood ratio test. All models were adjusted for age at POMC methylation measurement.



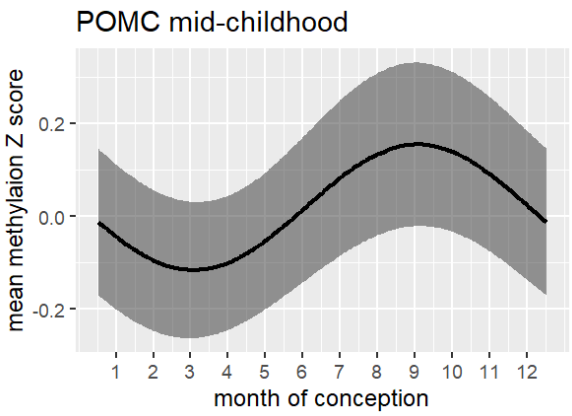


Figure 6.8. Model plot of Fourier regression model at age 2 years (left) and mid-childhood (right) demonstrating the seasonal change in POMC methylation across the year and a reduction in seasonal amplitude in mid-childhood. Black line is mean with 95% CI in grey.

6.3.6 Maternal nutritional biomarkers

Paired *POMC* methylation and maternal biomarker data were available for 186 (mid-childhood) and 166 (2 year).

There were no associations between mean *POMC* methylation z score and any one carbon metabolites at 2 years or in mid-childhood (see Tables 6.3 and 6.5 respectively).

Glycine and uracil were associated with mean *POMC* methylation at 2 years (see Table 6.4), such that for every unit SD increase in glycine there was an associated increase of 0.15 mean *POMC* methylation z score (p=0.03) and for every unit SD increase in uracil there was an associated -0.19 reduction in mean *POMC* methylation z score (p=0.02).

These effects were not seen in mid-childhood where aspartate was the only biomarker significantly associated with mean *POMC* methylation z score (see table 6, β=-0.20, p=0.0029).

One-carbon metabolites covary and are inter-dependent through complex interactions[34], therefore it is difficult to know if and how tests should be corrected for multiple testing. However, after a conservative Bonferroni correction for the number of biomarkers tested, none of these associations remained significant; glycine (p=0.88), uracil (p=0.50) and aspartate (p=0.08).

Dependent variable: Mean *POMC* methylation z score at age 2 years **Predictor variables: One-carbon metabolites**				
Covariate	**Standardised**	**Standard Error**	**t-value**	**p-value**
Hcy (µmol/L)	-0.01	0.07	-0.17	0.86
Methionine (µmol/L)	0.03	0.07	0.45	0.65
Cysteine (µmol/L)	-0.09	0.07	-1.23	0.22
Choline (µmol/L)	-0.01	0.07	-0.08	0.93
Betaine (µmol/L)	0.06	0.06	0.86	0.39
DMG (µmol/L)	0.01	0.07	0.11	0.92
B12 (pmol/L)	-0.06	0.07	-0.91	0.36
Folate (nmol/L)	-0.01	0.07	-0.12	0.91
PLP (nmol/L)	-0.08	0.08	-1.09	0.28
Riboflavin (nmol/L)	-0.07	0.07	-1.08	0.28

Table 6.3 Maternal one-carbon metabolite biomarkers as predictors of age 2 year mean POMC methylation z score. Multiple linear regression models have mean POMC methylation z-score as the dependent variable with maternal one-carbon metabolites measured in maternal plasma and back-extrapolated to the time of conception as predictors. **Key**: Hcy = Homocysteine, DMG = dimethylglycine, PLP = Pyridoxal 5-phosphate (B6 vitamer)

Dependent variable: Mean *POMC* methylation z score at age 2 years **Predictor variable: Other biomarkers including amino acids**				
Covariate	**Standardised Coefficient**	**Standard Error**	**t-value**	**p-value**
AGP (g/L)	0.03	0.07	0.39	0.70
Aspartate (µmol/L)	-0.10	0.07	-1.48	0.14
Threonine (µmol/L)	-0.08	0.07	-1.12	0.26
Serine (µmol/L)	0.04	0.07	0.55	0.58
Glutamate (µmol/L)	-0.02	0.08	-0.30	0.77
Glycine (µmol/L)	**0.15**	**0.07**	**2.19**	**0.03**
Alanine (µmol/L)	0.11	0.07	1.46	0.15
Valine (µmol/L)	0.00	0.07	0.04	0.97
Isoleucine (µmol/L)	0.06	0.07	0.84	0.40
Leucine (µmol/L)	0.03	0.07	0.50	0.62
Tyrosine (µmol/L)	0.04	0.07	0.57	0.57
Phenylalanine (µmol/L)	0.06	0.07	0.78	0.44
Lysine (µmol/L)	-0.04	0.07	-0.59	0.56
Histidine (µmol/L)	-0.07	0.07	-1.10	0.27
Arginine (µmol/L)	0.09	0.07	1.30	0.20
Proline (µmol/L)	0.06	0.07	0.80	0.43
Uridine (µmol/L)	0.03	0.07	0.45	0.65
Uracil (nmol/L)	**-0.19**	**0.08**	**-2.41**	**0.02**

Table 6.4. Linear regression models for additional nutritional predictors of age 2 year mean POMC methylation z score. Multiple linear regression models have mean POMC methylation z-score as the dependent variable with maternal one-carbon metabolites measured in maternal plasma and back-extrapolated to the time of conception as predictors. . **Key**: AGP = Alpha-1-acid glycoprotein.

Dependent variable: Mean *POMC* methylation z score in mid-childhood **Predictor variables: One-carbon metabolites**				
Covariate	**Standardised Coefficient**	**Standard Error**	**t-value**	**p-value**
Hcy (µmol/L)	0.02	0.08	0.29	0.77
Methionine (µmol/L)	-0.01	0.07	-0.21	0.84
Cysteine (µmol/L)	-0.02	0.08	-0.29	0.78
Choline (µmol/L)	-0.03	0.07	-0.48	0.63
Betaine (µmol/L)	0.06	0.07	0.91	0.37
DMG (µmol/L)	0.00	0.08	-0.04	0.97
B12 (pmol/L)	-0.04	0.07	-0.59	0.56
Folate (nmol/L)	-0.02	0.07	-0.29	0.77
PLP (nmol/L)	-0.07	0.08	-0.89	0.37
Riboflavin (nmol/L)	-0.08	0.07	-1.12	0.26

Table 6.5. Maternal one-carbon metabolite biomarkers as predictors of mid-childhood mean POMC methylation z score. Multiple linear regression models have mean POMC methylation z-score as the dependent variable with maternal one-carbon metabolites measured in maternal plasma and back-extrapolated to the time of conception as predictors. All models adjusted for age of POMC methylation measurement. Key: Hcy = Homocysteine, DMG = dimethylglycine, PLP = Pyridoxal 5-phosphate (B6 vitamer)

Dependent variable: Mean *POMC* methylation z score in mid-childhood **Predictor variables: Other biomarkers including amino acids**				
Covariate	**Standardised**	**Standard Error**	**t-value**	**p-value**
AGP (g/L)	-0.02	0.08	-0.32	0.75
Aspartate (µmol/L)	**-0.20**	**0.07**	**-3.02**	**0.0029**
Threonine (µmol/L)	-0.04	0.07	-0.56	0.58
Serine (µmol/L)	-0.04	0.07	-0.53	0.60
Glutamate (µmol/L)	-0.02	0.08	-0.28	0.78
Glycine (µmol/L)	0.00	0.07	-0.03	0.98
Alanine (µmol/L)	0.02	0.07	0.28	0.78
Valine (µmol/L)	-0.02	0.07	-0.23	0.82
Isoleucine (µmol/L)	-0.01	0.07	-0.20	0.84
Leucine (µmol/L)	0.02	0.07	0.27	0.79
Tyrosine (µmol/L)	0.04	0.07	0.57	0.57
Phenylalanine (µmol/L)	0.03	0.07	0.38	0.70
Lysine (µmol/L)	0.00	0.07	0.05	0.96
Histidine (µmol/L)	-0.03	0.07	-0.50	0.62
Arginine (µmol/L)	0.12	0.07	1.70	0.09
Proline (µmol/L)	0.03	0.07	0.41	0.68
Uridine (µmol/L)	0.04	0.07	0.57	0.57
Uracil (nmol/L)	-0.11	0.08	-1.35	0.18

Table 6.6. Linear regression models for additional nutritional predictors of mid-childhood mean POMC methylation z score. Multiple linear regression models have mean POMC methylation z-score as the dependent variable with maternal one-carbon metabolites measured in maternal plasma and back-extrapolated to the time of conception as predictors. All models adjusted for age of POMC methylation measurement. **Key:** AGP = Alpha-1-acid glycoprotein.

6.3.7 Genotype

Four independent SNPs (r^2<0.8) in cis within 10kb of the *POMC* gene were selected for an assessment of their relationship with mean *POMC* methylation z score (see Methods and Table 6.7). Four hundred and fifty participants had genotype data and *POMC* methylation measurement in mid-childhood. The mid-childhood *POMC* methylation measure was used to assess the influence of genotype data due to the larger number of participants with methylation measurements.

To assess if SoC-methylation effect was potentially confounded by genotype SoC-Genotype associations were tested using a 2x2 table and chi-squared test.

SNP	Genotype frequency (count) [All*]	Genotype frequency (count) [Gambian in Western Division, Gambia#]
rs11892647	**C|C: 0.423 (1060)** **C|T: 0.439 (1100)** **T|T: 0.137 (344)**	**C|C: 0.279 (141)** **C|T: 0.483 (244)** **T|T: 0.238 (120)**
rs6751851	**G|G: 0.665 (1665)** **A|A: 0.051 (127)** **A|G: 0.284 (712)**	**G|G: 0.574 (290)** **A|A: 0.075 (38)** **A|G: 0.350 (177)**
rs6713532	**T|T: 0.304 (761)** **C|C: 0.240 (602)** **C|T: 0.456 (1141)**	**T|T: 0.244 (123)** **C|C: 0.275 (139)** **C|T: 0.481 (243)**
rs6545976	**G|G: 0.659 (1650)** **G|T: 0.265 (664)** **T|T: 0.076 (190)**	**G|G: 0.271 (137)** **G|T: 0.497 (251)** **T|T: 0.232 (117)**

Table 6.7 SNP genotype frequency. *Relates to all phase three individuals from 1000 Genomes project, #Relates to all participants in Gambian Genome Variation Project. Data taken from Ensembl database[64]. SNPs were chr2:25156150 (hg38) or rs11892647 (rs SNP ID), chr2:25158279 or rs6751851, chr2:25161964 or rs6713532 and chr2:25171781 or rs6545976 and mapped to hg38 genome.

SNP rs11892647

One participant had missing genotype data. The most common genotype at this SNP was T/C (n=214), followed by T/T (n=123) and C/C (n=112). Genotype T/T was associated with a significantly higher mean *POMC* methylation z score compared to T/C (p=0.012) and C/C (p=8.7 $x10^{-14}$) and genotype T/C had a significantly higher mean *POMC* methylation z score compared to C/C (p=3.3$x10^{-8}$), see figure 6.8. The variability of POMC methylation appeared similar across all genotypes. This SNP therefore demonstrates the properties of a canonical methylation quantitative trait loci (mQTL).

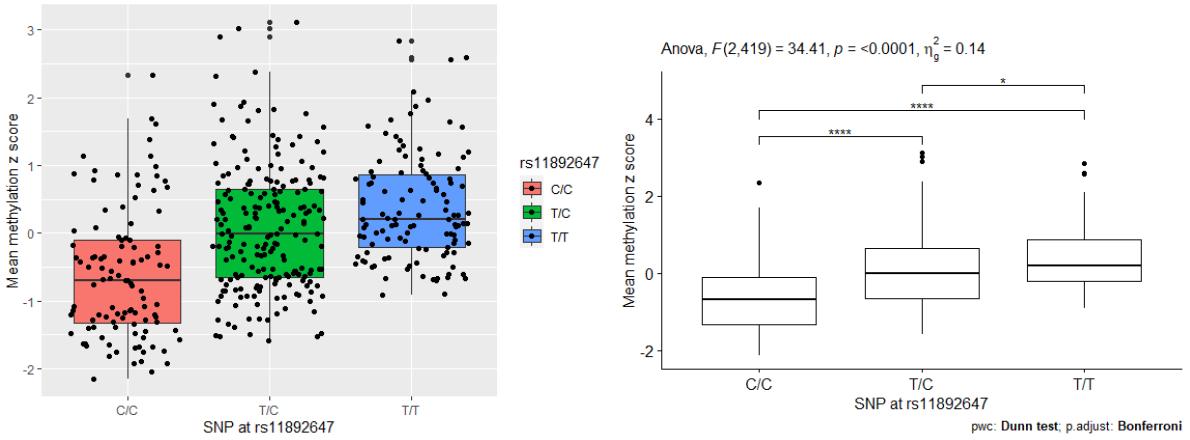


Figure 6.8. Boxplot of mean POMC methylation z score by genotype at SNP rs11892647. Distribution of individual methylation measurement by genotype is shown in left plot (black dots over box plot). Differences between groups were assessed by ANOVA test, post hoc analysis by Dunn test with Bonferroni correction (right plot). * p-value ≤ 0.05, ** p-value ≤0.01, *** p-value ≤0.001.

There was no association between SoC and genotype at rs11892647 (p =0.65, table 6.8) suggesting that genotype was not confounding the SoC effects reported in section 6.3.5.

SoC	**Genotype at rs11892647**		
	C/C	**T/C**	**T/T**
Dry	69	120	72
Rainy	43	92	55
Chi squared test, p-value = 0.65			

Table 6.8 . 2x3 Contingency table of SoC by Genotype at rs11892647. p-value calculated by Chi Squared test.

SNP rs6751851

The most common genotype at this loci was G/G (n=265), followed by G/A (n=155) and A/A (n=30). Genotype G/G had significantly higher mean *POMC* methylation z score compared to G/A ($p=3.7 \times 10^{-12}$) and A/A ($p=7.2 \times 10^{-15}$) genotypes. G/A had significantly higher methylation compared to A/A genotype ($p=8.6 \times 10^{-5}$), see figure 6.9. The variability of *POMC* methylation appeared similar across all genotypes. This SNP therefore demonstrated the properties of a canonical methylation quantitative trait loci (mQTL).

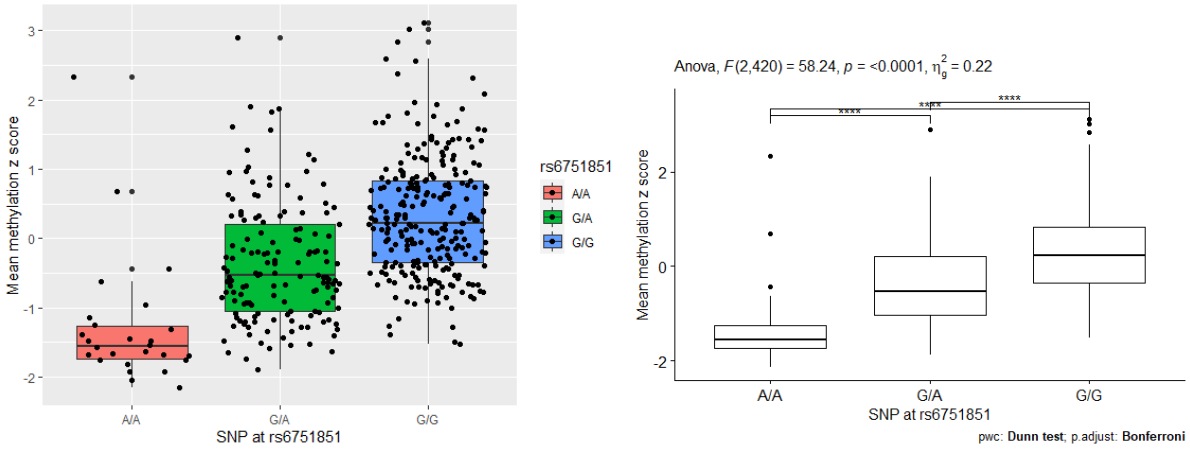


Figure 6.9. Boxplot of mean POMC methylation z score by genotype at SNP rs6751851. Distribution of individual methylation measurement by genotype is shown in left plot (black dots over box plot). Differences between groups were assessed by ANOVA test, post hoc analysis by Dunn test with Bonferroni correction (right plot). * p-value ≤ 0.05, ** p-value ≤0.01, *** p-value ≤0.001.

There was no association between SoC and genotype at rs6751851 (p=0.76, Table 6.9) suggesting that genotype was not confounding the SoC effects reported in section 6.3.5.

SoC	**Genotype at rs6751851**		
	A/A	**G/A**	**G/G**
Dry	18	82	162
Rainy	12	73	105
Chi squared test, p-value = 0.76			

Table 6.9 2x3 Contingency table of SoC by Genotype at rs6751851. P-value calculated by Chi Squared test.

SNP rs6713532

The most common genotype at this SNP was C/T (n=214), followed by T/T (n=119) and C/C (n=117). Genotype T/T had significantly higher mean POMC methylation z score compared to C/T ($p=1.8 \times 10^{-3}$) and C/C ($p=1.3 \times 10^{-8}$) genotypes. C/T had significantly higher methylation compared to genotype C/C (3.5×10^{-3}), see figure 6.10. The variability of *POMC* methylation appeared similar across all genotypes. This SNP therefore demonstrated the properties of a canonical methylation quantitative trait loci (mQTL).

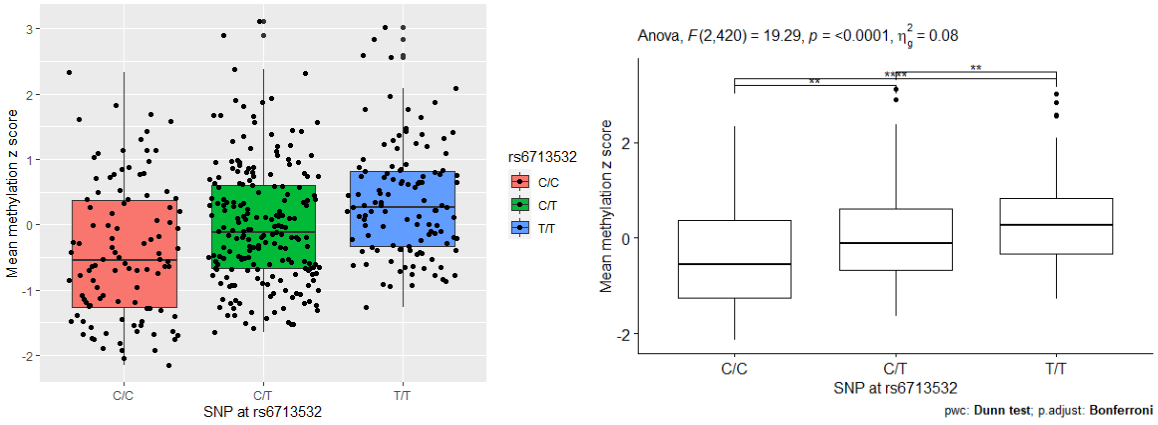


Figure 6.10. Boxplot of mean POMC methylation z score by genotype at SNP rs6713532. Distribution of individual methylation measurement by genotype is shown in left plot (black dots over box plot). Differences between groups were assessed by ANOVA test, post hoc analysis by Dunn test with Bonferroni correction (right plot). * p-value ≤ 0.05, ** p-value ≤0.01, *** p-value ≤0.001.

There was no association between SoC and genotype at rs6713532 (p=0.43, Table 6.10) suggesting that genotype was not confounding the SoC effects reported in section 6.3.5.

SoC	**Genotype at rs6713532**		
	C/C	**C/T**	**T/T**
Dry	73	118	71
Rainy	44	96	50
Chi squared test, p-value = 0.43			

Table 6.10 2x3 Contingency table of SoC by Genotype at rs6713532 . P-value calculated by Chi Squared test.

SNP rs6545976

One participant had missing genotype data. The most common genotype at this SNP was G/T (n=205), followed by G/G (n=137) and T/T (n=107). Genotype G/G was associated with significantly higher mean *POMC* methylation z score compared to genotypes G/T ($p=1.37 \times 10^{-4}$) and T/T ($p=6.5 \times 10^{-8}$), see figure 6.14. The variability of *POMC* methylation appeared similar across all genotypes. This SNP genotype G/G demonstrated the properties of a canonical methylation quantitative trait loci (mQTL).

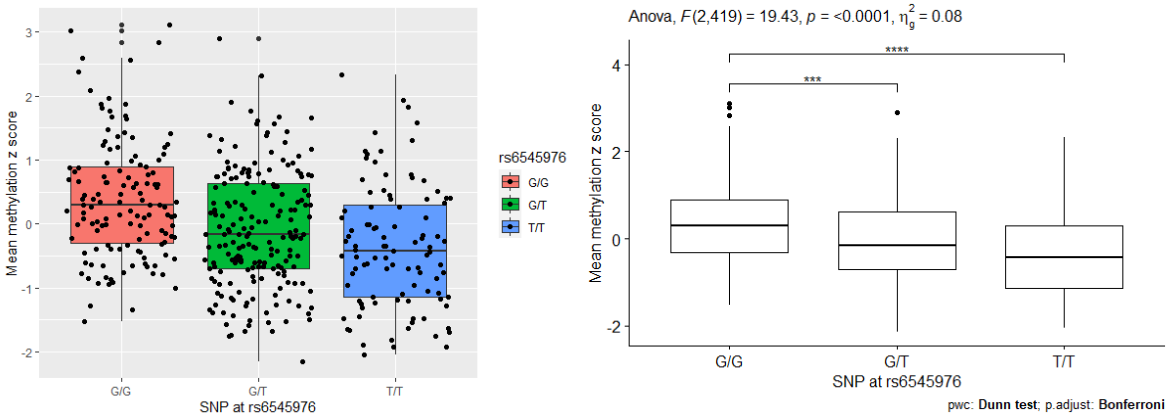


Figure 6.11. Boxplot of mean POMC methylation z score by genotype at SNP rs6545976. Distribution of individual methylation measurement by genotype is shown in left plot (black dots over box plot). Differences between groups were assessed by ANOVA test, post hoc analysis by Dunn test with Bonferroni correction (right plot). * p-value ≤ 0.05, ** p-value ≤0.01, *** p-value ≤0.001.

There was no association between SoC and genotype at rs6545976 (p=0.76, Table 6.11) suggesting that genotype was not confounding the SoC effects reported in section 6.3.5.

SoC	**Genotype at rs6545976**		
	G/G	**G/T**	**T/T**
Dry	84	116	61
Rainy	55	88	47
Chi squared test, p-value = 0.76			

Table 6.11 2x3 Contingency table of SoC by Genotype at rs6545976. P-value calculated by Chi Squared test.

6.3.8 Intergenerational associations

99 fathers (mean age (SD, range) = 53.1 years (10.5, 30.3 to 75.8) were recruited to the study with data available from 95 complete family trios. Complete family trios were used to ensure equitable comparison between mothers and fathers. There was a significant correlation between both mother's and father's mean *POMC* methylation z score and their offspring's. Mothers had a similar correlation coefficient (Pearson coefficient = 0.26, p=0.01) to fathers (Pearson coefficient = 0.24, p=0.018), see Figure 6.12.

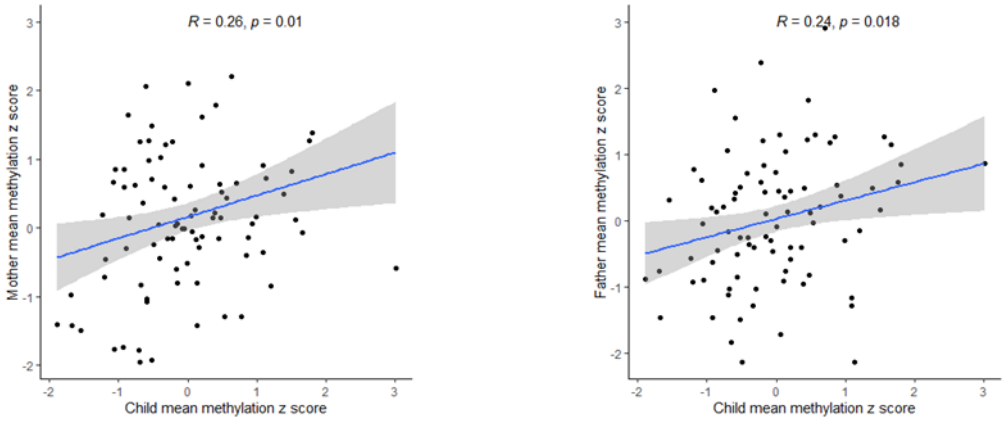


Figure 6.12 Scatterplots demonstrating the relationship between parent and child mean POMC methylation z score (left: Mother – Child; right: Father – Child). Linear regression line with 95% CI shown. Pearson's correlation coefficient reported with p-value.

6.3.9 Stability though childhood

265 children had paired *POMC* methylation measurements aged 2 and in mid-childhood. Mean *POMC* methylation z scores were significantly correlated between these two time points (Pearson R = 0.45, p = $1x10^{-12}$) demonstrating stability through early childhood (see Figure 6.16).

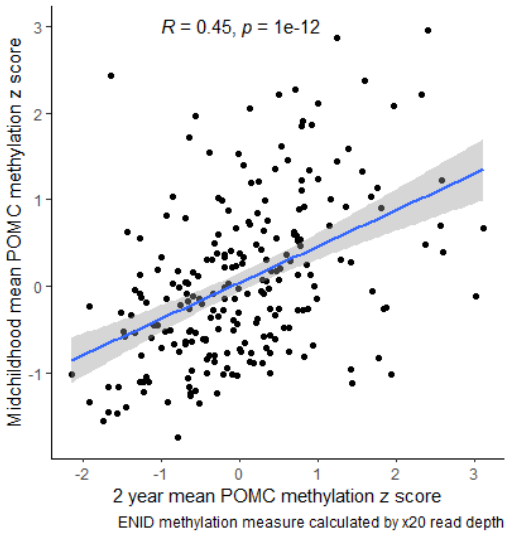


Figure 6.13. Scatterplot demonstrating the relationship between mean POMC methylation z score at age 2 years and in mid-childhood. Linear regression line with 95% CI shown. Pearson correlation coefficient reported with p-value.

6.4 Discussion

The early embryonic period is thought to represent a key window of developmental plasticity whereby environmental exposures can influence the developing offspring's DNA methylation[4–6,16,19,36]. The results presented in chapter 6 present a number of findings that point towards a periconceptional environmental effect on *POMC* methylation.

SoC was significantly associated with *POMC* methylation at age 2 years with a peak of mean methylation between August-September and February-March seeing the lowest mean methylation. The SoC association with *POMC* methylation has been demonstrated before in Gambian infants (aged between 2-8 months, mean age 3.6 months)[34]. In this earlier study, SoC was dichotomised into dry and rainy seasons, with dry season conceptions showing a 15.2% lower methylation compared to wet season conceptions. In the *POMC* study, data from children conceived throughout the year (not limited to peaks of season as with previous studies in this population[28,31,32]) was analysed and used regression models with Fourier terms to model conceptions across the year. This approach makes no assumptions about where potential peaks or nadirs of methylation ay occur. The finding of higher methylation during the rainy season is consistent with previous studies in The Gambia[28,31,32]. The timings of *POMC* methylation peak and nadirs from the *POMC* study coincide with seasonally driven methylation patterns seen across a 259 SoC sensitive CpGs recently

reported[65]; a peak between August-September (corresponding with the peak of The Gambian rainy season) and a nadir between February-March. A SoC – *POMC* methylation association was only statistically significant in the 2 year old children and not the children in mid-childhood. This suggests SoC association is attenuated through early childhood and has been reported previously where children aged 8-9 years showed an reduced amplitude of the SoC effect compared to ENID children at age 2 years[62]. This finding may suggest further epigenetic change through childhood. Longitudinal analysis of individual's *POMC* methylation at age 2 and mid-childhood showed that methylation was significantly correlated (see Figure 6.13) and aligned with similar findings in German children from birth to early adolescence[34].

Understanding the causes of any seasonal effect on methylation is a challenge. The Gambian rainy season is associated with changes in energy balance (brought about by nutritional scarcity and increased farm work), infection disease burden and dietary constituents[54,66,67]. Previous work in The Gambia has identified seasonal fluctuations in i) one-carbon metabolite intake such as riboflavin, folate, choline and betaine and ii) dietary constituents of protein, fat and carbohydrate. Furthermore, seasonal fluctuations of circulating levels of B6 and folate were identified in women of childbearing age[66]. Associations between one carbon metabolites concentrations and offspring DNA methylation at MEs have been reported[28,29]. In the *POMC* study, there was no evidence of an association between offspring *POMC* methylation and any circulating level of periconceptional one-carbon metabolite. A previous study found periconceptional levels of betaine (positively associated), SAM and SAM:SAH ratio (negatively associated) predicted offspring *POMC* methylation[34] though these associations were not reproduced in these data. Unfortunately, SAM and SAM:SAH ratio could not be measured in the plasma from the MDEG2 study (where periconceptional biomarkers derived from) due to sample degradation. However, significant associations between 2 year *POMC* methylation and periconceptional levels of the amino acids glycine (positively associated) and uracil (negatively associated) were found. Amino acids and one-carbon metabolites are bidirectionally linked: amino acids can provide cells with one carbon metabolites[68] and levels of one carbon metabolites can influence amino acid metabolism[69]. Glycine acts as the substrate for Glycine N-methyltransferase (GNMT) which is an enzyme that catalyses the methylation of glycine with SAM to produce SAH and sarcosine. Glycine and GNMT therefore have important functions in regulating the levels of SAM and SAM:SAH ratio and therefore influence DNA methylation potential[70,71]. Uracil was also associated with *POMC* methylation, though its role in influencing DNA methylation remains unclear. Uracil-DNA glycosylase (UNG) enzyme is used to excise uracil from the DNA strand[72]. However, there is evidence in a murine model that this enzyme may also be involved in Tet (ten eleven translocation enzyme)-mediated DNA demethylation[73]. One explanation for the negative

association between uracil and *POMC* methylation could be if plasma levels of uracil were correlated with UNG activity. Thus, higher uracil levels would be associated with higher UNG activity and an increase in offspring DNA demethylation. Further work is needed to elucidate any biological pathways.

Previous related studies have not corrected for multiple tests as many of the biomarkers considered here are likely to covary. It should be noted that no biomarker reached significance after correction for multiple tests in this study.

Maternal periconceptional BMI showed no association with offspring *POMC* methylation which is consistent with previous work[34]. From these data, periconceptional BMI clustered around a BMI of 20 with relatively few women in the obese category (BMI >30). This narrow range of BMI may explain why a BMI effect on *POMC* methylation was not seen. In settings where obesity prevalence is high, there may be an association with offspring *POMC* methylation and this cannot be excluded from findings from this study in this population. Maternal BMI has been shown to influence offspring DNA methylation in previous studies[12–15] though not at the POMC VMR. Offspring DNA methylation profiles from pregnancies before and after bariatric surgery showed widespread between sibling differences [13]. Sharp and colleagues, using data from the ALSPAC (Avon Longitudinal Study of Parents and Children) cohort, demonstrated that differential methylation was seen in cord blood at numerous CpGs in offspring born to mothers at the extremes of BMI (obese and underweight)[12] with underweight having a larger effect. Furthermore, associations between cord methylation and maternal obesity far outweighed associations between cord methylation and paternal obesity suggesting the role for the intrauterine environment. Meta-analysed data of 9340 mother-child pairs, across 19 cohorts from diverse ethnic groups examining the association of maternal BMI and offspring's DNA methylation suggests that the effect of maternal BMI on offspring's DNA methylation may have previously been overstated[74] once cell composition is adjusted for.

There was no evidence of a difference in *POMC* methylation across all ENID pregnancy and infant supplementation arms. This is consistent with findings from a multicentre (Gambia and India) pregnancy supplementation trial (EMPHASIS)[75], where offspring's *POMC* methylation showed no association with the supplement intervention arm.

There is evidence of a relationship between inflammatory processes and adult DNA methylation from the literature[76]. A recent meta-analysis in adults showed that DNA methylation at 218 CpGs were associated with c-reactive protein (CRP) inflammatory marker[77]. Furthermore, there may be a complex relationship between inflammatory processes and DNA methylation as shown in a recent study exploring methylation associations with over 160 different inflammatory markers[76]. To

investigate the possible influence of inflammation and infection on offspring methylation alpha-1-acid glycoprotein (AGP) was measured in the periconceptional maternal blood sample. AGP is an acute phase reactant and a marker of inflammation[78]. There was no association between *POMC* methylation and AGP. Obesity is associated with increased levels of inflammatory markers such as TNFα, interleukin 6 and CRP. This study only examined AGP and therefore has limited scope to explore potential maternal inflammation – offspring DNA methylation associations. Adiposity driven alterations in inflammatory markers may explain some of the BMI effects on offspring DNA methylation reported in other studies and would warrant further study. Exploring associations between offspring DNA methylation and other maternal inflammatory biomarkers (e.g. TNFα, interleukin 6 and CRP) would be of interest for future study.

There are other seasonal factors that have been linked to epigenetic alterations that were not explored in this study. These include circulating levels of vitamin D known to influence the epigenome on multiple levels[79], vitamin A (shown to influence offspring *POMC* methylation in rats[80]), prenatal stress (shown influence offspring *POMC* methylation in rats[81]) and heat stress[82].

Sex dimorphism in *POMC* methylation was reported in this study with females having higher methylation than males. There have been reported sex-effects on ME methylation[28] and indeed this was also seen at the *PAX8* ME as reported in chapter 7. *PAX8* methylation, in a similar way to *POMC* methylation, was higher in girls compared to boys. In Bangladesh, colleagues have observed a similar effect with methylation in the *POMC* VMR found to be 4.8% higher in girls compared to boys([83], unpublished). Sex dimorphism in hypothalamic feeding circuits have been described in animal models[84], raising the possibility that sex differences in methylation of orexigenic genes such as *POMC* could offer an explanation for the observed differences.

Four SNPs in cis demonstrated an independent mQTL effect on *POMC* methylation i.e. SNPs that effect mean DNA methylation levels. Previously, Kühnen and colleagues explored the influence of a single POMC associated SNP (rs713586) associated with BMI[85] and found no effect on DNA methylation[34]. The *POMC* study goes further and explores all local genetic variants (<10kb of the gene). Though MEs are sensitive to the periconceptional environment, methylation at these loci have also been shown to be associated with genotype in humans[65,86]. Silver and colleagues reported half of SoC-CpGs had at least one mQTL with 92% occurring in *cis.* An advantage of the utilising the Gambian seasonality experimental model is that conceptions are randomised to occur at different timepoints across the year. For genetic variants to confound the SoC effect on *POMC* methylation, these variants would need to be either associated with seasonal timing of conception or to influence survival of the developing embryo/foetus under a particular seasonal stress. None of the 4 SNPs

analysed were associated with SoC. A previous study from the Dutch Hunger Famine cohort suggested that both prenatal famine and genetic variation influenced methylation at the imprinted *IGF2/H19* region and proposed that their effects were additive[87]. More recently, a genome-wide analysis of methylation identified that those loci with the greatest DNA methylation variability are influenced predominantly by either genotype and prenatal environment interactions or additive effects of genotype and environment[88]. There could be other SNPs in *cis* or *trans* that influence methylation. To characterise the genetic influence further would require greater power and to look genome wide.

A review of phenotypic data associated with the four SNPs explored in the *POMC* study yielded some interesting results. These studies did not examine any genotype-methylation effects so the phenotypic associations described relate to the genotype not methylation. SNP rs6545976, an mQTL for *POMC* VMR, has been associated with substance misuse[89]. Recent studies have implicated a role for appetite regulating neuropeptides (including POMC) in alcohol dependence (AD) and craving[90] and *POMC* methylation has been associated with craving in alcohol dependence[91,92] and higher in those with AD compared to controls[93]. A candidate gene study of BMI in black south African adolescents found no association with BMI and genotype at SNP rs6713532 (an mQTL identified in the *POMC* study)[94].

Intergenerational epigenetic inheritance represents the transmission of epigenetic marks, such as DNA methylation, from one generation to the next. The mode of transmission i.e. genetic (DNA-based) transmission or epigenetic mechanisms of transmission (e.g. micro RNA, or environmentally sensitive mechanisms) has wide implications for human health[95]. It is also important that paternal influence on DNA methylation is investigated[96]. A previous study found that offspring *POMC* methylation was correlated with the father's and not the mother's *POMC* methylation[34]. The *POMC* study reproduced the finding of a relationship between father-child *POMC* methylation yet in a disparate ethnic group. The family trios reported by Kühnen and colleagues were from Germany and my study reports trios from The Gambia[34]. However, I observed a similar effect between mother and child while Kühnen and colleagues did not. The *POMC* study presents data from a larger number of trios (n=95 trios) compared to what was reported by Kühnen and colleagues (n=47 trios) and therefore may have greater power to detect an association between mothers and children. Inheritance of *POMC-related* mQTLs from parents to child may explain some of the methylation similarities. Only children were genotyped in this study and therefore assessment of parent-child genotype similarities could not be assessed.

There is provisional evidence from animal models of foetal alcohol exposure, which suggest epigenetic marks related to *POMC* may be transmitted across generations mediated by the paternal line. Govorko and colleagues[48], explored this concept in a rat model by establishing two germlines; 1) breeding males foetal alcohol exposed rats and their male offspring with non-exposed females and 2) breeding female foetal alcohol exposed rats and their female offspring with normal males. Hypermethylation of the *POMC* promoter with associated reduced POMC expression were seen in both female and male offspring in F1 generation, but this pattern continued in male progeny in F2 and F3, thus demonstrating an apparent transgenerational effect from the male germline only. An assessment of transgenerational epigenetic inheritance is beyond the scope of the *POMC* study as only F0 and F1 generations were recruited. Studies of larger multigenerational family pedigrees would help aid understanding of familial patterns of *POMC* DNA methylation.

Govorko and colleagues[48] reported *POMC* promoter methylation was higher in sperm of male rats (F1-F3) from the male exposed germline suggesting a possible mechanism of epigenetic inheritance via methylation differences in the sperm. However, examining human sperm from 17 males from the 47 German family trios, Kühnen and colleagues[34] found lower DNA methylation in sperm compared to blood consistent with the erasure of epigenetics marks as part of germline differentiation. The potential for passage of epigenetic marks from father to child is of great scientific interest though the exact mechanisms are yet to be elucidated[97]. There is emerging evidence that epigenetic processes in spermatozoa are responsive to environmental factors, some of which may influence the human offspring's epigenetic landscape and phenotype[98]. Sperm RNAs could possibly be a conduit for acquired epigenetic inheritance in sperm as they escape epigenetic germ line reprogramming[97]. A study of 13 lean and 10 obese men demonstrated differential expression of piwi-interacting RNA (piRNA)[40] and identified differences sperm DNA methylation before and after bariatric surgery suggesting that these epigenetic processes are sensitive the father's nutrition and body composition. In the *POMC* study, *POMC* methylation was correlated with both fathers and mothers. If intergenerational epigenetic inheritance does occur in humans at the *POMC* VMR, it would be important to elucidate the mechanisms involved. If non-genetic processes are occurring and these were under environmental influence, this could have wide implications for human health with potential for intergenerational interventions in both mothers *and* fathers to mitigate health risks e.g. obesity for the next generation.

6.5 Conclusion

This study provides evidence that *POMC* methylation is associated to both environmental and genetic factors. Using Fourier terms, SoC effects on *POMC* methylation have been further characterised beyond a binary dry vs. rainy season analysis. Previously identified associations between *POMC* methylation and betaine have not been replicated though novel associations with certain amino acids have been identified. Identified POMC mQTLs were not associated with SoC, suggesting that they do not confound the observed association between SoC and methylation. Parent-child *POMC* methylation levels were correlated confirming a relationship between father and child methylation reported in a previous German family study but further identifying a novel finding of a mother – child correlation.

6.6 References

1. Roseboom T, de Rooij S, Painter R. The Dutch famine and its long-term consequences for adult health. *Early Hum Dev*. 2006;82(8):485-491. doi:10.1016/j.earlhumdev.2006.07.001

2. Moore SE. Early-Life Nutritional Programming of Health and Disease in the Gambia. *Ann Nutr Metab*. 2017;70(3):179-183. doi:10.1159/000456555

3. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr*. 2007;27:363-388. doi:10.1146/annurev.nutr.27.061406.093705

4. Heard E, Martienssen RA. Transgenerational Epigenetic Inheritance : Myths and Mechanisms. *Cell*. 2014;157(1):95-109. doi:10.1016/j.cell.2014.02.045

5. Cavalli G, Heard E. Advances in epigenetics link genetics to the environment and disease. *Nature*. 2019;571(7766):489-499. doi:10.1038/s41586-019-1411-0

6. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. *Nat Rev Genet*. 2007;8(4):253-262. doi:10.1038/nrg2045

7. Chen P, Piaggi P, Traurig M, et al. Differential methylation of genes in individuals exposed to maternal diabetes in utero. *Diabetologia*. 2017;60(4):645-655. doi:10.1007/s00125-016-4203-1

8. Bouchard L, Thibault S, Guay S-P, et al. Leptin Gene Epigenetic Adaptation to Impaired

Glucose Metabolism During Pregnancy. *Diabetes Care*. 2010;33(11):2436-2441. doi:10.2337/dc10-1024

9. Lesseur C, Armstrong DA, Paquette AG, Li Z, Padbury JF, Marsit CJ. Maternal obesity and gestational diabetes are associated with placental leptin DNA methylation. *Am J Obstet Gynecol*. 2014;211(6):654.e1-9. doi:10.1016/j.ajog.2014.06.037

10. Yang I V., Zhang W, Davidson EJ, Fingerlin TE, Kechris K, Dabelea D. Epigenetic marks of in utero exposure to gestational diabetes and childhood adiposity outcomes: the EPOCH study. *Diabet Med*. 2018;35(5):612-620. doi:10.1111/dme.13604

11. Hjort L, Martino D, Grunnet LG, et al. Gestational diabetes and maternal obesity are associated with epigenome-wide methylation changes in children. *JCI Insight*. 2018;3(17):1-14. doi:10.1172/jci.insight.122572

12. Sharp GC, Lawlor DA, Richmond RC, et al. Maternal pre-pregnancy BMI and gestational weight gain , offspring DNA methylation and later offspring adiposity : findings from the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2015;44(4):1288-1304. doi:10.1093/ije/dyv042

13. Guenard F, Deshaies Y, Cianflone K, Kral JG, Marceau P, Vohl M-C. Differential methylation in glucoregulatory genes of offspring born before vs. after maternal gastrointestinal bypass surgery. *Proc Natl Acad Sci*. 2013;110(28):11439-11444. doi:10.1073/pnas.1216959110

14. Berglind D, Müller P, Willmer M, et al. Differential methylation in inflammation and type 2 diabetes genes in siblings born before and after maternal bariatric surgery. *Obesity*. 2016;24(1):250-261. doi:10.1002/oby.21340

15. Khouja JN, Simpkin AJ, O'Keeffe LM, et al. Epigenetic gestational age acceleration: A prospective cohort study investigating associations with familial, sociodemographic and birth characteristics. *Clin Epigenetics*. 2018;10(1):1-9. doi:10.1186/s13148-018-0520-1

16. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci U S A*. 2008;105(44):17046-17049. doi:10.1073/pnas.0806560105

17. Tobi EW, Slieker RC, Stein AD, et al. Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. *Int J Epidemiol*. 2015;44(4):1211-1223. doi:10.1093/ije/dyv043

18. Shen L, Li C, Wang Z, et al. 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*. 2019;11(1):88. doi:10.1186/s13148-019-0676-3

19. Tobi EW, Lumey LH, Talens RP, et al. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. 2009;18(21):4046-4053. doi:10.1093/hmg/ddp353

20. Tobi EW, Goeman JJ, Monajemi R, et al. DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nat Commun*. 2014;5(5592):1-13. doi:10.1038/ncomms6592

21. Wang Z, Song J, Li Y, Dong B, Zou Z, Ma J. Early-Life Exposure to the Chinese Famine Is Associated with Higher Methylation Level in the INSR Gene in Later Adulthood. *Sci Rep*. 2019;(January):1-9. doi:10.1038/s41598-019-38596-6

22. Tobi EW, Slieker RC, Luijk R, et al. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci Adv*. 2018;4(1):eaao4364. doi:10.1126/sciadv.aao4364

23. Joubert BR, Håberg SE, Nilsen RM, et al. 450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns Related to Maternal Smoking during Pregnancy. *Environ Health Perspect*. 2012;120(10):1425-1431. doi:10.1289/ehp.1205412

24. Bouwland-both MI, Mil NH Van, Tolhoek CP, et al. Prenatal parental tobacco smoking , gene specific DNA methylation , and newborns size : the Generation R study. *Clin Epigenetics*. 2015:1-11. doi:10.1186/s13148-015-0115-z

25. Küpers LK, Xu X, Jankipersadsing SA, et al. DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *Int J Epidemiol*. 2015;44(4):1224-1237. doi:10.1093/ije/dyv048

26. Richmond RC, Suderman M, Langdon R, Relton CL, Davey Smith G. DNA methylation as a marker for prenatal smoke exposure in adults. *Int J Epidemiol*. 2018;47(4):1120-1130. doi:10.1093/ije/dyy091

27. Drake AJ, Mcpherson RC, Godfrey KM, et al. An unbalanced maternal diet in pregnancy associates with offspring epigenetic changes in genes controlling glucocorticoid action and foetal growth. *Clin Endocrinol (Oxf)*. 2012;77:808-815. doi:10.1111/j.1365-2265.2012.04453.x

28. Dominguez-Salas P, Moore SE, Baker MS, et al. Maternal nutrition at conception modulates

DNA methylation of human metastable epialleles. *Nat Commun*. 2014;5:3746. doi:10.1038/ncomms4746

29. James PT, Dominguez-Salas P, Hennig BJ, Moore SE, Prentice AM, Silver MJ. Maternal One-Carbon Metabolism and Infant DNA Methylation between Contrasting Seasonal Environments: A Case Study from The Gambia. *Curr Dev Nutr*. 2019;3(1). doi:10.1093/cdn/nzy082

30. James P, Sajjadi S, Tomar AS, et al. Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism. *Int J Epidemiol*. 2018;47(6):1910-1937. doi:10.1093/ije/dyy153

31. Waterland RA, Kellermayer R, Laritsky E, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet*. 2010;6(12):1-10. doi:10.1371/journal.pgen.1001252

32. Silver MJ, Kessler NJ, Hennig BJ, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol*. 2015;16:118. doi:10.1186/s13059-015-0660-y

33. Van Baak TE, Coarfa C, Dugué P, et al. Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol*. 2018;19(1):2. doi:10.1186/s13059-017-1374-0

34. Kühnen P, Handke D, Waterland RA, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016;24(3):502-509. doi:10.1016/j.cmet.2016.08.001

35. Pauwels S, Ghosh M, Duca RC, et al. Dietary and supplemental maternal methyl-group donor intake and cord blood DNA methylation. *Epigenetics*. 2017;12(1):1-10. doi:10.1080/15592294.2016.1257450

36. Pauwels S, Ghosh M, Duca RC, et al. Maternal intake of methyl-group donors affects DNA methylation of metabolic genes in infants. *Clin Epigenetics*. 2017;9(1):16. doi:10.1186/s13148-017-0321-y

37. Godfrey KM, Sheppard A, Gluckman PD, et al. Epigenetic Gene Promoter Methylation at Birth Is Associated With Child's Later Adiposity. *Diabetes*. 2011;60(5):1528-1534. doi:10.2337/db10-0979

38. Azzi S, Sas TCJ, Koudou Y, et al. Degree of methylation of ZAC1 (PLAGL1) is associated with

prenatal and post-natal growth in healthy infants of the EDEN mother child cohort. *Epigenetics*. 2014;9(3):338-345. doi:10.4161/epi.27387

39. Soubry A, Schildkraut JM, Murtha A, et al. Paternal obesity is associated with IGF2hypomethylation in newborns: results from a Newborn Epigenetics Study (NEST) cohort. *BMC Med*. 2013;11(1):29. doi:10.1186/1741-7015-11-29

40. Donkin I, Versteyhe S, Ingerslev LR, et al. Obesity and Bariatric Surgery Drive Epigenetic Variation of Spermatozoa in Humans. *Cell Metab*. 2016;23(2):369-378. doi:10.1016/j.cmet.2015.11.004

41. Noor N, Cardenas A, Rifas-Shiman SL, et al. Association of Periconception Paternal Body Mass Index With Persistent Changes in DNA Methylation of Offspring in Childhood. *JAMA Netw Open*. 2019;2(12):e1916777-e1916777. doi:10.1001/jamanetworkopen.2019.16777

42. Rosenbaum M, Leibel RL. Role of leptin in energy homeostasis in humans. *J Endocrinol*. 2014;223(1):T83-T96. doi:10.1530/JOE-14-0358

43. Potabattula R, Dittrich M, Böck J, et al. Allele-specific methylation of imprinted genes in fetal cord blood is influenced by cis-acting genetic variants and parental factors. *Epigenomics*. 2018;10(10):1315-1326. doi:10.2217/epi-2018-0059

44. Potabattula R, Dittrich M, Schorsch M, Hahn T, Haaf T, El Hajj N. Male obesity effects on sperm and next-generation cord blood DNA methylation. *PLoS One*. 2019;14(6):e0218615.

45. Jamieson E, Korologou-Linden R, Wootton R, et al. Smoking, DNA methylation and lung function: a Mendelian randomization analysis to investigate causal relationships. 2019. doi:10.1101/19003335

46. Juvinao-Quintero DL, Hivert M-F, Sharp GC, Relton CL, Elliott HR. DNA Methylation and Type 2 Diabetes: the Use of Mendelian Randomization to Assess Causality. *Curr Genet Med Rep*. 2019;7(4):191-207. doi:10.1007/s40142-019-00176-5

47. Candler T, Kühnen P, Prentice AM, Silver M. Epigenetic regulation of POMC; implications for nutritional programming, obesity and metabolic disease. *Front Neuroendocrinol*. 2019;54. doi:10.1016/j.yfrne.2019.100773

48. Govorko D, Bekdash RA, Zhang C, Sarkar DK. Male germline transmits fetal alcohol adverse effect on hypothalamic proopiomelanocortin gene across generations. *Biol Psychiatry*. 2012;72(5):378-388. doi:10.1016/j.biopsych.2012.04.006

49. Mead EA, Sarkar DK. Fetal alcohol spectrum disorders and their transmission through genetic and epigenetic mechanisms. *Front Genet*. 2014;5(JUN):1-10. doi:10.3389/fgene.2014.00154

50. Sarkar DK. Male germline transmits fetal alcohol epigenetic marks for multiple generations: A review. *Addict Biol*. 2016;21(1):23-34. doi:10.1111/adb.12186

51. Moore SE, Fulford AJC, Darboe MK, Jobarteh ML, Jarjou LM, Prentice AM. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: The ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth*. 2012. doi:10.1186/1471-2393-12-107

52. Okala SG, Darboe MK, Sosseh F, et al. Impact of nutritional supplementation during pregnancy on antibody responses to diphtheria-tetanus-pertussis vaccination in infants: A randomised trial in The Gambia. *PLoS Med*. 2019;16(8):e1002854. doi:10.1371/journal.pmed.1002854

53. James PT, Jawla O, Mohammed NI, et al. A novel nutritional supplement to reduce plasma homocysteine in nonpregnant women: A randomised controlled trial in The Gambia. *PLoS Med*. 2019;16(8):e1002870. doi:10.1371/journal.pmed.1002870

54. Moore SE, Cole TJ, Collinson AC, Poskitt EM, McGregor IA, Prentice AM. Prenatal or early postnatal events predict infectious deaths in young adulthood in rural Africa. *Int J Epidemiol*. 1999;28(6):1088-1095.

55. Alicia Alonso MPF, Kacmarczyk TJ. *Agilent SureSelectXT Methyl-Seq Applications with Low-Input DNA and Smaller Capture Libraries*.; 2017.

56. Technologies A. SureSelectXT Methyl-Seq Target Enrichment System for Illumina Multiplexed Sequencing. 2018:1-81.

57. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011;27(11):1571-1572. doi:10.1093/bioinformatics/btr167

58. Ramsay M, Sankoh O, Consortium as members of the A-G study and the H. African partnerships through the H3Africa Consortium bring a genomic dimension to longitudinal population studies on the continent. *Int J Epidemiol*. 2016;45(2):305-308. doi:10.1093/ije/dyv187

59. Howe KL, Achuthan P, Allen J, et al. Ensembl 2021. *Nucleic Acids Res*. 2021;49(D1):D884-D891. doi:10.1093/nar/gkaa942

60. Slatkin M. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*. 2008;9(6):477-485. doi:10.1038/nrg2361

61. VanLiere JM, Rosenberg NA. Mathematical properties of the r2 measure of linkage disequilibrium. *Theor Popul Biol*. 2008;74(1):130-137. doi:10.1016/j.tpb.2008.05.006

62. Silver MJ, Saffari A, Kessler NJ, et al. Environmentally sensitive hotspots in the methylome of the early human embryo. *Elife*. 2022;11:e72031. doi:10.7554/eLife.72031

63. Gunasekara CJ, Scott CA, Laritsky E, et al. A genomic atlas of systemic interindividual epigenetic variation in humans. *Genome Biol*. 2019;20(1):105. doi:10.1186/s13059-019-1708-1

64. Cunningham F, Allen JE, Allen J, et al. Ensembl 2022. *Nucleic Acids Res*. 2022;50(D1):D988-D995. doi:10.1093/nar/gkab1049

65. Silver MJ, Saffari A, Kessler NJ, et al. Environmentally sensitive hotspots in the methylome of the early human embryo. *bioRxiv*. January 2021:777508. doi:10.1101/777508

66. Dominguez-Salas P, Moore SE, Cole D, et al. DNA methylation potential: Dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am J Clin Nutr*. 2013;97(6):1217-1227. doi:10.3945/ajcn.112.048462

67. Dugdale AE, Payne PR. A model of seasonal changes in energy balance. *Ecol Food Nutr*. 1987;19(3):231-245. doi:10.1080/03670244.1987.9990966

68. Brosnan ME, MacMillan L, Stevens JR, Brosnan JT. Division of labour: how does folate metabolism partition between one-carbon metabolism and amino acid oxidation? *Biochem J*. 2015;472(2):135-146. doi:10.1042/BJ20150837

69. Rios-Avila L, Coats B, Chi Y-Y, et al. 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 . *J Nutr*. 2015;145(1):87-95. doi:10.3945/jn.114.201095

70. Zhang N. Role of methionine on epigenetic modification of DNA methylation and gene expression in animals. *Anim Nutr*. 2018;4(1):11-16. doi:https://doi.org/10.1016/j.aninu.2017.08.009

71. Rowling MJ, Schalinske KL. Retinoid compounds activate and induce hepatic glycine N-methyltransferase in rats. *J Nutr*. 2001;131(7):1914-1917. doi:10.1093/jn/131.7.1914

72. Krokan HE, Bjørås M. Base excision repair. *Cold Spring Harb Perspect Biol*. 2013;5(4):a012583-a012583. doi:10.1101/cshperspect.a012583

73. Xue J-H, Xu G-F, Gu T-P, et al. Uracil-DNA Glycosylase UNG Promotes Tet-mediated DNA Demethylation. *J Biol Chem*. 2016;291(2):731-738. doi:10.1074/jbc.M115.693861

74. Sharp GC, Salas LA, Monnereau C, et al. Maternal BMI at the start of pregnancy and offspring epigenome-wide DNA methylation : findings from the pregnancy and childhood epigenetics (PACE) consortium. 2017;26(20):4067-4085. doi:10.1093/hmg/ddx290

75. Saffari A, Shrestha S, Issarapu P, et al. Effect of maternal preconceptional and pregnancy micronutrient interventions on children's DNA methylation: Findings from the EMPHASIS study. *Am J Clin Nutr*. 2020;112(4):1099-1113. doi:10.1093/ajcn/nqaa193

76. Myte R, Sundkvist A, Van Guelpen B, Harlid S. Circulating levels of inflammatory markers and DNA methylation, an analysis of repeated samples from a population based cohort. *Epigenetics*. 2019;14(7):649-659. doi:10.1080/15592294.2019.1603962

77. Ligthart S, Marzi C, Aslibekyan S, et al. DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biol*. 2016;17(1):255. doi:10.1186/s13059-016-1119-5

78. Gannon BM, Glesby MJ, Finkelstein JL, Raj T, Erickson D, Mehta S. A point-of-care assay for alpha-1-acid glycoprotein as a diagnostic tool for rapid, mobile-based determination of inflammation. *Curr Res Biotechnol*. 2019;1:41-48. doi:https://doi.org/10.1016/j.crbiot.2019.09.002

79. Fetahu IS, Höbaus J, Kállay E. Vitamin D and the epigenome. *Front Physiol*. 2014;5:164. doi:10.3389/fphys.2014.00164

80. Sánchez-Hernández D, Cho CE, Kubant R, et al. Increasing vitamin A in post-weaning diets reduces food intake and body weight and modifies gene expression in brains of male rats born to dams fed a high multivitamin diet. *J Nutr Biochem*. 2014;25(10):991-996. doi:10.1016/j.jnutbio.2014.05.002

81. Boersma GJ, Liang N, Lee RS, et al. Failure to upregulate Agrp and Orexin in response to activity based anorexia in weight loss vulnerable rats characterized by passive stress coping and prenatal stress experience. *Psychoneuroendocrinology*. 2016;67:171-181. doi:10.1016/j.psyneuen.2016.02.002.Failure

82. Wu J, Zhang W, Li C. Recent Advances in Genetic and Epigenetic Modulation of Animal

Exposure to High Temperature. *Front Genet*. 2020;11:653. doi:10.3389/fgene.2020.00653

83. Kristin Kohlmann G. Season of conception influences epigenetic modifications of the pro-opiomelanocortin gene in children in rural Bangladesh. 2019.

84. Chowen JA, Freire-Regatillo A, Argente J. Neurobiological characteristics underlying metabolic differences between males and females. *Prog Neurobiol*. 2019;(179):18-32. doi:10.1016/j.pneurobio.2018.09.001

85. Speliotes EK, Willer CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet*. 2010;42(11):937-948. doi:10.1038/ng.686

86. Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. *Sci Adv*. 2018;4(7):eaat2624. doi:10.1126/sciadv.aat2624

87. Tobi EW, Slagboom PE, van Dongen J, et al. Prenatal famine and genetic variation are independently and additively associated with DNA methylation at regulatory loci within IGF2/H19. *PLoS One*. 2012;7(5):e37933. doi:10.1371/journal.pone.0037933

88. Czamara D, Eraslan G, Page CM, et al. Integrated analysis of environmental and genetic influences on cord blood DNA methylation in new-borns. *Nat Commun*. 2019;10(1):2548. doi:10.1038/s41467-019-10461-0

89. Zhang H, Kranzler HR, Weiss RD, et al. Pro-opiomelanocortin gene variation related to alcohol or drug dependence: evidence and replications across family- and population-based studies. *Biol Psychiatry*. 2009;66(2):128-136. doi:10.1016/j.biopsych.2008.12.021

90. Hillemacher T, Kahl KG, Heberlein A, et al. Appetite- and volume-regulating neuropeptides: Role in treating alcohol dependence. *Curr Opin Investig Drugs*. 2010;11(10):1097-1106.

91. Muschler MAN, Hillemacher T, Kraus C, Kornhuber J, Bleich S, Frieling H. DNA methylation of the POMC gene promoter is associated with craving in alcohol dependence. *J Neural Transm*. 2010;117(4):513-519. doi:10.1007/s00702-010-0378-7

92. Muschler MAN, Lenz B, Hillemacher T, et al. CAGn repeat of the androgen receptor is linked to proopiomelanocortin promoter methylation - Relevance for craving of male alcohol-dependent patients? *Psychopharmacology (Berl)*. 2014;231(10):2059-2066. doi:10.1007/s00213-013-3349-5

93. Zhang H, Herman AI, Kranzler HR, et al. Array-based profiling of dna methylation changes associated with alcohol dependence. *Alcohol Clin Exp Res*. 2013. doi:10.1111/j.1530-0277.2012.01928.x

94. Lombard Z, Crowther NJ, van der Merwe L, Pitamber P, Norris SA, Ramsay M. Appetite regulation genes are associated with body mass index in black South African adolescents: a genetic association study. *BMJ Open*. 2012;2(3). doi:10.1136/bmjopen-2012-000873

95. Perez MF, Lehner B. Intergenerational and transgenerational epigenetic inheritance in animals. *Nat Cell Biol*. 2019;21(2):143-151. doi:10.1038/s41556-018-0242-9

96. Sharp GC, Lawlor DA, Richardson SS. It's the mother!: How assumptions about the causal primacy of maternal effects influence research on the developmental origins of health and disease. *Soc Sci Med*. 2018;213:20-27. doi:10.1016/j.socscimed.2018.07.035

97. Chen Q, Yan W, Duan E. Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nat Rev Genet*. 2016;17(12):733-743. doi:10.1038/nrg.2016.106

98. Donkin I, Barrès R. Sperm epigenetics and influence of environmental factors. *Mol Metab*. 2018;14:1-11. doi:10.1016/j.molmet.2018.02.006

Chapter 7 Examining the potentially modifiable effect of differential PAX-8 methylation on thyroid function in children in The Gambia

Summary of the chapter

This chapter includes presents all the methods and results for the PAX8 study. The chapter reports that offspring *PAX8* methylation is influenced by a number of maternal biomarkers and child's genotype. Furthermore, I report a marked difference between thyroid function (free T4) and thyroid volume in those with high and low *PAX8* methylation. I also demonstrate that variability in free T4 (even within the population reference range) is associated with fat and bone measures in children. The chapter represents the final accepted manuscript. The order of subheadings reflects the journal requirements. The open access typeset version can be found at DOI: 10.1126/sciadv.abj1561.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	1801141	**Title**	Dr
First Name(s)	Toby		
Surname/Family Name	Candler		
Thesis Title	Assessing the influence of maternal nutrition-sensitive epigenetic signatures in the POMC and PAX-8 genes on health-related outcomes in offspring		
Primary Supervisor	Dr Matt Silver		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Science Advances		
When was the work published?	2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	n/a		
Have you retained the copyright for the work?*	**No**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I conceived the idea to examine the relationship between PAX8 methylation and thyroid outcomes measures in Gambian Children. I was Co-PI and authored the successful grant application and ethics proposals I led the study design with Dr Silver and implemented the study activity on the ground in Keneba. I organised the thyroid ultrasound training in Keneba and performed all of the ultrasound scans. I organised the blood sample preparation and sample shipment from MRC The Gambia to the University of Cardiff where the thyroid function analysis took place. I performed all of the data analysis for the paper, though I had support with identifying the participants from Noah Kessler (University of Cambridge) and thus I remained blinded to the study group when performing the ultrasound. Noah Kessler also identified the genotype from methyl-seq data and performed the TCGA methylation-expression analysis. I collaborated with Prof Waterland (Baylor College of Medicine, Texas USA) once the initial results were analysed to explore the potential of using the GTEx biobank to examine PAX8 gene expression and methylation relationships. This led to the GTEx results being used in the paper. I worked with Dr Kate Ward to be trained on DXA scanning and scan quality control. She made suggestion for the analysis plan used for the DXA measures. Dr Phil James kindly shared the MDEG2 maternal biomarker data which I analysed further in relation to PAX8 methylation. Dr Silver supported the project throughout and helped with particular guidance with analysis of the biomarkers and PAX8 methylation and the genotype analysis. All the authors reviewed and approved the manuscript.

SECTION E

Student Signature	
Date	11/07/2022

Supervisor Signature	[redacted]
Date	13/7/22

Full Title:

DNA methylation at a nutritionally sensitive region of the *PAX8* gene is associated with thyroid volume and function in Gambian children

Short Title:

PAX8 methylation influences thyroid volume and function

Authors:

T Candler[1], NJ Kessler[2], CJ Gunasekara[3], KA Ward[1,4], P James[5], E Laritsky[3], MS Baker[3], R Dyer[6], R Elango[6], D Jeffries[1], RA Waterland[3,7], S Moore[1,8], M Ludgate[9], AM Prentice[1], MJ Silver[1*]

Affiliations:

[1]MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, London, UK

[2]Department of Genetics, University of Cambridge, Cambridge, UK

[3]USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, USA

[4]MRC Lifecourse Epidemiology, University of Southampton, UK

[5]Department of Population Health, London School of Hygiene and Tropical Medicine, London, UK

[6] British Columbia Children's Hospital Research Institute, Vancouver, BC, Canada

[7]Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, USA

[8]Department of Women & Children's Health, King's College London, London, UK

[9]Thyroid Research Group, School of Medicine, Cardiff University, Cardiff, UK

[*]**Corresponding author:** Dr Matt Silver, matt.silver@lshtm.ac.uk

Abstract

PAX8 is a key thyroid transcription factor implicated in thyroid gland differentiation and function, and *PAX8* gene methylation is reported to be sensitive to the periconceptional environment. Using a novel recall-by-epigenotype study in Gambian children, we found that *PAX8* hypomethylation at age 2 years is associated with a 21% increase in thyroid volume and an increase in free thyroxine (T4) at 5-8 years, the latter equivalent to 8.4% of the normal range. Free T4 was associated with a decrease in DXA-derived body fat and bone mineral density. Furthermore, offspring *PAX8* methylation was associated with periconceptional maternal nutrition and methylation variability was influenced by genotype suggesting sensitivity to environmental exposures may be under partial genetic control. Taken together, our results demonstrate a possible link between early environment, *PAX8* gene methylation and thyroid gland development and function, with potential implications for early embryonic programming of thyroid-related health and disease.

Teaser

Children's PAX8 gene methylation is associated with thyroid volume and function and is influenced by maternal periconceptional nutrition.

Introduction

Thyroid hormones regulate a wide range of physiological processes and influence various outcomes related to cognition, growth, skeletal, cardiovascular, and metabolic health[1]. While clinical sequelae of severe perturbations of thyroid hormone production are well documented[2], variation of thyroid function within the normal population reference range is also associated with a range of phenotypic traits including blood pressure, lipids, obesity, cardiovascular mortality, bone mineral density (BMD) and cancer risk[3].

The function of the hypothalamic-pituitary-thyroid axis is clinically assessed by measurement of pituitary derived thyrotropin (TSH), free T4 (free thyroxine) and free T3 (free tri-iodothyronine). Higher TSH and lower free T4 are associated with adverse pregnancy outcomes[3] and increased body mass index (BMI) in adults[4] and children[5]. Lower TSH and higher free T4 are associated with an increased risk of osteoporosis and fracture[3], and levels of free T4 are correlated with the concentration of various lipoproteins[6].

Heritability of thyroid function has been reported to be between 32-65% (free T4), 23-67% (free T3) and 32-65% (TSH)[7]. However, current identified genetic variants associated with thyroid function contribute only a small amount to interindividual variation in thyroid hormone concentrations. For example, although a recent genome-wide association study (GWAS) identified 74 loci associated with TSH, together these explained just 13.3% of TSH variance, leaving much of the reported heritability unexplained[8]. Moreover, congenital hypothyroidism (CHT), the most common endocrinopathy of childhood, is generally not inherited (less than 5% of cases have an identifiable genetic cause[9]), with 98% of cases non-familial[10] and a high discordance rate (92%) in monozygotic twins[11]. In addition, seasonal variation in CHT incidence has been reported in the UK and Japan[12,13], and a recent study reported that prenatal famine exposure has been associated with higher TSH in adulthood[14]. Together, these observations suggest that environmental factors and epigenetic or unknown genetic mechanisms may play a role in thyroid development and/or function.

PAX8 (paired-box 8) protein is one of four known thyroid transcription factors (TTFs) essential for thyroid development and function (others include NKX2-1, FOXE1 and HHEX)[15]. These transcription factors regulate expression of thyroid specific genes related to thyroid hormone production and storage such as TPO (thyroid peroxidase), Tg (thyroglobulin) and the sodium-iodide symporter[15,16], and are important for development and differentiation of the thyroid gland[17]. PAX8 has been described as the "master regulator", with a role in regulating the activity of other TTFs[16]. *PAX8* knockout mice demonstrate thyroid hypoplasia, low birth weight and growth retardation[18], and genetic mutations in *PAX8* can cause CHT in humans[19].

Epigenetic processes, including DNA methylation, histone modification, protein binding of DNA, chromatin remodelling and RNA-based mechanisms can affect gene expression[20]. DNA methylation at CpG 'islands' within promoter regions may regulate gene transcription in a tissue-specific manner[21]. Promoter DNA methylation is usually associated with condensed heterochromatin and transcriptional down-regulation[22]. Metastable epialleles (MEs) are epigenetic loci that demonstrate systemic (i.e. not tissue-specific) methylation with substantial variation between individuals[23]. The methylation patterns are thought to be established early in embryonic development[24] and are not driven by genetic variation[23,24]. Methylation at MEs is influenced by maternal diet around conception[25–29], and putative human MEs with stable methylation levels have been linked to disease-related phenotypes[25,26]. These characteristics position MEs as potential epigenetic mediators of the effect of early environmental exposures in the developmental origins of health and disease[30].

The *PAX8* gene contains a putative human ME with evidence of systemic interindividual variation[23,25,31] and sensitivity to periconceptional and prenatal environment (see Supplementary Table 1). Indeed, there is evidence that *PAX8* promoter methylation patterns in leucocytes and thyroid are concordant in children[32]. Data from The Gambia showed that leucocyte *PAX8* methylation was higher in children conceived in the annual rainy (or 'hungry') season[25,31]. In Bangladesh, higher methylation at the *PAX8* gene was

reported in offspring following gestational famine exposure[27]. A set of interlocking pathways, collectively known as one-carbon metabolism, provides methyl groups for DNA methylation, and is dependent on multiple nutritional factors which act as substrates and essential co-factors[33]. Seasonally-driven variations in maternal circulating levels of one-carbon metabolites have been reported in The Gambia[34]. Maternal supplementation with folic acid (a synthetic source of folate, a one-carbon metabolite) has been associated with differential *PAX8* methylation in adult offspring[35] and maternal preconception multiple micronutrient supplementation was associated with differential methylation at *PAX8* in Gambian children[36].

Epigenetic influence on thyroid function or development has been little explored. Considering its key roles in thyroid development and regulation of the mature thyroid gland, and its epigenetic sensitivity to the early environment, *PAX8* is a prime candidate for study. Using a recall-by-epigenotype design, we examined links between *PAX8* DNA methylation measured at 2 years of age in peripheral blood, and thyroid gland function and development in the same Gambian children aged 5-8 years. In addition, by examining body composition and BMD using Dual Energy X-ray Absorptiometry (DXA) scans, we explored how *PAX8* methylation (via its putative effect on thyroid hormone production) may influence measures of adiposity and BMD. Using maternal biomarker data (including measures of one-carbon metabolites) we also explored how a child's *PAX8* methylation status may be influenced by their mother's nutritional status around the time of conception, and assessed the influence of genetic variation in *cis*. Finally, using data from The Cancer Genome Atlas (TCGA) we investigated the relationship between *PAX8* methylation and gene expression in thyroid tissue, and assessed correlations between thyroid and whole blood methylation in samples from the Genotype-Tissue Expression (GTEx) Project.

Results

PAX8 region of interest, participant selection and baseline characteristics

Children from the "ENID" (Early Nutrition and Immune Development) trial[37] now aged between 5 and 8 years (n=493) were recruited into "high" and "low" groups, according to their DNA methylation status in peripheral blood at age 2 years at a region of the *PAX8* (see Figure 1). This region was selected based on evidence of systemic inter-individual variation and sensitivity to early environment from previous studies[23,25,27]. DNA methylation was measured at 4 CpGs which were highly correlated with adjacent CpGs and had sufficient coverage on the methyl-seq platform used (see Methods and Supplementary Figure 10 for further details).

One hundred and eighteen children were recruited (high *PAX8* methylation group n=58 (mean methylation (SD) = 0.96 (0.036)) and low *PAX8* methylation group n=60 (0.50 (0.088)) with a median age (IQR) of 7.18 years (1.67). Children in the low *PAX8* methylation group had a significantly lower BMI z-score, but there were no significant differences in age, sex, weight-for-age z-score (WAZ), height-for-age z-score (HAZ), or pregnancy and infant supplementation received between the two groups (see Table 1).

Associations between *PAX8* methylation and thyroid outcomes

Thyroid volume

Thyroid ultrasound scans were performed blinded to *PAX8* methylation group on 114 children (4 children did not attend for scans). Of these, 2 had technical difficulties in ascertaining accurate measurements and were removed from subsequent analysis (both were in the high methylation group). A single case of right thyroid lobe hemi-agenesis in the high *PAX8* methylation group was the only abnormality identified. This participant was retained in the analysis as total thyroid volume was within normal range.

In crude unadjusted analyses, total thyroid volume was elevated in the low methylation group (3.24 vs 2.87 cm^3 in the high PAX8 methylation group, p=0.035; see Table 2). In a multiple linear regression model adjusting for age, sex, BMI and UIC, the association between *PAX8* methylation group and total thyroid volume was strengthened, yielding a 0.61cm^3 [SE=0.15] or 21% higher thyroid volume in the low vs high group (p=0.0001; see Table 3). This analysis also revealed significant associations between total thyroid volume and age, sex, BMI, and UIC (see Supplementary Table 2).

Thyroid function

In crude comparisons of thyroid function measures (free T4, free T3, TSH and Tg) between the two groups, free T4 was observed to be significantly lower in the high *PAX8* methylation group (13.3 vs 13.9 pmol/L, p=0.009, see Table 2).

Iodine insufficiency was evaluated by two methods: plasma Tg level and UIC. Based on Tg level there was no difference in iodine insufficiency between the groups, but significantly more children were classified with iodine insufficiency by UIC in the high *PAX8* methylation group compared to the low *PAX8* methylation group (21/56 (37.5%) vs 10/60 (16.7%), p=0.02), and UIC was also higher in the low *PAX8* methylation group (170 vs 128 ug/L, p=0.04, see Supplementary Table 3). Since iodine concentration and/or insufficiency can affect thyroid function, UIC was included as an adjustment covariate in regression models to ensure appropriate adjustment was made for these group differences.

In a multiple linear regression model with adjustment for age, sex, and UIC, free T4 level was 0.85 pmol/L [SE=0.24] higher in the low *PAX8* methylation group (p=0.0007; see Table 3 and Supplementary Table 4). The laboratory normal range for free T4 is between 9.0 and 19.1 pmol/L, so this difference between the two groups – 8.4% of the normal range – is substantial. No significant associations between *PAX8* methylation and TSH, free T3 or Tg

were observed after adjustment for age, sex, and UIC (see Table 3 and Supplementary Table 4).

Association between free T4 and body composition and bone mineral outcomes

We hypothesized that *PAX8* methylation (via its putative effect on thyroid hormone production) may influence measures of adiposity and BMD. To explore potential associations between free T4 and fat mass index (FMI), lean mass index (LMI) and BMD, whole body DXA scans were performed on 113 of the 5-8 year old children.

In multiple regression models adjusted for relevant covariates, log FMI was associated with free T4 (β=-0.04 [SE 0.02], p=0.033; see Table 4 and Supplementary Table 5), so that for every pmol/L increase in free T4 there was a 4.3% reduction in FMI. Free T4 was not associated with LMI (see Supplementary Table 6). Free T4 was inversely associated with log total-body-less-head (TBLH) BMD (β=-0.008 [0.004], p=0.044), so that TBLH BMD was reduced by 0.8% for every pmol/L increase in free T4 (see Table 4 and Supplementary Table 7).

Causal Mediation Analysis

Since we demonstrated significant associations between *PAX8* methylation group and both thyroid volume and free T4, we postulated that the methylation state set early in embryonic development (which we measured at age 2 of age) could influence thyroid development and volume via regulation of *PAX8* expression, and that this change in thyroid volume could in turn affect thyroid function and free T4 measured at 5-7 years of age. We therefore conducted a causal mediation analysis to test this, but this provided no evidence that the effect of differential methylation at PAX8 on thyroid function was mediated by thyroid volume (ACME or mediated effect = 0.031, 95% CI -0.15-0.22, p=0.74, see Supplementary Figure 1).

As there was no significant effect of *PAX8* methylation group on either BMD or any fat measure (data not shown), we did not perform causal mediation analysis for these pathways.

Predictors of *PAX8* methylation

We performed an analysis of the potential influence of sex, and child and maternal BMI on *PAX8* methylation in n=521 ENID-recruited children with 2-year *PAX8* methylation measurements (i.e. not restricted to the n=118 high and low methylation groups analysed above). In simple linear regression models, sex was significantly associated with *PAX8* methylation, with males having 0.24 (SE=0.09) lower mean logit methylation compared to females (p=0.005, see Supplementary Figure 2 for comparison of equivalent untransformed mean % methylation difference). There were no significant associations between *PAX8* methylation and child BMI z-score (β=-0.02 [SE=0.04], p=0.64), maternal BMI (β=0.0004 [0.01], p=0.98), infant WAZ (β= -0.05 [0.05], p=0.27) or SoC (β=0.09 [0.09], p=0.29).

303 of the 521 children with *PAX8* methylation data also had paired maternal biomarker data. In multiple linear regression models adjusted for sex, 4 one-carbon metabolites were associated with a decrease in *PAX8* methylation (see Table 5): homocysteine (standardised β=-0.11 [0.05], p=0.048), cysteine (β=-0.16 [0.05], p=0.003), vitamin B12 (β=-0.1 [0.05], p=0.05) and vitamin B6/PLP (β=-0.12 [0.06], p=0.03).

We also tested an extended panel of maternal biomarkers including amino acid levels (see Supplementary Table 8). Two of these biomarkers were associated with an increase in *PAX8* methylation: uracil (standardised β=0.21 [0.05], p=0.0001) and arginine (β=0.16 [0.06], p=0.006). Valine was associated with a decrease in *PAX8* methylation (β=-0.13 [0.05], p=0.02).

***PAX8* methylation, gene expression and methylation tissue concordance**

We next investigated the relationship between *PAX8* methylation and gene expression in thyroid tissue by using data from The Cancer Genome Atlas (TCGA). TCGA provides gene expression and DNA methylation data on a variety of non-cancerous tissues including 50 thyroid samples. The methylation data is derived from the Illumina HM450 array, which includes two of the *PAX8* CpGs of interest in this study. Since this region overlaps transcripts from sense (*PAX8*) and anti-sense (*PAX8-AS1*) genes, we considered the expression of both genes. Using FPKM as an expression metric, *PAX8* and *PAX8-AS1* expression were positively correlated (Spearman R=0.477, p=0.0006, see Supplementary Figure 3 left), but we found only weak evidence of an inverse relationship between *PAX8* methylation and *PAX8-AS1* or *PAX8* expression (see Supplementary Figure 4 top). However, with RSEM as the metric of gene expression, we found strong evidence that *PAX8* methylation was negatively correlated with *PAX8-AS1* expression (Spearman R=-0.70, $p=6x10^{-8}$, see Supplementary Figure 4 bottom right), although there was no correlation between *PAX8* and *PAX8-AS1* expression (Supplementary Figure 3 right). There was no association between *PAX8* methylation and expression using RSEM (Supplementary Figure 4 bottom left).

We analysed mean DNA methylation across the 4 CpGs of interest in paired peripheral blood cell and thyroid tissue from 86 adult samples from the Genotype-Tissue Expression (GTEx) Project[38] and found no correlation (see Supplementary Figure 5). Public GTEx expression data from 54 tissues shows that both *PAX8* and *PAX8-AS1* are predominantly expressed in thyroid (see Supplementary Figure 6). Note that in both the TGCA and GTEx dataset, thyroid methylation levels of the *PAX8* region of interest fall only within the high methylation group from our Gambian analysis in peripheral blood. Together, our interpretation is that although individual variation in PAX8 methylation in the early embryo may affect thyroid development, at some stage of thyroid development this region becomes uniformly highly methylated in thyroid of all individuals.

Stability of *PAX8* methylation from 7 to 17 years of age

We measured *PAX8* methylation in 49 Gambian children from stored peripheral blood DNA taken at age 7 and 17 years by pyrosequencing. Methylation was highly correlated between the two age groups (R=0.76, $p<2.2x10^{-16}$; see Supplementary Figure 7), indicating that blood cell methylation at this locus is highly stable across this age range.

Exploring the relationship between *PAX8* methylation, genotype, and thyroid function

We focused on a single SNP (rs10193733) proximal to our *PAX8* region of interest for genetic analyses (see Figure 2 and Methods for rationale for selecting this SNP). We observed higher methylation for homozygotes for the alternate allele (C/C, mean methylation = 0.97) compared to heterozygotes (T/C = 0.82 or wild type alleles (T/T = 0.66; see Figure 4). However, most striking was that homozygotes for the alternate allele had a markedly limited methylation range (C/C, range = 0.93 - 1.00) compared to heterozygotes (C/T, range= 0.50 - 0.99) or homozygotes for the reference allele (T/T, range = 0.18 - 1.00; see Figure 4).

This led us to postulate that the genotype with the greatest variability (T/T) could be more sensitive to periconceptional environment, offering a potential explanation for the lack of evidence for a SoC effect at this locus. Indeed, methylation distributions stratified by genotype suggested a potential interaction between SoC and genotype (see Supplementary Figure 8), although this was not statistically significant, possibly due to a lack of power. We also found no evidence for an interaction between maternal nutritional biomarkers and genotype (data not shown).

We next assessed the direct effect of rs10193733 genotype on free T4 and total thyroid volume in multiple linear regression models with free T4 and total thyroid volume as the dependent variables regressed against the number of rs10193733 C alleles and adjusted for age, sex and UIC. The C allele had a significant effect on both free T4 (β (per C allele) = -0.61 [SE=0.16], p=0.0002) and thyroid volume (β (per C allele) = -0.25 [0.11], p=0.02). As

expected, (see Figures 3 and 4), rs10193733 genotype was also significantly associated with *PAX8* methylation group (chi-squared test $p=4.24x10^{-13}$). However, the effect of *PAX8* methylation group on thyroid volume and free T4 did not appear to be purely driven by genotype, as T/C and T/T individuals were well represented in both the high and low PAX8 methylation groups (see Supplementary Figure 9).

Discussion

Previous studies in The Gambia and elsewhere have shown a consistent association between DNA methylation at a genomic region in the *PAX8* gene and the maternal periconceptional environment[25,27,29,31]. Here we have demonstrated an association between *PAX8* methylation and thyroid function in Gambian children, specifically free T4 and thyroid volume, with a relatively large effect size. Thyroid volume differed by >20% between low and high *PAX8* methylation groups after adjusting for covariates. Free T4 demonstrated a difference of 0.85 pmol/L between groups or 8.4% of the normal range.

A recent GWAS of thyroid function found 4 SNPs (with MAF >1%) associated with free T4. The largest reported SNP effect was 0.22 pmol/L per variant allele and overall, common genetic variants together explained just over 20% of the variance in free T4[39]. The difference in free T4 between high and low *PAX8* methylation groups in our study is therefore much greater than the largest individual genetic effect previously observed (0.85 vs 0.22 pmol/L). Similarly, a recent GWAS found four variants associated with thyroid volume explaining just over 3% of thyroid volume variance[40]. Here, the largest SNP effect size was 0.093 cm^3 per allele, again far smaller than the effect size seen in our study (0.61 cm^3 after adjustment for covariates).

Few studies have analysed associations between DNA methylation and thyroid function. A recent epigenome-wide association study (EWAS) did not find any associations with free T4, but did identify 2 differentially methylated positions (DMPs) associated with TSH and 6

DMPs associated with free T3[41]. There were no DMPs from the *PAX8* gene associated with thyroid function in this EWAS, although the study analysed data from older European children (aged 14-17 years), and they may not have analysed CpGs from our region of interest which has limited coverage on the Illumina HM450k array used in the study.

Our previous work in The Gambia demonstrated that *PAX8* methylation is sensitive to the periconceptional environment, with rainy season conceptions having higher methylation compared to conceptions in the dry season[25,31]. There is evidence to suggest the prenatal window where environmental exposures could influence *PAX8* gene methylation may not be limited to the periconceptional period. Famine exposure (for at least 7 months prenatally)[27] and pregnancy folic acid supplementation[35] have been associated with *PAX8* gene methylation in offspring. We did not find a significant SoC effect in this cohort. Our CpGs of interest only partially overlap the SoC-sensitive *PAX8* differentially methylated region highlighted by Silver et al[25] and they are adjacent to (but do not include) the SoC-sensitive region reported by Waterland et al[31]. Thus, while they are correlated with regions previously identified as being sensitive to SoC effects (see Supplementary Figure 10), the *PAX8* CpGs in this study were different and may be less sensitive to SoC effects. Additionally, seasonal effects vary from year to year (see Figure S12 of Waterland et al.[31]), so that it is possible that there may have been a diminished seasonal effect during the period of ENID trial recruitment compared to previous years. However, we did find associations with circulating levels of several one-carbon metabolites and amino acids measured in maternal plasma and back-extrapolated to the time of conception, and these have previously been shown to vary seasonally in this population[34]. One-carbon metabolites covary through complex interactions[34] so that, as previously, we have presented their nominal associations, unadjusted for multiple comparisons[27,56].

Interestingly, there is evidence that the *PAX8* region of interest is an ME[23,25,27], and circulating levels of one-carbon metabolites at the time of conception have previously been associated with offspring DNA methylation at several MEs in our population[26,27,30]. In our study increased cysteine, homocysteine and PLP (a B6 vitamer) were associated with lower

DNA methylation at the *PAX8* region of interest, as previously observed at other MEs[29,42]. Furthermore, we identified the amino acids arginine and valine as potential predictors of *PAX8* methylation in a wider panel of peri-conceptional amino acid data. Amino acids and one-carbon metabolites are bidirectionally linked: amino acids can supply cells with one carbon metabolites[43] and levels of one carbon metabolites can influence amino acid metabolism[44].

We observed a strong effect of genotype on DNA methylation variability, raising the possibility of a genotype-early environment interaction effect on methylation, as has been observed in a number of studies[45,46]. We did not find evidence for a genotype-SoC interaction at the *PAX8* region studied, although we may have had limited power to detect this. Interestingly, we noted the presence of an indel polymorphism within the *PAX8* variably methylated region, and in very strong LD with the *cis*-SNP associated with *PAX8* methylation variability. It is plausible that this polymorphism influences the binding of a protein which affects methylation variability. Future work should explore this and other mechanisms underpinning the strong variability effect that we observed, and its potential link to differential sensitivity to environmental factors.

In this study *PAX8* methylation was measured at 2 years of age in peripheral blood, with phenotypic measures of thyroid function measured at 5-8 years. There is evidence that methylation in this region is systemic, with similar methylation patterns across tissues derived from endoderm, ectoderm and mesoderm lineages, suggesting methylation is established prior to gastrulation[23,29,31,47]. We therefore speculate that thyroid *PAX8* methylation, influenced by environmental factors and established early in gestation, sets a trajectory for thyroid gland development that is reflected in differential thyroid morphology and function in mid-childhood (see Figure 5).

PAX8 is expressed early in embryological development (from day 20-22 in humans[15]). Therefore, methylation patterns established in the cleavage-stage embryo could influence

PAX8 gene expression from the beginning of thyroid growth and development. Furthermore, methylation-driven differential *PAX8* expression in the differentiated gland could explain the difference in free T4 between the groups if high *PAX8* methylation downregulates thyroid specific genes associated with free T4 production. Further work is required to elucidate the cellular and molecular mechanisms underpinning these relationships.

Although our finding that this region is highly methylated in adult thyroid tissue from GTEx and TCGA appears to contradict this model, it is possible that thyroid methylation follows the systemic pattern in gestation and early life, followed by aging-associated hypermethylation specifically in thyroid. This is consistent with a previous study in children that compared DNA methylation in the *PAX8* promoter between leucocytes and thyroid tissue and found them to be correlated[32]. A similar lineage-specific effect was observed at the murine AxinFused ME[48].

Our work also highlights a potential epigenetic contribution to thyroid gland development that has implications for understanding the aetiology of non-heritable thyroid pathologies such as CHT. CHT is one of the top treatable causes of neurodevelopmental delay and is the commonest endocrinopathy of childhood (incidence between 1:2000–1:4000 newborns)[49], however only a small proportion of CHT cases are attributed to known genetic mutations. Previous studies in Europe and Asia have found that the incidence of CHT displays a seasonal pattern[12,13], in support of an environmental component in the aetiology of CHT. Further characterisation of environmentally sensitive epigenetic regions (such as *PAX8*) associated with thyroid development could thus have public health implications in identifying environmental drivers for CHT. Sexually dimorphic methylation effects such as those we observed at *PAX8* (and in a previous Gambian study[29]) may also help to explain the preponderance for CHT in females (2:1)[49].

We identified one child from the high methylation group with thyroid hemiagenesis, a rare developmental thyroid anomaly (prevalence reported to be between 0.05-0.5%) that has been associated with *PAX8* gene mutations[50]. While we cannot attribute *PAX8* hypermethylation as the direct cause of this anomaly, the possible contribution of epigenetic alterations to abnormal thyroid development warrants further research.

In this study we have also demonstrated associations between free T4 and body fat and BMD in Gambian children. In common with our findings, there is growing evidence that free T4 levels within the normal range are inversely correlated with measures of body fat[51]. Data from Korean children found free T4 was inversely correlated with waist circumference, BMI and markers of insulin resistance[51]. A study in UK children demonstrated that free T4 was negatively associated with FMI (from DXA) and BMI. To explore the direction of causation, the authors used Mendelian randomisation and found that whilst BMI and body fat caused an increase in free T3, they did not seem to influence levels of free T4[52]. Thyroid volume and TSH have been shown to change in response to weight loss, though free T4 remains unchanged, further supporting the idea that free T4 is not influenced by body composition[53]. Data from African children is limited, though a positive correlation between free T3/T4 ratio and BMI has been reported in Nigerian children[54]. We found no association between free T4 and measures of lean mass, though an inverse relationship between free T4 and lean mass has been reported previously[55].

Free T4 has been reported to be negatively associated with BMD in adults[56], and higher free T4 (within the population reference range) is associated with an increased incidence of fractures in older people[57]. Chondrocytes, osteoblasts and osteoclasts express TSH receptor and thyroid hormone receptor and thus there is a biological target for thyroid hormone to affect bone modelling[58]. Our novel finding that BMD is associated with free T4 in mid-childhood is of great interest as this reflects an important period of bone accrual and attainment of peak bone mass which is related to future fracture risk. We found no overall effect of *PAX8* methylation on BMD and so did not investigate the possibility that the effect of methylation on BMD is mediated by free T4. However, it is possible that *PAX8*

methylation may exert countervailing effects on BMD such that the net effect is non-significant despite their being a significant causal pathway between *PAX8* methylation, free T4 and BMD. *PAX8* has been reported to be hypomethylated in human cartilage in individuals with osteoarthritis[59], which is of interest considering our finding that low *PAX8* methylation is associated with higher free T4, which in turn is associated with lower BMD.

The associations between *PAX8* methylation, thyroid phenotype and BMI, fat and bone measures could reflect an adaptive process. For example, if higher PAX8 methylation (associated with a negative maternal energy balance in the Gambian rainy season[34]) is associated with a reduced level of thinness (as shown by a higher BMI in the high methylation group) this could be an example of the organism developing greater fat stores in response to a predicted nutritionally adverse postnatal environment.

We found an inverse relationship between *PAX8* methylation and *PAX8-AS1* expression (measured by RSEM) in normal thyroid biopsies using data from the TCGA. However, we obtained contradictory results when *PAX-AS1* expression was measured using the FPKM metric (both expression measures are provided in the TCGA dataset). In general, these metrics are well correlated except when the exons are short (as for *PAX8-AS1*). There may be difficulty in accurately assessing the expression with short transcripts due to limitations with the sequencing technology and normalisation methods[60], therefore we present both metrics for consideration. DNA methylation in gene promoters is generally associated with condensed heterochromatin and reduced gene expression[22], while methylation within genes, downstream from the transcription start site, does not have as clear a correlation with gene expression[61]. Our region of interest is intragenic to *PAX8*, but is located within a promoter region of the antisense long non-coding RNA (lncRNA) *PAX8-AS1* (also referred to as *LOC654433*). At loci such as this, it is possible for complex regulatory interactions to exist between DNA methylation, coding gene expression, and non-coding RNA expression[61]. Importantly, a study in fibroblasts demonstrated that the relationship between methylation and gene expression at *PAX8* is region-specific, with a positive correlation between methylation and expression reported at CpGs close to the *PAX8* transcription start site (TSS)

and a strong negative correlation reported at intragenic CpGs located near the *PAX8-AS1* TSS, including one of the CpGs in our region of interest[62]. Furthermore, this study demonstrated a positive correlation between *PAX8* and *PAX8-AS1* gene expression and postulated a potential chromatin-activation linked role of *PAX8-AS1* in regulating *PAX8* expression. Previous studies have demonstrated that polymorphisms in *PAX8-AS1* are potentially expression quantitative trait loci (eQTL) for *PAX8*[63] and therefore it appears that our region could be a putative intragenic regulatory region. The specific biological consequence of the *PAX8-AS1* lncRNA is still yet to be fully elucidated but variants in *PAX8-AS1* are associated with cancer risk[64].

A previous study identified moderate iodine deficiency in Gambian children aged 8-12 years[65] and mothers recruited for the ENID trial whose offspring were followed up for this study were found to have moderate iodine deficiency[66]. We therefore attempted to account for this in our analysis as iodine levels can influence thyroid volume and function. We found no difference between the *PAX8* methylation groups when a thyroglobulin cut-off was used to define iodine insufficiency, but we did find a significant difference between the groups when using UIC and therefore adjusted for this in the analysis. We recognise that a measure of urinary iodine can be a useful tool to understand iodine insufficiency in a population, but may be less useful for characterising an individual's iodine status[67]. However, we note that iodine levels were all taken at the same time of year, all in the morning from fasted individuals and where the local diet day to day is relatively consistent, suggesting that UIC may be informative. It is possible that differential *PAX8* methylation could be contributing to the observed differences in urinary iodine between the two groups possibly by influencing expression of the sodium-iodide symporter.

Prior evidence that *PAX8* methylation is a putative ME, with systemic methylation established in the early embryo, supports the notion that interindividual variation in *PAX8* methylation or related epigenetic marks may drive the phenotypic differences observed in this study. This highlights the potential benefits of studying links between regions of systemic interindividual variation and risk of disease[68]. In this study we cannot rule out the

possibility of a reverse causation effect of thyroid function measures on methylation or other related epigenetic factors at 2 years of age. Elucidation of causal pathways linking environmental exposure, methylation and other epigenetic factors, gene expression and postnatal phenotype will require mechanistic investigations in cell or animal models. For example, a mouse model has demonstrated an association between high oestrogen exposure in early pregnancy and higher free T4, *PAX8* promoter hypomethylation and increased PAX8 expression in offspring[69]. However we note that our region of interest is at least partially absent from several species commonly used as models of development including mouse, zebrafish, and *Xenopus tropicalis*. Furthermore, the *PAX8-AS1* lncRNA present in humans does not appear to exist in other species. Its putative role in regulating *PAX8* expression and thyroid development may therefore only exist in humans.

In summary, we have demonstrated that individual variation in DNA methylation at a region of the *PAX8* gene sensitive to periconceptional nutrition is significantly associated with total thyroid volume and free T4 levels in Gambian children. Our work has potential implications for understanding the foetal origins of health and disease and may contribute to our understanding of the epigenetic drivers of thyroid development and function.

Materials and Methods

Experimental Design

We used a 'recall-by-epigenotype' design to examine links between *PAX8* DNA methylation measured at 2 years of age and thyroid gland function and development in the same Gambian children aged 5-8 years. We used an existing longitudinal cohort of Gambian children (children from the "ENID" (Early Nutrition and Immune Development[37]) study now aged between 5 and 8 years, n=493) to recruit the top ("high") and bottom ("low") quantiles for DNA methylation at a region of the PAX8 gene previously identified as sensitive to the periconceptional environment from banked DNA at 2 years of age. Participants were assessed for thyroid volume, function (free T3, free T4, TSH and Tg), urinary iodine

concentration (UIC), and body composition and bone measures by whole body DXA scan. Figure 1 provides an overview of the study and further details are provided below.

ENID Trial and cohort

The ENID trial (ISRCTN49285450)[37] was a combined trial of pregnancy and infancy nutritional supplementation conducted in West Kiang, a rural region of The Gambia, recruiting pregnant mothers between January 2010 and February 2014. In brief, women were recruited in early pregnancy (10-20 weeks) and randomised to receive either i) Iron-Folate (standard care) ii) multiple micronutrient (MMN) iii) Energy, protein, and lipid with Iron-Folate; or iv) energy, protein, and lipid with MMN supplements for the remainder of their pregnancy. There were no differences in maternal (BMI, age, parity) or infant characteristics (birthweight, birth length, sex or gestational age) across the study arms[70]. From 6 to 18 months of age, infants were further randomized to a lipid-based nutritional supplement, with or without additional MMN. A total of 875 women were randomised in pregnancy to one of the four study arms and 686 participants completed follow-up to 2 years of age[71]. Routine blood samples were collected from mothers and infants at various time points including a peripheral blood cell DNA samples stored at age 2 years. At the time of this current study, children from the ENID study, now aged between 5 and 8 years of age (n=493), were being followed up monthly in a separate longitudinal observational study (see Figure 1).

PAX8 Methylation measurement at 2 years of age

Children from the original ENID trial had DNA isolated from whole blood at age 2 years. DNA was enriched for a panel of candidate regions and bisulfite-converted using a custom Agilent SureSelect Methyl-seq targeted capture system on a subset of these children with sufficient DNA available for processing (n=521)[72]. Target-enriched DNA, including the *PAX8* region of interest, was sequenced using the Illumina NovaSeq platform at the Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA. Reads were mapped to the human genome (hg38) using Bismark v0.20.0[73] with default options, which was also

used to extract methylation values after mapping. Methylation calls from opposite strands of the same CpG site were combined. Within each individual, CpG sites were considered 'covered' if they had a read depth of at least 20x; un-covered sites were excluded from analyses.

Selection of PAX8 region of interest

We selected a subset of CpGs in the *PAX8* region for which there was independent evidence of systemic inter-individual variation and sensitivity to early environment from previous studies[23,25,27]. From these, four CpGs of interest, with coverage in a large number of samples and strong correlation with nearby CpGs, were chosen (see Supplementary Figure 10): chr2:113,235,186; chr2:113,235,228; chr2:113,235,251; and chr2:113,235,267. All of these CpGs lie in intron 9 of *PAX8* and in a promoter region of antisense *PAX8-AS1* (see Figure 2; all genomic coordinates hg38).

Participant selection

From the 493 children being followed up in the ENID cohort at age 5-8yrs (see Figure 1), a "recall by epigenotype" design was used whereby participants were selected by methylation level at CpG chr2:113,235,228. This CpG was chosen due to it having the best overall correlations with nearby CpG methylation levels (see Supplementary Figure 10). Individuals with at least one other informative CpG among the four CpGs of interest and without large differences (≥0.2) in quantile among the CpGs of interest (n=217), were then placed into "high" or "low" groups based on chr2:113,235,228 methylation level (see Figure 3). In total 125 participants (low *PAX8* methylation group n=64, high *PAX8* methylation group n=61) were identified for potential recruitment, with 118 participants (94%) consenting to participate in the study (n=7 declined to participate).

Thyroid volume assessment

Thyroid ultrasound was conducted by TC who was blinded to the participant's *PAX8* methylation group. The length (l), width (w), and depth (d) of each thyroid lobe (in cm) were measured on transverse and longitudinal scans using Sonosite MicroMaxx (10Hz probe). All measurements were made in triplicate and the mean used for analysis. The volume (Vol) of each lobe (in ml) was estimated by the modified formula for an ellipsoid; Vol(ml) = (0.479 x d x w x l)[74] and the total thyroid volume calculated as the sum of the volumes of both lobes.

Mid-childhood blood collection and biochemical assessments

Recruited children had a morning (between 8am – 10.30am) venous blood sample taken into a serum sample collection tube in April 2019. Aliquoted serum was frozen at -70°C. The cellular fraction was discarded. During the same study visit as the blood sample and thyroid ultrasound, a fasted morning urine sample was collected in iodine-free tubes and stored at -20°C. The urine and serum samples were shipped to the University Hospital of Wales, UK for thyroid hormone and urinary iodine measurements. Serum TSH, free T3 and free T4 were measured by the automated ALINITY® System (ABBOTT Laboratories, USA) and Tg was measured by the Beckman Access DxI. Urinary iodine and urinary creatinine were measured by Inductively Coupled Plasma Mass spectrometry. Iodine sufficiency was defined as a Tg <40 μg/L and/or a UIC >100 μg/L[65].

Other phenotypic measures

Standing height was calculated as the mean of measures taken in triplicate to the nearest millimetre using a portable stadiometer (Seca 213). Weight was similarly calculated from measures in triplicate to the nearest 0.1kg using electronic scales (Seca 803), with participants clothed, but with shoes and coat removed. BMI was calculated as weight (kg) divided by height^2 (m^2). HAZ, WAZ and BMI standard deviation score (SDS) for each participant were calculated using WHO reference ranges[75]. A whole body DXA scan was performed using the GE-Lunar Prodigy scanner (GE Medical, Waltham, MA; software version 13.60.033) on 113 children (5 did not attend for DXA scan). Bone related outputs included

areal bone mineral density (BMD, g/cm^2), bone mineral content (BMC, g), and bone area (BA, cm^2). For children, it is recommended to use total body less head (TBLH) for bone-related measurements[76]. The analysed outcome measure for bone was TBLH BMD and was calculated as TBLH BMD (g/cm^3) = TBLH bone mineral content (g)/ TBLH bone area (cm^2). The analysed outcome measure for body fat was FMI, which was calculated as fat mass (kg) divided by height^2 (m^2). The analysed outcome measure for lean mass was LMI, which was calculated as lean mass (kg) divided by height^2 (m^2).

Maternal Biomarker data

Women recruited to the ENID trial provided a 10ml venous blood sample at the time of enrolment in early pregnancy (<20 weeks gestation). Plasma was stored at -70^{o}C. A subset of 350 women were previously selected for biomarker analysis on the stored samples; n=303 women with maternal biomarker data had *PAX8* methylation data available for their offspring. These women were selected to give an even distribution by month of enrolment and to provide data from the earliest gestational age i.e. sample collection nearest to time of conception (median 12.1 weeks, IQR 3.5). In our analysis we examined a core set of ten nutritional biomarkers involved in one-carbon metabolism (homocysteine, methionine, cysteine, choline, betaine, dimethylglycine (DMG), vitamin B12, folate, pyridoxal-5'-phosphate (PLP, a vitamin B6 vitamer), and vitamin B2 (riboflavin)). We also considered an extended panel of biomarkers to capture other metabolic pathways potentially influencing one carbon metabolism (alpha-1-acid glycoprotein (AGP), aspartate, threonine, serine, glutamate, glycine, alanine, valine, isoleucine, leucine, tyrosine, phenylalanine, lysine, histidine, arginine, proline, uridine and uracil). Biomarkers were analysed at the BC Children's Hospital, Canada, using liquid chromatography-tandem mass spectrometry (choline, betaine, DMG, homocysteine, cysteine, methionine, PLP, riboflavin, uracil, uridine), Abbott AxSYM autoanalyzer (folate, vitamin B12) and Hitachi L-8900 amino acid analyser (additional amino acids: serine, glycine, alanine, arginine, aspartic acid, glutamic acid, histidine, isoleucine, leucine, lysine, phenylalanine, proline, threonine, tryptophan, tyrosine and valine). The inflammatory marker AGP was measured using the Cobas Integra 400 plus autoanalyser at the MRC Unit The Gambia, Keneba field station.

Stability of PAX8 methylation from mid-childhood to after puberty/adolescence

Stability of the four selected *PAX8* CpGs with age was assessed by analysing blood samples from 49 Gambians with samples collected in mid-childhood (7 years old) and adolescence (17 years old)[25], by pyrosequencing (see Supplementary Methods).

PAX8 gene expression and DNA methylation in blood and thyroid tissues

We accessed data from The Cancer Genome Atlas (TCGA), an open access resource providing genomic data for over 20,000 cancer and non-cancerous matched samples obtained from biopsy from live individuals. We downloaded *PAX8* gene expression data on 50 individuals with non-cancerous thyroid samples, reported as both fragments per kilobase per million mapped reads – upper quartile (FPKM-UQ; from GDC v18.0 PANCAN HTSeq hosted by Xena, https://gdc.xenahubs.net) and RNAseq by Expectation-Maximization (RSEM; data from TCGA Wanderer, http://maplab.imppc.org/wanderer/) from Illumina HiSeq RNA-seq data. TCGA methylation data for the same samples measured on the Illumina HumanMethylation450 array was available for two of our CpGs of interest (chr2:113,235,186 and 113,235,267) from TCGA Wanderer.

A comparison of *PAX8* methylation levels in blood and thyroid tissue was carried out through quantitative analysis of the four *PAX8* CpGs of interest by bisulfite pyrosequencing in paired thyroid tissue and whole blood samples from 86 donors obtained from the GTEx post mortem tissue database[38] (see Supplementary Methods for sample details).

Genotyping from methyl-seq data

We used the targeted methyl-seq data to call individual genotypes within the *PAX8* gene. A modified version of BS-Snper[77,78] was used to call SNP genotypes from the bisulfite-converted reads. Genotypes with a non-reference allele frequency (AF) of >=0.05 and with a call rate (proportion of individuals with called genotypes) of 20% or higher were considered for the analysis (n=14 SNPs). Of these, SNP rs10193733 (chr2:113,235,047 T>C; MAF 0.25)

was the closest to the *PAX8* variably methylated region analysed here (see Figure 2), and had a high call rate in the methyl-seq dataset (99%), enabling analysis of interactions between methylation, genotype, and thyroid phenotypes. Furthermore, we verified that this SNP is in linkage disequilibrium (LD) with several other SNPs on either side of the variably-methylated region using Gambian reference data from the Gambian Genome Variation Project (GGVP)[79] and 1000 Genomes Project, Phase3 (Gambia West Division - GWD)[80], and confirmed this in our study population (see Supplementary Figure 11). This SNP, and the nearby SNPs in LD all had similar AFs to the GWD reference population from GGVP, demonstrating the reliability of the methyl-seq-derived genotypes (with the exception of rs7576384 chr2:113235808 C>G, which had a much lower AF than expected, likely due to the loss of informative reads after bisulfite conversion). A 2 bp indel polymorphism not detectable by the genotype caller, rs35724515 (chr2:113235224 CCC>C), is located within our region of interest and is in high LD with rs10193733 (1000 Genomes GWD LD $r^2 = 0.976$).

Statistical analysis

Baseline characteristics and crude, between-group comparisons

Differences between high and low *PAX8* methylation groups for the categorical variables of sex, pregnancy and infant supplementation, iodine sufficient (yes/no) were assessed using Pearson chi-squared tests. Birth weight, WAZ, BMI z-score, HAZ, total thyroid volume, free T4 and free T3 were normally distributed and group differences were assessed using student t-test with mean and SD (standard deviation) reported. Age, UIC, urinary iodine:creatinine ratio, TSH and Tg were not normally distributed and group differences were assessed by Mann-Whitney U tests with median and IQR (interquartile range) reported.

Multiple linear regression models

Unless otherwise indicated, all outcome variables were normally distributed.

Thyroid volume as the dependent variable was regressed against *PAX8* methylation group and adjusted for sex, age, BMI, and UIC. BMI was included to ensure an adjustment for body size as i) there was a significant difference between the two groups at recruitment, and ii) BMI also provided the best model fit (indicated by lowest Akaike Information Criterion) compared to other measures of body size (e.g. WAZ, HAZ, body surface area). UIC was included as there was a significant difference between the groups at recruitment and iodine deficiency can be associated with goitre/thyroid size.

Measures of thyroid function (free T4, free T3, TSH and Tg) were regressed against *PAX8* methylation group and adjusted for sex, age, and UIC. TSH and Tg were normally distributed after log transformation prior to regression analysis. Iodine deficiency can be associated with hypothyroidism and an increase in Tg and therefore UIC was included in all models. Correction for multiple tests was not made as the markers of thyroid function are not independent.

DXA-derived measures: FMI was normally distributed after log transformation. FMI was regressed against free T4 and adjusted for age, sex, and weight (to adjust for lean mass). LMI was regressed against free T4, with LMI adjusted for age, sex, and weight (to adjust for fat mass). TBLH BMD was normally distributed after log transformation and regressed against free T4, adjusted for age, sex, weight, and height.

Predictors of PAX8 methylation analysis

In analyses where *PAX8* methylation was considered as an outcome, 2-year methylation measurements were taken from the wider ENID cohort (i.e. not just the high and low *PAX8* methylation groups; n=303-521 depending on the predictor considered), and treated as a continuous variable. Methylation across the four CpGs of interest was highly correlated (Pearson R between 0.871 – 0.952; Supplementary Figure 10). A univariate composite measure was therefore used in all regression models, calculated as the mean z-score (over all CpG sites) of the logit-transformed methylation level at each CpG site (referred to as

PAX8 mean logit methylation z-score). Methylation levels of 1.0 were reduced to 0.99 to prevent infinite values after logit transformation.

All maternal biomarkers were preadjusted for gestational age, maternal BMI, maternal age and inflammation (AGP) and then back extrapolated to date of conception using previously described methods[29]. Those biomarkers not normally distributed were log-transformed and all biomarkers were scaled and centred to enable comparison of standardised coefficients. Multiple linear regression models with *PAX8* mean logit methylation z-score as the dependent variable were fitted individually with each biomarker as a predictor and adjusted for sex.

For other predictors, *PAX8* mean logit methylation z-score was regressed against maternal BMI, infant BMI z-score, infant WAZ, infant sex and season of conception (SoC) in separate models. SoC was defined as 'rainy' (January-June) and 'dry' (July-December) as previously described[81] with the conception date calculated from a gestational age estimation obtained from antenatal ultrasound at ENID trial recruitment.

All model covariates were assessed for multicollinearity, and standard tests were performed to ensure that linear modelling assumptions were met. Where reported, coefficients (β) associated with log transformed dependent variables were back transformed using (exp(β) - 1) x 100, to represent percentage change in dependent variable per unit increase in the corresponding predictor.

Causal mediation analysis

We performed a causal mediation analysis to test the hypothesis that the observed effect of *PAX8* methylation group on free T4 is mediated by its effect on thyroid volume. We used the Mediation package (v4.5.0) in R, with confidence intervals for direct and indirect effects calculated using a non-parametric bootstrap with 10,000 simulations.

All statistical analysis were performed using R version 3.6.2[82].

Analyses by genotype

Genetic analyses focused on a single SNP (rs10193733) that tagged an LD block proximal to our *PAX8* region of interest (see above for justification for choosing this SNP). We performed two sets of analyses stratified by rs10193733 genotype. First, student t-tests were used to compare the *PAX8* mean logit methylation z-score between SoC (rainy vs dry) separately for individuals with each genotype. Second, we fitted multiple linear regression models with total thyroid volume and free T4 as dependent variables adjusted for relevant co-variables, again stratified by rs10193733 genotype. We also assessed the influence of the rs10193733 'C' allele on free T4 and thyroid volume (outcome variables) in multiple regression models with allelic dosage (predictor) coded as C/C=2, C/T=1, T/T=0.

Ethics

Ethical approvals for the ENID trial and for this study were given by The Gambia Government/MRC Joint Ethics Committee (SCC1126v2, SCC1640, L2015.51v2). Consent was gained by signature or thumb print from mothers for their own participation and that of their child. All data were anonymized before analysis.

Acknowledgments

We thank the women and children of West Kiang who patiently participated in the ENID trial and the participants in this study. We acknowledge the enthusiastic work of the study teams, especially the fieldworkers, village assistants, midwives, clinical staff, data office staff and laboratory technicians who collected the data and samples. We would like to thank Dr Ato Quansah (Consultant Radiologist, Komfo Anokye Teaching Hospital, Ghana), Dr Philippa Lewis (Royal College of Radiologists, UK), Dr Joanna Kasznia-Brown (Royal College of Radiologists, UK) and Dr Elizabeth Joekes (World Wide Radiology) for support in arranging thyroid ultrasound training for TC in The Gambia. We would like to thank Dr Carol Evans and team at the University Hospital of Wales for their work processing the thyroid function tests; Yumei Li and Rui Chen (Human Genome Sequencing Center, Baylor College of Medicine) for performing and troubleshooting the target capture and Illumina sequencing; Professor Ann Prentice and her team (Head of Calcium, Vitamin D, Bone Health group at

University of Cambridge and MRC Unit in the Gambia at LSHTM) for facilitating the use of the DXA facilities at Keneba Field Station; Mustapha Ceesay, Michael Mendy, Mariama Jammeh and Kissima Sawo at MRC The Gambia, Keneba Field station for performing the DXA scans; Ebrima Sise (MRC The Gambia at LSHTM) who performed the AGP analysis in Keneba; Professor Anne Ferguson-Smith for discussions regarding the functional implications of indel and SNP variants and Dr Peter Kuhnen for discussions regarding *PAX8* methylation and CHT. This publication uses data from the Gambian Genome Variation Project, a collaboration between the MRC Unit in the Gambia (http://www.mrc.gm), the Wellcome Sanger Institute (https://www.sanger.ac.uk), the MRC Centre for Genomics and Global Health (https://www.cggh.org/collaborations/mrc-unit-the-gambia) at the University of Oxford, and the MalariaGEN Resource Centre (https://www.malariagen.net). The *PAX8* and *PAX8-AS1* methylation and expression results here are in part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Funding

TC is funded by an MRC Clinical Research Training Fellowship (MR/S006516/1). The British Society of Paediatric Endocrinology and Diabetes provided funding for this work to TC through their Research and Innovation Award 2018.

The Gambian studies are supported by core funding from the UK Medical Research Council (MRC) MC-A760-5QX00 and by MRC grant MR/ M01424X/1. Programme funding for Professor Ann Prentice supported data acquisition and analysis of the whole body DXA scans (U105960371, U123261351).

Author contributions

Conceptualization: TC, ML, AMP, MJS

Methodology: TC, NJK, KAM, DJ, MJS

Investigation: TC, CJG, PJ, EL, MSB, RD, RE, RAW, SEM

Formal analysis: TC, NJK

Visualization: TC, NJK, CJG, RAW, AMP, MJS

Supervision: ML, AMP, MJS

Writing—original draft: TC, MJS

Writing—review & editing: all authors reviewed and contributed to the final manuscript

Competing interests: Authors declare that they have no competing interests

Data and materials availability: Data from TCGA (https://www.cancer.gov/tcga) and GTEx (https://www.gtexportal.org/home/) used in this analysis are publicly available. NJK's code for the modified version of BS-Snper can be found at https://github.com/knowah/bssnper2/tree/v0.1. Other Gambian data are available on request.

Figures and Tables

	All (n=118)	High *PAX8* methylation (n=58)	Low *PAX8* methylation (n=60)	p-value
Sex				
Male	71	30	41	0.10
Female	47	28	19	
Age (years)				
Median [IQR]	7.18 [1.67]	7.28 [1.49]	7.09 [1.61]	0.29
Range	5.22 to 8.71	5.22 to 8.68	5.23 to 8.71	
Body size measures[$]				
Birth Weight (kg) [SD]	3.03 [0.43]	3.01 [0.46]	3.05[0.41]	0.62
Mean WAZ [SD]	-1.25 [0.85]	-1.18 [0.87]	-1.32 [0.82]	0.36
Mean HAZ [SD]	-0.75 [0.83]	-0.81 [0.86]	-0.70 [0.79]	0.43
Mean BMI Z Score [SD]	-1.20 [0.87]	-1.01 [0.87]	-1.38 [0.84]	**0.02**
ENID Pregnancy Supplementation				
PE	31	16	15	0.90
FeFol	28	12	16	
MMN	26	13	13	
PE & MMN	33	17	16	
ENID Infant supplementation				
MMN	54	27	27	1.00
No MMN	64	31	33	

Table 1. Baseline characteristics between high and low *PAX8* methylation groups. Group differences in normally distributed variables (WAZ, HAZ, BMI) assessed by student t-test, non-normally distributed variables (age) by Mann-Whitney U test and categorical variables (sex, ENID supplementation group) by chi-squared test. [$]Body size measures reported were measured at recruitment for this current study (i.e. between 5 and 8 years of age).

Key: IQR = interquartile range, SD = standard deviation, WAZ = weight-for-age z-score, HAZ = height-for-age z-score, BMI = body mass index, ENID = Early Nutrition and Immune Development Trial, PE = Protein Energy supplementation, FeFol = Iron and Folate supplementation, Multiple micronutrient supplementation = MMN.

	All	High *PAX8* methylation	Low *PAX8* methylation	p-value
Thyroid Volume				
Mean total thyroid volume [SD] cm^3	3.06 [0.93] n=112	2.87 [0.83] n=53	3.24 [0.98] n=59	**0.035**
Thyroid function				
Mean free T4 [SD] pmol/L	13.6 [1.34] n=114	13.3 [1.33] n=55	13.9 [1.29] n=59	**0.009**
Mean free T3 [SD] pmol/L	6.09 [0.68] n=114	6.08 [0.63] n=55	6.09 [0.74] n=59	0.92
Median TSH [IQR] mU/L	1.82 [1.01] n=116	1.86[0.92] n=56	1.73 [1.08] n=60	0.34
Median Tg [IQR] ug/L	19.2 [10.70] n=105	18.5 [7.80] n=49	19.4 [15.1] n=56	0.46

Table 2. Thyroid volume and function comparison by *PAX8* methylation group; crude (unadjusted) analyses.

Group differences in normally distributed variables (thyroid volume, free T4, free T3) were assessed by student t-test and non-normally distributed variables (TSH, Tg) by Mann-Whitney U test.

Key: IQR = interquartile range, SD = standard deviation, Free T4 = free thyroxine, free T3 = free tri-iodothyronine, TSH = thyroid stimulating hormone, Tg = thyroglobulin.

Outcome	Number of individuals	Low *PAX8* methylation group coefficient [standard error]	P-value
Total Thyroid Volume (cm^3)	112	0.61 [0.15]	**0.0001**
Free T4 (pmol/L)	113	0.85 [0.24]	**0.0007**
Free T3 (pmol/L)	113	-0.02 [0.13]	0.88
Log TSH (mU/L)	115	-0.11 [0.09]	0.26
Log Tg (μg/L)	94	0.02 [0.12]	0.85

Table 3. *PAX8* methylation group as a predictor of thyroid volume and function; multiple linear regression (adjusted) analyses.

Low *PAX8* methylation group coefficient gives the mean increase (decrease if negative) relative to the high *PAX8* methylation group. *Adjustment covariates for the multiple linear regression model for total thyroid volume are age, sex, BMI, urinary iodine (see Supplementary Table 2). Adjustment covariates for multiple linear regression model for free T4, free T3, logTSH, logTg are age, sex, urinary iodine (see Supplementary Table 3).*

Key: Free T4 = free thyroxine, free T3 = free tri-iodothyronine, TSH = thyroid stimulating hormone, Tg = thyroglobulin

Outcome	Number of individuals	Free T4 (pmol/L) coefficient [standard error]	% change in outcome per one unit increase in free T4[$]	P-value
Log FMI (Fat Mass (Kg)/m^2)	113	-0.04 [0.02]	-4.30	**0.033**
Log TBLH BMD (g/cm^2)	113	-0.008 [0.004]	-0.80	**0.044**

Table 4. Free T4 as a predictor of FMI and BMD; multiple linear regression (adjusted) analyses. *Adjustment covariates for the multiple linear regression model for logFMI are age, sex, weight [see Supplementary Table 4]. Adjustment covariates for multiple linear regression models for log TBLH BMD and log TBLH BMC are age, sex, height, and weight [see Supplementary Table 6].*

Key: FMI = fat mass index, TBLH = total body less head, BMD = bone mineral density, $ back transformation if outcome log transformed using [exp[β]- 1] x 100

Covariate	Standardised Coefficient	Standard Error	t-value	p-value
Hcy (µmol/L)	-0.11	0.05	-1.99	**0.048**
Methionine (µmol/L)	-0.01	0.06	-0.13	0.89
Cysteine (µmol/L)	-0.16	0.05	-2.95	**0.003**
Choline (µmol/L)	-0.01	0.06	-0.20	0.84
Betaine (µmol/L)	-0.07	0.05	-1.29	0.20
DMG (µmol/L)	-0.03	0.06	0.52	0.61
B12 (pmol/L)	-0.10	0.05	-1.96	**0.050**
Folate (nmol/L)	-0.07	0.05	-1.37	0.17
PLP (nmol/L)	-0.12	0.06	-2.17	**0.031**
Riboflavin (nmol/L)	-0.11	0.06	-1.93	0.055

Table 5. Maternal one-carbon metabolite biomarkers as predictors of *PAX8* Methylation.

Multiple linear regression models have *PAX8* mean logit methylation z-score as the dependent variable with maternal one-carbon metabolites measured in maternal plasma and back-extrapolated to the time of conception as predictors, adjusted for sex.

Key: Hcy = Homocysteine, DMG = dimethylglycine, PLP = Pyridoxal 5-phosphate (B6 vitamer)

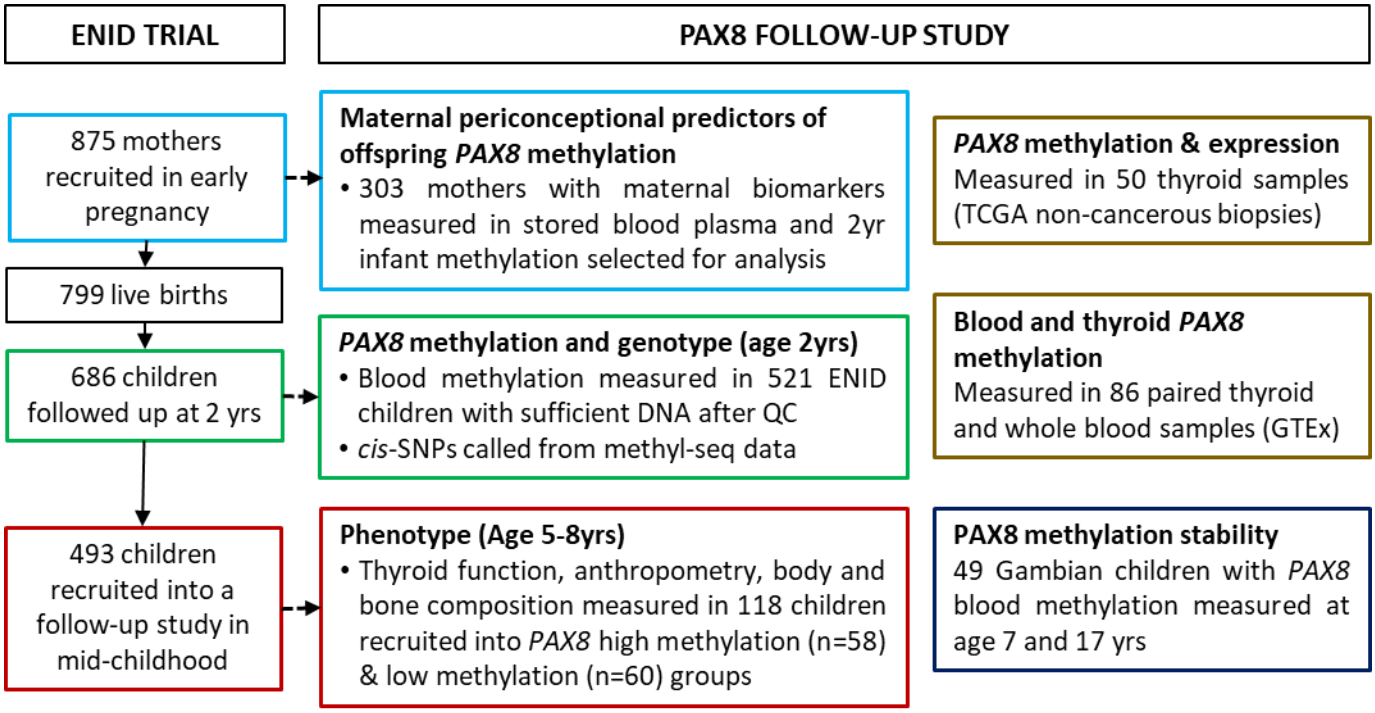


Figure 1. Study overview.

Key: ENID =Early Nutrition and Immune Development, QC = Quality Control, TCGA = The Cancer Genome Atlas, GTEx = Genotype-Tissue Expression biobank.

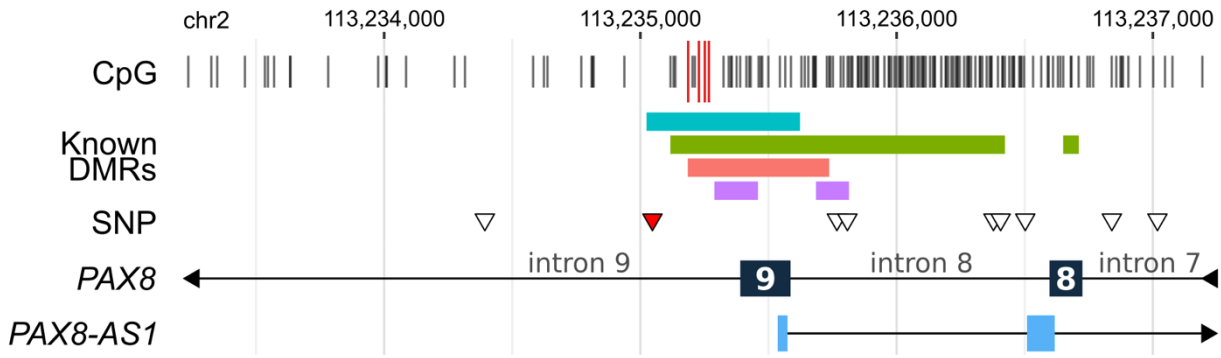


Figure 2. *PAX8* region of interest.

The *PAX8* gene extends from chr2:113,215,997-113,278,921 (hg38) and contains 12 exons and 11 introns. Exons 8 and 9 of *PAX8* are shown, as well as the first two exons of an isoform of the *PAX8-AS1* antisense lncRNA. CpGs are shown in the top track. The four CpGs highlighted in red were analysed in this study (chr2:113,235,186; 113,235,228; 113,235,251; and 113,235,267). The 'Known DMRs' track highlights regions identified in the following studies: putative metastable epialleles displaying systemic interindividual variation identified in Silver et al[25] and Kessler et al[23] (blue and green regions); DMR associated with gestational famine[27], Gambian season of conception and maternal folic acid supplementation[35] (pink region); and DMRs associated with Gambian season of conception in an additional study[31] (purple region). See Methods and Supplementary Figure 10 for further details on how the 4 CpGs analysed in this study were selected. The SNP track denotes variants within 2000bp of the CpGs of interest that were called from the methyl-seq data. The SNP highlighted in red is close to our region of interest and tags an LD block encompassing it (see Supplementary Figure 11). This is the SNP used for the genotype analyses (rs10193733; chr2:113,235,047 T>C).

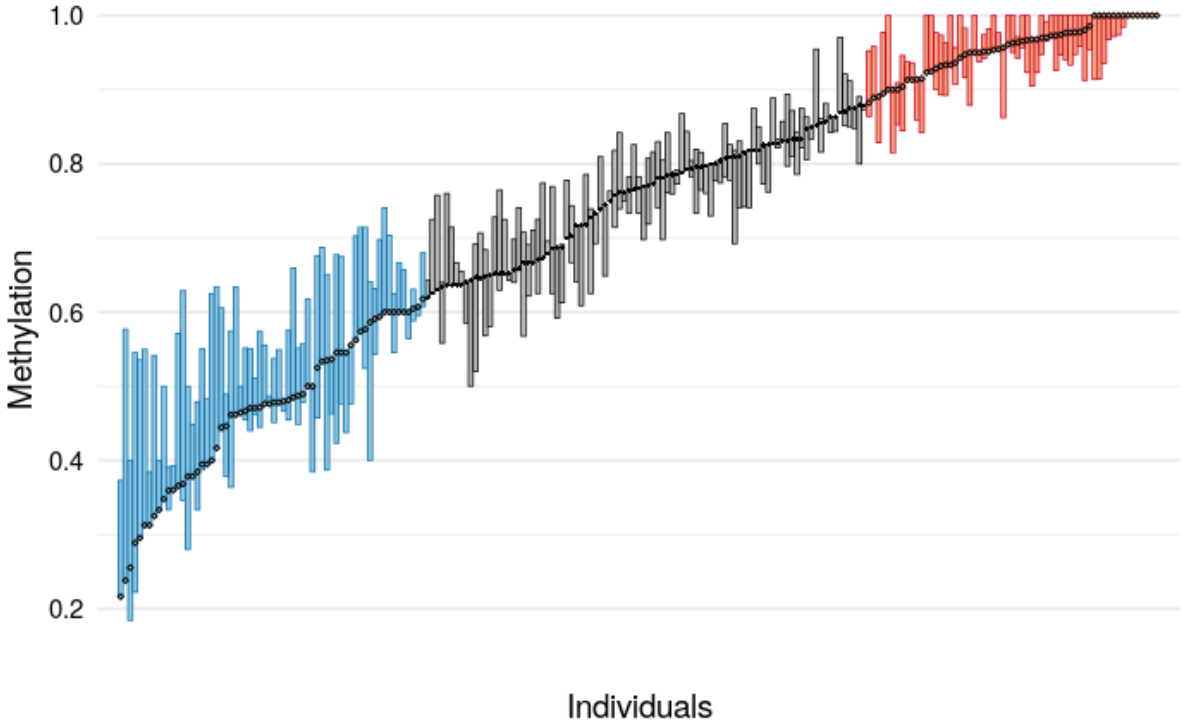


Figure 3. DNA methylation in the study population at the *PAX8* CpGs of interest.

Individuals were sorted by methylation level at CpG chr2:113,235,228 (indicated by dots), which was used to select the high (red) and low (blue) *PAX8* methylation groups. Methylation range across the four CpGs of interest is indicated by the boxes.

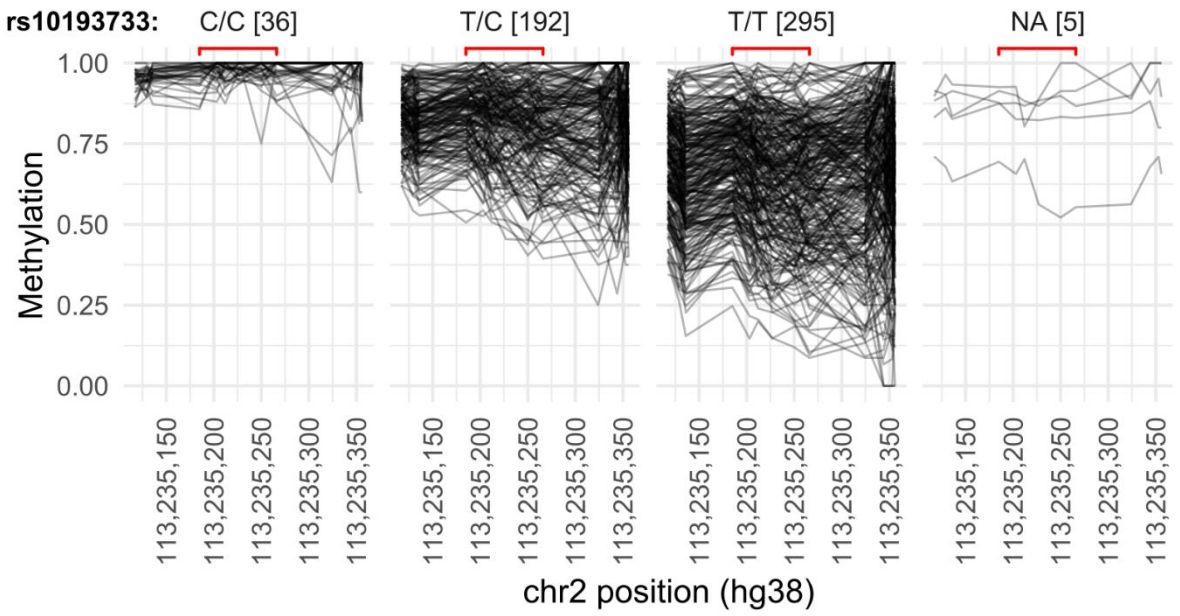


Figure 4. C alleles at rs10193733 are associated with decreased DNA methylation variation at the *PAX8* region of interest.

DNA methylation at CpGs is shown for the *PAX8* region of interest (delimited by the red bracket) and flanking 200bp region, split by rs10193733 genotype. Numbers in brackets denote the counts of individuals with each genotype at the SNP (total N=528). Five individuals (1%) did not have sufficient unambiguous methyl-seq reads to call a genotype and are marked as NA.

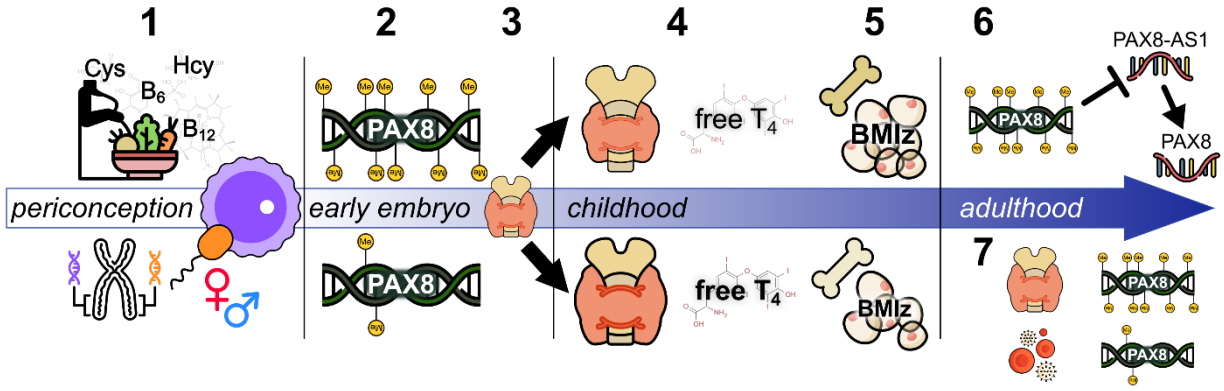


Figure 5. Proposed model and summary of evidence linking *PAX8* methylation and expression, thyroid function and development, and phenotype.

1. ***PAX8* methylation is influenced by periconceptional nutritional factors, sex, and genotype.** Previous studies have reported links between season of conception[25,31], prenatal famine exposure[27], periconceptional micronutrient supplementation[36] and *PAX8* methylation. We present evidence of associations between *PAX8* methylation and maternal biomarkers back extrapolated to conception (homocysteine, cysteine, B12 and pyridoxal 5-phosphate (B6 vitamer)). We also demonstrate associations between *PAX8* methylation and sex, and between *cis* genotype and *PAX8* methylation variability.

2. ***PAX8* methylation is set early in embryonic development.** Evidence from Gambian[25,29] and European children[32]; and from Vietnamese [29,31] and Caucasian[23] adults that *PAX8* methylation is concordant across germ layers suggests that methylation is set early in embryological development (before gastrulation).

3. ***PAX8* methylation sets a trajectory for thyroid gland development**. *PAX8* is expressed from gestational day 20-22 in humans, around the same time that thyroid progenitor cells begin specification in the endoderm[15]. *PAX8* methylation state (or related epigenetic factors) alters *PAX8* expression which influences thyroid gland development.

4. ***PAX8* methylation is inversely associated with thyroid size and free T4 in mid-childhood.** We show, with a large effect size, that *PAX8* methylation at age 2 is inversely associated with thyroid gland size and free T4 in Gambian children.

5. **Free T4 is inversely associated with fat and bone mineral density in mid-childhood.** We show that free T4 is inversely associated with all measures of overall fat mass and bone mineral density in Gambian children.

6. ***PAX8* methylation is inversely associated with PAX8-AS1 expression.** We present evidence from public data that CpG methylation in our region of interest is inversely correlated with PAX8-AS1 expression in thyroid. There is evidence that PAX8-AS1 may have a regulatory role on PAX8 expression[62,63]. Our region of interest is an intragenic putative regulatory region of *PAX8* which may act as a promoter for *PAX8-AS1*.

7. **There is a divergence between blood and thyroid methylation in adults.** We present evidence that blood and thyroid methylation at the *PAX8* gene is not correlated in adults. This suggests that the systemic methylation concordance between thyroid and blood is not maintained into adulthood.

Science Advances
AAAS

Supplementary Materials for

DNA methylation at a nutritionally sensitive region of the *PAX8* gene is associated with thyroid volume and function in Gambian children

T Candler, NJ Kessler, CJ Gunasekara, KA Ward, P James, E Laritsky, MS Baker, R Dyer, R Elango, D Jeffries, RA Waterland, S Moore, M Ludgate, AM Prentice, MJ Silver*

*Corresponding author. Email: matt.silver@lshtm.ac.uk

This PDF file includes:

Supplementary Tables

Study	Exposure	Timing of prenatal exposure	Offspring *PAX8m* methylation outcome	Age of offspring's *PAX8m* measurement	Methylation platform
Waterland et al, 2010[1]	Gambian Season of Conception	Periconceptional	*Rainy season conceptions associated with ↑ leucocyte *PAX8*m	Children (~9 years of age)	Pyrosequencing
Silver et al, 2015[2]	Gambian Season of Conception	Periconceptional	*Rainy season conceptions associated with ↑ leucocyte *PAX8*m	Infants (~3 months of age)	Illumina Infinium Human Methylation 450 array
Finer et al, 2016[3]	Famine Exposure in Bangladesh	At least 7 months of pregnancy	Gestational famine exposure associated with ↑ leucocyte *PAX8*m	Adults (27-32 years of age)	Illumina Infinium Human Methylation 450 array
Richmond et al, 2018[4]	Maternal folic-acid supplementation in pregnancy	Folic acid supplementation or placebo given between ~17 weeks GA (mean GA at study recruitment) to ~40 weeks GA	Maternal folic acid supplementation associated ↓salivary *PAX8*m	Adults (46-48 years of age)	Illumina Infinium Human Methylation 450 array
Saffari et al, 2020[5]	Micronutrient supplementation (UNIMMAP) in pregnancy	Preconception to positive pregnancy test	Micronutrient supplementation associated with ↑ leucocyte *PAX8*m	Children (~9 years of age)	Pyrosequencing

Supplementary Table 1. Summary of previous studies investigating associations between prenatal environmental or nutritional exposures and *PAX8* methylation. *Rainy season conceptions occur in the context of lower maternal calorific intake and altered concentrations of circulating C1 metabolites[6,7].

Key: PAX8m = *PAX8* gene methylation, C1 = one-carbon, GA=gestational age, UNIMMAP= United Nations International Multiple Micronutrient Antenatal Preparation

	Total Thyroid Volume (cm^3)
Age (Years)	0.38*** [0.07]
Sex (Male)	-0.32* [0.16]
BMI Z score	0.22* [0.09]
***PAX8* Methylation Category (Low)**	0.61*** [0.15]
Urinary Iodine (μg/L)	-0.0002* [0.0001]

Supplementary Table 2. Multiple linear regression models for predictors of total thyroid volume.

Coefficients from regression model are given with [standard error]. *p<0.05; **p<0.01; ***p<0.001.

Key: BMI = body mass index

Iodine Status				
	All	**High *PAX8* methylation**	**Low *PAX8* methylation**	**p-value**
Median UIC [IQR] ug/L	152 [141] n=116	128 [161] n=56	170 [140] n=60	**0.04**
Median urinary iodine:creatinine ratio [IQR] nmol/mmol	189 [113] n=116	190 [100.96] n=56	186.71 [153.79] n=60	0.65
Iodine insufficient by Tg level	9 [8.6%]	1 [2%]	8 [14.3%]	0.06
Iodine sufficient by Tg level	96 [91.4%]	48 [98%]	48 [85.7%]	
Iodine insufficient by UIC	31 [26.7%]	21 [37.5%]	10 [16.7%]	**0.02**
Iodine sufficient by UIC	85 [73.3%]	35 [62.5%]	50 [83.3%]	

Supplementary Table 3. Iodine status by *PAX8* methylation groups

Group differences in non-normally distributed variables (TSH, Tg, urinary iodine, urinary iodine:creatinine ratio) were assessed by Mann-Whitney U test, categorical variables (iodine sufficiency categories) by chi-squared test.

Key: Tg = thyroglobulin, UIC = Urinary Iodine concentration. Iodine insufficiency by Tg is defined as a Tg level >40ug/L. Iodine insufficiency by UIC is defined as a UIC <100ug/L.

	Free T4 (pmol/L)	Free T3 (pmol/L)	Log TSH (mU/L)	Log Tg (μg/L)
Age (Years)	-0.04 [0.12]	-0.19** [0.06]	-0.02 [0.05]	-0.05 [0.06]
Sex (Male)	-0.73** [0.25]	-0.08 [0.13]	-0.1* [0.10]	-0.001 [0.12]
***PAX8* Methylation Category (Low)**	0.85*** [0.24]	-0.02 [0.13]	-0.11 [0.09]	0.02 [0.12]
Urinary Iodine (μg/L)	-0.0002 [0.0001]	0.00003 [0.0001]	0.0001 [0.00004]	-0.0001** [0.0001]

Supplementary Table 4. Multiple linear regression models for predictors of free T4, free T3, TSH and Tg.

Coefficients from regression models are given with [standard error]. *p<0.05; **p<0.01; ***p<0.001.

Key: Free T4 = free thyroxine, free T3 = free tri-iodothyronine, TSH = thyroid stimulating hormone, Tg = thyroglobulin.

	Log FMI (fat mass in kg)/m²
Age (Years)	-0.14*** [0.04]
Sex (Male)	-0.53*** [0.06]
Weight (Kg)	0.06*** [0.01]
Free T4 (pmol/L)	-0.04* [0.02]

Supplementary Table 5. Multiple linear regression models for predictors of fat measures as assessed by DXA.

Coefficients from regression model are given with [standard error]. *p<0.05; **p<0.01; ***p<0.001.

Key: FMI = fat mass index

	LMI (Lean Mass (Kg)/m^2)
Age (Years)	-0.21* [0.08]
Sex (Male)	0.82*** [0.13]
Weight (Kg)	0.19*** [0.03]
Free T4 (pmol/L)	-0.02 [0.05]

Supplementary Table 6. Multiple linear regression models for predictors of lean measures as assessed by DXA.

Coefficients from regression model are given with [standard error]. *p<0.05; **p<0.01; ***p<0.001.

Key: LMI = lean mass index

	Log TBLH BMD (g/cm^2)
Sex (Male)	-0.01 [0.01]
Age (Years)	0.02 [0.01]
Height (cm)	0.003 [0.002]
Weight (Kg)	0.01* [0.003]
Free T4 (pmol/L)	-0.008* [0.004]

Supplementary Table 7. Multiple linear regression models for predictors of BMD as assessed by DXA.

Coefficients from regression model are given with [standard error]. *p<0.05; **p<0.01; ***p<0.001.

Key: TBLH = total body less head, BMD = bone mineral density.

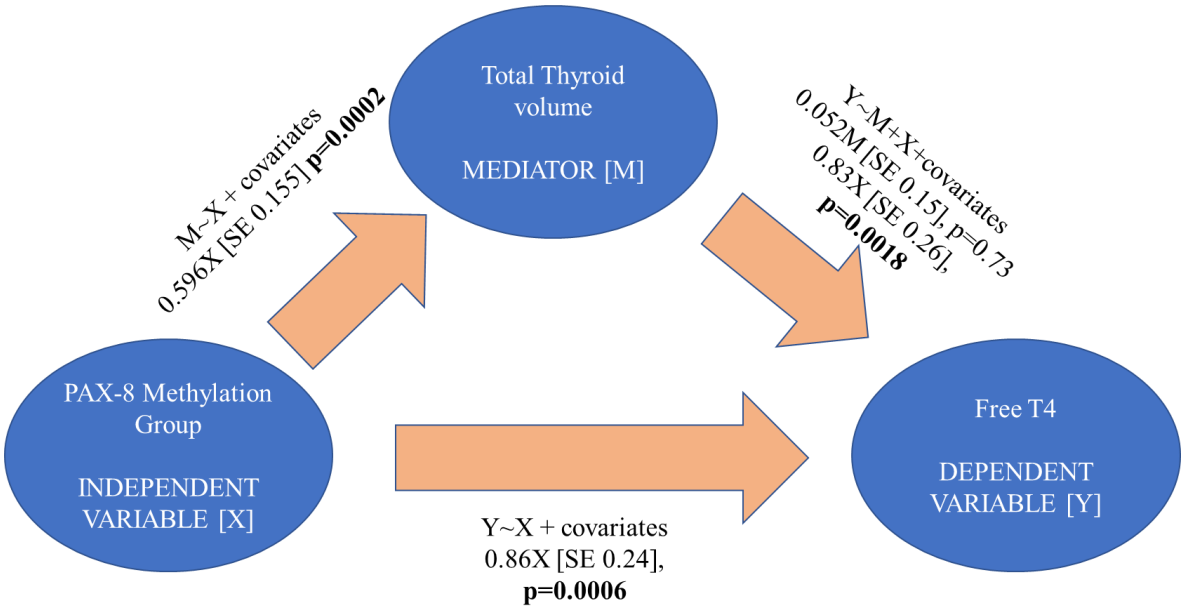
Covariate	Standardised Coefficient	Standard Error	t-value	p-value
AGP (g/L)	0.02	0.05	0.42	0.63
Aspartate (µmol/L)	0.05	0.06	0.91	0.37
Threonine (µmol/L)	-0.05	0.06	-0.84	0.40
Serine (µmol/L)	-0.08	0.06	-1.40	0.16
Glutamate (µmol/L)	0.08	0.06	1.38	0.17
Glycine (µmol/L)	-0.05	0.06	-0.93	0.35
Alanine (µmol/L)	0.05	0.06	0.79	0.43
Valine (µmol/L)	-0.13	0.05	-2.37	**0.02**
Isoleucine (µmol/L)	0.01	0.06	-0.26	0.80
Leucine (µmol/L)	0.02	0.06	0.32	0.75
Tyrosine (µmol/L)	0.04	0.06	0.79	0.43
Phenylalanine (µmol/L)	0.10	0.06	1.78	0.08
Lysine (µmol/L)	-0.01	0.06	-0.20	0.85
Histidine (µmol/L)	0.03	0.06	0.55	0.58
Arginine (µmol/L)	0.16	0.06	2.77	**0.006**
Proline (µmol/L)	0.04	0.06	0.72	0.47
Uridine (µmol/L)	-0.04	0.06	-0.78	0.43
Uracil (nmol/L)	0.21	0.05	3.87	**0.0001**

Supplementary Table 8. Linear regression models for additional nutritional predictors of *PAX8* Methylation.

Dependent Variable = *PAX8* Mean Logit Methylation z-score. All biomarkers measured in maternal plasma and back-extrapolated to conception and adjusted for sex.

Key: AGP = Alpha-1-acid glycoprotein.

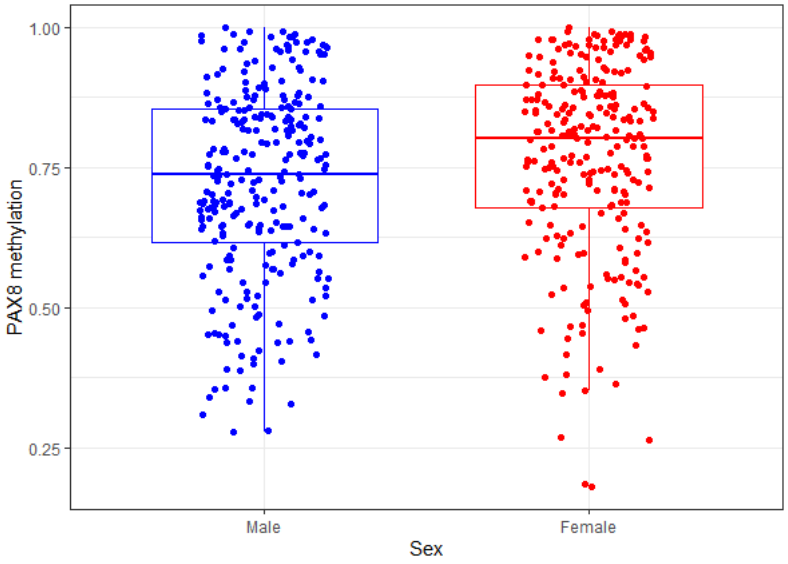
Supplementary Figures



	Estimate	95% CI Lower	95% CI Upper	p-value
Average causal mediation effect	0.03	-0.15	0.22	0.74
Average direct effect	0.83	0.34	1.34	**0.001**
Total effect	0.86	0.39	1.35	**0.0002**
Proportion mediated (%)	0.035	-0.22	0.30	0.74

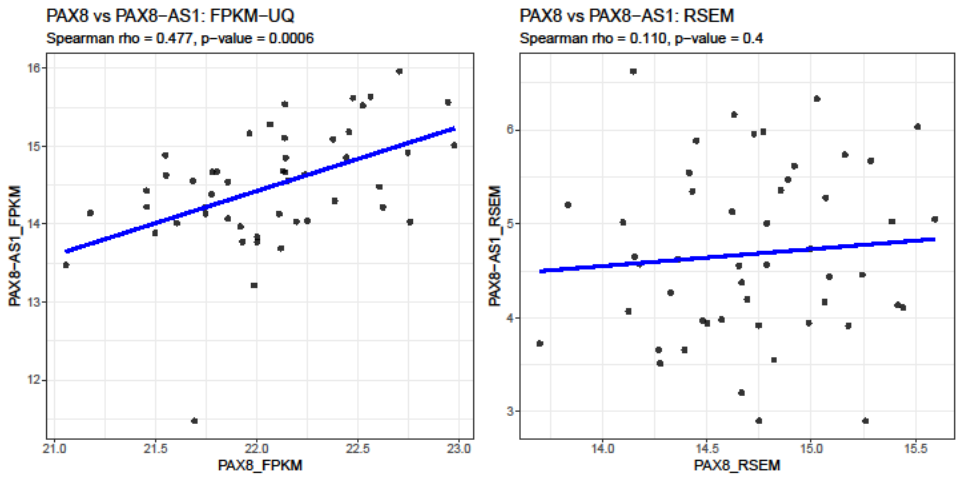
Supplementary Figure 1. Causal Mediation Analysis of the effect of *PAX8* methylation (X) on free T4 (Y), mediated by thyroid volume (M).

In each case only coefficients [SE] for relevant predictors (X or M) are given.



Supplementary Figure 2. Boxplot of mean *PAX8* methylation by Sex.

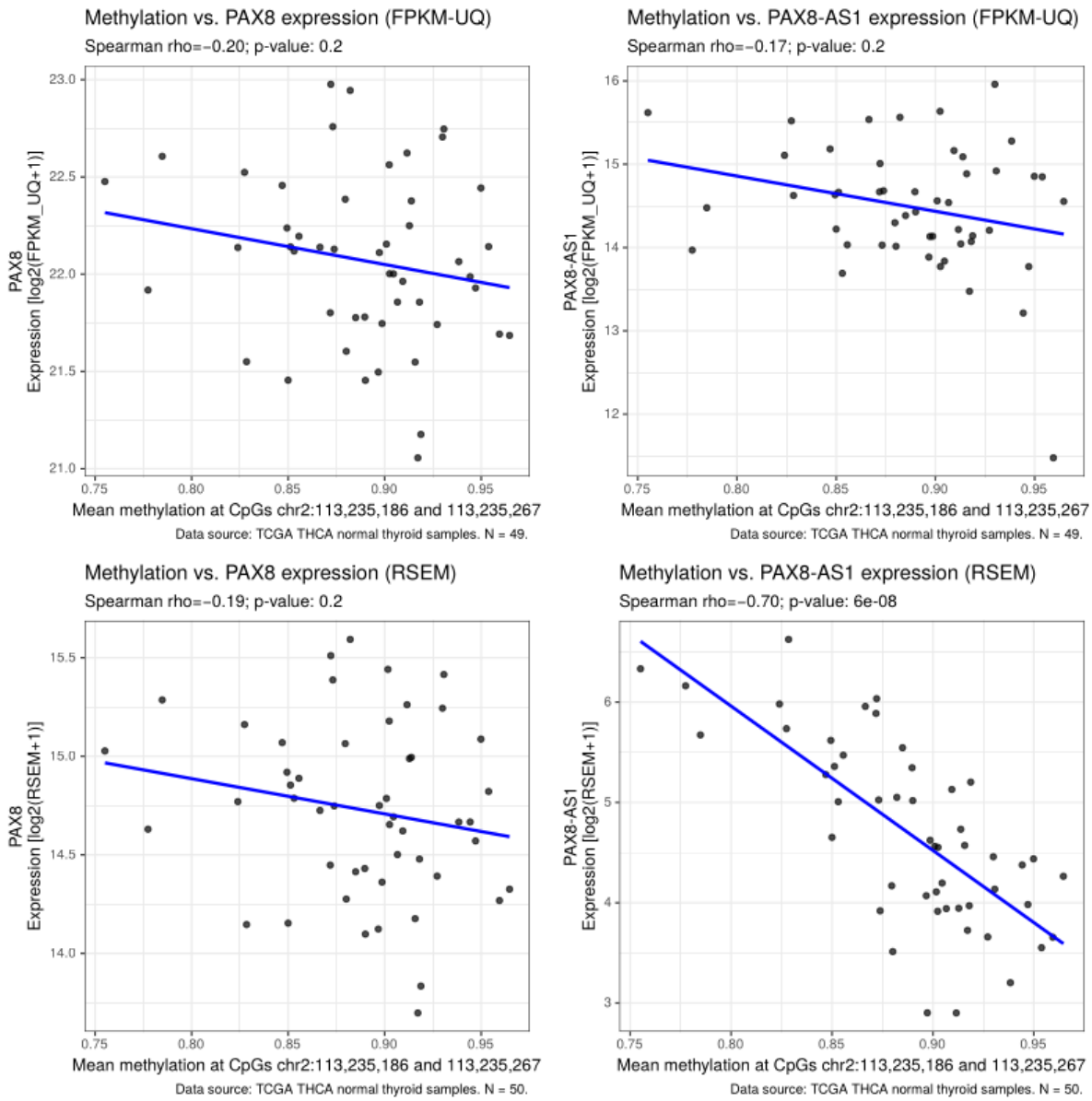
Mean methylation is calculated across the 4 CpGs in the *PAX8* region of interest. N=521 children.



Supplementary Figure 3. Relationship between *PAX8* and *PAX-AS1* expression.

Correlation between *PAX8* and *PAX8-AS1* expression in normal thyroid tissue (n=50) using data downloaded from The Cancer Genome Atlas (TCGA). Spearman correlation coefficients and p-values are shown in the graphs.

Key: FPKM = fragments per kilobase per million mapped reads, RSEM= RNAseq by Expectation-Maximization, *PAX AS1 = PAX8 antisense*.

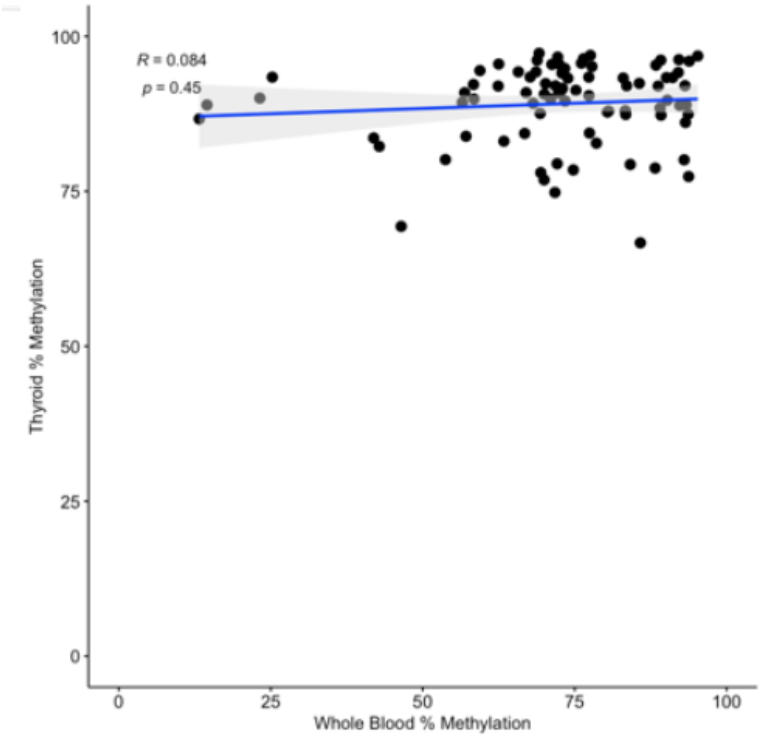


Supplementary Figure 4. Relationship between methylation and expression for *PAX8* and *PAX8-AS1* in TCGA thyroid samples.

Top: Correlation between *PAX8* and *PAX8-AS1* gene expression (using FPKM-UQ method) and mean methylation from normal thyroid tissue samples (n=49).

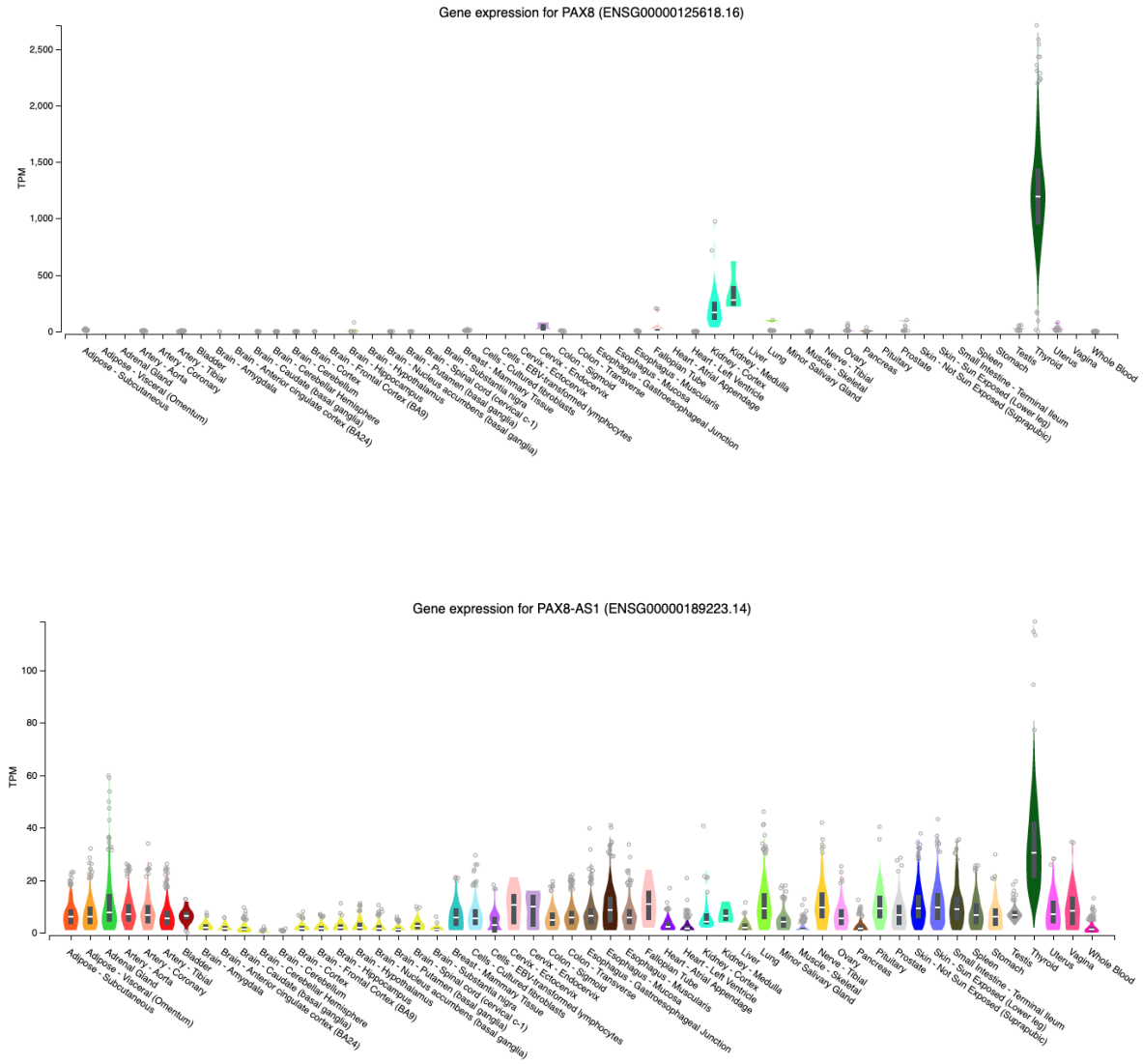
Bottom: Correlation between *PAX8* and *PAX8-AS1* gene expression (using RSEM method) and mean methylation from normal thyroid tissue samples (n=50). Data is downloaded from The Cancer Genome Atlas (TCGA), and covers 2 CpGs in our region of interest. Spearman correlation coefficients and p-values are shown in the graphs. Plots show mean methylation across the 2 CpGs.

Key: FPKM-UQ = fragments per kilobase per million mapped reads – upper quartile; RSEM = RNAseq by Expectation-Maximization.



Supplementary Figure 5. Relationship between *PAX8* methylation in adult blood and thyroid samples from GTEx.

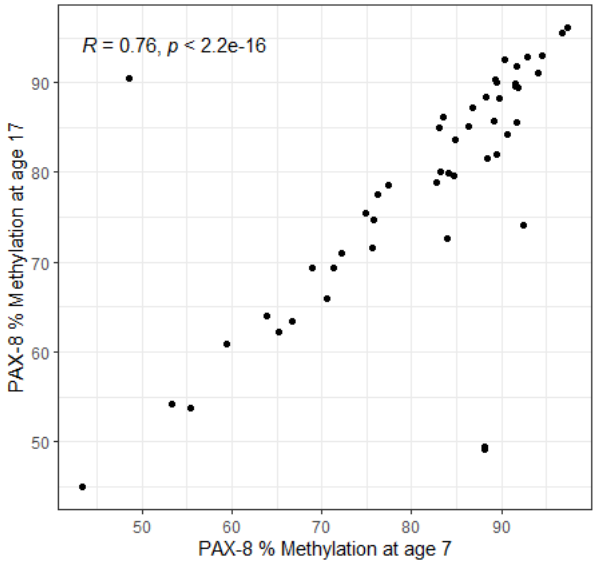
Correlation between mean whole blood and thyroid tissue methylation at 4 CpGs in the *PAX8* region of 86 paired samples from the GTEx biobank. Spearman correlation and p-value is shown.



Supplementary Figure 6. Expression of PAX8 and PAX8-AS1 by tissue type.

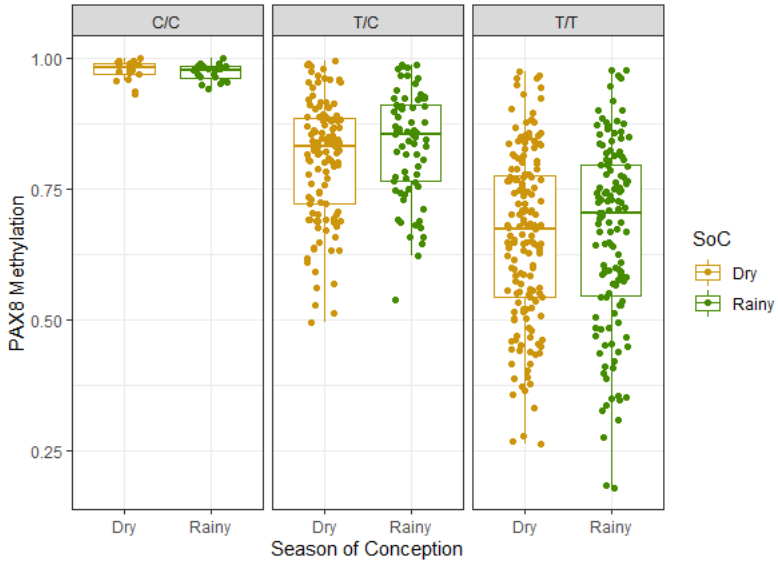
PAX8 (top) and PAX8-AS1 (bottom) gene expression plots from the GTEx (Genotype-Tissue Expression) Project portal (https://www.gtexportal.org/).

Key: Transcripts per million = TPM, *PAX8-AS1* = *PAX8* antisense.

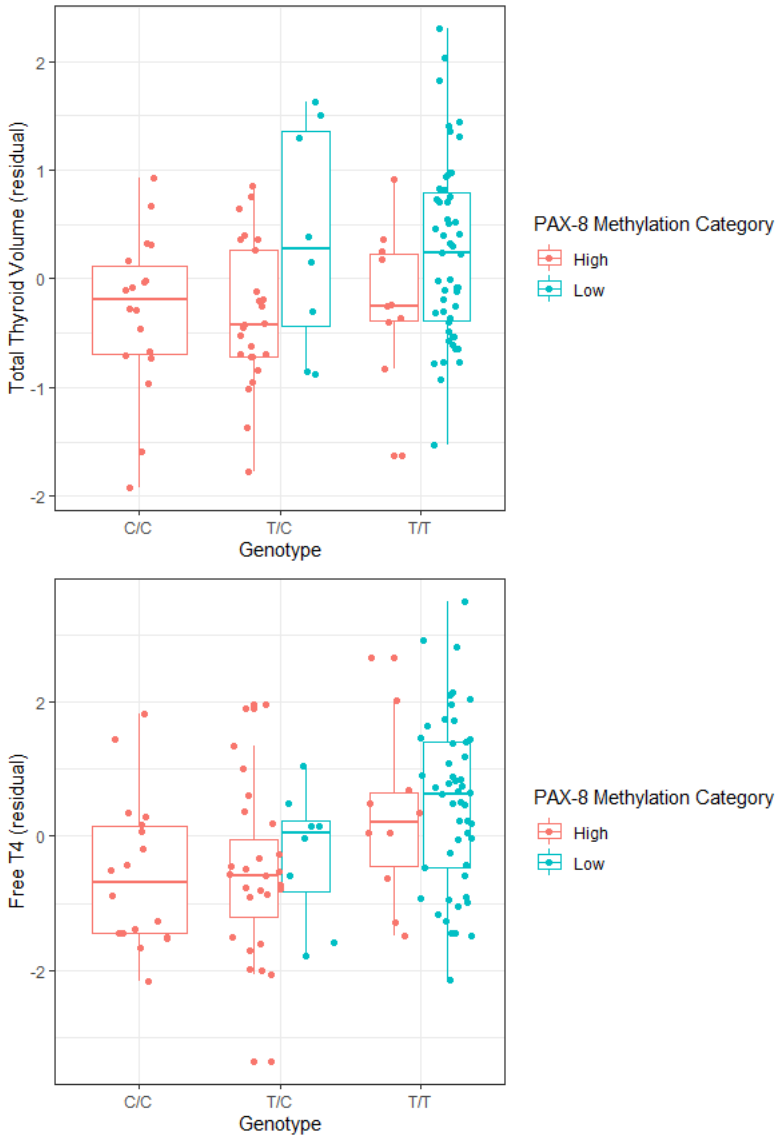


Supplementary Figure 7. Scatterplot of mean *PAX8* methylation measured in peripheral blood DNA from Gambian children in mid-childhood (aged 7 years) and in young adulthood (aged 17 years).

Spearman correlation coefficient and p-value are shown. N=49.



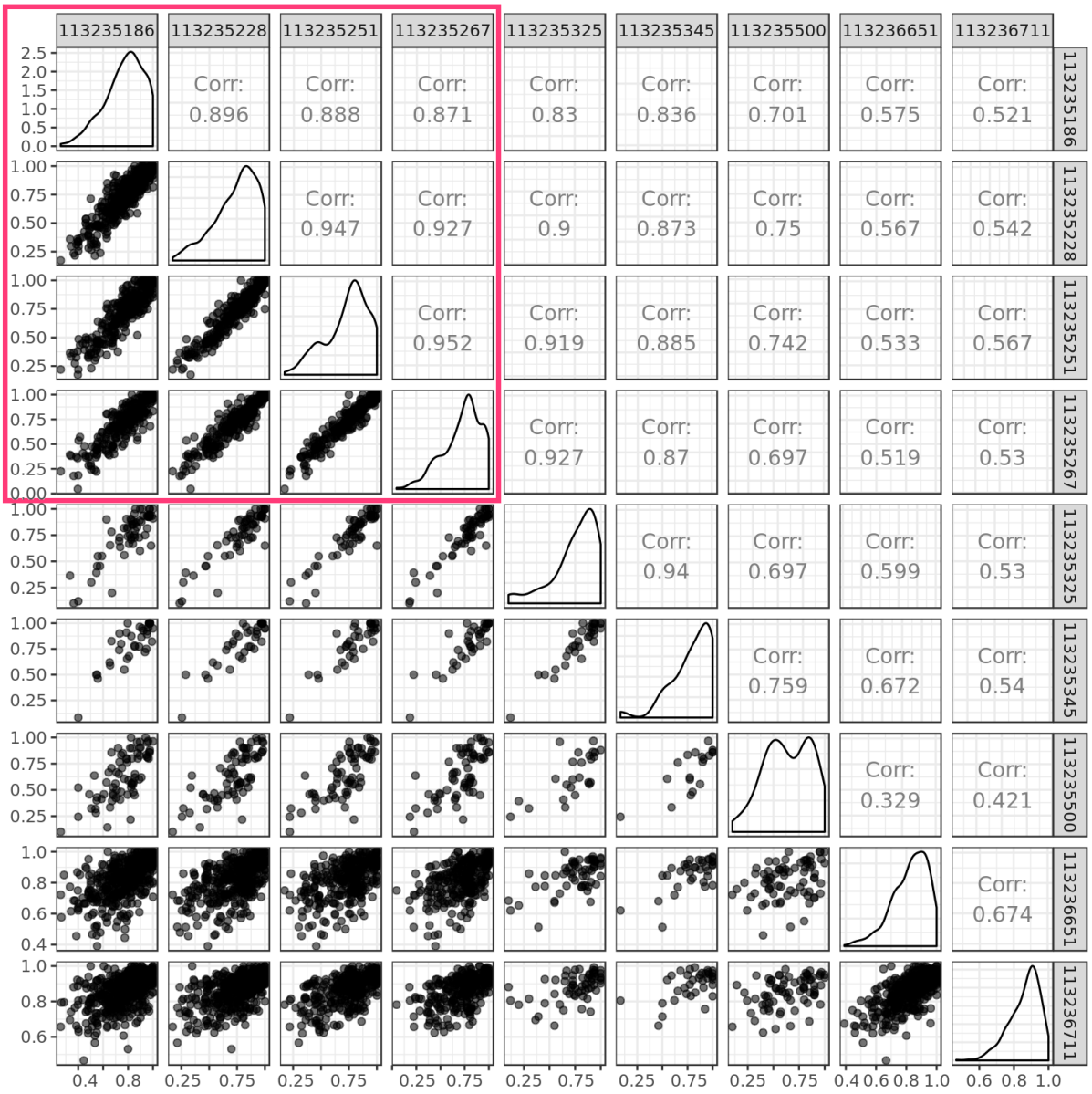
Supplementary Figure 8. Effect of season of conception on mean methylation at the *PAX8* region of interest, stratified by rs10193733 genotype.



Supplementary Figure 9. Relationship between *PAX8* methylation group, total thyroid volume and free T4, stratified by rs10193733 genotype.

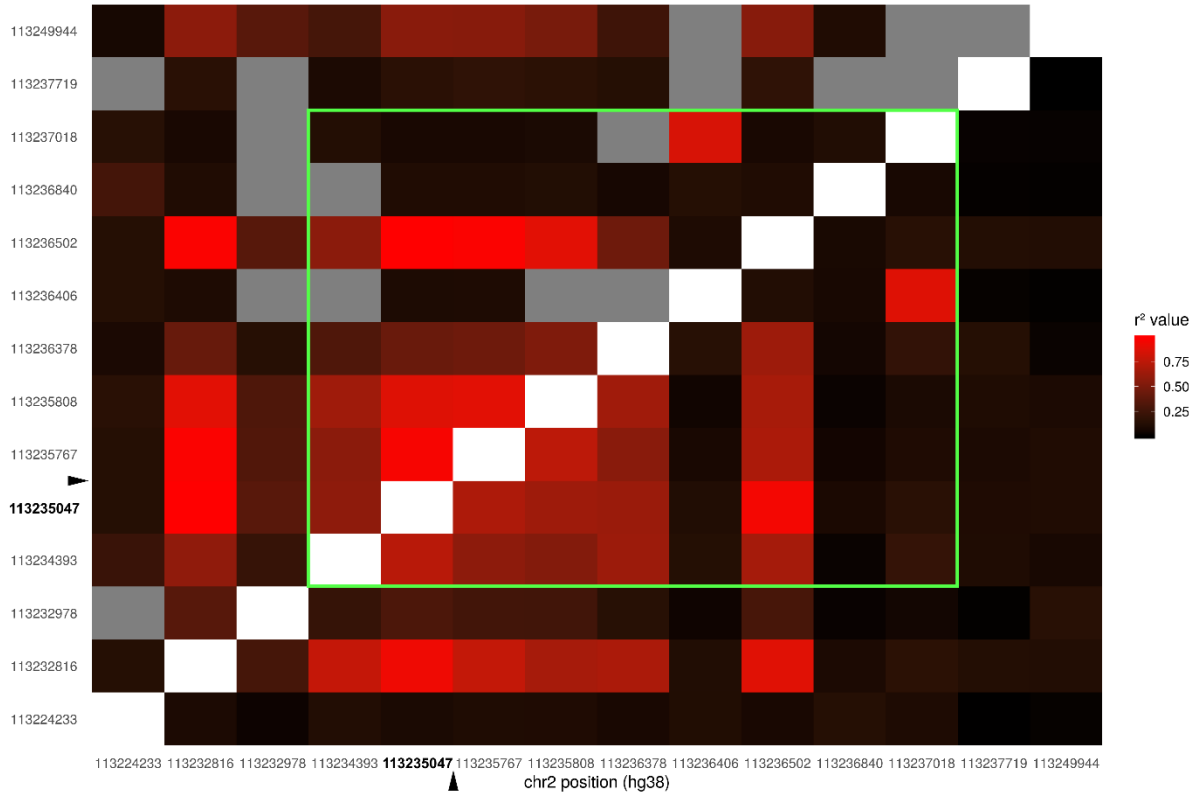
Top: Total thyroid volume is expressed as a residual adjusted for age, sex, BMI z-score and urinary iodine concentration (UIC).

Bottom: Free T4 is expressed as a residual adjusted for age, sex, and UIC.



Supplementary Figure 10. CpG – CpG correlations across the *PAX8* region.

Genomic regions are mapped to hg38 on chromosome 2. Waterland et al, 2010[1] region: 113,235,685 - 113,235,814 and 113,235,289 - 113,235,459. Finer et al, 2016[3] and Silver et al, 2015[2] region: 113,235,185 - 113,235,736. Kessler et al, 2018[8] region: 113,235,117 - 113,236, 423. The final region of interest selected for analysis is marked by the red box.



Supplementary Figure 11. Pairwise linkage disequilibrium (LD) of 14 SNPs located near the *PAX8* gene called from methyl-seq data.

Genomic regions are mapped to hg38 on chromosome 2. LD r^2 from reference populations (GGVP or 1000 Genomes where missing; above and left of diagonal); Pearson R^2 reported for ENID methyl-seq data (bottom and left of diagonal). Grey boxes indicate LD values missing from both reference datasets. Green box encloses SNPs denoted in Figure 2. The SNP (rs10193733; chr2:113235047) used in genotype analyses is indicated in bold, the approximate location of the PAX8 region of interest analysed in this study is indicated by the black triangles.

Supplementary Table and Figure References:

1. Waterland, R. A. *et al.* Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genetics* **6**, 1–10 (2010).

2. Silver, M. J. *et al.* Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biology* **16**, 118 (2015).

3. Finer, S. *et al.* 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 (2016).

4. Richmond, R. C. *et al.* The long-term impact of folic acid in pregnancy on offspring DNA methylation: Follow-up of the Aberdeen folic acid supplementation trial (AFAST). *International Journal of Epidemiology* **47**, 928–937 (2018).

5. Saffari, A. *et al.* Effect of maternal preconceptional and pregnancy micronutrient interventions on children's DNA methylation: Findings from the EMPHASIS study. *The American journal of clinical nutrition* **112**, 1099–1113 (2020).

6. Dominguez-Salas, P. *et al.* DNA methylation potential: Dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *American Journal of Clinical Nutrition* **97**, 1217–1227 (2013).

7. Dominguez-Salas, P. *et al.* Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nature Communications* **5**, 1–7 (2014).

8. Kessler, N. J., Waterland, R. A., Prentice, A. M. & Silver, M. J. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. *Science Advances* **4**, eaat2624 (2018).

Supplementary Methods

Pyrosequencing assay details

>hg38_dna range=chr2:113235183-113235270 5'pad=3 3'pad=3 strand=+ repeatMasking=none

Cccggcacccctacagcatccgccccttccgagcatgtcttccccgtcacagagaacttcatgttggcgcctccaaaagttgccggag

Primer Set 2/2020			**Score: 85** **Quality: Medium**		
Primer	**Id**	**Sequence**	**Nt**	**Tm, ºC**	**%GC**
PCR	PAX8-2/20-F1	TATGGGGTTTTGGGGTGGT	19	63.2	52.6
PCR	PAX8-2/20-R1	CCCCCTCTCTAACCTCAATCTCA	23	63.6	52.2
Sequencing	PAX8-2/20-S1	TTTGGGGTGGTGTAT	15	46.7	46.7
Target Polymorphisms	Position1, Position2, Position3, Position4, Position5				
Sequence to Analyze	T T TYGGTATT TTTATAGTAT TYGTTTTTTT YGAGTATGTT TTTTTYGTTA TAGAGAATTT TATGTTGGYG TTTTTAAAAG TTGT				
Primer Pair					
Amplicon length	161				
Score	94				
Amplicon %GC	34.8				

Details of post-mortem GTEx thyroid and whole blood samples

	Number
Sex	
Male	63
Female	23
Age range (years)	
20-29	5
30-39	6
40-49	15
50-59	33
60-69	24
70-79	3
Death Classification	
0	40
1	3
2	30
3	5
4	8

Death classification is based on the 4-point Hardy Scale:

0) Ventilator Case; All cases on a ventilator immediately before death.

1) Violent and fast death Deaths due to accident, blunt force trauma or suicide, terminal phase estimated at < 10 min.

2) Fast death of natural causes Sudden unexpected deaths of people who had been reasonably healthy, after a terminal phase estimated at < 1 hr (with sudden death from a myocardial infarction as a model cause of death for this category)

3) Intermediate death Death after a terminal phase of 1 to 24 hrs (not classifiable as 2 or 4); patients who were ill but death was unexpected

4) Slow death Death after a long illness, with a terminal phase longer than 1 day (commonly cancer or chronic pulmonary disease); deaths that are not unexpected

References

1. Boelaert K, Franklyn JA. Thyroid hormone in health and disease. *J Endocrinol*. 2005;187(1):1-15. doi:10.1677/joe.1.06131

2. Hanley P, Lord K, Bauer AJ. Thyroid Disorders in Children and Adolescents: A Review. *JAMA Pediatr*. 2016;170(10):1008-1019. doi:10.1001/jamapediatrics.2016.0486

3. Taylor PN, Razvi S, Pearce SH, Dayan CM. Clinical review: A review of the clinical consequences of variation in thyroid function within the reference range. *J Clin Endocrinol Metab*. 2013;98(9):3562-3571. doi:10.1210/jc.2013-1315

4. Mosso L, Mart??nez A, Rojas MP, et al. Early pregnancy thyroid hormone reference ranges in Chilean women: the influence of body mass index. *Clin Endocrinol (Oxf)*. 2016;85(6):942-948. doi:10.1111/cen.13127

5. Taylor PN, Richmond R, Davies N, et al. Paradoxical Relationship Between Body Mass Index and Thyroid Hormone Levels: A Study Using Mendelian Randomization. *J Clin Endocrinol Metab*. 2015;101(2):730-738. doi:10.1210/jc.2015-3505

6. Duntas LH. Thyroid Disease and Lipids. *Thyroid*. 2002;12(4):287-293. doi:10.1089/10507250252949405

7. Panicker V. Genetics of thyroid function and disease. *Clin Biochem Rev*. 2011;32(4):165-175.

8. Zhou W, Brumpton B, Kabil O, et al. GWAS of thyroid stimulating hormone highlights pleiotropic effects and inverse association with thyroid cancer. *Nat Commun*. 2020;11(1):3981. doi:10.1038/s41467-020-17718-z

9. Peters C, van Trotsenburg ASP, Schoenmakers N. DIAGNOSIS OF ENDOCRINE DISEASE: Congenital hypothyroidism: update and perspectives. *Eur J Endocrinol*. 2018;179(6):R297-R317. doi:10.1530/EJE-18-0383

10. Castanet M, Lyonnet S, Bonaiti-Pellie C, Polak M, Czernichow P, Leger J. Familial forms of thyroid dysgenesis among infants with congenital hypothyroidism. *N Engl J Med*. 2000;343(6):441-442. doi:10.1056/NEJM200008103430614

11. Perry R, Heinrichs C, Bourdoux P, et al. Discordance of monozygotic twins for thyroid dysgenesis: implications for screening and for molecular pathophysiology. *J Clin Endocrinol Metab*. 2002;87(9):4072-4077. doi:10.1210/jc.2001-011995

12. Hall SK, Hutchesson ACJ, Kirk JM. Congenital hypothyroidism, seasonality and consanguinity

in the West Midlands, England. *Acta Paediatr*. 1999;88(2):212-215. doi:https://doi.org/10.1111/j.1651-2227.1999.tb01084.x

13. Gu Y-H, Kato T, Harada S, Inomata H, Saito T, Aoki K. Seasonality in the incidence of congenital hypothyroidism in Japan: gender-specific patterns and correlation with temperature. *Thyroid*. 2007;17(9):869-874. doi:10.1089/thy.2006.0317

14. Guo J, Teng D, Shi X, et al. Exposure to the Chinese Great Famine in Early Life and Thyroid Function and Disorders in Adulthood: A Cross-Sectional Study. *Thyroid*. 2021;31(4):563-571. doi:10.1089/thy.2020.0325

15. Fernández LP, López-Márquez A, Santisteban P. Thyroid transcription factors in development, differentiation and disease. *Nat Rev Endocrinol*. 2015;11(1):29-42. doi:10.1038/nrendo.2014.186

16. Ruiz-Llorente S, de Pau EC, Sastre-Perona A, et al. Genome-wide analysis of Pax8 binding provides new insights into thyroid functions. *BMC Genomics*. 2012;13(1):147. doi:10.1186/1471-2164-13-147

17. De Felice M, Di Lauro R. Thyroid Development and Its Disorders: Genetics and Molecular Mechanisms. *Endocr Rev*. 2004;25(5):722-746. doi:10.1210/er.2003-0028

18. Mansouri A, Chowdhury K, Gruss P. Follicular cells of the thyroid gland require Pax8 gene function. *Nat Genet*. 1998;19(1):87-90. doi:10.1038/ng0598-87

19. Komatsu M, Takahashi T, Takahashi I, Nakamura M, Takahashi I, Takada G. Thyroid dysgenesis caused by PAX8 mutation: The hypermutability with CpG dinucleotides at codon 31. *J Pediatr*. 2001;139(4):597-599. doi:10.1067/mpd.2001.117071

20. Gibney ER, Nolan CM. Epigenetics and gene expression. *Heredity (Edinb)*. 2010;105(1):4-13.

21. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes Dev*. 2011. doi:10.1101/gad.2037511

22. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13(7):484-492.

23. Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. *Sci Adv*. 2018;4(7):eaat2624. doi:10.1126/sciadv.aat2624

24. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals.

Trends Genet. 2002;18(7):348-351. doi:10.1016/S0168-9525(02)02709-9

25. Silver MJ, Kessler NJ, Hennig BJ, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol*. 2015;16:118. doi:10.1186/s13059-015-0660-y

26. Kühnen P, Handke D, Waterland RA, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016;24(3):502-509. doi:10.1016/j.cmet.2016.08.001

27. Finer S, Iqbal MS, Lowe R, et al. 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*. 2016;6(11):e011768.

28. Waterland R a, Jirtle RL. Transposable Elements : Targets for Early Nutritional Effects on Epigenetic Gene Regulation. *Mol Cell Biol*. 2003;23(15):5293-5300. doi:10.1128/MCB.23.15.5293

29. Dominguez-Salas P, Moore SE, Baker MS, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat Commun*. 2014;5:3746. doi:10.1038/ncomms4746

30. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr*. 2007;27:363-388. doi:10.1146/annurev.nutr.27.061406.093705

31. Waterland RA, Kellermayer R, Laritsky E, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet*. 2010;6(12):1-10. doi:10.1371/journal.pgen.1001252

32. Abu-Khudir R, Magne F, Chanoine JP, Deal C, Van Vliet G, Deladoëy J. Role for tissue-dependent methylation differences in the expression of FOXE1 in nontumoral thyroid glands. *J Clin Endocrinol Metab*. 2014;99(6). doi:10.1210/jc.2013-4414

33. James P, Sajjadi S, Tomar AS, et al. Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism. *Int J Epidemiol*. 2018;47(6):1910-1937. doi:10.1093/ije/dyy153

34. Dominguez-Salas P, Moore SE, Cole D, et al. DNA methylation potential: Dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am J Clin Nutr*. 2013;97(6):1217-1227. doi:10.3945/ajcn.112.048462

35. Richmond RC, Sharp GC, Herbert G, et al. 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*. 2018;47(3):928-937. doi:10.1093/ije/dyy032

36. Saffari A, Shrestha S, Issarapu P, et al. Effect of maternal preconceptional and pregnancy micronutrient interventions on children's DNA methylation: Findings from the EMPHASIS study. *Am J Clin Nutr*. 2020;112(4):1099-1113. doi:10.1093/ajcn/nqaa193

37. Moore SE, Fulford AJC, Darboe MK, Jobarteh ML, Jarjou LM, Prentice AM. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: The ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth*. 2012. doi:10.1186/1471-2393-12-107

38. Consortium Gte. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580-585. doi:10.1038/ng.2653

39. Taylor PN, Porcu E, Chew S, et al. Whole-genome sequence-based analysis of thyroid function. *Nat Commun*. 2015;6(1):5681. doi:10.1038/ncomms6681

40. Teumer A, Rawal R, Homuth G, et al. Genome-wide association study identifies four genetic loci associated with thyroid volume and goiter risk. *Am J Hum Genet*. 2011;88(5):664-673. doi:10.1016/j.ajhg.2011.04.015

41. Lafontaine N, Campbell PJ, Castillo-Fernandez JE, et al. Epigenome-Wide Association Study of Thyroid Function Traits Identifies Novel Associations of fT3 With KLF9 and DOT1L. *J Clin Endocrinol Metab*. January 2021. doi:10.1210/clinem/dgaa975

42. James PT, Dominguez-Salas P, Hennig BJ, Moore SE, Prentice AM, Silver MJ. Maternal One-Carbon Metabolism and Infant DNA Methylation between Contrasting Seasonal Environments: A Case Study from The Gambia. *Curr Dev Nutr*. 2019;3(1). doi:10.1093/cdn/nzy082

43. Brosnan ME, MacMillan L, Stevens JR, Brosnan JT. Division of labour: how does folate metabolism partition between one-carbon metabolism and amino acid oxidation? *Biochem J*. 2015;472(2):135-146. doi:10.1042/BJ20150837

44. Rios-Avila L, Coats B, Chi Y-Y, et al. 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 . *J Nutr*. 2015;145(1):87-95. doi:10.3945/jn.114.201095

45. Howe CG, Cox B, Fore R, et al. Maternal Gestational Diabetes Mellitus and Newborn DNA Methylation: Findings From the Pregnancy and Childhood Epigenetics Consortium. *Diabetes Care*. 2019;16:dc190524. doi:10.2337/dc19-0524

46. Teh AL, Pan H, Chen L, et al. The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Res*. 2014;24(7):1064-1074. doi:10.1101/gr.171439.113

47. Gunasekara CJ, Scott CA, Laritsky E, et al. A genomic atlas of systemic interindividual epigenetic variation in humans. *Genome Biol*. 2019;20(1):105. doi:10.1186/s13059-019-1708-1

48. Waterland RA, Dolinoy DC, Lin JR, Smith CA, Shi X, Tahiliani KG. Maternal methyl supplements increase offspring DNA methylation at Axin fused. *Genesis*. 2006;44(9):401-406. doi:10.1002/dvg.20230

49. Rastogi M V, LaFranchi SH. Congenital hypothyroidism. *Orphanet J Rare Dis*. 2010;5(1):17. doi:10.1186/1750-1172-5-17

50. Szczepanek-Parulska E, Zybek-Kocik A, Wartofsky L, Ruchala M. Thyroid Hemiagenesis: Incidence, Clinical Significance, and Genetic Background. *J Clin Endocrinol Metab*. 2017;102(9):3124-3137. doi:10.1210/jc.2017-00784

51. Ma CG, Shim YS. Association of Thyroid-Stimulating Hormone and Thyroid Hormones with Cardiometabolic Risk Factors in Euthyroid Children and Adolescents Aged 10 – 18 Years : A Population-Based Study. *Sci Rep*. 2019:1-9. doi:10.1038/s41598-019-51963-7

52. Taylor PN, Richmond R, Davies N, et al. Paradoxical Relationship Between Body Mass Index and Thyroid Hormone Levels: A Study Using Mendelian Randomization. *J Clin Endocrinol Metab*. 2016;101(2):730-738. doi:10.1210/jc.2015-3505

53. Licenziati MR, Valerio G, Vetrani I, De Maria G, Liotta F, Radetti G. Altered Thyroid Function and Structure in Children and Adolescents Who Are Overweight and Obese: Reversal After Weight Loss. *J Clin Endocrinol Metab*. 2019;104(7):2757-2765. doi:10.1210/jc.2018-02399

54. Emokpae MA, Obazelu PA. The Association of Triiodothyronine-to-Thyroxine Ratio with Body Mass Index in Obese Nigerian Children and Adolescents. *Med Sci (Basel, Switzerland)*. 2017;5(4). doi:10.3390/medsci5040036

55. Barjaktarovic M, Korevaar TIM, Gaillard R, et al. Childhood thyroid function, body composition and cardiovascular function. *Eur J Endocrinol*. 177(4):319-327. doi:10.1530/EJE-

17-0369

56. Taylor PN, Razvi S, Pearce SH, Dayan CM. Clinical review: A review of the clinical consequences of variation in thyroid function within the reference range. *J Clin Endocrinol Metab*. 2013;98(9):3562-3571. doi:10.1210/jc.2013-1315

57. Wirth CD, Blum MR, da Costa BR, et al. Subclinical thyroid dysfunction and the risk for fractures: a systematic review and meta-analysis. *Ann Intern Med*. 2014;161(3):189-199. doi:10.7326/M14-0125

58. Bassett JHD, Williams GR. Role of Thyroid Hormones in Skeletal Development and Bone Maintenance. *Endocr Rev*. 2016;37(2):135-187. doi:10.1210/er.2015-1106

59. Alvarez-Garcia O, Fisch KM, Wineinger NE, et al. Increased DNA Methylation and Reduced Expression of Transcription Factors in Human Osteoarthritis Cartilage. *Arthritis Rheumatol*. 2016;68(8):1876-1886. doi:10.1002/art.39643

60. Guo Y, Sheng Q, Li J, Ye F, Samuels DC, Shyr Y. Large Scale Comparison of Gene Expression Levels by Microarrays and RNAseq Using TCGA Data. *PLoS One*. 2013;8(8):e71462.

61. Kulis M, Queirós AC, Beekman R, Martín-Subero JI. Intragenic DNA methylation in transcriptional regulation, normal differentiation and cancer. *Biochim Biophys Acta - Gene Regul Mech*. 2013;1829(11):1161-1174. doi:https://doi.org/10.1016/j.bbagrm.2013.08.001

62. Wagner JR, Busche S, Ge B, Kwan T, Pastinen T, Blanchette M. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biol*. 2014;15(2):R37. doi:10.1186/gb-2014-15-2-r37

63. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22(9):1790-1797. doi:10.1101/gr.137323.112

64. Bahari G, Hashemi M, Naderi M, Sadeghi-Bojd S, Taheri M. Long non-coding RNA PAX8-AS1 polymorphisms increase the risk of childhood acute lymphoblastic leukemia. *Biomed reports*. 2018;8(2):184-190. doi:10.3892/br.2017.1028

65. Zimmermann MB, Aeberli I, Andersson M, et al. Thyroglobulin is a sensitive measure of both deficient and excess iodine intakes in children and indicates no adverse effects on thyroid function in the UIC range of 100-299 µg/L: A UNICEF/ICCIDD study group report. *J Clin Endocrinol Metab*. 2013. doi:10.1210/jc.2012-3952

66. Eriksen KG, Andersson M, Hunziker S, Zimmermann MB, Moore SE. Effects of an Iodine-

Containing Prenatal Multiple Micronutrient on Maternal and Infant Iodine Status and Thyroid Function: A Randomized Trial in The Gambia. *Thyroid*. 2020;30(9):1355-1365. doi:10.1089/thy.2019.0789

67. Soldin OP. Controversies in urinary iodine determinations. *Clin Biochem*. 2002;35(8):575-579. doi:10.1016/s0009-9120(02)00406-x

68. Gunasekara CJ, Waterland RA. A new era for epigenetic epidemiology. *Epigenomics*. 2019;11(15):1647-1649. doi:10.2217/epi-2019-0282

69. Lv P-P, Tian S, Feng C, et al. Maternal High Estradiol Exposure is Associated with Elevated Thyroxine and Pax8 in Mouse Offspring. *Sci Rep*. 2016;6:36805. doi:10.1038/srep36805

70. Okala SG, Darboe MK, Sosseh F, et al. Impact of nutritional supplementation during pregnancy on antibody responses to diphtheria-tetanus-pertussis vaccination in infants: A randomised trial in The Gambia. *PLoS Med*. 2019;16(8):e1002854. doi:10.1371/journal.pmed.1002854

71. Eriksen KG, Johnson W, Sonko B, Prentice AM, Darboe MK, Moore SE. Following the World Health Organization's Recommendation of Exclusive Breastfeeding to 6 Months of Age Does Not Impact the Growth of Rural Gambian Infants. *J Nutr*. 2017;147(2):248-255. doi:10.3945/jn.116.241737

72. Alicia Alonso MPF, Kacmarczyk TJ. *Agilent SureSelectXT Methyl-Seq Applications with Low-Input DNA and Smaller Capture Libraries*.; 2017.

73. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011;27(11):1571-1572. doi:10.1093/bioinformatics/btr167

74. Brunn J, Block U, Ruf G, Bos I, Kunze WP, Scriba PC. [Volumetric analysis of thyroid lobes by real-time ultrasound]. *Dtsch Med Wochenschr*. 1981. doi:10.1055/s-2008-1070506

75. WHO (World Health Organization). WHO Anthro for personal computers. Software for assessing growth and development of the world's children. 2011.

76. Crabtree NJ, Arabi A, Bachrach LK, et al. Dual-energy X-ray absorptiometry interpretation and reporting in children and adolescents: the revised 2013 ISCD Pediatric Official Positions. *J Clin Densitom Off J Int Soc Clin Densitom*. 2014;17(2):225-242. doi:10.1016/j.jocd.2014.01.003

77. Kessler N. knowah/bssnper2: Initial release. March 2021. doi:10.5281/ZENODO.4647013

78. Gao S, Zou D, Mao L, et al. BS-SNPer: SNP calling in bisulfite-seq data. *Bioinformatics*.

2015;31(24):4006-4008. doi:10.1093/bioinformatics/btv507

79. Band G, Le QS, Clarke GM, et al. Insights into malaria susceptibility using genome-wide data on 17,000 individuals from Africa, Asia and Oceania. *Nat Commun*. 2019;10(1):5732. doi:10.1038/s41467-019-13480-z

80. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi:10.1038/nature15393

81. Moore SE, Cole TJ, Collinson AC, Poskitt EM, McGregor IA, Prentice AM. Prenatal or early postnatal events predict infectious deaths in young adulthood in rural Africa. *Int J Epidemiol*. 1999;28(6):1088-1095.

82. R Core Team. R: A Language and Environment for Statistical Computing. *R Found Stat Comput Vienna, Austria*. 2013;0:{ISBN} 3-900051-07-0. doi:http://www.R-project.org/

Chapter 8 Conclusions

Summary of the chapter

In this last chapter, I provide an overview of the results presented and contextualise them within the objectives of the thesis. I discuss challenges faced and limitations of the work. I propose future directions for research and discuss implications for the findings within a range of scientific fields.

8.1 Summary of findings

There is burgeoning body of evidence that the prenatal intrauterine environment can impact an offspring's long term health[1–4]. DNA methylation sits as a potential mediator between environmental exposures and later health outcomes[5]. The overarching theme of this thesis has been to characterise phenotypic sequelae related to maternal nutrition-sensitive epigenetic signatures.

8.1.1 Periconceptional environment and DNA methylation

A number of studies have identified Gambian season of conception (SoC) as a predictor of offspring DNA methylation[6–11]. The results presented in this thesis adds to the evidence of a SoC effect on offspring methylation[6,7,9].

The finding of higher methylation in rainy season conceptions at *POMC* is consistent with previous studies[12] and the effect seen at other MEs [6,7,9]. A SoC effect was not seen in *PAX8* as previously observed[6,9]. This could be because in the *PAX8* study, seasonality was modelled as a dichotomised variable of rainy/dry season of conception rather than as continuously varying sinusoidal function using Fourier analysis. The latter approach was employed with the *POMC* study seasonal modelling and is arguably more powerful than splitting the year into two seasons. Using Fourier analysis the modelling approach makes no assumptions about where potential peaks or nadirs of methylation may occur. Furthermore, it draws on data from children conceived throughout the year (not limited to peaks of season as with previous studies in this population[6,7,9]). At *POMC* there was an observed attenuation of the SoC effect on DNA methylation from 2 years to mid childhood. A reduced amplitude of the SoC effect between 2 years and 8-9 years of age has been reported in ENID children at other MEs[11]. This finding may suggest further epigenetic change through childhood. However, the consistently reported early SoC effect on DNA methylation, even if attenuated through childhood, may set a developmental course with phenotypic consequences.

One-Carbon metabolites provide methyl groups for methylation reactions including the methylation of cytosine bases. Women's dietary intake and maternal circulating levels of one-carbon metabolites (riboflavin, folate, choline, and betaine) has been shown to fluctuate throughout the year in The Gambia[13]. Furthermore, there is an association between seasonally-driven differences in maternal circulating one-carbon metabolites and offspring DNA methylation at several MEs[7]. Higher

periconceptional circulating levels of cysteine, homocysteine and PLP (a B6 vitamer) were associated with lower DNA methylation at *PAX8.* These same associations have been observed at other MEs[7,8].

Additionally, data presented in this thesis demonstrated novel associations between offspring DNA methylation and the levels of a number of circulating periconceptional amino acids. There was an association between uracil and both *POMC* (positively associated) and *PAX8* (negatively associated) methylation, though the effect was directionally inconsistent. Uracil is excised from the DNA strand by Uracil-DNA glycosylase (UNG) enzyme[14] and there is evidence in a murine model that this enzyme is involved in Tet (ten-eleven-translocation enzyme)-mediated DNA demethylation[15]. The negative association between uracil and *POMC* methylation could be explained if plasma levels of uracil were correlated with UNG activity. Thus, higher uracil levels would be associated with higher UNG activity and an increase in offspring DNA demethylation. This theory does not support the finding in *PAX8* and thus further work exploring this potential driver of methylation is certainly warranted.

8.1.2 Genetics and ME methylation

There are two main findings to come from this thesis with regard to the relationship between genotype and DNA methylation.

A number of mQTL were identified at the *POMC* gene. These have not been previously characterised. In animal models, MEs have been characterised as being independent of genotype[16]. However, MEs were first described in isogenic mice and therefore the influence of different genetic backgrounds was not explored. However, more recent studies have suggested that methylation at human MEs may to some degree be under genetic influence[11,17].

There was an interesting finding with regard to the genotype – methylation relationship at PAX8. In this case, genotype was associated with methylation variability. This raises the prospect of a genotype-early environment interaction effect on methylation, as previously described[18,19]. There was not an interaction between genotype and season of conception identified in the *PAX8* study, though likely there was limited power to detect this. If there were a genotype-environment interaction at this locus, this raises the interesting possibility that the *PAX8* ME could have evolved as an environmental sensor that facilitates adaptation in phenotype accordingly. This is consistent with Feinberg's model of the adaptive benefit of (genetically-entrained) epigenetic variance which describes how genetic variants could enhance fitness[20]. Future work should explore the mechanisms

surrounding the strong variability effect observed, and its potential link to differential sensitivity to environmental factors.

8.1.3 DNA methylation and phenotypic associations

For the first time in The Gambia, DNA methylation, at MEs sensitive to the periconceptional environment, have been associated with phenotype.

The results from the *PAX8* study demonstrated that DNA methylation at a region of the P*AX8* gene sensitive to periconceptional nutrition, was significantly associated with total thyroid volume and free thyroxine (T4) levels in Gambian children. Furthermore the effect size seen was considerable. Compared to *PAX8* hypermethylation, *PAX8* hypomethylation was associated with a 21% increase in thyroid volume and an increase in free thyroxine equivalent to 8.4% of the normal range. The methylation – phenotype association reported for both thyroid volume and free T4 appeared much higher than the reported association with genetic variants. Results from a recent GWAS of thyroid function found only 4 SNPs associated with free T4 with the largest reported SNP effect was 0.22 pmol/L per variant allele[21]. By comparison, the difference in free T4 between high and low *PAX8* methylation groups in the *PAX8* study was 0.85 pmol/L i.e. much greater than the largest individual genetic effect observed in the GWAS. Similarly, a recent GWAS found only 4 variants associated with thyroid volume[22]. In that study, the largest SNP effect size was 0.093 cm^3 per allele, again far smaller than the effect size seen in the *PAX8* study (0.61 cm^3 after adjustment for covariates).

Understanding the consequence of variation in *PAX8* methylation on bone density and body fat was also studied. There was no overall effect of *PAX8* methylation on either bone mineral density (BMD) or fat mass index (FMI) but there were significant associations between levels of free T4 and BMD and FMI. As there was not an overall effect of *PAX8* methylation on BMD and FMI formal mediation analysis could not be performed. However, PAX8 may still be affecting bone and body fat phenotype via a complex and/or indirect effect. For example, *PAX8* methylation could be exerting complex countervailing effects on BMD and FMI such that the net overall effect of *PAX8* on these phenotypes is non-significant. However, there may be one or more significant causal pathways between *PAX8* methylation, free T4 and BMD and/or FMI.

The *POMC* study used seasonally driven alterations in energy balance to model weight and FMI change over one calendar year. The *POMC* study yielded an association between *POMC* methylation and maternal seasonal FMI change. *POMC* methylation was negatively associated with maternal FMI amplitude meaning that for every SD increase in *POMC* methylation z score the amplitude of FMI

change reduced by 0.045 kg/m^2. In other words, higher methylation was associated with a smaller change (or more stable) in fat mass over the year. However, there was no evidence of an association between *POMC* methylation and maternal satiety measures or energy consumption. It is possible that the observed change in FMI could be driven by direct or indirect POMC-mediated effects on adipose tissue or energy expenditure. Alternatively it could be the case that *POMC* methylation mediated effects on appetite, satiety, and energy intake that could still be mediating minor alterations in energy balance that were not detectable due to limitations already described (see chapter 5.7 and 5.7.1). There was no effect of *POMC* methylation on maternal or child weight, or on FMI. The null result in children is of interest and one can postulate the reason for the significant effect seen in mothers but not children in the *POMC* study. Genetic influence on body weight and BMI changes over the life course. For example, the polygenic effect on weight has been reported to emerge in infancy and increases in adulthood[23]. DNA methylation may also differentially influence body weight, fat mass or BMI across the life course.

Kühnen et al, reported an association between *POMC* methylation and obesity in both women and children[24,25]. The effect of *POMC* on body weight and fat may be situational. For example, in rural Gambia, *POMC* methylation may have limited scope to influence body weight due to nutritional scarcity or high infectious disease burden across the population. However, in more obesogenic environments such as found in Germany[26], the opportunity to consume excess calories which in turn effects longstanding positive energy balance (leading to obesity) is far greater.

8.2 Limitations

8.2.1 Epigenome wide vs. candidate gene approach

Whereas epigenome wide association studies (EWAS) explore associations between phenotype and epigenetic alterations across large numbers of loci across the genome, a candidate gene approach limits analysis to a select number of genes. This thesis employed a candidate gene approach exploring phenotypic associations with just two genes: *PAX8* and *POMC.* These genes were selected based on prior evidence of an association with season of conception and maternal biomarkers. Experimental design to explore phenotype was based on prior knowledge of biological pathways associated with thyroid development and function (*PAX8*) and the role within the melanocortin system in the regulation of appetite and energy balance (*POMC*). Candidate gene approaches can

have greater power to detect associations compared to EWAS. Furthermore, experimental study design can be formulated to evaluate a specific hypothesis related to known biological mechanisms as opposed to EWAS which does not look to test a specific gene association or test a hypothesised mechanism.

Taking a candidate gene approach has limitations as it may give undue focus on a select group of genes and does not identify novel variants or epigenetic loci associated with phenotype. Many associations with phenotype identified in candidate gene approaches are not replicated in genome wide studies. For example, associations between *POMC* methylation and weight/BMI reported in candidate gene studies by Kühnen [12,25]and Crujeiras[27] have not been replicated in EWAS studies[28,29], although this might be due to the fact that associated *POMC* loci are not well covered by array-based methylation platforms (see below). Candidate gene approaches have been criticised for reporting a high level of false positive associations and wider association studies are needed to replicate findings[30].

Most EWAS studies report data from Illumina methylation arrays. Whole-genome bisulphite sequencing provides a methylation profile across all CpGs in the human genome (~28 million CpGs): this approach is currently prohibitively expensive for population-based studies. Illumina array platforms such as the 450k and EPIC array cover at most around 850,000 CpGs which is a small fraction of the total number in the genome, meaning that most studies fail to capture most of the variation in the methylome. However arrays continue to provide a cost-effective and easy to perform technology. In the *POMC* study, methylation was measured by pyrosequencing of multiple CpGs in the gene that are not covered by arrays. This may allow detection of important associations not detected in array and EWAS approaches. Furthermore, EWAS studies are often cross sectional and therefore the direction of causation between epigenetic marks and phenotype can be hard to interpret.

The work in this thesis employed a hypothesis-driven study with focussed experimental design to assess associations between DNA methylation and phenotype. Replication of the findings are essential to further build a body of evidence for the associations identified.

8.2.2 Causality

It is important to adopt a critical approach when interpreting the results from epigenetic studies. Establishing causal relationships between exposures, DNA methylation and subsequent postnatal

phenotype is challenging. Often times, a unidirectional relationship between exposure, methylation and phenotype is assumed. This assumption fails to acknowledge the potential for confounding, genetic influences or reverse causation effects where phenotypic effects drive epigenetic changes[31].

There are a number of approaches for strengthening causal inference[32]. One approach to reduce the potential for reverse causation is to measure the epigenetic mark prior to disease or phenotype development. In some studies this has meant collecting neonatal blood cord methylation and relating this to phenotype later in childhood[33]. Replicating associations between methylation and phenotype in disparate cohorts can be powerful as this tests the association against different genetic backgrounds and with different environmental confounding factors e.g. socioeconomic difference between West African and Europe. Exploring epigenetic differences in monozygotic twins discordant for a disease of interest is another study design that can be powerful to control for genetic confounding[34]. Integrating epigenomic, genetic and transcriptome data from biological samples can help understand the association between methylation and expression and acts as a useful molecular basis to understand causative pathways. In some studies, statistical approaches including mediation analysis and/or Mendelian randomisation have been used to infer direction of causation. Some EWAS using causal inference techniques have suggested that BMI (or obesity) is the main driver of epigenetic changes (rather than epigenetic changes driving BMI phenotype)[28,35] and this is an important consideration when interpreting the results from this study.

Both the *POMC* and *PAX8* study employed some of these approaches to help assess the case for a causal pathway between methylation and phenotype.

In the *POMC* study, methylation was assessed at the start of the seasonal cycle and the subsequent weight and fat change tested i.e. methylation measured prospectively before measuring phenotype. However, repeated patterns of seasonal weight or fat change in previous years could have driven DNA methylation changes. An association was only seen in women and not children and unfortunately there are no banked DNA samples from these women when they were younger (or as children) to assess methylation much earlier in life.

As outlined above, an important part of building a causal case is to test associations in different cohorts and settings. Kühnen et al identified an association between obesity and *POMC* methylation in women and children[24,25]. This study was from a German population and used a cross sectional design, although Kühnen also reported that *POMC* methylation was consistent from the neonatal period at least through to adolescence. These findings are similar to those reported in this thesis with regard to finding an association between *POMC* methylation and a measure of body composition (e.g. BMI (Germany) and FMI(Gambia)) and therefore provide some supporting

evidence of possible causative link in adult women. Furthermore, the results in this thesis provide evidence of an association in a prospective study and thus have an advantage over cross sectional designs in terms of the ability to build evidence of a causal relationship.

Though Kühnen had reported an inverse relationship between methylation and expression in leucocytes[25], gene expression data was not assessed in the Gambian women. Importantly, methylation-expression data was not assessed in human hypothalamic tissue because of the obvious challenges of obtaining such tissue samples. Though genotype was associated with methylation it did not appear to be a confounding factor for the season of conception association reported. Mendelian Randomisation or other causal inference techniques were not performed though would be a useful tools to help understand causation. Unfortunately, this analysis could not be completed due to time restraints related to thesis submission but is something that will be taken forward in preparation for future publication.

While the data presented in this thesis cannot prove a causal relationship between methylation and seasonally driven changes in FMI, these findings build on previous evidence of a potential causal pathway between *POMC* methylation and adipose related phenotype in adults.

In the *PAX8* study, DNA methylation was assessed much earlier than the measured phenotype i.e. DNA methylation was measured in 2 year old children and the phenotype was measured 3-6 years later in mid childhood. However, there could be thyroid volume or function differences at age 2 years that alters DNA methylation at *PAX8* at the time of measurement (2 years) i.e. phenotype drives methylation changes, not the reverse. In the *PAX8* study, there was significant associations between *PAX8* methylation and free T4, and between free T4 and both BMD and FMI. However, there was no *direct* association identified between *PAX8* methylation and BMD and FMI and therefore statistical tools such as mediation analysis could not be performed. An assessment of the relationship between methylation and expression was obtained from TCGA data. This demonstrated an inverse relationship between *PAX8* methylation and *PAX8-AS1* expression which help builds the case for a causal link between methylation, gene expression and thyroid-function related phenotypes. Certainly, further study should look to replicate the *PAX8* study findings in different cohorts. As with the *POMC* study, causal pathways cannot be proved from the results in the thesis but provide initial evidence of a possible pathway between periconceptional environment, *PAX8* methylation and childhood thyroid function and size.

8.2.3 Assessing tissue of interest

Methylation patterns measured in blood may be a poor proxy for methylation in the tissue relevant to the phenotype of interest. DNA methylation at metastable epialleles is thought to be systemic and not tissue-specific[17] and therefore methylation in accessible tissue such as blood should be correlated with the tissue of interest. There is evidence of consistent methylation patterns across tissues from different embryonic germ layers in *POMC*[12] and *PAX8*[6,9]. At *POMC*, methylation between leucocytes and the tissue of interest of MSH (melanocyte stimulating hormone) neurons of the arcuate nucleus, have been shown to be correlated[12]. There are limited data exploring correlation between leucocyte and normal thyroid tissue methylation at *PAX8* and assessing how DNA methylation at the *PAX8* region of interest is correlated between leucocyte and thyroid tissue in children is important for future study. Measuring methylation in normal thyroid tissue in children is challenging though there is some evidence of concordant methylation between leucocytes and thyroid tissue discovered in a small number of samples examining a genomic region close to the *PAX8* region of interest[36]. Interestingly, there was no evidence of a correlation between thyroid and leucocyte methylation in adults (chapter 7.2 supplementary table 5).

Cell composition differences were not accounted for in analysis as *POMC* and *PAX8*. The argument for correcting for cell composition is that individual cell types may have distinct methylation profiles. Not correcting for cell composition can lead to hidden confounding, since a disease or phenotype could influence cell type proportions within a tissue and therefore the measured DNA methylation levels. Evidence of systemic methylation at both *PAX8* and *POMC* suggests that any cell composition effects should be minimal. Furthermore the study design should reduce the potential for confounding since methylation was measured prior to phenotype, many years before in the case of the *PAX8* study.

8.2.4 Characterising environmental exposures

This thesis explored the relationship between *periconceptional* environmental and nutritional factors and offspring's methylation. There is good reason for this focus as the periconceptional period is a key period in epigenetic reprogramming[37] and there is a burgeoning body of evidence that early embryo methylation is sensitive to maternal health, nutritional and the environmental milieu[38,39].

Maternal blood samples were taken soon after mother's first missed menses when pregnancy was confirmed by ultrasound scan (~12 weeks gestation). Estimations of maternal periconceptional circulating biomarker were made by back extrapolating levels to the date of conception using previously described methods[7]. This approach was a practical compromise to assess nutritional state at the time of conception as it is impractical to take repeated blood draws from mothers over many months in case they conceive. Therefore, estimations of maternal periconceptional circulating biomarkers may not reflect the exact nutritional milieu at the time of conception or soon afterwards.

A focus on the periconceptional window meant that other prenatal (i.e. mid or late gestational) and postnatal effects were not explored. Finer et al reported that as opposed to the effects seen in early pregnancy, there was no late gestational effects of famine on *PAX8* methylation[40]. However, apparent later gestational effects on *PAX8* methylation have been reported. Richmond et al, identified a differential effect of folic acid supplementation on *PAX8* methylation in adults whose mothers had taken folic acid from 17 weeks gestation until the end of pregnancy[41]. Therefore, *PAX8* methylation may potentially undergo further change after the periconceptional period. In the *PAX8* study, when assessing the predictors of methylation, neither mid nor late gestational factors were considered and there remain uncharacterised.

There is evidence that the *POMC* ME region may be stable from periconception to late adolescence in humans from previous studies[12,42], and in the data in chapter 6.3.9. This stability suggests that the methylation levels at the ME region appear somewhat resistant to post-natal factors up until late adolescence. However, there is also evidence to suggest that methylation at the *POMC* gene maybe sensitive to specific nutritional factors in adulthood. In the Lipogain study, adults were randomised to a 7 week diet high in either PUFA (polyunsaturated fatty acids) or SFA (saturated fatty acid). *POMC* (though not specifically the ME region) was one of a number of genes where mean methylation in adipose tissue increased only in response to PUFA and not SFA treatment[43]. There is also evidence that adverse child hood events such as emotional, physical or sexual abuse, family violence, family mental illness or substance abuse are associated with- epigenetic alterations that persist into adulthood[44] and are linked to a range of health outcomes[45]. Hecker et al[46] showed that high exposure to child abuse was associated with hypermethylation of the *POMC* gene promotor in saliva from Tanzanian children. The association between salivary *POMC* methylation and adverse childhood events may not be apparent in MSH neurons and may be tissue specific. Furthermore, though not specific to *PAX8* or *POMC,* Han et al found that adolescence was a period that observed significant change in DNA methylation landscape with changes in DNA methylation reported at more than 15K CpGs[47]. Therefore progression through puberty may also represent a window of epigenetic change and remodelling.

Historically research has been focussed on maternal factors that influence offspring DNA methylation. Sharp et al, reported that "maternal effect" DoHAD related publications out number "paternal effects" publications by 17 to 1[48]. Sharp argues that maternal 'causal primacy' is driven by a number of *assumptions* namely "1) early life exposures are primarily transmitted via maternal exposures, 2) maternal factors around the time of pregnancy and early infancy are particularly important, and 3) paternal and postnatal factors, have relatively minor impact in comparison". The causal primacy has driven a research and publication agenda which has primarily focussed on maternal effects over paternal effects and reinforced the three assumptions without critique.

There is evidence to suggest that paternal exposures may also influence DNA methylation in offspring and influence later phenotype.

Data presented in this thesis (see chapter 6.3.8) does show that offspring *POMC* methylation was correlated with their father's methylation as reported previously by Kühnen et al[12]. There was no further exploration as to the reasons for this.

The potential for paternal transmission of intergenerational methylation patterns is an emerging topic in epigenetic study. Paternal health in the years just before puberty have been linked to later health outcomes in subsequent generations. The Överkalix cohort utilised harvest records between 1890 – 1920 in Överkalix, Sweden to make assumptions about historical individual food availability[49]. Poor food availability during the father's slow growth phase (SGP) before puberty was associated with protection against cardiovascular disease in their sons[50]. Interestingly, if paternal grandfather had relatively plentiful nutrition during the SGP their grandchildren had a 4-fold increase in diabetes risk [50]. Despite these data coming from small numbers of individuals (~300) and the findings identified in secondary analysis, some of these findings are supported by a recent, larger multigenerational Swedish study[51]. An association between paternal smoking before puberty and body fat in their sons has been reported [52]. Body fat was between 5-10kg higher in boys aged between 13 and 17 years if their father smoked before the age of 11. Paternal body fat has been associated with body fat in their prepubertal daughters[53] and paternal obesity and has been associated with an increased risk of an offspring born small for gestational age[54]. Paternal Betel nut use has been associated with increased risk of metabolic syndrome in offspring[55]. In all these studies, epigenetic mechanisms were not studied. Therefore the biological processes whereby paternal health and environmental exposures are translated into phenotype in their offspring were not identified.

Transmission of paternal environmentally-sensitive epigenetic marks between generations suggests a mechanism driven by alterations to the developing gametes, although widespread erasure of most

epigenetic marks during germ cell development and at conception mean inter-generational transmission is likely to be rare[34]. Kühnen et al examined sperm *POMC* methylation and found that methylation was significantly lower than in leucocytes, suggesting that the apparent paternal transmission of epigenetic marks seen in the offspring is unlikely to be mediated through sperm methylation. Similar patterns of apparent inheritance of epigenetic marks via the paternal line at *POMC* have been reported in animal studies[56,57] where DNA methylation patterns associated with foetal alcohol exposure persisted in male progeny in F2 and F3 in the male germline only[58]. However, a separate human study examined sperm DNA methylation before and after bariatric surgery (and hence after considerable weight loss) and identified multiple methylation differences including sites mapping to *FTO* and *MC4R* genes[59].

Due to the erasure of much of the epigenetic landscape of developing gametes, other processes such as sperm RNA have been identified as candidates for epigenetic transmission[60]. Differential expression of piwi-interacting RNA (piRNA) has been identified in a study comparing lean and obese men[59]. Differences in expression of four piRNAs were identified which targeted the *CART* gene, implicated in appetite regulation. The impact of altered DNA methylation between obese and lean men on their children was not assessed.

In summary, the influence of season of conception and a range maternal periconceptional nutritional markers on DNA methylation at *POMC* and *PAX8* ME regions has been explored. However, later gestational and many postnatal factors and their relationship with DNA methylation were not investigated in this thesis. Father-offspring methylation correlations were identified at *POMC* though potential mechanisms of transmission of these epigenetic marks were not studied.

8.3 Challenges

There were a number of challenges during my PhD which relate to the issues with the pyrosequencer and the COVID-19 pandemic. These two issues significantly delayed my thesis submission.

8.3.1 Pyrosequencer

As outlined in section 4.2.7, I experienced a significant challenge to get accurate methylation output on the Q48 pyromark in The Gambia. This meant samples were processed to produce PCR product in The Gambia and then shipped to Germany for pyrosequencing. Eight months were spent trying to get meaningful data on the Q48 machine in The Gambia which led to delays in getting methylation data for statistical analysis and modelling. The experience was however valuable as I i) developed the laboratory processes with regard to bisulphite conversion and polymerase chain reaction (PCR) ii) learnt key laboratory techniques iii) learnt analytical laboratory skills. I initially planned to measure *POMC* methylation in the subset of samples taken at midline and endline. The reason for this was to explore any seasonally driven changes in individual's methylation and thus assess stability of methylation across the year. Due to the delays described above, I focussed on completing the baseline methylation samples to allow the main hypotheses to be tested.

8.3.2 COVID-19 pandemic

The COVID-19 pandemic had a major effect on the health and geopolitical landscape of the world. Not withstanding the catastrophic health impacts, the pandemic also led to limitations in travel, movement of people and global trade. In March 2020, The Gambian government closed their international border. The Foreign and Commonwealth Office gave minimal notice of this event and therefore myself and my family made the decision at short notice to be repatriated to the UK. This led to some of my lab work needing to be completed by scientific officers in The Gambia. There was a further delay in shipping my samples from The Gambia to Germany due to reduced global movement of cargo.

My regional health authority (Severn Deanery, UK) requested those seconded to research to halt their research activity and return to the National Health Service (NHS) to support the pandemic response. I returned to full time clinical duties between March 2020 – September 2020 and therefore all my research activity ceased during this period. I had planned to visit Germany to work on the pyrosequencing for the POMC samples, however ongoing government mandated restrictions on international travel meant I could not travel for this work and it was completed in my absence.

Two scheduled statistical courses at the University of Bristol were also cancelled due to the pandemic.

8.4 Future Directions and Public Health interventions

8.4.1 *PAX8:* Future directions for study

Developing molecular models

It was beyond the scope of this thesis to explore mechanistic models at a molecular or cellular level. Taking this approach certainly has merit and would allow controlled experiments to explore factors associated with *PAX8* methylation and/or thyroid development.

Developing a cellular model to assess the influence of levels of one-carbon metabolites on methylation at *PAX8* would be informative. To replicate the association between *PAX8* methylation and cysteine, homocysteine and PLP (a B6 vitamer) levels would seem a logical next step to take forward in a cell model.

One approach to test this *in vitro* would be through an induced pluripotent stem cell (iPSC) model. Embryonic stem cells have the potential to differentiate into a range of tissue types including thyrocytes[61,62]. iPSCs technology allows the development of stem cells from any somatic cells e.g. human skin fibroblasts. Recent work has successfully differentiated human iPSCs into human thyroid follicular cells[63]. The proposed work would look to vary concentrations of methyl donor substrates (e.g. cysteine, homocysteine and PLP (a B6 vitamer)) in the growth medium of iPSCs as they differentiate towards thyroid cells. Methylation at *PAX8* could then be measured and associations with levels of methyl donors assessed. Furthermore, similar work has been completed by collaborators led by Dr Peter Kuhnen who successfully developed an *in vitro* cell model examining the influence of one carbon metabolites on *POMC* methylation (results yet to be published).

Technologies designed to manipulate methylation at specific genes has an application for further research related to methylation at the *PAX8* gene. Technologies such as zinc finger proteins (ZFP) and transcription activator like effector proteins (TALEs) have been used to recognise specific nucleotide sequences and induce epigenetic editing at single loci. Clustered regulatory interspaced short palindromic repeats(CRISPR)-based editing technologies provide a far less laborious method of epigenetic manipulation than ZFP or TALE approaches[64]. The DNA editing CRISPR technology uses a

Cas9 nuclease and a hybridised CRISPR RNA to alter the DNA sequence as required. The Cas9 nuclease acts as a DNA binding domain and can be directed at any part of the genome. For epigenetic editing the DNA strand is not required to be cleaved so the Cas9 nuclease catalytic activity is therefore deactivated[65]. In brief, for methylation manipulation, the CRISPR and deactivated Cas9 protein combine with an effector protein (DNMT (DNA methyltransferase) 3A for methylation and TET (ten-eleven-twelve) protein for demethylation) to alter methylation at a particular genomic region[66]. This technology could be utilised to alter DNA methylation at the *PAX8* region. This would have a number of uses. It would allow assessment of *PAX8* methylation-expression relationship in thyrocytes. Furthermore, it could also be used to assess how differential *PAX8* methylation alters thyroid cell development from iPSCs.

***PAX8* methylation and CHT**

Twin studies, especially in monozygotic (MZ) twins, are a powerful tool in the study of epigenetic drivers of disease. These studies are useful as they control for major confounding factors including genetic variation, age and sex[34]. Assessing MZ twins discordant for congenital hypothyroidism would be interesting to ascertain if they are epigenetically dissimilar at *PAX8.* The vast majority of CHT cases have no known genetic cause, raising the possibility of epigenetic dissimilarity as a possible aetiology.

A challenge with this regard would be to find MZ twins with CHT as the incidence of CHT is 1 in 3000 births[67]. Nonetheless, organisations such as Twins UK[68] may provide a useful organisation to identify MZ twins discordant for CHT.

Replication

It is very important that replication of the findings in this thesis are explored in other studies. One avenue would be to explore thyroid data taken from different longitudinal cohorts, such as Avon Longitudinal Study of Parents and Children (ALSPAC)[69]. Here there may be the opportunity to obtain early life epigenetic measurements together with the opportunity to characterise later thyroid phenotype. It would be interesting to see if the findings are replicated in a different cohort from disparate ethic groups and genetic background.

8.4.2 POMC: Future directions for study

As discussed in the POMC chapters (chapters 5-6) it may be that the effect of *POMC* methylation on appetite, satiety, energy balance and weight and fat regulation is context dependent i.e. different populations may have different results. In more obesogenic environments *POMC* methylation might be a key factor in weight gain or indeed weight loss.

One group of patients to explore this in future would be those who are obese. The finding that prospectively *POMC* methylation was associated with fat mass change over a year suggests it may have a prognostic indicator for fat change in other scenarios e.g. weight loss programmes. Methylation is already used for precision medicine in fields such as oncology but far less so in areas of metabolic health[70], and understanding how *POMC* methylation could be associated with the response to weight loss interventions would be of great interest and could help guide precision medicine.

Such research would focus on *POMC* methylation and subsequent response to weight loss programmes, pharmacological interventions such as glucagon-like peptide agonists (e.g. Lirglutide) or MC4R agonists (e.g. Setmelanotide), and bariatric surgery. Crujeiras et al identified that lower *POMC* promotor methylation predicted weight loss maintenance[27], offering a proof of principle that *POMC* methylation could be used as a prognostic biomarker for treatment by highlighting specific therapies dependent on methylation state.

8.4.3 Periconceptional intervention

Let one assume that there is sufficient evidence linking DNA methylation with periconceptional nutrition and postnatal phenotypes. This would suggest the possibility of periconceptional nutritional interventions aimed at correcting aberrant methylation patterns with the aim of promoting a more healthy phenotype in the offspring.

The idea that a periconceptional nutrition intervention could be used to improve the health of future offspring is of interest for improving public health and disease susceptability[71]. There is evidence to suggest that DNA methylation in offspring can be manipulated by altering maternal diet during pregnancy. For example, pregnancy micronutrient supplementation has been associated with differential methylation in offspring [72] and maternal folate supplementation in pregnancy has been associated with *PAX8* methylation in adult offspring[41]. Maternal circulating levels of homocysteine has been identified as a key predictor of offspring methylation[8]. James et al, produced a nutritional

supplement that could alter the level of homocysteine in women of child bearing age[73], though there was no assessment of offspring methylation in this study as it concerned non-pregnant women. These studies provide a proof of concept that there is potential to influence offspring methylation by nutritional interventions in pregnancy.

However, there is much to consider before any such intervention is implemented or indeed trialled. Firstly, more research is needed to understand the effect of nutrition on methylation across wider genomic regions and other MEs. The specific nutritional drivers of methylation need to be further characterised to understand the constituents of any nutritional intervention. Additionally, further investigation is required to understand clear phenotypic associations of altered methylation across a range of genes.

There would be much to consider when contemplating an intervention to alter methylation even at just one gene. Take the example of the *PAX8*. In this study, alterations in *PAX8* were associated with thyroid volume and free T4. Changes in free T4 (even in the normal range) were associated with fat and bone mass. Therefore, any intervention that may alter free T4 leads to a phenotypic trade-off between different fat mass and bone density i.e. higher free T4 may lead to lower body fat but also lower bone mineral density. The long term effects of these changes remains unknown but it may equate to a trade-off between poorer cardiometabolic health and risk of fracture. Therefore, any intervention may need to balance a range of phenotypic consequences.

Furthermore, let us propose that on balance we opted to develop a nutritional intervention that lowered *PAX8* methylation. This intervention may have numerous o*ff target effects* i.e. effects on methylation at genes in addition to *PAX8.* The effect of altered methylation at the off target genes may not be known or may be associated with adverse health outcomes. These off target effects need to be characterised. Phenotypic consequences of altering methylation across many genes needs to be assessed in order to make an informed and balanced decision regarding a nutritional intervention.

In summary, any prenatal nutritional intervention requires further research into the effects on offspring methylation and phenotype. However, with prenatal adversity consistently associated with poorer health outcomes for future generations, a focus on understanding the optimal prenatal nutrition and environment is certainly warranted.

8.5 References

1. Barker DJ. The fetal and infant origins of adult disease. *BMJ*. 1990;301(6761):1111. doi:10.1136/bmj.301.6761.1111

2. Barker DJ, Osmond C, Law CM. The intrauterine and early postnatal origins of cardiovascular disease and chronic bronchitis. *J Epidemiol Community Health*. 1989;43(3):237-240. doi:10.1136/jech.43.3.237

3. Lumey LH, Stein AD, Kahn HS, et al. Cohort Profile: The Dutch Hunger Winter Families Study. *Int J Epidemiol*. 2007;36(6):1196-1204. doi:10.1093/ije/dym126

4. Heindel JJ, Vandenberg LN. Developmental origins of health and disease: a paradigm for understanding disease cause and prevention. *Curr Opin Pediatr*. 2015;27(2):248-253. doi:10.1097/MOP.0000000000000191

5. Stevenson K, Lillycrop KA, Silver MJ. Fetal programming and epigenetics. *Curr Opin Endocr Metab Res*. 2020;13:1-6. doi:https://doi.org/10.1016/j.coemr.2020.07.005

6. Waterland RA, Kellermayer R, Laritsky E, et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet*. 2010;6(12):1-10. doi:10.1371/journal.pgen.1001252

7. Dominguez-Salas P, Moore SE, Baker MS, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat Commun*. 2014;5:3746. doi:10.1038/ncomms4746

8. James PT, Dominguez-Salas P, Hennig BJ, Moore SE, Prentice AM, Silver MJ. Maternal One-Carbon Metabolism and Infant DNA Methylation between Contrasting Seasonal Environments: A Case Study from The Gambia. *Curr Dev Nutr*. 2019;3(1). doi:10.1093/cdn/nzy082

9. Silver MJ, Kessler NJ, Hennig BJ, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol*. 2015;16:118. doi:10.1186/s13059-015-0660-y

10. Gunasekara CJ, Scott CA, Laritsky E, et al. A genomic atlas of systemic interindividual epigenetic variation in humans. *Genome Biol*. 2019;20(1):105. doi:10.1186/s13059-019-1708-1

11. Silver MJ, Saffari A, Kessler NJ, et al. Environmentally sensitive hotspots in the methylome of

the early human embryo. *Elife*. 2022;11:e72031. doi:10.7554/eLife.72031

12. Kühnen P, Handke D, Waterland RA, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016;24(3):502-509. doi:10.1016/j.cmet.2016.08.001

13. Dominguez-Salas P, Moore SE, Cole D, et al. DNA methylation potential: Dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am J Clin Nutr*. 2013;97(6):1217-1227. doi:10.3945/ajcn.112.048462

14. Krokan HE, Bjørås M. Base excision repair. *Cold Spring Harb Perspect Biol*. 2013;5(4):a012583-a012583. doi:10.1101/cshperspect.a012583

15. Xue J-H, Xu G-F, Gu T-P, et al. Uracil-DNA Glycosylase UNG Promotes Tet-mediated DNA Demethylation. *J Biol Chem*. 2016;291(2):731-738. doi:10.1074/jbc.M115.693861

16. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. *Trends Genet*. 2002;18(7):348-351. doi:10.1016/S0168-9525(02)02709-9

17. Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. *Sci Adv*. 2018;4(7):eaat2624. doi:10.1126/sciadv.aat2624

18. Howe CG, Cox B, Fore R, et al. Maternal Gestational Diabetes Mellitus and Newborn DNA Methylation: Findings From the Pregnancy and Childhood Epigenetics Consortium. *Diabetes Care*. 2019;16:dc190524. doi:10.2337/dc19-0524

19. Teh AL, Pan H, Chen L, et al. The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Res*. 2014;24(7):1064-1074. doi:10.1101/gr.171439.113

20. Feinberg AP, Irizarry RA. Evolution in health and medicine Sackler colloquium: Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. *Proc Natl Acad Sci U S A*. 2010;107 Suppl(Suppl 1):1757-1764. doi:10.1073/pnas.0906183107

21. Taylor PN, Porcu E, Chew S, et al. Whole-genome sequence-based analysis of thyroid function. *Nat Commun*. 2015;6(1):5681. doi:10.1038/ncomms6681

22. Teumer A, Rawal R, Homuth G, et al. Genome-wide association study identifies four genetic loci associated with thyroid volume and goiter risk. *Am J Hum Genet*. 2011;88(5):664-673. doi:10.1016/j.ajhg.2011.04.015

23. Khera A V., Chaffin M, Wade KH, et al. Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. *Cell*. 2019;177(3):587-596.e9. doi:10.1016/j.cell.2019.03.028

24. Kühnen P, Handke D, Waterland R a, et al. Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab*. 2016:1-8. doi:10.1016/j.cmet.2016.08.001

25. Kuehnen P, Mischke M, Wiegand S, et al. An alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet*. 2012;8(3):1-12. doi:10.1371/journal.pgen.1002543

26. Präger M, Kurz C, Böhm J, Laxy M, Maier W. Using data from online geocoding services for the assessment of environmental obesogenic factors: a feasibility study. *Int J Health Geogr*. 2019;18(1):13. doi:10.1186/s12942-019-0177-9

27. Crujeiras AB, Campion J, Díaz-Lagares A, et al. Association of weight regain with specific methylation levels in the NPY and POMC promoters in leukocytes of obese men: A translational study. *Regul Pept*. 2013;186:1-6. doi:10.1016/j.regpep.2013.06.012

28. Dick KJ, Nelson CP, Tsaprouni L, et al. DNA methylation and body-mass index: a genome-wide analysis. *Lancet*. 2014;383(9933):1990-1998. doi:10.1016/S0140-6736(13)62674-4

29. Demerath EW, Guan W, Grove ML, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum Mol Genet*. 2015;24(15):4464-4479. doi:10.1093/hmg/ddv161

30. Sullivan PF. Spurious genetic associations. *Biol Psychiatry*. 2007;61(10):1121-1126. doi:10.1016/j.biopsych.2006.11.010

31. Relton CL, Davey Smith G. Is epidemiology ready for epigenetics? *Int J Epidemiol*. 2012;41(1):5-9. doi:10.1093/ije/dys006

32. Sharp GC, Relton CL. Epigenetics and noncommunicable diseases. *Epigenomics*. 2017;9(6):789-791. doi:10.2217/epi-2017-0045

33. Sharp GC, Lawlor DA, Richmond RC, et al. Maternal pre-pregnancy BMI and gestational weight gain , offspring DNA methylation and later offspring adiposity : findings from the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2015;44(4):1288-1304. doi:10.1093/ije/dyv042

34. Candler T, Kühnen P, Prentice AM, Silver MJ. Chapter 7 - Twin and family studies on epigenetics and obesity. In: Li S, Hopper JLBT-T and FS of E, eds. *Translational Epigenetics*. Vol 27. Academic Press; 2021:119-147. doi:https://doi.org/10.1016/B978-0-12-820951-6.00012-0

35. Wahl S, Drong A, Lehne B, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature*. 2017;541(7635):81-86. doi:10.1038/nature20784

36. Abu-Khudir R, Magne F, Chanoine JP, Deal C, Van Vliet G, Deladoëy J. Role for tissue-dependent methylation differences in the expression of FOXE1 in nontumoral thyroid glands. *J Clin Endocrinol Metab*. 2014;99(6). doi:10.1210/jc.2013-4414

37. Smallwood SA, Kelsey G. De novo DNA methylation: A germ cell perspective. *Trends Genet*. 2012. doi:10.1016/j.tig.2011.09.004

38. Roseboom T, de Rooij S, Painter R. The Dutch famine and its long-term consequences for adult health. *Early Hum Dev*. 2006;82(8):485-491. doi:10.1016/j.earlhumdev.2006.07.001

39. Roseboom TJ, Painter RC, van Abeelen AFM, Veenendaal MVE, de Rooij SR. Hungry in the womb: what are the consequences? Lessons from the Dutch famine. *Maturitas*. 2011;70(2):141-145. doi:10.1016/j.maturitas.2011.06.017

40. Finer S, Iqbal MS, Lowe R, et al. 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*. 2016;6(11):e011768.

41. Richmond RC, Sharp GC, Herbert G, et al. 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*. 2018;47(3):928-937. doi:10.1093/ije/dyy032

42. Yoo JY, Lee S, Lee HA, et al. Can proopiomelanocortin methylation be used as an early predictor of metabolic syndrome? *Diabetes Care*. 2014;37(3):734-739. doi:10.2337/dc13-1012

43. Perfilyev A, Dahlman I, Gillberg L, et al. Impact of polyunsaturated and saturated fat overfeeding on the DNA-methylation pattern in human adipose tissue: A randomized controlled trial. *Am J Clin Nutr*. 2017;105(4):991-1000. doi:10.3945/ajcn.116.143164

44. Lang J, McKie J, Smith H, et al. Adverse childhood experiences, epigenetics and telomere length variation in childhood and beyond: a systematic review of the literature. *Eur Child Adolesc Psychiatry*. April 2019. doi:10.1007/s00787-019-01329-1

45. Hao G, Youssef NA, Davis CL, Su S. The role of DNA methylation in the association between childhood adversity and cardiometabolic disease. *Int J Cardiol*. 2018;255:168-174. doi:10.1016/j.ijcard.2017.12.063

46. Hecker T, Radtke KM, Hermenau K, Papassotiropoulos A, Elbert T. Associations among child abuse, mental health, and epigenetic modifications in the proopiomelanocortin gene (POMC): A study with children in Tanzania. *Dev Psychopathol*. 2016;28(4):1401-1412. doi:10.1017/S0954579415001248

47. Han L, Zhang H, Kaushal A, et al. Changes in DNA methylation from pre- to post-adolescence are associated with pubertal exposures. *Clin Epigenetics*. 2019;11(1):176. doi:10.1186/s13148-019-0780-4

48. Sharp GC, Lawlor DA, Richardson SS. It's the mother!: How assumptions about the causal primacy of maternal effects influence research on the developmental origins of health and disease. *Soc Sci Med*. 2018;213:20-27. doi:10.1016/j.socscimed.2018.07.035

49. Pembrey ME, Bygren LO, Kaati G, et al. Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet*. 2006;14(2):159-166. doi:10.1038/sj.ejhg.5201538

50. Kaati G, Bygren L, Edvinsson S. Cardiovascular and diabetes mortality determined by nutrition during parents' and grandparents' slow growth period. *Eur J Hum Genet*. 2002;10(11):682-688. doi:10.1038/sj.ejhg.5200859

51. Vågerö D, Pinger PR, Aronsson V, van den Berg GJ. Paternal grandfather's access to food predicts all-cause and cancer mortality in grandsons. *Nat Commun*. 2018. doi:10.1038/s41467-018-07617-9

52. Northstone K, Golding J, Davey Smith G, Miller LL, Pembrey M. Prepubertal start of father's smoking and increased body fat in his sons: further characterisation of paternal transgenerational responses. *Eur J Hum Genet*. 2014;22(12):1382-1386. doi:10.1038/ejhg.2014.31

53. Figueroa-Colon R, Arani RB, Goran MI, Weinsier RL. Paternal body fat is a longitudinal predictor of changes in body fat in premenarcheal girls. *Am J Clin Nutr*. 2000;71(3):829-834. doi:10.1093/ajcn/71.3.829

54. McCowan LM, North RA, Kho EM, et al. Paternal Contribution to Small for Gestational Age Babies: A Multicenter Prospective Study. *Obesity*. 2011;19(5):1035-1039. doi:10.1038/oby.2010.279

55. Chen TH-H, Chiu Y-H, Boucher BJ. Transgenerational effects of betel-quid chewing on the development of the metabolic syndrome in the Keelung Community-based Integrated Screening Program. *Am J Clin Nutr*. 2006;83(3):688-692. doi:10.1093/ajcn.83.3.688

56. Govorko D, Bekdash RA, Zhang C, Sarkar DK. Male germline transmits fetal alcohol adverse effect on hypothalamic proopiomelanocortin gene across generations. *Biol Psychiatry*. 2012;72(5):378-388. doi:10.1016/j.biopsych.2012.04.006

57. Mead EA, Sarkar DK. Fetal alcohol spectrum disorders and their transmission through genetic and epigenetic mechanisms. *Front Genet*. 2014;5(JUN):1-10. doi:10.3389/fgene.2014.00154

58. Sarkar DK. Male germline transmits fetal alcohol epigenetic marks for multiple generations: A review. *Addict Biol*. 2016;21(1):23-34. doi:10.1111/adb.12186

59. Donkin I, Versteyhe S, Ingerslev LR, et al. Obesity and Bariatric Surgery Drive Epigenetic Variation of Spermatozoa in Humans. *Cell Metab*. 2016;23(2):369-378. doi:10.1016/j.cmet.2015.11.004

60. Chen Q, Yan W, Duan E. Epigenetic inheritance of acquired traits through sperm RNAs and sperm RNA modifications. *Nat Rev Genet*. 2016;17(12):733-743. doi:10.1038/nrg.2016.106

61. Kurmann AA, Serra M, Hawkins F, et al. Regeneration of Thyroid Function by Transplantation of Differentiated Pluripotent Stem Cells. *Cell Stem Cell*. 2015;17(5):527-542. doi:10.1016/j.stem.2015.09.004

62. Ma R, Morshed SA, Latif R, Davies TF. TAZ Induction Directs Differentiation of Thyroid Follicular Cells from Human Embryonic Stem Cells. *Thyroid*. 2017;27(2):292-299. doi:10.1089/thy.2016.0264

63. Ma R, Shi R, Morshed SA, Latif R, Davies TF. Derivation and 97% Purification of Human Thyroid Cells From Dermal Fibroblasts . *Front Endocrinol* . 2020;11.

64. Adli M. The CRISPR tool kit for genome editing and beyond. *Nat Commun*. 2018;9(1):1911. doi:10.1038/s41467-018-04252-2

65. Enríquez P. CRISPR-Mediated Epigenome Editing. *Yale J Biol Med*. 2016;89(4):471-486.

66. Urbano A, Smith J, Weeks RJ, Chatterjee A. Gene-Specific Targeting of DNA Methylation in the Mammalian Genome. *Cancers (Basel)*. 2019;11(10). doi:10.3390/cancers11101515

67. Klett M. Epidemiology of congenital hypothyroidism. *Exp Clin Endocrinol diabetes Off journal, Ger Soc Endocrinol [and] Ger Diabetes Assoc*. 1997;105 Suppl:19-23. doi:10.1055/s-0029-

1211926

68. UK T. No Title. https://twinsuk.ac.uk/. Accessed July 9, 2022.

69. Lawlor DA, Lewcock M, Rena-Jones L, et al. The second generation of The Avon Longitudinal Study of Parents and Children (ALSPAC-G2): a cohort profile. *Wellcome open Res*. 2019;4:36. doi:10.12688/wellcomeopenres.15087.2

70. Werner RJ, Kelly AD, Issa J-PJ. Epigenetics and Precision Oncology. *Cancer J*. 2017;23(5):262-269. doi:10.1097/PPO.0000000000000281

71. Remely M, Stefanska B, Lovrecic L, Magnet U, Haslberger AG. Nutriepigenomics: the role of nutrition in epigenetic control of human diseases. *Curr Opin Clin Nutr Metab Care*. 2015;18(4).

72. Saffari A, Shrestha S, Issarapu P, et al. Effect of maternal preconceptional and pregnancy micronutrient interventions on children's DNA methylation: Findings from the EMPHASIS study. *Am J Clin Nutr*. 2020;112(4):1099-1113. doi:10.1093/ajcn/nqaa193

73. James PT, Jawla O, Mohammed NI, et al. A novel nutritional supplement to reduce plasma homocysteine in nonpregnant women: A randomised controlled trial in The Gambia. *PLoS Med*. 2019;16(8):e1002870. doi:10.1371/journal.pmed.1002870

Annexes

Annex 1.1 Supervisors, advisory committee members and collaborators

Name	Position and title	Role in thesis
Dr Matt Silver	Senior Investigator Scientist and Associate Professor MRC The Gambia at London School of Hygiene and Tropical Medicine (MRCG @ LSHTM)	Primary supervisor for my PhD Senior Author on all outputs from the thesis
Professor Andrew Prentice	Professor of International Nutrition and Nutrition Theme Leader MRCG @ LSHTM	Associate supervisor Co-author on all outputs from the thesis
Dr Peter Kühnen	Paediatric Endocrinologist Institute for Experimental Pediatric Endocrinology, Charite Universitatsmedizin Berlin, Germany	Advisory committee member Co-author on outputs related to *POMC* studies Technical advise and help with pyrosequencing
Professor Marian Ludgate	Professor of Molecular Endocrinology Department of Immunology, University of Cardiff	Advisory committee member Co-author on *PAX8* paper
Professor Tim Cole	Professor of Medical Statistics UCL Institute of Child Health	Support and advice with multilevel modelling of growth, weight and adiposity data for *POMC* study
Professor John Blundell	Professor and Research Chair of PsychoBiology University of Leeds	Support and advice with designing the revised appetite and satiety test for *POMC* study
Noah Kessler	Bioinformatics Analyst University of Cambridge	Bioinformatic analysis of genotype/epigenotype for *PAX8* study

Dr Sophie Moore	Senior Lecturer Kings College London / MRCG @ LSHTM	Established the ENID trial and longitudinal cohort studies
Professor Ann Prentice	Programme Leader of the MRC Nutrition and Bone Health (NBH) Research Group, Cambridge, University of Cambridge Head of the Calcium, Vitamin D and Bone Health (CDBH) research team at MRCG@LSHTM	Coordinated training on DXA scans and scan scrutiny Facilitated use of DXA scans for both *POMC* and *PAX8* studies
Dr Kate Ward	Associate Professor of Global Musculo-skeletal heath University of Southampton	Support and advice on analysis of DXA outcomes Co-author on *PAX8* paper
Dr Landing Jarjou	Senior Investigator Scientist CDBH subtheme lead MRC Unit, The Gambia @ LSHTM	Facilitated use of DXA scans for both *POMC* and *PAX8* studies Provided radiation safety training
Professor Robert Waterland	Professor of Epigenetics Baylor College of Medicine Houston, USA	Collaborated with the *PAX8* study to use GTEx samples for methylation-expression analysis
Dr Abdul Karim Cessay	Head of Genomics MRCG@LSHTM	Line manager during my work in the genomics department
Abdoulie Kanteh	Scientific Officer MRCG@LSHTM	Trained me in basic lab skills, DNA QC and quantification, bisulfite conversion, PCR and worked together on pyrosequencing optimisation for *POMC* study
Ebrima Bah	Scientific Officer MRCG@LSHTM	Completed *POMC* PCR after I returned to UK to return to

		clinical work to support COVID-19 pandemic
Lamin Sanyang Kebba Bajo	POMC and PAX8 study Field supervisors	Led East and West WK teams, helped coordinate study visits, Supervised field workers and field activity
Mustapha Saidy Wally Camara Foday Bah Alasan Sey Famara Barjo Isatou Jammeh Fekebba Camara Lamin Jarju Ebrima K Cessay	POMC and PAX8 study Field workers	Implemented field work
Tida Sammateh	POMC study cook for appetite tests and later field worker	Cook for appetite testing
Edrisa Sinjanka Mustapha Saidy Sainey Beyai	POMC and PAX8 study nurses	Performed blood draws and collected urine samples for study participants
Ebrima Danso	Scientific Officer Keneba Laboratory MRCG@LSHTM	Trained me an ELISA for leptin analysis. Supported and advised regarding laboratory processes for study samples and DNA extraction
Bakary Sonko	Data manager MRCG@LSHTM	Supported and advised regarding study database setup and management and REDCap interface.

Activity	Dec-18	Apr-18	Jul- 18	Oct-18	Jan-19	Apr-19	Jul - 19	Oct-19	Jan-20	Apr-20 NHS	Jul-20 NHS	Oct-20	Jan-21	Apr-21	Jul-21	Oct-21 NHS	Jan-22 NHS
***POMC* Study**																	
Literature Review																	
Review article writing																	
Review article submission																	
MRCG/LSHTM SCC & EC	MRCG				LSHTM												
Study Baseline																	
Appetite test redesign																	
Study Midline																	
Study Endline																	
Pyrosequencing testing																	
Pyrosequencing *POMC* samples																	
Leptin ELISA																	
DXA scan scrutiny																	
Import DXA data																	
Data cleaning																	
Data Analysis and modelling																	
Paper writing																	
***PAX-8* Study**																	
Grant submission and award																	
MRCG/LSHTM SCC & EC																	
Ultrasound training																	
Identifying participants																	
Study activity																	
Pyrosequencing *PAX-8*																	

samples																	
Samples to Cardiff for thyroid measures																	
Thyroid results from Cardiff																	
Data analysis																	
Paper writing																	
PhD milestones and training																	
Upgrading																	
Thesis writing																	
Epigenetic Epidemiology Course - Bristol											Cancelled due to COVID						
Longitudinal Modelling Course - Bristol											Cancelled due to COVID						
R Course – online & face to face MRCG@LSHTM					F2F	Online	Online	Online	Online								
Submission of thesis																	
Activity	**Dec-18**	**Apr-18**	**Jul- 18**	**Oct-18**	**Jan-19**	**Apr-19**	**Jul - 19**	**Oct-19**	**Jan-20**	**Apr-20**	**Jul-20**	**Oct-20**	**Jan-21**	**Apr-21**	**Jul-21**	**Oct-21**	

Annex 1.2 PhD Timeline

Annex 2.1 POMC study ethics approval

MRC Unit The Gambia @ LSHTM, Fajara
P.O. Box 273, Banjul
The Gambia, West Africa
Fax: +220 – 4495919 or 449951
Tel: +220 – 4495442-6 Ext: 2309
Email: ethics@mrc.gm

The Gambia Government/MRC Joint
ETHICS COMMITTEE

16 March 2018

Dr Matt Silver
MRCG at LSHTM
Keneba

Dear Dr Silver

SCC 1588v1.1, Inter-generational risk factors for obesity: a path to prevention in low and middle-income countries based on a modifiable epigenetic signature in the POMC gene

Thank you for submitting your proposal dated 14 February 2018 for consideration by The Gambia Government/MRC Joint Ethics Committee at its meeting held on 9 March 2018.

The committee is pleased to approve your proposed study.

With best wishes

Yours sincerely

[redacted]
Mr Malamin Soniko
Chairman, Gambia Government/MRC Joint Ethics Committee

Documents submitted for review:

- SCC Approval letter – 16 February 2018
- SCC application form, version 1.1 – 14 February 2018
- Response letter – 14 February 2018
- ICD (Child), version 1.1 – 14 February 2018
- ICD (Child with DEXA scans), version 1.1 – 14 February 2018
- ICD (Father), version 1.0 – 14 February 2018
- ICD (Mother), version 1.1– 14 February 2018
- ICD (Mother with DEXA scan), version 1.1 – 14 February 2018
- CV – Toby Philip Candler

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636

www.lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Observational / Interventions Research Ethics Committee

Dr Toby Candler
LSHTM

26 February 2019

Dear Toby

Submission Title: Inter-generational risk factors for obesity: a path to prevention in low and middle-income countries based on a modifiable epigenetic signature in the POMC gene

LSHTM Ethics Ref: 16439

Thank you for responding to the Observational Committee Chair's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved is as follows:

Document Type	File Name	Date	Version
Local Approval	SCC_ Approval_L2010 77_Darboe_11Oct10	11/10/2010	2
Local Approval	Ethics Approval_L2010 77_Darboe_3Nov10	03/11/2010	2
Consent form	ENID-Growth_SCCL2010.77-Information Sheet and Consent form -V 02 -04-12-2011	04/12/2011	2
Consent form	Gong_SCC_140613	14/06/2013	1
Local Approval	POMC SCC Application_accepted by SCC	22/01/2018	1
Local Approval	SCC 1588_Silver_Approved (pending clarifications)_12Feb18 (1)	12/02/2018	1
Local Approval	POMC_Response to SCC 14.2.18	14/02/2018	1
Local Approval	POMC ICD Child SCC 140218	14/02/2018	1
Local Approval	POMC ICD Mother SCC 140218	14/02/2018	1
Local Approval	POMC ICD Fathers SCC 140218	14/02/2018	1
Local Approval	POMC ICD Child with DEXA SCC 140218	14/02/2018	1
Local Approval	POMC ICD Mothers with DEXA scan SCC 140218	14/02/2018	1
Information Sheet	POMC ICD Child SCC 140218	14/02/2018	1
Information Sheet	POMC ICD Fathers SCC 140218	14/02/2018	1
Information Sheet	POMC ICD Mother SCC 140218	14/02/2018	1
Information Sheet	POMC ICD Mothers with DEXA scan SCC 140218	14/02/2018	1
Information Sheet	POMC ICD Child with DEXA SCC 140218	14/02/2018	1
Local Approval	SCC 1588v1.1_Silver (Approved) (1)	16/03/2018	1
Protocol / Proposal	POMC Study_FieldManual_v3	01/04/2018	1
Consent form	POMC ICD Child SCC 140218	01/04/2018	2
Consent form	POMC ICD Fathers SCC 140218	01/04/2018	1
Consent form	POMC ICD Mother SCC 140218	01/04/2018	1
Consent form	POMC ICD Child with DEXA SCC 140218	01/04/2018	1

Annex 2.2 PAX8 study ethics approval

The Gambia Government/MRCG Joint
ETHICS COMMITTEE

C/o MRC Unit: The Gambia @ LSHTM, Fajara
P.O. Box 273, Banjul
The Gambia, West Africa
Fax: +220 – 4495919 or 4496513
Tel: +220 – 4495442-6 Ext. 2308
Email: ethics@mrc.gm

5 December 2018

Dr. Toby Candler,
MRCG at LSHTM,
Keneba Field Station.

Dear Dr. Candler,

SCC 1640v1.1, Epigenetic Regulation Of Thyroid Function: Examining The Potentially Modifiable Effect Of Differential PAX-8 Methylation On Thyroid Function In Children In The Gambia

Thank you for submitting your proposal dated 12 November 2018 for consideration by the Gambia Government/MRCG Joint Ethics Committee at its meeting held on 22 November 2018.

Our Committee is pleased to approve your proposed study. However, you are advised to indictate in the SCC form (page 11, section B4 c) that treatment will be provided if abnormal thyroid function is detected.

With best wishes,

Yours sincerely,

[redacted]
Dr. Mohammadou Kabir Cham
Chair, Gambia Government/MRCG Joint Ethics Committee

Documents submitted for review:

- Response letter – 12 November 2018
- SCC Application form, version 1.1 – 12 November 2018
- SCC Approval letter – 14 November 2018
- SCC Reply letter – 8 November 2018
- ICD (Child), version 1.0 – 22 October 2018

The Gambia Government/MRCG Joint Ethics Committee:

Dr Mohammadou Kabir Cham, Chair
Prof Ousman Nyan, Scientific Advisor
Dr Kalifa Bojang
Dr Ahmadou Lamin Samateh
Dr Pamela Esanghedo
Dr Jane Achan
Rev Gabriel L. Allen
Prof Umberto D'Alessandro
Dr Mamady Cham
Mr Momodou YM Sallah
Prof Martin Antonio
Dr Assan Jaye
Ms Naffie Jobe, Secretary

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT
United Kingdom
Switchboard: +44 (0)20 7636 8636
www.lshtm.ac.uk

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Observational / Interventions Research Ethics Committee

Dr Toby Candler
LSHTM

26 February 2019

Dear Toby

Submission Title: Epigenetic regulation of thyroid function: examining the potentially modifiable effect of differential PAX-8 methylation on thyroid function in children in The Gambia

LSHTM Ethics Ref: 16280

Thank you for responding to the Observational Committee Chair's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

Approval is dependent on local ethical approval having been received, where relevant.

Approved documents

The final list of documents reviewed and approved is as follows:

Document Type	File Name	Date	Version
Local Approval	SCC_ Approval_ENIDL2010 77_Darboe_11Oct10	11/10/2010	2
Local Approval	Ethics Approval_ENIDL2010 77_Darboe_3Nov10	03/11/2010	2
Consent form	ENID-Growth_SCCL2010.77-Information Sheet and Consent form -V 02 -04-12-2011	04/12/2011	2
Consent form	Gong_SCC_140613 (1)	14/06/2013	1
Local Approval	POMC SCC Application_accepted by SCC	22/01/2018	1
Local Approval	SCC 1588_Silver_Approved (pending clarifications)_12Feb18 (1)	12/02/2018	1
Local Approval	POMC_Response to SCC 14.2.18	14/02/2018	1
Consent form	POMC ICD Child SCC 140218	14/02/2018	1
Consent form	POMC ICD Child with DEXA SCC 140218	14/02/2018	1
Local Approval	POMC_Response to SCC 14.2.18	14/02/2018	1
Local Approval	SCC 1588v1.1_Silver (Approved)	16/03/2018	1
Local Approval	SCC 1588v1.1_Silver (Approved)	16/03/2018	1.1
Protocol / Proposal	PAX-8 ICD for submission	22/10/2018	1
Information Sheet	PAX-8 ICD for submission	22/10/2018	1.0
Local Approval	SCC 1640_Candler_Approved (pending clarifications)_8Nov18 signed (1)	08/11/2018	1
Local Approval	PAX8_SCCresponse12112018	12/11/2018	1
Local Approval	SCC 1640v1.1_Candler_Approved (with comment)_5Dec18_signed (2)	05/12/2018	1
Protocol / Proposal	SCC Application form PAX-8 for submission_revisedSCC_revisedEC_V2	13/12/2018	2
Investigator CV	Toby Candler CV December 2018	13/12/2018	1
Local Approval	SCC Application form PAX-8 for submission_revisedSCC_revisedEC_V2	13/12/2018	2
Investigator CV	Indgate cv Dec 2018	17/12/2018	1
Protocol /	POMC Study_FieldManual_v4	30/01/2019	4

Proposal			
Covering Letter	LSHTM ethics response_POMC&PAX-8	14/02/2019	1

After ethical review

The Chief Investigator (CI) or delegate is responsible for informing the ethics committee of any subsequent changes to the application. These must be submitted to the committee for review using an Amendment form. Amendments must not be initiated before receipt of written favourable opinion from the committee.

The CI or delegate is also required to notify the ethics committee of any protocol violations and/or Suspected Unexpected Serious Adverse Reactions (SUSARs) which occur during the project by submitting a Serious Adverse Event form.

An annual report should be submitted to the committee using an Annual Report form on the anniversary of the approval of the study during the lifetime of the study

At the end of the study, the CI or delegate must notify the committee using the End of Study form.

All aforementioned forms are available on the ethics online applications website and can only be submitted to the committee via the website at: http://leo.lshtm.ac.uk

Further information is available at: www.lshtm.ac.uk/ethics.

Yours sincerely,

Professor John DH Porter
Chair

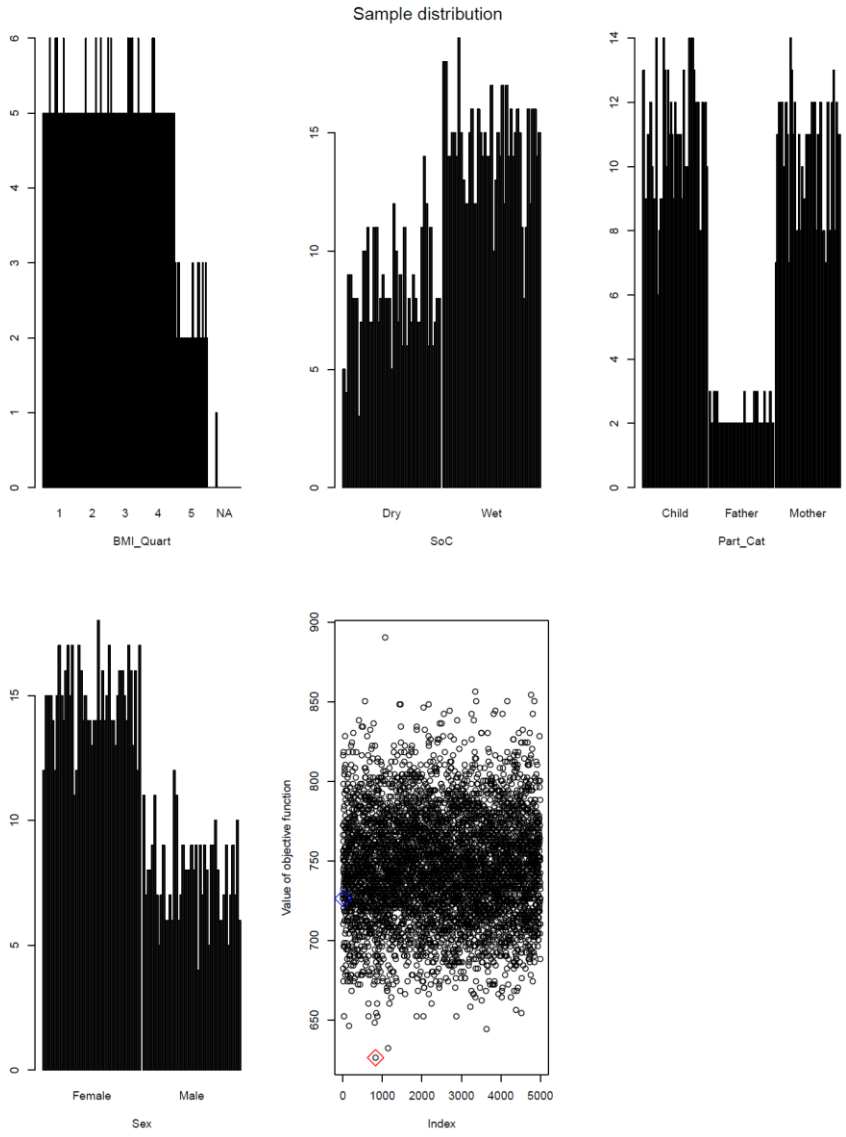
ethics@lshtm.ac.uk
http://www.lshtm.ac.uk/ethics/

Improving health worldwide

Annex 4.1 OSAT PCR plate randomisation

Variable	X-squared	df	p.value
1 BMI Quartile	48.96	220	1.00
2 Season of conception	41.41	44	0.58
3 Participant Category	39.45	88	0.99
4 Sex	25.49	44	0.99

No significant differences across the plates between BMI quartile, season of conception, participant category nor sex.



Annex 4.2 POMC PCR and sequencing primers

hg19_dna range=chr2:25384508-25384832 strand = -

Primers listed 5'-3'

Forward primer:

chr2:25384811-25384832

GTGGTAAGATTTTAGATGTTTA

Reverse primer:

chr2:25384508-25384529

Btn**AAAATAACCCATAACRTACTTC**

Sequencing primer (CpG -1 to +7):

chr2:25384637-25384653

GGTTGTTTTTATGTTTT

GTGGCAAGATCTTAGATGCCCACGAGTGCCAAGAAAGCAGGTGGGCAGAC
CTGCTCTGTAGGGAGGCCTCGACGCTTGACACGCCCGACACTGTGCCCTG
TGTCCTCGGCACGTGGCGAGGGCGGCCAGGGCCTAGGCGCAGTGACGGGC
GCGGCAGCCGGGCCGGGGTGCGGGGCACG**GGCTGCCCTCATGCCCT**CGCG
TCTTCCCCCAGGAGTGCATCCGGGCCTGCAAGCCCGACCTCTCGGCCGAG
ACTCCCATGTTCCCGGGAAATGGCGACGAGCAGCCTCTGACCGAGAACCC
CCG**GAAGTACGTCATGGGCCACTTC**
(Amplicon length - 325 bp)

Annex 4.3 POMC methylation by CpG and the association with obesity and periconceptional nutritional factors

CpG	h38 coordinates	Significant association with obesity? (Kuhnen, 2012[7] and 2016[21])	Relationship with key one carbon metabolites (Kuhnen et al, 2016[21])								
			SAH			SAM:SAH ratio			Betaine		
			Coeff	z	p	Coeff	z	p	Coeff	z	p
-2	chr2:25161767	N	-0.29	-2.99	**0.003**	0.21	2.11	**0.036**	0.04	0.46	0.647
-1	chr2:25161765	N	-0.33	-2.95	**0.004**	0.23	2.08	**0.039**	0.07	0.64	0.525
+1	chr2:25161748	Y	-0.41	-3.58	**< 0.001**	0.33	2.81	**0.006**	0.10	0.89	0.378
+2	chr2:25161743	Y	-0.32	-2.30	**0.023**	0.24	1.66	0.099	0.34	2.63	**0.010**
+3	chr2:25161740	Y	-0.34	-2.76	**0.007**	0.28	2.23	**0.028**	0.32	2.69	**0.008**
+4	chr2:25161736	N	-0.32	-2.63	**0.009**	0.23	1.83	0.070	0.35	3.03	**0.003**
+5	chr2:25161732	Y	-0.59	-3.58	**< 0.001**	0.58	3.50	**0.001**	0.28	1.72	0.088
+6	chr2:25161729	N	-0.40	-3.34	**0.001**	0.35	2.92	**0.004**	0.37	3.23	**0.002**
+7	chr2:25161721	N	-0.32	-2.00	0.048	0.25	1.58	0.117	0.16	1.09	0.278

Key: SAM = S-adenosyl methionine, SAH = S-adenosylhomocysteine, Chr=Chromosome

Annex 5.1 Baseline Ad libitum test protocol

Ad libitum breakfast

After their blood is taken they should be taken to the study area next to the CDBH building and provided with a breakfast of Tiakere Churo.

Tiakere Churo will be prepared by a set recipe by the cook to the following recipe:

Rice 705g, Groundnut 665g, Sugar 555g and Milk 865g by the cook.

The field worker will ask if the participant has any allergies prior to starting the test and when the last time they ate anything (this information is recorded on the form).

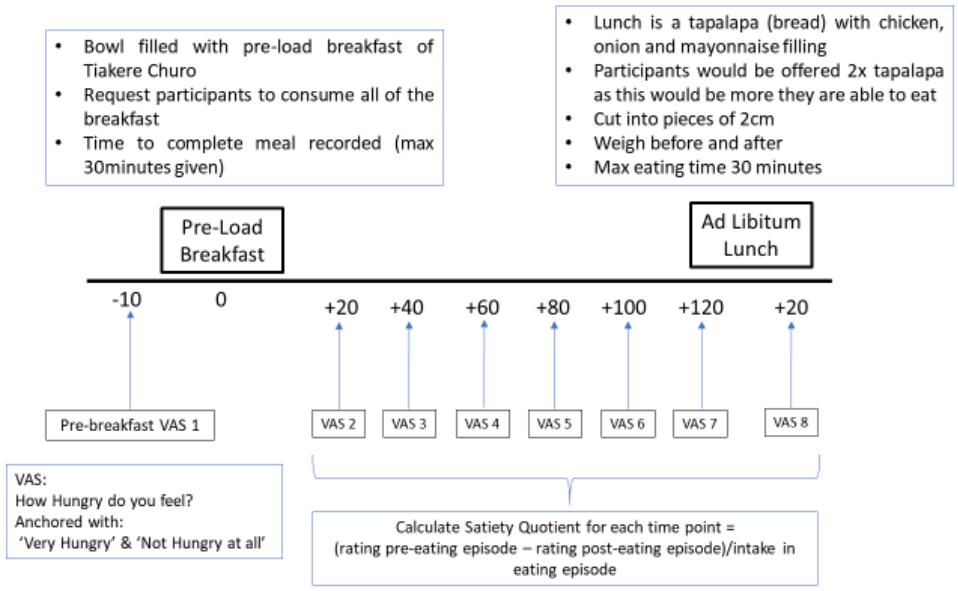
Check the recipe to see if the participant allergic to any ingredients.

1. The cook will provide the Tiakere Churo by 8am on the day of testing
2. The field worker puts the breakfast into the specific *POMC* bowls and fills up to the fill level line.
3. The bowl with the breakfast in is weighed, and the weight recorded on the form.
4. Children are shown the visual analogue scale (VAS). The participant asked to score "how hungry are you?" as per the VAS (picture on far left 1 (not at all) to picture on far right 5 (extremely hungry)). The result is recorded in the form.
5. The bowl is given to the participant and they invited to eat as much as they wish, they are to be informed that if they finish the bowl it will be refilled until they have eaten what the amount they would like.
6. The participant is asked that when the mother or child has finished eating and wants no more food they should inform the field worker.
7. A stop clock will be given to each mother and child pair and started when they commence eating.
8. When the participant has indicated that they have stopped eating and want no more food, the field worker will record the following information on the form
    - the time taken to eat should be recorded
    - the bowl weighed
9. After 5 minutes, the participant should be offered the bowl again, if they want to eat more then the bowl is refilled, weighed and given to the participant again.
10. Ten minutes after finishing eating, the VAS is shown to the children and the result recorded.

    N.B. If during the test the bowl is completed, and the participant wants more then the following should be done and recorded on the form

    - The empty bowl is weighed, and weight recorded
    - The bowl is refilled to the line again with Tiakere Churo
    - The bowl with the breakfast in is weighed
    - The refilled bowl is given to the participant (and repeat instructions for point 5 and 6).

Annex 5.2 Midline and Endline revised appetite test protocol



Outline of appetite and satiety test midline and endline

After their blood is taken the participant will be taken to the study room 4 and 5 (this is a private study room where 2 tables with 4 chairs each are arranged in each room) for the appetite test. The field worker will ask if the participant has any allergies prior to starting the test and when the last time they ate anything (this information is recorded on the form) check the recipe to see if the participant allergic to any ingredients.

The preload breakfast of Tiakere Churo is made with the following recipe; Rice 705g, Groundnut 665g, Sugar 555g and Milk 865g by the cook.

When the participant attends the appetite testing area, their height and weight is entered into the tablets and the amount of Tiakere Churo to give each participant is calculated. The bowls are weighed on electronic scales to the nearest gram.

Due to logistical considerations mothers and children are sat in pairs 4 pairs in each room.

The preload bowl weight (with the breakfast in) is inputted into the electronic CRF with the lid on.

Participants are shown the visual analogue scale (VAS) appropriate for age group and gender (preloaded into electronic CRF).

The participant is asked 'how hungry do you feel?' and they input a score as per the VAS (picture on far left 1 (very hungry) to picture on far right 5 (not hungry at all)). The result is recorded in the form.

The bowl is given to the participant and they are asked to finish the breakfast served. They are told they have up to 30 minutes to eat, after which the bowl will be removed.

The participant is told that when the mother or child has finished eating and wants no more food they should inform the field worker (a field worker is present in the room to observe when the participant has finished eating).

A stop clock will be given to each mother and child pair and started when they commence eating.

When the participant has indicated that they have stopped eating and want no more food, the field worker will record the following information on the form

- the time taken to eat should be recorded
- the bowl weighed empty

VAS is recorded every 20 minutes until 2 hours post breakfast until the ad libitum lunch.

The recipe for the ad libitum lunch is a standard sized tapalapa with 100g of filling. Each participant is given 2 tapalapas. The tapalapa is cut into 2 cm pieces and put in the bowl.

The tapalapa filling is made up with 3x 198g tins of Chicken, 768g of onion, 1020g of mayonnaise. Then 100g of this filling goes into each tapalapa. The participant is invited to eat as much as they wish. The participants are told they have up to 30 minutes to eat, after which the remaining food will be removed and weighed.

The field worker weighs the remaining food and records this in the CRF.

VAS is recorded after 20 minutes after lunch.